DAGSTUHL REPORTS

Volume 12, Issue 5, May 2022

Dagstuhl Reports, Vol. 12, Issue 5 ISSN 2192-5283

ISSN 2192-5283

Published online and open access by
Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH, Dagstuhl Publishing, Saarbrücken/Wadern, Germany. Online available at
https://www.dagstuhl.de/dagpub/2192-5283

Publication date
December, 2022

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at `https://dnb.d-nb.de`.

Digital Object Identifier: 10.4230/DagRep.12.5.i

Aims and Scope
The periodical *Dagstuhl Reports* documents the program and the results of Dagstuhl Seminars and Dagstuhl Perspectives Workshops.
In principal, for each Dagstuhl Seminar or Dagstuhl Perspectives Workshop a report is published that contains the following:
- an executive summary of the seminar program and the fundamental results,
- an overview of the talks given during the seminar (summarized as talk abstracts), and
- summaries from working groups (if applicable).

This basic framework can be extended by suitable contributions that are related to the program of the seminar, e. g. summaries from panel discussions or open problem sessions.

Contact
Schloss Dagstuhl – Leibniz-Zentrum für Informatik
Dagstuhl Reports, Editorial Office
Oktavie-Allee, 66687 Wadern, Germany
reports@dagstuhl.de
https://www.dagstuhl.de/dagrep

Report from Dagstuhl Seminar 22181

Computational Metabolomics: From Spectra to Knowledge

Corey Broeckling*[1], Timothy Ebbels*[2], Ewy Mathé*[3], Nicola Zamboni*[4], and Cecilia Wieder†[5]

1 **Colorado State University – Fort Collins, US.** corey.broeckling@colostate.edu
2 **Imperial College London, GB.** t.ebbels@imperial.ac.uk
3 **National Institutes of Health – Bethesda, US.** ewy.mathe@nih.gov
4 **ETH Zürich, CH.** zamboni@imsb.biol.ethz.ch
5 **Imperial College London, UK.** cecilia.wieder19@imperial.ac.uk

Abstract

The fourth edition of the Computational Metabolomics seminars, Dagstuhl Seminar 22181, brought together a wide range of computational and experimental experts to share state-of-the art methodologies and push our collective understanding of how to interpret and maximise insight of metabolomic data. With increasing amounts of metabolomic data being generated, including large-scale epidemiological studies, and increasing sensitivity of instrumentation, development of sophisticated and robust computational solutions is required. Further, community agreement on which data standards should be used and which data sets are most apt for benchmarking computational tools is needed in the field. Building upon the previous successful formats of previous seminars (17491, 15492, and 20051) on this topic, attendees gathered each morning to collectively agree on the number of sessions and topics to discuss. A summary of the daily sessions were shared amongst all participants after dinner during each day's final formal session. Further, informal evening sessions were spontaneously created to further dive into specific topics. As with past seminars, this format was very well received and enabled all participants to weigh in. Of particular note, this seminar was delayed and travel was complicated due to the pandemic. Despite these setbacks, this seminar brought together a balanced number of previous and new, seasoned and early career participants. All participants were active in these discussions, and a true sense of renewed energy ensued from the seminar. This report provides highlights of formal and informal evening sessions, including future anticipated research directions rooted from this seminar. Possible future workshops, such as a next phase of this Computational Metabolomics Dagstuhl seminar in late 2023 or 2024 were also discussed and will be applied for.

Seminar May 1–6, 2022 – http://www.dagstuhl.de/22181

2012 ACM Subject Classification Applied computing → Life and medical sciences

Keywords and phrases bioinformatics, cheminformatics, computational mass spectrometry, metabolite identification, computational metabolomics, machine learning, data integration, pathway analysis

Digital Object Identifier 10.4230/DagRep.12.5.1

* Editor / Organizer
† Editorial Assistant / Collector

1 Executive Summary

Corey Broeckling (Colorado State University – Fort Collins, US
Timothy Ebbels (Imperial College London, GB)
Ewy Mathé (National Institutes of Health – Bethesda, US)
Nicola Zamboni (ETH Zürich, CH)

Metabolomics is the study of small molecules in living systems, including those which generate the energy to sustain life, those that form the building blocks of macromolecules such as DNA, as well as some originating outside the living system such as pollutants. Biologically, this field is of increasing importance due to its strong connection to organism function. Metabolomics is rapidly expanding with significant advances in both measurement technology (e.g. mass spectrometry, chromatography, NMR spectroscopy) and informatics approaches. The amount and complexity of data routinely exceeds the capacity of typical software and other computational systems used in bioanalytical labs and there is an ongoing and increasingly acute need for improvements in computational, informatics and statistical/machine learning approaches to make sense of it all.

This seminar, the fourth in the series on computational metabolomics, continued some themes previously well developed, and explored many new ones. A good example of the former is the problem of how to use mass spectral data to annotate (putatively identify) the 1000s of unknown metabolites typically observed in routine assays. Another example would be the discussion of new developments in dealing with Data Independent Acquisition which has diversified considerably in the last 5 years. Many new directions were also discussed. For instance, the question of pathway analysis – how to generate semi-automated interpretation of metabolomics data on the level of groups of molecules working together in biological processes – is becoming more prominent as larger annotated datasets become available. Another new direction was "metaboproteomics", looking at the diverse array of interactions between metabolites and proteins, in particular in how metabolite derived post-translational modifications of proteins can be picked up in annotation pipelines. Other discussions focused on software aspects such as visualization of chemical space (a key problem in designing effective software tools) and the generation/curation of high quality data for benchmarking new informatics algorithms. A session on extended metabolic models looked at ways to link data and prediction tools from protein function studies to metabolites in order to gain new knowledge of unknown metabolic pathways. From a data generation technology perspective, while mass spectrometry (MS) dominated as expected (e.g. sessions on MS spectral quality requirements, fragmentation trees etc), the seminar extended beyond previous ones into a discussion of NMR data processing and modeling. Open databases, repositories and knowledge representation also featured their own discussions including Wikidata, CxSMILEs and Wikipathways/RaMP-DB. Finally the important issue of integrating metabolomic data with other relevant data types (e.g. genomics, proteomics etc) was discussed.

The seminar organization followed a similar flexible format to the previous one, where topics were both suggested in advance and brainstormed on the Monday. The whole group participated in brainstorming and prioritization and this was further refined each morning of the meeting. Parallel discussions were organized with the aim to minimize clashes in individual interests and at the end of each morning/afternoon session a plenary feedback session was held to disseminate the main discussion points to the whole group. Evening sessions were very popular and covered a wide range of topics. Overall the seminar was felt to be one of the most successful yet, highlighting the growing importance of computational metabolomics as a field in its own right and emphasizing the need for further meetings to address the important problems in this exciting area of research.

2 Table of Contents

3 Overview of Talks

3.1 How to use unlabeled mass spectra for machine learning

Sebastian Böcker (Universität Jena, DE)

The chemical space of molecules of biological interest is very large, and we have the problem that labelled datasets only span a fraction of this space; but more importantly regions of this space are not covered at all. The current trend is to use transformers for this purpose, and there are already several tools available. We discussed approaches that were brought up in other areas (natural language processing) but we also discussed the importance of not repeating what other fields have done. We discussed using fragmentation trees to a) annotate peaks, allowing us to work with molecular formulas instead of masses, b) data augmentation, removing subtrees, and c) for generating decoys, we also discussed different representations of molecular structures, and the merits and dangers of providing information on collision energies.

3.2 Pathway analysis in metabolomics

Timothy Ebbels (Imperial College London, UK)

The session started by collecting the different definitions of network and pathway analysis. One of the approaches is statistical analysis of sets of metabolites against a (difficult-to-come-by) background distribution, noting that a reference metabolome does not exist. Generally, a pathway is defined as a set/collection of molecules, and statistical analyses include over-representation analysis, set enrichment analysis, and topology-based approaches. The input to these approaches is typically two-fold: 1) molecules of interest, and optionally a metric that defines their relevance (e.g. p-value, effect size); 2) prior annotations (e.g. biological, chemical) from knowledge sources associated with these molecules. The session focused on over-representation and set enrichment-based approaches.

Several broad categories of challenges/questions were consistently brought up: 1) ability to apply broad methods, including those developed in genomics, to metabolomics without tailoring to specific peculiarities of metabolomics data; 2) metabolite identification and annotation (structural resolution); 3) ability to map metabolites to pathways and coverage of metabolite annotations; 4) ability to develop tools that are not strongly affected by missing annotations and different levels of uncertainty; 5) difficulty in defining a common language across communities and how this affects the utility of knowledge sources across different fields (e.g. human diseases, natural products, etc.). Further, the notion of single-sample pathway analysis was only mentioned briefly, but could hold great potential to move from univariate biomarker-type analysis towards more biochemically informed metabolic state analysis.

Despite the title "Pathway analysis in metabolomics/lipidomics and multi-omics" the session focused on the pathway and network analysis, and the multi-omics aspect was left for a later occasion.

3.3 Benchmarking data for metabolomics

Ewy Mathé (National Institutes of Health – Bethesda, US)

Reference samples are critical to evaluate reproducibility of metabolomic measurements across different platforms or laboratories. Publicly available reference datasets are useful to develop, test, and benchmark new methods and tools. The goal of this informal evening session was to survey the various reference samples and datasets that are used by the metabolomics community and to identify gaps in availability of these samples/datasets.

Several key reference samples include NIST 1950, COMETS reference samples, and MetQual, that are used to evaluate comparability of metabolite measurements across different groups/labs. Other efforts, e.g. Metabolomics Workbench and CASMI provide available key datasets that can be used to develop algorithms, including those that identify metabolites. Synthetic and simulated datasets are also being used.

There was general consensus that it is difficult to find useful benchmarking datasets, and that benchmarking data used to develop algorithms often do not publish the benchmarking data or the process used to generate the data. Possible solutions here include appropriate tags to datasets so that they could be searchable, or to create a repo of benchmarking datasets. Also the consensus is that more benchmarking datasets should be made available to cover various use cases (e.g. time-series, categorical/continuous outcomes, identification, normalization methodologies), and these should include commonly used platforms. Other gaps identified in this area include the scarcity of resources that bridge multi-omic data and the creation of a DREAM-like challenge.

3.4 Extended metabolic models

Juho Rousu (Aalto University, FI)

The aim of the session was to discuss the ways to link data and prediction tools from protein (enzyme) function studies and metabolomics in order to gain new knowledge of unknown metabolic pathways. Integration of enzyme function data with MS/MS and NMR measurements from untargeted metabolomics experiments is recognized as a field of great potential, for example, for enhancing structural annotation accuracy of unknown metabolites. Given the complexity of the overall task, it is important to identify subtasks that are solvable given the current state-of-the-art.

The discussions covered several aspects that make the prediction of enzymatic reactions in a fine-grained resolution challenging and approaches to overcome these aspects. In particular, the topology of enzyme function spaces is challenging due to the activity cliffs that make interpolation in the function space hard. Other key challenges include the scarcity and poor findability of data on the substrate specificity, promiscuity, and mode of action of enzymes as well as the lack of suitable representations of promiscuous enzyme function. In addition, the current deep learning approaches are likely not sufficient to cover the needs of enzyme function prediction at the fine-grained level. It was concluded that work is required on several fronts. First, writing a review article on the state of the art computational approaches (especially deep learning), on enzyme function prediction in metabolomics was seen as important. The review should aim to provide a common vocabulary and collect information on data and software availability. Secondly, development of graph neural networks tailored to the enzyme-function prediction task, was identified as an important task. Thirdly, efforts will be made to collect useful software packages for tackling the problem.

3.5 Quality control in untargeted metabolomics

María Eugenia Monge (CIBION – Buenos Aires, AR)

The session covered various aspects and levels of quality control (QC) along the metabolomics workflow. A number of QC metrics can be calculated on the raw spectral data files. Several implementations exist, and can be calculated retrospectively, but also during the measurement process. This allows an instrumental- and lab health dashboard visualization and alerting, and is a first step to having a digital twin for the physical setup. Good experimental design includes multiple runs of standard reference materials and pooled sample QC. These are commonly used in the data analysis stages at the end, to ensure and sometimes correct data in the matrix prior or as part of the statistical analysis.

3.6 NMR computational approaches in Metabolomics

Panteleimon Takis (Imperial College London, UK)

NMR spectroscopy is one of the dominant analytical approaches for the deconvolution of complex matrices such as biofluids, which is widely used in the field of metabolomics with the caveat of low sensitivity. Consequently, there are plenty of challenges to overcome with computational approaches to interpret and extract as much metabolite information as possible from the 1H-NMR profile of complex mixtures. In this session, a detailed introduction to NMR and its application to biofluids was discussed. In addition, a brief explanation of what could be observed in a typical 1H-NMR profile of serum/plasma samples was discussed along with the basics of NMR.

3.7 Wikidata: empowering metabolomics research

Egon Willighagen (Maastricht University, NL)
Adriano Rutz (University of Geneva, CH)
Ewy Mathé (National Institutes of Health – Bethesda, US)

Joint work of Egon Willighagen, Adriano Rutz, Ewy Mathé

Wikidata (https://wikidata.org/) was founded in 2012 as a platform for collaborative data collection [1]. Similar to other wikis like Wikipedia, it crowdsources knowledge, but unlike Wikipedia, this new project is data-driven not narrative-driven. Another difference is the liberal license/waiver, CCZero, which mimics the concept of public domain. This allows many research domains across the world to adopt Wikidata to easily collect, curate, and integrate data, including the life sciences [2]. Because Wikidata has application programming interfaces to add, update (e.g. curate), and retrieve data many tools have been developed around it which empowers even more use cases. Notable for data editing are Mix'n'match, QuickStatements, OpenRefine, and Author Disambiguator. For visualization the Wikidata Query Service (WDQS, https://query.wikidata.org/) took Wikidata to a new level, empowered by a continuously updated SPARQL endpoint, which now hosts over 13 billion RDF triples. Based on the ideas of the WDQS many tools have been developed, including Scholia (https://scholia.toolforge.org/) which visualizes data from Wikidata using SPARQL queries [3] around so-called aspects for genes, proteins, metabolites, but also researchers, institutes, and journals and publishers. Scholia encourages knowledge discovery. In this session, we discussed how Wikidata organizes the development of the data in Wikidata and we discussed its mechanisms to ensure quality including the notion of collaborating crowds around WikiProject, like the one for Chemistry (https://www.wikidata.org/wiki/Wikidata_talk:WikiProject_Chemistry). For metabolomics research Wikidata certainly has potential. It already includes 1.3 million common chemicals (see https://scholia.toolforge.org/chemical/) and natural products [4] most with external identifiers to compound databases like PubChem, EPA CompTox dashboard, KEGG, HMDB, ChEBI, Chemical Abstracts' Common Chemistry, LIPID MAPS, and SwissLipids. But it also includes links to databases with experimental data, like MassBank, NMRShiftDB, the PDB database, or to the primary literature. Databases that have been using Wikidata include WikiPathways [1], Reactome, Complex Portal [6], and LOTUS [4].

References

1 Denny Vrandečić. Wikidata: A new platform for collaborative data collection. In *Proceedings of the 21st International Conference on World Wide Web*, WWW '12 Companion, page 1063–1064, New York, NY, USA, 2012. Association for Computing Machinery.
2 Andra Waagmeester, Gregory Stupp, Sebastian Burgstaller-Muehlbacher, Benjamin M Good, Malachi Griffith, Obi L Griffith, Kristina Hanspers, Henning Hermjakob, Toby S Hudson, Kevin Hybiske, Sarah M Keating, Magnus Manske, Michael Mayers, Daniel Mietchen, Elvira Mitraka, Alexander R Pico, Timothy Putman, Anders Riutta, Nuria Queralt-Rosinach, Lynn M Schriml, Thomas Shafee, Denise Slenter, Ralf Stephan, Katherine Thornton, Ginger Tsueng, Roger Tu, Sabah Ul-Hasan, Egon Willighagen, Chunlei Wu, and Andrew I Su. Science forum: Wikidata as a knowledge graph for the life sciences. *eLife*, 9:e52614, mar 2020.
3 Finn Årup Nielsen, Daniel Mietchen, and Egon Willighagen. Scholia, scientometrics and wikidata. In Eva Blomqvist, Katja Hose, Heiko Paulheim, Agnieszka Ławrynowicz, Fabio

Ciravegna, and Olaf Hartig, editors, *The Semantic Web: ESWC 2017 Satellite Events*, pages 237–259, Cham, 2017. Springer International Publishing.

4 Adriano Rutz, Maria Sorokina, Jakub Galgonek, Daniel Mietchen, Egon Willighagen, Arnaud Gaudry, James G Graham, Ralf Stephan, Roderic Page, Jiří Vondrášek, Christoph Steinbeck, Guido F Pauli, Jean-Luc Wolfender, Jonathan Bisson, and Pierre-Marie Allard. The lotus initiative for open knowledge management in natural products research. *eLife*, 11:e70780, may 2022.

5 Marvin Martens, Ammar Ammar, Anders Riutta, Andra Waagmeester, Denise N Slenter, Kristina Hanspers, Ryan A. Miller, Daniela Digles, Elisson N Lopes, Friederike Ehrhart, Lauren J Dupuis, Laurent A Winckers, Susan L Coort, Egon L Willighagen, Chris T Evelo, Alexander R Pico, and Martina Kutmon. WikiPathways: connecting communities. *Nucleic Acids Research*, 49(D1):D613–D621, 11 2020.

6 Birgit H M Meldal, Livia Perfetto, Colin Combe, Tiago Lubiana, João Vitor Ferreira Cavalcante, Hema Bye-A-Jee, Andra Waagmeester, Noemi del Toro, Anjali Shrivastava, Elisabeth Barrera, Edith Wong, Bernhard Mlecnik, Gabriela Bindea, Kalpana Panneerselvam, Egon Willighagen, Juri Rappsilber, Pablo Porras, Henning Hermjakob, and Sandra Orchard. Complex Portal 2022: new curation frontiers. *Nucleic Acids Research*, 50(D1):D578–D586, 10 2021.

3.8 Visualization and graphical user interfaces

Carolin Huber (UFZ – Leipzig, DE)

Visualization for MS data analysis is necessary for quality control and evaluation steps and it is preferred outside of vendor software applications. Different approaches for visualization and their advantages and disadvantages were discussed. Especially for MS scientists and coworkers without previous programming skills, the use of these tools needs to be as easily accessible as possible. As a final step, we discussed different workflow frameworks beyond the command line, and how these can also integrate more interactive visualizations.

3.9 Metaboproteomics

Lennart Martens (Ghent University, BE)

The advent of open modification search engines in proteomics, which allow the discovery of arbitrarily modified peptides from shotgun proteomics data sets, has enabled a new view on the proteome. And while specificity of these analyses at first remained an issue, the use of powerful predictors of peptide behaviour in liquid chromatography and tandem mass spectrometry (e.g., using DeepLC [1] and MS2PIP [2]) has now led to a new type of search engine that is entirely machine learning driven (notably ionbot: `https://ionbot.cloud`, [3]) and that can utilize these predictors to increase both sensitivity and specificity. By applying ionbot to a billion spectra from human samples, and 600 million spectra from mouse samples, both obtained from the PRIDE database [4], the most comprehensive map so far of proteome-wide protein modification is obtained for these two organisms. Upon inspection of the most commonly encountered modifications, several known artefacts of sample processing are encountered (e.g., carbamidomethylation, oxidation), alongside modifications of known biological relevance (e.g., phosphorylation), but also modifications that appear tightly linked to basic metabolism, such as certain acylations. This session discussed these findings, and looked at some of the modifications and ways to analyse these in a broader context, notably their mapping onto pathways (e.g., WikiPathways [5]), and the potential effect of the exposome on proteins as measured by the resulting induced protein modifications, coining the tentative term "metaboproteomics" to describe this potential new interface between the metabolomics and proteomics fields.

References

1 Robbin Bouwmeester, Ralf Gabriels, Niels Hulstaert, Lennart Martens, and Sven Degroeve. DeepLC can predict retention times for peptides that carry as-yet unseen modifications. *Nature Methods*, 18(11):1363–1369, 2021.

2 Ralf Gabriels, Lennart Martens, and Sven Degroeve. Updated MS^2PIP web server delivers fast and accurate MS^2 peak intensity prediction for multiple fragmentation methods, instruments and labeling techniques. *Nucleic Acids Research*, 47(W1):W295–W299, 04 2019.

3 Sven Degroeve, Ralf Gabriels, Kevin Velghe, Robbin Bouwmeester, Natalia Tichshenko, and Lennart Martens. ionbot: a novel, innovative and sensitive machine learning approach to LC-MS/MS peptide identification. *bioRxiv*, 2022.

4 Lennart Martens, Henning Hermjakob, Philip Jones, Marcin Adamski, Chris Taylor, David States, Kris Gevaert, Joël Vandekerckhove, and Rolf Apweiler. PRIDE: The proteomics identifications database. *PROTEOMICS*, 5(13):3537–3545, 2005.

5 Ryan A Miller, Martina Kutmon, Anwesha Bohler, Andra Waagmeester, Chris T Evelo, and Egon L Willighagen. Understanding signaling and metabolic paths using semantified and harmonized information about biological interactions. *PLOS ONE*, 17(4):1–21, 2022.

3.10 Data independent analysis (DIA)

Corey Broeckling (Colorado State University – Fort Collins, US)

Data independent analysis (DIA) is an MS/MS acquisition approach by which fragmentation is performed for many precursors simultaneously; the benefit of this approach is that more (all) precursors can be sampled. Some approaches may sacrifice sensitivity, and all methods will sacrifice selectivity to enable broad sampling. Discussion revolved around these trade-offs, and how DIA is being used and perceived across the community. The data analysis challenges were discussed, noting that target applications are generally a solved problem, but untargeted analysis remains challenging. The various flavors of DIA were summarized, including ion mobility spectrometry as a component of the process. Plans to develop R-based tools for a more "universal" DIA product ion assignment were outlined.

3.11 Visualization of chemical space

Justin van der Hooft (Wageningen University, NL)
Rui Pinto (Imperial College London, UK)

Joint work of Justin van der Hooft, Rui Pinto

This session started by asking what chemical space actually means. Naturally, depending on the context, this can mean different things: all known chemicals, all chemicals identified in one experiment or in a number of samples, etc. Furthermore, chemical space can be built based on structural (fingerprints) and analytical data (mass spectral similarity).

Following this, Rui Pinto demonstrated his approach to embedding metabolite features based on MS1 feature correlations. Rui showed how metabolites could be mapped onto a UMAP embedding, and how chemical compound classes were separated. Then, Justin van der Hooft showed various ways to visualize mass spectral embeddings, using treemap, t-MAP and t-SNE. Also, a 3-dimensional molecule app was demonstrated to explore the mass spectral embedding.

The group finished the session with a brainstorm on how to connect molecules using chemical and biological information. We identified many types and realized that both mass spectral similarity scores and correlation distances could be used to construct a "base network" or embedding on which other information is then mapped. This is an interesting area for future research.

3.12 MS/MS spectral quality (part 1)

Michael Andrej Stravs (Eawag – Dübendorf, CH)

Multiple topics concerning fragmentation spectra and their computational analysis were discussed. Initial discussion reflected critically a recent publication in the field proposing entropy metrics as a measure for spectral quality and spectral similarity. On a closer inspection, the metric penalizes evenness without a stringent justification. Further discussion sought to clarify the meaning of "spectral quality", as a measure to express discriminatory power of a spectrum versus a measure of cleanness i.e. freedom from artefacts/noise.

Combination (merging) and combined acquisition (ramping/stepping) of spectra is anecdotally said to enhance separation metrics, but it remains unclear whether experimental (ramping/ stepping) combined spectra fulfill the promise expected from ideal, computationally combined (merged) spectra due to loss of signal. Evaluations are planned to shed light on the matter.

Mass range: Combinatorially, we expect low mass and high-mass ions to be uninformative, due to limited combinatorial space. However in practice low-mass ions provide an information rich structural fingerprint as evidenced by machine learning methods. Weighting curves based on mass and entropy expectations have shown little use; a curve only penalizing high mass is worth trying. Further, the "quality" of a spectrum could be approximated by how well a fragmentation tree, its model, can represent it. We note that reproducible unannotated peaks still are frequent in multiple libraries. Finally, correlation between scans across an entire dataset (as opposed to correlations in a single measurement) show promise for fragment deconvolution both for targeted and broad-range fragmentation.

3.13 MS/MS spectral quality (part 2)

Adriano Rutz (University of Geneva, CH)

A visualization tool to explore possibly unexplained fragments (i.e. no molecular formula but possibly good intensity correlation with other peaks) was showcased. This tool was presented along with a method to minimize overlap when acquiring reference spectra to build libraries. The influence of this minimized overlap could help better explain (or minimize) the previously unexplained fragments. Multiple origins of these unexplained fragments were discussed, such as complex multi-charged multimers, or fragments originating from instrumental issues such as "peak ringing" for ToF spectrometers.

A need for a new metric enabling the estimation of the percentage of fragmentation independently from the collision energy was discussed. This will be a great help for further machine learning applications. The highest obstacle to the implementation of such a metric resides in the "left-censored" spectra (e.g. acquisition starting at 50 m/z). Ways to implement such a metric and overcome this limitation were discussed and cumulative distribution of intensities from precursor down was proposed.

Finally, some new fragmentation methods (e.g. EAD, UVPD) supposed to improve spectral quality were discussed. While some of them seem promising, there is actually not enough data fully evaluated to draw any conclusions at this time.

3.14 CxSMILES: computation ready representation for compound classes

Egon Willighagen (Maastricht University, NL)
Adriano Rutz (University of Geneva, CH

Joint work of Egon Willighagen, Adriano Rutz

The 2017 Dagstuhl Seminar 17491 [1] meeting posed the problem that experimental characterization of measured metabolites often has an uncertainty in the annotation of their chemical identity. However, computational metabolomics often requires a computer representation to be able to do calculations with the chemical structures. During the 2020 Dagstuhl Seminar 20051 [2] meeting cheminformatics was explored in more detail. ChemAxon Extended SMILES (CxSMILES) and Markush Structures were identified as possible solutions to represent compound classes.

In this 1.5 hour session a group of seven Seminar participants continued to explore the power and limitations of CxSMILES and continued working on a write-up to explore the use and cheminformatics implementation in the Chemistry Development Kit [3], online available at `https://egonw.github.io/cdk-cxsmiles/`. Particularly, we added a new "Classes of compounds and where to find a CxSMILES" section with example CxSMILES for a number of compound classes. Another notable improvement during this Dagstuhl Seminar was that the CxSMILES property P10718 [4] in Wikidata was accepted, allowing us to archive and disseminate CxSMILES representations in Wikidata. Enumeration of CxSMILES, their integration into chemoinformatic pipelines for properties calculations, or conversion to InChI(-Keys) remains to be explored.

References

1 Theodore Alexandrov, Sebastian Böcker, Pieter Dorrestein, and Emma Schymanski. Computational Metabolomics: Identification, Interpretation, Imaging (Dagstuhl Seminar 17491). *Dagstuhl Reports*, 7(12):1–17, 2018.
2 Sebastian Böcker, Corey Broeckling, Emma Schymanski, and Nicola Zamboni. Computational Metabolomics: From Cheminformatics to Machine Learning (Dagstuhl Seminar 20051). *Dagstuhl Reports*, 10(1):144–159, 2020.
3 Egon L Willighagen, John W Mayfield, Jonathan Alvarsson, Arvid Berg, Lars Carlsson, Nina Jeliazkova, Stefan Kuhn, Tomáš Pluskal, Miquel Rojas-Chertó, Ola Spjuth, Gilleain Torrance, Chris T Evelo, Rajarshi Guha, and Christoph Steinbeck. The Chemistry Development Kit (CDK) v2.0: atom typing, depiction, molecular formulas, and substructure searching. *Journal of Cheminformatics*, 9(1):33, 2017.
4 CXSMILES (P10718).

3.15 Estimating concentration from untargeted MS data

Anneli Kruve (Stockholm University, SE)
Steffen Neumann (IPB – Halle, DE)

Joint work of Anneli Kruve, Steffen Neumann

Currently mostly relative quantification of untargeted mass spectrometry data is used on e.g. follow up studies with identification and quantification. In relative quantification isotopically labeled C13 for cell cultures, or CO2 for plants is a good measure. Absolute quantification is used for giving results which are comparable between peaks/chemicals with different structures. For this, structurally similar standards are used but these can yield very inaccurate results due to very different behavior in MS. There are efforts to model this, but more data is needed e.g. calibration graphs. This highlights the importance of data exchange. It was also discussed that NMR is quantitative, and there have been NMR-guided MS approaches. This might open up possibilities to create quantification curves for chemicals which cannot be quantified in any other way.

3.16 WikiPathways and RaMP-DB

Egon Willighagen (Maastricht University, NL)
Ewy Mathé (National Institutes of Health – Bethesda, US)

Joint work of Egon Willighagen, Ewy Mathé

WikiPathways is an expert-curated, collaborative biological pathway database [1]. It describes biological processes for multiple species at many levels. During the meeting we discussed various aspects of the database: data format (GPML), export formats (RDF, PNG, SBML, etc), data curation, automated testing (Jenkins), and how it links to other knowledge bases (via identifier mapping databases using BridgeDb [2]).

WikiPathways is also used by other databases, e.g. for pathway enrichment and data visualization. RaMP-DB is one of those tools [3, 4]. RaMP-DB aggregates multiple types of annotations on human metabolites and proteins/genes from multiple sources to provide a comprehensive and up-to-date source of biological and chemical annotations. RaMP-DB provides user-friendly means of interacting with the database for batch queries and enrichment analyses.

Through the session, thoughts on future maintenance and development of these resources were brought up, including the need for standards of reporting to ease harmonization across resources, assessment of the quality of annotations, and the ability to use these resources as benchmarking data for method development.

References

1 Marvin Martens, Ammar Ammar, Anders Riutta, Andra Waagmeester, Denise N Slenter, Kristina Hanspers, Ryan A. Miller, Daniela Digles, Elisson N Lopes, Friederike Ehrhart, Lauren J Dupuis, Laurent A Winckers, Susan L Coort, Egon L Willighagen, Chris T Evelo, Alexander R Pico, and Martina Kutmon. WikiPathways: connecting communities. *Nucleic Acids Research*, 49(D1):D613–D621, 11 2020.

2 Martijn P van Iersel, Alexander R Pico, Thomas Kelder, Jianjiong Gao, Isaac Ho, Kristina Hanspers, Bruce R Conklin, and Chris T Evelo. The BridgeDb framework: standardized access to gene, protein and metabolite identifier mapping services. *BMC Bioinformatics*, 11(1):5, 2010.

3 Bofei Zhang, Senyang Hu, Elizabeth Baskin, Andrew Patt, Jalal K. Siddiqui, and Ewy A. Mathé. RaMP: A comprehensive relational database of metabolomics pathways for pathway enrichment analysis of genes and metabolites. *Metabolites*, 8(1), mar 2018.

4 John Braisted, Andrew Patt, Cole Tindall, Tara Eicher, Timothy Sheils, Jorge Neyra, Kyle Spencer, and Ewy A Mathé. RaMP-DB 2.0: a renovated knowledgebase for deriving biological and chemical insight from genes, proteins, and metabolites. *bioRxiv*, 2022.

3.17 Data collection considerations for MS^n spectral libraries

Tomáš Pluskal (The Czech Academy of Sciences – Prague, CZ)

We discussed many issues and parameters to consider when building a multi-stage MSn spectral library, such as ionization modes, chromatography conditions, scanning mass resolution, and acquiring preliminary data to inform and optimize data collection. The flow injection method was debated in more detail to extend the peak width for obtaining more fragmentation experiments per compound. Here, applying a second LC pump was mentioned as a way to add an additional makeup flow, which broadens the peak width as the first pump delivers the sample with a minimal flow rate. Furthermore, we discussed the optimization of the parameters of the MSn data acquisition method. The first parameter was the mass resolution and we discussed the community's expectations for high-quality fragmentation tree libraries. A lower mass resolution enables fast scanning but loses accuracy. Accordingly, a higher mass resolution needs more analysis time. As a compromise, the resolution of 30000 for MS1 and 15000 for MSn were suggested as a compromise, because the filling time is the more crucial step to enhance signal intensities for deeper MSn spectra. Since MS2 data are already available for many compounds in the library that we are planning to acquire, these MS2 scans should be limited, spending more time on MS3, MS4, and maybe MS5 experiments. Furthermore, the existing MS2 data can be used to plan the optimal strategy of MSn data collection for each compound. The depth and breadth of the spectral trees should be evaluated for multiple compounds as proof of principle and to avoid empty scans.

In this context, the application of multiple collision energies for one precursor needs to be examined for its usefulness. As a consequence, we agreed that the first data of hundreds of different compounds are needed to evaluate the multi-step fragmentation method.

3.18 RT, adduct formation, and calibration curve sharing

Michael Witting (Helmholtz Zentrum München, DE)

We have discussed GitHub repositories for sharing different types of data, including RT, adduct formation, and calibration curves. This data will make it possible to develop new machine learning methods for the prediction of RT, sodium and other adduct formation as well as ionization efficacy. The repository can be found at `https://github.com/michaelwitting/RtPredTrainingData`.

3.19 Metabolomics data integration

Justin van der Hooft (Wageningen University, NL)

The session started with a collection of scenarios and approaches used by the discussion group that include metabolomics-based data integration. Based on the group's input, it turned out that mostly genomics data is linked to metabolomics data. The main aims mentioned are to find biomarkers in clinical metabolomics or link genes to molecules in natural product processing. Sometimes these integrative analyses are further enriched by the addition of proteome or transcriptome data to assist in forming a systems-wide view of the data. If clinical or phenotype data is available, predictive analysis is often targeted, for example in the form of MWAS. Common challenges the group identified include the available covariates (samples), the number of false positives in correlation analysis, and the biological interpretation (of data links). The group identified that we usually link multiple omics using data tables/feature tables, which feed into correlation analysis or machine learning/multivariate statistical approaches. To make links between molecules, it is important to use consistent identifiers, and we discussed FAIR integration of data.

We concluded the session with defining the various levels of data integration: early (concatenate matrices, then model), intermediate (statistical approach with multiple input matrices to make predictions), and late (model each matrix separately and combine predictions). Finally, several dimensions of integration were discussed: across omics, within omics (e.g. different metabolomics assays), and across organisms.

Participants

- Sebastian Böcker
Universität Jena, DE
- Corey Broeckling
Colorado State University – Fort Collins, US
- Roman Bushuiev
The Czech Academy of Sciences – Prague, CZ
- Timothy Ebbels
Imperial College London, GB
- Soha Hassoun
Tufts University – Medford, US
- Carolin Huber
UFZ – Leipzig, DE
- Katerina Kechris
University of Colorado – Aurora, US
- Oliver Kohlbacher
Universität Tübingen, DE
- Anneli Kruve
Stockholm University, SE
- Tytus Mak
NIST – Gaithersburg, US
- Lennart Martens
Ghent University, BE
- Ewy Mathé
National Institutes of Health – Bethesda, US
- María Eugenia Monge
CIBION – Buenos Aires, AR
- Steffen Neumann
IPB – Halle, DE
- Louis-Felix Nothias
University of Geneva, CH
- Rui Pinto
Imperial College London and UK-Dementia Research Institute – London, GB
- Tomáš Pluskal
The Czech Academy of Sciences – Prague, CZ
- Hannes Röst
University of Toronto, CA
- Juho Rousu
Aalto University, FI
- Francesco Russo
SSI – Copenhagen, DK
- Adriano Rutz
University of Geneva, CH
- Michael Andrej Stravs
Eawag – Dübendorf, CH
- Panteleimon Takis
Imperial College London, GB
- Justin van der Hooft
Wageningen University, NL
- Cecilia Wieder
Imperial College London, GB
- Egon Willighagen
Maastricht University, NL
- Michael Anton Witting
Helmholtz Zentrum München, DE
- Mitja Zdouc
Wageningen University, NL

Report from Dagstuhl Seminar 22182

Estimation-of-Distribution Algorithms: Theory and Applications

Josu Ceberio Uribe[*1], **Benjamin Doerr**[*2], **Carsten Witt**[*3], **and Vicente P. Soloviev**[†4]

1 **University of the Basque Country – Donostia, ES.** josu.ceberio@ehu.eus
2 **Ecole Polytechnique – Palaiseau, FR.** doerr@lix.polytechnique.fr
3 **Technical University of Denmark – Lyngby, DK.** cawi@dtu.dk
4 **Polytechnic University of Madrid, ES.** vicente.perez.soloviev@alumnos.upm.es

Abstract

The Dagstuhl seminar 22182 *Estimation-of-Distribution Algorithms: Theory and Practice* on May 2–6, 2022 brought together 19 international experts in estimation-of-distribution algorithms (EDAs). Their research ranged from a theoretical perspective, e.g., runtime analysis on synthetic problems, to an applied perspective, e.g., solutions of industrial optimization problems with EDAs. This report documents the program and the outcomes of the seminar.

Seminar May 1–6, 2022 – http://www.dagstuhl.de/22182
2012 ACM Subject Classification Computing methodologies → Search methodologies; Theory of computation → Design and analysis of algorithms
Keywords and phrases estimation-of-distribution algorithms, heuristic search and optimization, machine learning, probabilistic model building
Digital Object Identifier 10.4230/DagRep.12.5.17

1 Executive Summary

Josu Ceberio Uribe (University of the Basque Country – Donostia, ES)
Benjamin Doerr (Ecole Polytechnique – Palaiseau, FR)
Carsten Witt (Technical University of Denmark – Lyngby, DK)

The seminar "Estimation-of-Distribution Algorithms: Theory and Practice" on May 2–6, 2022 brought together 19 international experts in estimation-of-distribution algorithms (EDAs). Their research ranged from a theoretical perspective, e.g., runtime analysis on synthetic problems, to an applied perspective, e.g., solutions of industrial optimization problems with EDAs.

The main aim of the seminar was to narrow the gap between theory and practice in EDAs by bringing together researchers from both sides and stimulating interaction. We facilitated this interaction through longer introductory talks, e. g., on theoretical analyses of EDAs, regular conference-style talks, short flash talks presenting open problems and stimulating discussions and, last but not least, through various breakout sessions and group work. After each talk, we scheduled ample time for discussion, and even adapted the schedule when

* Editor / Organizer
† Editorial Assistant / Collector

discussions had gained momentum and took longer than expected. On the last day of the seminar, all participants joined a 1-hour plenum discussion summarizing the findings of the seminar, discussing open problems and identifying further research topics.

We believe that the seminar has achieved its main aims by making a step towards narrowing the gap between theory and practice in EDAs. This is witnessed by the high number of spontaneous talks given by the participants, allowing almost everyone to present his/her perspective on EDAs, high and stable attendance of participants at talks and group sessions, and lively discussions after basically every talk as well as in the dedicated discussion fora. Several participants shared the feedback with us that they learned new aspects of EDAs during the seminar and had increased their understanding of what the other side (theory/practice) was interested in.

The seminar also identified several open problems related to the design, analysis, and application of EDAs. Details can be found in the summary of the concluding plenum discussion (see further below in the report).

Finally, several participants reported to the organizers that the seminar had helped them to understand current challenges in the theory and practice of EDAs, that they had extended their professional network, and drawn inspiration from the seminar for new research ideas. They also expressed interest in attending a similar seminar in the future.

The only downside of this seminar was that several colleagues, most notably from China, could not attend because of the still ongoing sanitary crisis.

2 Table of Contents

3 Overview of Talks

3.1 An application of EDAs on a random difference equation of tumour growth

Carlos Andreu Vilarroig (Technical University of Valencia, ES) and Josu Ceberio Uribe (University of the Basque Country – Donostia, ES)

In this contribution, a discussion of two possible implementations of an EDA on a tumour growth model has been presented. This model describes the size of a tumour X_t in each time instant t through a simple difference equation – $X_{t+1} = (1 + k)X_t$ – with two parameters: the growth rate k and the initial condition (size) X_0 [1]. Since the parameters have been considered as Gaussian random variables, the aim of the problem is to find the best probability distributions of the parameters that capture the uncertainty of the real tumour size data [2] using an EDA. A first, more classical, approach is to estimate by EDA the means and variances parameters of the probability distributions. In contrast, in a second, more alternative approach, the final distributions of the EDA are the parameter distributions themselves. On this second case, a possible implementation has been developed, in which the cost function is the RMSE between the expected value of the difference equation and the data and, as a stopping criterion to avoid convergence of the variance of the distributions to 0, the algorithm stops at the iteration in which the data starts to fall outside the 95% confidence interval of the model solution.

References

1 Kathleen MC Tjørve and Even Tjørve. The use of gompertz models in growth analyses, and new gompertz model approach: An addition to the unified-richards family. *PloS one*, 12(6):e0178691, 2017.

2 Andrea Worschech, Nanhai Chen, Yong A Yu, Qian Zhang, Zoltan Pos, Stephanie Weibel, Viktoria Raab, Marianna Sabatino, Alessandro Monaco, Hui Liu, Vladia Monsurró, R Mark Buller, David F Stroncek, Ena Wang, Aladar A Szalay, and Francesco M Marincola. Systemic treatment of xenografts with vaccinia virus GLV-1h68 reveals the immunologic facet of oncolytic therapy. *BMC Genomics*, 10(1):301, 2009.

3.2 Comparing EDAs with quantum or quantum-inspired solvers

Mayowa Ayodele (Fujitsu Research of Europe – Slough, GB)

According to the observation made by Gordon Moore that the number of transistors in a dense integrated circuit doubles about every two years [1], it has been projected that general purpose computers may struggle to cope with computational demand in the future. There has therefore been research interests in the use of other specialised hardware such as quantum computers. Within the context of solving optimisation problems, quantum computers and quantum-inspired hardware have been developed. Quadratic Unconstrained Binary Optimisation (QUBO) is a common formulation used by quantum and quantum-inspired solvers. Examples of quantum or quantum-inspired QUBO solvers are D-wave's

Advantage released in 2020, which uses quantum hardware and has the capability of 5,000+ quantum bits (qubits) [2], IBM's Eagle which is being extended from a 27-qubit solver to 127-qubit [3] and Fujitsu's Digital Annealer (DA) which uses application specific CMOS to find optimum or sub-optimum solution to problems formulated with up to 100,000 bits [4]. The DA has advantages over other quantum-based algorithms such as scale (up to 100,000 bits), precision (64bit gradations), stability (stable operation at room temperature with digital circuits) and connectivity (easy to use with total bit coupling) [4].

Estimation of Distribution Algorithms (EDAs) have been applied to many problem classes including QUBOs. Quantum-inspired EDA (QiEDA) [5] have also been designed for and applied to problems formulated as QUBO e.g., TSP [6] and feature selection [7]. It is however unclear how these algorithms compare to other QUBO solvers such as the DA. The need to design EDAs that are not problem specific and that take advantage of specialised hardware is therefore motivated.

References

1 Schaller, R.R., 1997. *Moore's law: past, present and future*. IEEE spectrum, 34(6), pp.52-59.
2 McGeoch, C. and Farré, P., 2020. *The D-wave advantage system: An overview*. D-Wave Systems Inc., Burnaby, BC, Canada, Tech. Rep.
3 Chow, J., Dial, O. and Gambetta, J., 2021. *IBM Quantum breaks the 100-qubit processor barrier*. IBM Research Blog, available in https://research.ibm.com/blog/127-qubit-quantum-process-or-eagle.
4 Nakayama, H., Koyama, J., Yoneoka, N. and Miyazawa, T., 2021 *Third Generation Digital Annealer Technology*. Fujitsu, available in https://www.fujitsu.com/jp/documents/digitalannealer/researcharticles\/DA_WP_EN_20210922.pdf
5 Chen, M. and Quan, H., 2007, September. *Quantum-inspired evolutionary algorithm based on estimation of distribution*. In 2007 Second International Conference on Bio-Inspired Computing: Theories and Applications (pp. 17-19). IEEE.
6 Soloviev, V.P., Bielza, C. and Larranaga, P., 2021, June. *Quantum-Inspired Estimation Of Distribution Algorithm To Solve The Travelling Salesman Problem*. In 2021 IEEE Congress on Evolutionary Computation (CEC) (pp. 416-425). IEEE.
7 Soliman, O.S. and Rassem, A., 2012, December. *Correlation based feature selection using quantum bio inspired estimation of distribution algorithm*. In International Workshop on Multi-disciplinary Trends in Artificial Intelligence (pp. 318-329). Springer, Berlin, Heidelberg.

3.3 The question of structure in Markov Network EDAs

Alexander Brownlee (University of Stirling, GB)

Joint work of Alexander E. I. Brownlee, John A. W. McCall, Qingfu Zhang, Lee A. Christie, Martin Pelikan, D. Brown
Main reference Alexander E. I. Brownlee, John A. W. McCall, Qingfu Zhang: "Fitness Modeling With Markov Networks", IEEE Trans. Evol. Comput., Vol. 17(6), pp. 862–879, 2013.
URL https://doi.org/10.1109/TEVC.2013.2281538

This talk is on the long-running theme of "structure" in EDAs. I will introduce Markov network EDAs, which use an undirected probabilistic graphical model relating binary decision variables to the fitness distribution. This model's accuracy and, in turn, algorithm efficiency, depend on the neighbourhood structures that are captured within it. That is, does the model treat variables independently, or are bivariate or higher order interactions considered. I will briefly recap some early results looking at the impact of problem and algorithm structure in this context, covering well-known benchmarks including Onemax, Trap, and MAX3SAT, before looking at a systematic study covering 2 and 3 bit problems which shows that not all structure is essential to problem solution.

3.4 General univariate EDAs

Benjamin Doerr (Ecole Polytechnique – Palaiseau, FR)

Joint work of Benjamin Doerr, Marc Dufay

We propose a general model for univariate EDAs that encompasses the four classic EDAs. We prove a genetic drift result for the general model that implies the corresponding known results for the particular algorithms. Finally, we show that the general class of algorithm contains EDAs which are more powerful than the existing ones (with optimal parameters).

3.5 Genetic drift and a smart restart scheme

Benjamin Doerr (Ecole Polytechnique – Palaiseau, FR)

Joint work of Benjamin Doerr, Weijie Zheng
Main reference Benjamin Doerr, Weijie Zheng: "Sharp Bounds for Genetic Drift in Estimation of Distribution Algorithms", IEEE Trans. Evol. Comput., Vol. 24(6), pp. 1140–1149, 2020.
URL https://doi.org/10.1109/TEVC.2020.2987361

Genetic drift, that is, the random movement of sampling frequencies not justified by a fitness signal, usually leads to an inferior performance of EDAs. In this talk, we quantify how the parameters of the four common univariate EDAs lead to genetic drift. This allows to set the parameters in a way that during the desired runtime, the genetic drift is low enough to avoid the undesired effects. We also describe how to use this quantification to define a smart-restart mechanism that, provably, leads to a performance asymptotically equal to the one obtainable from optimal parameter values.

3.6 A Gentle Introduction to Information-Geometric Optimization (IGO)

Nikolaus Hansen (INRIA Saclay – Palaiseau, FR)

Main reference Yann Ollivier, Ludovic Arnold, Anne Auger, Nikolaus Hansen: "Information-Geometric Optimization Algorithms: A Unifying Picture via Invariance Principles", J. Mach. Learn. Res., Vol. 18, pp. 18:1–18:65, 2017.
URL https://dl.acm.org/doi/10.5555/3122009.3122027

With invariance as major design principle, we present a canonical way to turn any smooth parametric family of probability distributions (on an arbitrary search space) into a continuous-time black-box optimization method and into an "IGO algorithm" through time discretization. The construction is based upon (i) the natural gradient over the family of distributions to express (intrinsic) update directions and (ii) the cumulative distribution function of the f-values under the current sample distribution to express (invariant) preference weights. The construction recovers well-known algorithms like cGA, PBIL or Natural Evolution Strategies and major aspects of CMA-ES.

3.7 Generating permutations

Ekhine Irurozki (Telecom Paris, FR)

In this tutorial we show some known and new results in statistical properties and methods for permutations

- Uniform sampling : in this section we show how to generate permutations uniformly (where all the possible permutations have the same probability of being generated) among all those at a given distance for the identity.
- Distributions on permutations: We show the different and most common distribution for permutation data, motivation, main statistics and inference algorithms.
- Depth function: we show a typical function for real-valued data (its novel adaptation for permutation spaces)
- Experiments: in order to contribute we present an easy to use code with the above mentioned methods and operations.

3.8 Universal permutation aggregation

Ekhine Irurozki (Telecom Paris, FR)

The problem of permutation-aggregation (PA) is the main optimization problem in statistical community. It is equivalent to the median computation in the real-valued vector spaces. It can be posed under different distances and the complexity varies for these cases: it is known to be NP-hard for some distances, polynomial for others and unknown (although conjetured to be NP-hard for other)

This talk introduces a current work on the problem of aggregating permutations. We introduce an aggregation algorithm that is valid for several distance. This is an approximated algorithm for PA for which we can characterize the conditions for convergence. Moreover, we bound the expected error and the expected run-time. We also introduce simulations that confirm the theoretical results.

3.9 Theoretical Analyses of EDAs – Introduction and Overview

Martin S. Krejca (Sorbonne University – Paris, FR)

The theoretical analysis of estimation-of-distribution algorithms (EDAs) has gained a lot of momentum during the last decade. In this talk, I introduce the basic concepts and terminology used in the analysis of EDAs, and I summarize most of the latest results. The aim is to provide the audience with a solid foundation for understanding theoretical analyses of EDAs and to give it a good overview about the state of the art.

3.10 Do univariate EDAs cope well with deception/epistasis?

Per Kristian Lehre (University of Birmingham, GB)

Recently, there have been mixed theoretical and empirical evidence as to whether univariate estimation of distribution algorithms (EDAs) such as the UMDA perform worse than classical EAs on multimodal/deceptive problems. We discuss empirical investigations of a range of EAs and EDAs on NK-landscapes, as well as theoretical analyses of UMDA end EAs on the Deceptive Leading Ones problem. These results point to the limitation of univariate EDAs under certain parameter regimes.

3.11 EDAs and more noise

Johannes Lengler (ETH Zürich, CH)

This talk pickes up the topic of Andrew Sutton's previous discussion of EDAs on noise. I discussed several instances in which Evolutionary Algorithms (EAs) struggle in noisy or dynamic environments. All examples have the fact at their heart that the mutation generator of EAs is intrinsincally biased. In contrast, EDAs are unbiased (except for the boundaries, which are exceptional for EDAs). This is the reason why EDAs are often much more robust to noise than EAs.

3.12 Where and How Does the Brain Use EDAs?

Johannes Lengler (ETH Zürich, CH)

In this flash talk, I will present some recent interpretations of low-level conscious processing that resemble the way EDAs operate. Unconscious perceptions have local inconsistencies that can be interpreted as a probability distribution. A conscious perception is then simply a sample from this distribution. Episodic memory seems to be based (or to consist) purely on these samples, while the probabilistic model is also updated by unconscious perceptions.

3.13 Expensive black-box optimization problem in permutation spaces

Manuel López-Ibáñez (University of Málaga, ES)

Although there is a large amount of research in black-box optimization problems in the continuous domain, there is far fewer algorithms designed for permutation spaces, where the input points in the decision space represent a ranking or an ordering, in particular, when the number of evaluations is extremely limited, e.g., to a number of function evaluations between 100 and 1000. We describe two alternative approaches: CEGO, which is a Bayesian (i.e., surrogate-based) optimiser using Kriging (i.e., Gaussian Processes) regression and a distance metric between permutations, and UMM, an Estimation-of-Distribution algorithm based on the Mallows model. We discuss results on classical combinatorial problems and a real-world-like benchmark inspired by an application to Space exploration and on the effect of starting from a very good initial solution versus completely random ones. We conclude by discussing that there is still room for improvement in existing Bayesian and EDA optimisers in this particular context.

References

1. Manuel López-Ibáñez, Francisco Chicano, and Rodrigo Gil-Merino. *The Asteroid Routing Problem: A Benchmark for Expensive Black-Box Permutation Optimization*. In J. L. Jiménez Laredo et al., editors, Applications of Evolutionary Computation, volume 13224 of Lecture Notes in Computer Science. Springer Nature, Switzerland, 2022.
2. Ekhine Irurozki and Manuel López-Ibáñez. *Unbalanced Mallows Models for Optimizing Expensive Black-Box Permutation Problems*. In Proceedings of the Genetic and Evolutionary Computation Conference, GECCO 2021. ACM Press, New York, NY, 2021.

3.14 EDAs, epistatis, and structure

John McCall (The Robert Gordon University – Aberdeen, GB)

An EDA constructs a probabilistic model that estimates the distribution of good solutions / fitness. Typically the model is a probabalistic graphical model (PGM) composed of structure components with weights. Bayesian and Markov networks are two widely used PGMs [1, 2]. EDAs select good quality solutions from the current population, use them to estimate the topology and parameters of a PGM and them sample the PGM to create a successor population.

That part of the modelling process that relates to learning the topology of the PGM is known as structure learning. Structure components relate directly to subsets of variables (graph components of the PGM) and can be thought of as indicating patterns that make a contribution to fitness depending on the values of the variables. This is strongly related to the concept of epistasis which means strong interactions in determining individual fitness between separated genes. Repeated runs of an EDA may learn different structure and different structure elements may be present in models generated at different stages of a single run. A mainstream outcome of EDA research was to develop a series of EDAs for a wide range of problems that used PGMs with increasingly complex structure to tackle challenging problems with high degrees of epistasis and with a range of representations e.g. MAXSAT, Ising problems, NK landscapes, continuous benchmark sets, QAP, PFSP, TSP, etc. EDAs that use PGMs are classified as univariate, bivariate or multivariate according to the structural complexity of their models.

A key observation is that fitness may be thought of as a mass distribution and so by normalising, or other means, we can create "true" probabilistic graphical models of the fitness distribution, examine their structure and explore the relationship between this structure and the structure estimated by EDAs. While it is not always possible to determine all of the true structure of a fitness function, structure discovery algorithms exist – e.g. Heckendorn and Wright [3]. It is possible at reasonable cost in evaluations to "probe" for structure of a certain complexity by evaluating carefully selected solutions. One can ask for a given problem, what structure does an EDA need to discover to optimise efficiently?

A folklore intuition exists that, on optimisation problems with a complex true structure, algorithms that can detect and model complex structure will exhibit superior performance to algorithms that cannot do so. Many papers exist showing multivariate algorithms outperforming univariate or bivariate algorithms on benchmarks. In the talk, we present a real world cancer chemotherapy optimisation problem with a high degree of bivariate structure on which univariate EDAs, despite being incapable of detecting this structure, outperform hBOA, one of the most powerful and widely applied multivariate EDAs [4].

The result is important because a drawback of multivariate EDAs is that the computational cost of model building can quickly become significant compared to overall function evaluation costs as structural complexity increases. An understanding of when structure learning is unnecessary to, or even impedes, optimisation may lead to important insights for the design of EDAs that use PGMs. A key observation is that most PGM EDAs operate invariantly to order-preserving function transformations, that is to say they are invariant across welldefined function classes. In this talk we share some results from [5] where we algebraically characterise inessential structure for such EDAs on pseudo-boolean problems.

References

1 Larrañaga, P. and Lozano, J.A. 2002. Estimation of Distribution Algorithms: A New Tool for Evolutionary Computation. Kluwer Academic Publishers, Boston.

2 Chen, Y. 2010. Exploitation of Linkage Learning in Evolutionary Algorithms, Springer

3 Heckendorn, R.B. and Wright, A.H. 2004. Efficient Linkage Discovery by Limited Probing. Evolutionary Computation 12, 517-545.

4 Brownlee, Alexander EI; Pelikan, Martin; McCall, John AW; Petrovski, Andrei, 2008, An application of a multivariate estimation of distribution algorithm to cancer chemotherapy, Proceedings of the 10th annual conference on Genetic and evolutionary computation (GECCO), pp463-464. Full technical paper at arxiv.2205.08438 .

5 Christie, L. A., McCall, J. A. W. and Lonie, D. P., 2014. Minimal walsh structure and ordinal linkage of monotonicity-invariant function classes on bit strings. In: Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation (GECCO '14). 12-16 July 2014. New York: ACM. pp. 333-340.

3.15 EDA applications in real industrial problems

Vicente P. Soloviev (Polytechnic University of Madrid, ES)

In this talk four different applications of the Estimation of Distribution Algorithms (EDAs) have been shown. The first application [1] is a complex multivariate optimization problem where a Gaussian Bayesian network is used a probabilistic model where some prior knowledge is involved and fixed as evidence in the model to restrict the search space. The second one is a Feature Subset Selection problem with time series, where not only the optimum subset of variables are chosen, but also de ideal time series transformation that improves the model. An EDA as a wrapper approach is implemented. The third one is a Real Estate portfolio optimization problem where it is aimed to minimize the risk of the investment and to maximize the profit of the portfolio. It is presented how EDA deal with large scale data in this problem. The last optimization problem [2] consists of optimizing some hyperparameters of a concrete quantum circuit. This talk leads to discussion about the probabilistic model complexity needed to approach the real-world optimization problems.

References

1 Soloviev, V.P., Larranaga, P. and Bielza, C. , 2022, June. *Estimation of distribution algorithms using Gaussian Bayesian networks to solve industrial optimization problems constrained by environment variables*. In Journal of Combinatorial Optimization.
2 Soloviev, V.P., Larranaga, P. and Bielza, C. , 2022, June. *Quantum Parametric Circuit Optimization with Estimation of Distribution Algorithms*. In 2022 The Genetic and Evolutionary Computation Conference (GECCO). DOI: https://doi.org/10.1145/3520304.3533963

3.16 Semi-parametric Bayesian networks in EDAs

Vicente P. Soloviev (Polytechnic University of Madrid, ES)

Some real applications involve leading with variables that do not fit a normal distribution in the continuous space. Thus, some more complex algorithms should be used rather than the well-known Gaussian Bayesian network. A novel method is presented where EDA decides whether the variables will fit or not a Gaussian distribution during runtime. Some results are shown comparing the approach with some other state-of-the-art methods such as EGNA, JADE or CMA-ES for some specific benchmarks, and also explaining the performance of the approach during runtime.

3.17 A bit about algorithm configuration

Thomas Stützle (Free University of Brussels, BE)

This is a bit about automated algorithm configuration. It is divided in two aspects. The first highlights the main advantages of addressing algorithm design and configuration by algorithmic techniques and describes the main existing techniques. We show as one of the various available ones the irace algorithm for automated configuration. The second highlights the advantages of automated configuration. In particular, we show how flexible algorithm frameworks can support the automatic design of high-performing hybrid stochastic local search algorithms. We show this by the example of nine permutation flow shop problems for which we obtained in each one of these a new state-of-the-art algorithm even though they have on average around 500 parameters.

3.18 Different parameter regimes in the cGA and their effect on performance

Dirk Sudholt (Universität Passau, DE)

Understanding the dynamics of the underlying probabilistic model is a fundamental problem in estimation-of-distribution algorithms. We consider the compact Genetic Algorithm (cGA) on the function OneMax and show how the choice of the update strength $1/K$ affects performance. For small K updates have a large impact and genetic drift leads to chaotic behaviour and exponential times. For large K, genetic drift is negligible since updates are small and the algorithm slowly learns good bit values. For medium K, frequencies may hit their lower borders, but these wrong decisions can be reverted efficiently. The parameter landscape has a bimodal shape with two asymptotically optimal parameter values. A similar, but more extreme effect is observed for the multimodal Cliff function where the cGA is inefficient for all values of K and the best (exponential) runtime is obtained for exponential values of K that prevent genetic drift.

References

1 Johannes Lengler, Dirk Sudholt, and Carsten Witt. The complex parameter landscape of the compact genetic algorithm. *Algorithmica*, 83(4):1096–1137, 2021.
2 Frank Neumann, Dirk Sudholt, and Carsten Witt. A few ants are enough: ACO with iteration-best update. In *Proc. of GECCO '10*, pages 63–70. ACM Press, 2010.
3 Frank Neumann, Dirk Sudholt, and Carsten Witt. The compact genetic algorithm struggles on cliff functions. In *Proc. of GECCO '22*. ACM Press, 2022.
4 Dirk Sudholt and Carsten Witt. On the choice of the update strength in estimation-of-distribution algorithms and ant colony optimization. *Algorithmica*, 81(4):1450–1489, 2019.

3.19 EDAs, noise and graceful scaling

Andrew M. Sutton (University of Minnesota – Duluth, US)

A search algorithm scales gracefully with noise when a polynomial increase in noise intensity can be overcome by at most a polynomial increase in resource usage. In this talk, I outline a few results related to graceful scaling on additive posterior noise and estimation of distribution algorithms.

4 Working groups

4.1 Breakout session: Dealing with constraint search spaces

Josu Ceberio Uribe (University of the Basque Country – Donostia, ES) and Ekhine Irurozki (Telecom Paris, FR)

There are many problems whose search space is the space that results on applying constraints to $\{0,1\}^n$. For example, the graph partitioning problem on a graph of n edges. The set of solutions to this problem is the set of binary vectors of length n in which the number of 1s equals the number of 0s, $S = \{v : v \in \{0,1\}^n \wedge \sum_i^n v(i) = n/2\}$. Another example is the MaxCut problem.

We consider this set S as a running example but extensions to different setting with other constraints are straight forward.

In order to use EDAs on S the naive approach is to consider the whole set $\{0,1\}^n$ and then discard the infeasible solutions (aka those that do not belong to S). However, this approach is not efficient as the number of infeasible solutions can be a large proportion of the total space and the proportion of discarded solutions increases with n. This motivates the need for distributions with support on S.

Based on techniques described in the talks in this seminar we have described a probability model for S. We describe here the basic methods for EDAs, which are sampling and learning.

Given a sample of vectors $V \in S$ of cardinality m the learning process consists on a the estimation of a matrix $M \in R^{n\times n}$ for which M_{ij} is the ratio of the vectors for which $v(i) = 0$ and $v(j) = 1$ among those in which they have different values, i.e., set $w_{ij} = |\{v \in V \wedge v(i) = 0 \wedge v(j) = 1\}|$, $M_{ij} = \{w_{ij}/w_{ji}\}$

The sampling process of a vector v is as follows: (1) let i, j be random pair in $\{1, n\}$ and (2) $v[i] = 0$ and $v[j] = 1$.

For the development of an EDA the learning could be skipped (substitute by a lazy-learning) in which the learning is done on-line.

4.2 Breakout session: Univariate EDAs, deception and epistatis

Per Kristian Lehre (University of Birmingham, GB) and Johannes Lengler (ETH Zürich, CH)

We discussed the surprising similarity of the cGA and UMDA on OneMax, where the step size K of UMDA is equated with the population size μ of the UMDA, and the sample size λ is in $\Theta(\mu)$. The right scaling is to compare K steps of the cGA with one step of the UMDA. In this scaling, both algorithms use $\Theta(K)$ function evaluations per round. It turns out that for any distribution, the drift (expected change) in each frequency is essentially the same for UMDA and cGA. The most extremal possible change is also the same: it is just so possible to go from one boundary to the other in one round. However, we found out that the tail distribution for such extremal events is not the same.

The discussion took place with the whole seminar, after a talk on deception and epistasis by Per Kristian Lehre. Martin Krejca speculated that one might see a difference on such functions, because the tail distribution of large steps is important in this case. In the end, we identified several research questions in that direction:

1. For which other fitness functions than OneMax is the drift identical? Can we identify a class of such fitness functions?
2. While cGA and UMDA have the same expected change on OneMax, do they also have the same variance? If not, this should result in a stronger/weaker genetic drift that might be important in some settings
3. Can we formally or experimentally prove that the behaviour of the two algorithms is qualitatively different in the deceptive environments invented by Lehre?

4.3 Breakout session: Explainability

John McCall (The Robert Gordon University – Aberdeen, GB) and Josu Ceberio Uribe (University of the Basque Country – Donostia, ES)

In this breakout session, the explainability of evolutionary computation has been discussed. The chair explained that when it comes to approach real world problems, many times, the solutions that are obtained from the EC algorithm are difficult to accept by the decision makers of the projects (usually people who do not have a scientific background). Some participants of the session have discussed that it is not the responsibility of the scientist or engineer to make the algorithm understandable for the decision maker. On the other side, some participants have explained that when it comes to apply the proposed solution in critical scenarios (medical therapies, healthcare systems, or any other critical infrastructure configuration), then decision makers usually does not trust in solutions that they do not trust, as they are risky in the opinion. The goal of the explainability is to try to understand why the algorithm output a particular solution with rules that are understandable for the human.

In the board, a scheme of types of explainability was sketched. The first branching distinguishes between (a) risk assessment and (b) problem understanding, and from the second two other branches were obtained: (1) sensitivity analysis and (2) problem information from algorithm traces.

Thinking on an analogous problem on machine learning and explainability, it was mentioned that good explainability in general requires looking power in the model/algorithm. No direct link was done with EC.

The audience agreed that there could be scenarios like this, but there was no clue about how to do it. It will take place a workshop in GECCO 2022 on the topic under the acronym ECXAI, that could be a nice discussion point.

4.4 Breakout session: Do we need to model interactions in EDAs

Jonathan L. Shapiro (University of Manchester, GB)

This session motivated to some extent by a comment in a talk by John McCall, which said (paraphrasing) sometimes it is not worth the effort to model the interactions, but rather to just search fast using a univariate EDA.

Jon Shapiro started by pointing out that many combinatorial optimization problems exist on undirected connected graphs. These are computationally expensive to sample from, requiring hundreds or thousands of Markov chain Monte Carlo steps. These are also very difficult to model from data, because in complex situations, variables can be "frozen" into correlated or anti-correlated states even without there being direct interactions between them. So would dedicating computational resources to model building be fruitful, or would this computational effort be better used in search?

4.5 Breakout session: Genetic drift

Dirk Sudholt (Universität Passau, DE) and Johannes Lengler (ETH Zürich, CH)

We first tried to clarify our understanding of the term of genetic drift, and found a satisfactory definition: it is the difference between the mean from perfect distribution and empirical distribution, in other words it is the effect of finite samples. We also discussed what this implies about genetic drift in case of several optima, or when the fitness signal is zero, for example in the checkerboard benchmark. Finally, in Mallows models, it is less clear what the genetic drift is. Empirically, when the solution stop improving, then the variance steeply goes down. It is unclear which is the cause and which the effect.

We discussed several methods on how to observe it empirically. A conservative way would be to measure or compute for a fitness-neutral frequency whether there is a risk of hitting boundaries. Another option is to either compare a run with normal sample size with a run with a very high sample size, or several runs with normal sample sizes. If the runs are diverging, this is a sign of genetic drift.

We then turned to the effects of genetic drift. One question is whether it leads to a loss of variance, and that depends on the situation. We have also discussed several cases in which genetic drift might help, including tie-breaking and escaping from local optima.

Finally, we discussed how to control and prevent it, though it was debated whether we even need or want to do that, so we also discussed ways to increase it. One way would be to control parameters like learning rate or steps size, which may be interpreted as controlling the amount of genetic drift. It was suggested to design a cGA where we adapt K in order to control the amount of genetic drift. The sig-cGA achieves a related goal in a similar way, by using statistical test. The sig-cGA is rather artificial and extreme, but it might be promising to look at less extreme version of this algorithm. We discussed that this might be hard in continuous domains with a large or moderate number of dimensions.

In the end, we digressed into discussion what other methods one can use to deal with local optima or premature convergence.

5 Panel discussions

5.1 Concluding group discussion of the seminar

Josu Ceberio Uribe (University of the Basque Country – Donostia, ES), Benjamin Doerr (Ecole Polytechnique – Palaiseau, FR), Per Kristian Lehre (University of Birmingham, GB), and Carsten Witt (Technical University of Denmark – Lyngby, DK)

Chair: Per Kristian Lehre

The group discussion on the last day of the seminar involved all participants of the seminar and aimed at summarizing the findings of the seminar and discussing central questions related to the main aim of the seminar, bridging the gap between theory and practice in EDAs. Concretely, the following two main questions along with their subtopics were proposed as items for discussion:

1. What are practice-relevant EDAs?
- When is a univariate model enough?
- When do we have to model dependencies between variables?
- When is it too expensive to employ an EDA?

2. What are the barriers for theory of EDAs
- What is the right theory?
- What problems should theory study?
- What are the major challenges?

Regarding the first subject, practice-relevant EDAs, the seminar discussed the use of restart strategies. Also, the suggestion came up to start with very simple EDAs (e. g. UMDA) in the first place (an idea supported by talks of practitioners during the seminar). One would then proceed to more complex interaction models only if the results of simple EDAs, along with restart strategies, are unsatisfactory.

A challenge in using complex EDAs is the complex parameter space. On the one hand, empirical evidence indicates that parameter tuning for EDAs is more benign than expected [9]; on the other hand, theoretical research has identified examples where the parameter landscape is surprisingly complex [7]. As an interesting note, state-of-the-art automated parameter configuration tools (e. g. *irace*, [8]) conduct their search in a manner resembling an EDA. As a note of caution, participants pointed out that optimal parameter settings in EDAs and other search heuristics may also vary during the run.

As another domain where EDAs may be the preferred algorithm, the participants discussed the evolution of surrogate models of fitness functions. They pointed out that building these models can be computationally expensive so standard complexity measures like the number of fitness calls are no longer representative.

The seminar discussed briefly that common EDAs focus on building a model of the search space and do not rely on directed search operators like mutation vectors in continuous search spaces. However, no conclusion was reached on how to best model the notion of *direction* in discrete search spaces.

During the second half of the group discussion, the seminar turned to a complementary subject, the barriers for theory mentioned above. As theoretical benchmark problems for multivariate EDAs, the so-called concatenated trap functions [3], checkerboard functions [1], deceiving LeadingOnesBlocks [6], Ising models [4] and generalized models of multimodal problems (e. g., the *SparseLocalOpt* problem class, [2]) were suggested. Several of these may serve as examples where multivariate models may be required for efficient optimization times.

The participants discussed the purpose of simple benchmark functions like OneMax, which were defended as building blocks of more complex functions and a basic sanity check for new algorithms. If an algorithm does not optimize OneMax efficiently, it should be discarded.

Criticism with respect to the commonly analyzed search space of bit strings $\{0, 1\}^n$ led the discussion to permutation spaces, which were the subject of several talks during the seminar. Runtime analyses of EDAs in permutation spaces, using simple benchmark problems like sorting were suggested.

More generally, participants also demanded runtime analyses of EDAs on classical combinatorial optimization problems (e. g., minimum spanning trees) and the study of other performance measures than the first hitting time of an optimum. In particular, fixed-budget analyses as common with evolutionary algorithms [5] seem to be missing so far even though existing tools for the analyses should make results in this direction feasible.

Regarding the “right” theory, the suggestion came up to study general properties/characterizations of problems and algorithms that guarantee efficient runtimes. No final conclusion was reached how such a theory could look like.

Towards the end of the session, the use of surrogate models in EDAs was picked up again. Runtime analyses seem very complex; however, a starting point may be to discriminate the structure of surrogate models at two points in time t_1, t_2 using experimental studies of EDAs at these times or even using theoretical means. Such studies of the quality of surrogate models could possibly lead to estimations of the development of the model over time and even the final optimization time.

As a last point in the discussion, the participants talked about runtime analyses of EDAs in noisy optimization where the noise is not Gaussian; for example, dynamic noise that decreases over time or when getting closer to the optimum. This was linked to the observation that the variance of the sampling distribution in EDAs tends to decrease over time.

References

1. Shumeet Baluja and Scott Davies. Using optimal dependency-trees for combinational optimization. In Douglas H. Fisher, editor, *Proceedings of the Fourteenth International Conference on Machine Learning (ICML 1997), Nashville, Tennessee, USA, July 8-12, 1997*, pages 30–38. Morgan Kaufmann, 1997.
2. Duc-Cuong Dang, Anton V. Eremeev, and Per Kristian Lehre. Escaping local optima with non-elitist evolutionary algorithms. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021*, pages 12275–12283. AAAI Press, 2021.
3. Kalyanmoy Deb and David E. Goldberg. Analyzing deception in trap functions. In L. Darrell Whitley, editor, *Proceedings of the Second Workshop on Foundations of Genetic Algorithms. Vail, Colorado, USA, July 26-29 1992*, pages 93–108. Morgan Kaufmann, 1992.
4. Simon Fischer and Ingo Wegener. The ising model on the ring: Mutation versus recombination. In Kalyanmoy Deb, Riccardo Poli, Wolfgang Banzhaf, Hans-Georg Beyer, Edmund K. Burke, Paul J. Darwen, Dipankar Dasgupta, Dario Floreano, James A. Foster, Mark Harman, Owen Holland, Pier Luca Lanzi, Lee Spector, Andrea Tettamanzi, Dirk Thierens, and Andrew M. Tyrrell, editors, *Genetic and Evolutionary Computation – GECCO 2004, Genetic and Evolutionary Computation Conference, Seattle, WA, USA, June 26-30, 2004, Proceedings, Part I*, volume 3102 of *Lecture Notes in Computer Science*, pages 1113–1124. Springer, 2004.
5. Thomas Jansen and Christine Zarges. Fixed budget computations: a different perspective on run time analysis. In Terence Soule and Jason H. Moore, editors, *Genetic and Evolutionary Computation Conference, GECCO '12, Philadelphia, PA, USA, July 7-11, 2012*, pages 1325–1332. ACM, 2012.
6. Per Kristian Lehre and Phan Trung Hai Nguyen. On the limitations of the univariate marginal distribution algorithm to deception and where bivariate edas might help. In Tobias Friedrich, Carola Doerr, and Dirk V. Arnold, editors, *Proceedings of the 15th ACM/SIGEVO Conference on Foundations of Genetic Algorithms, FOGA 2019, Potsdam, Germany, August 27-29, 2019*, pages 154–168. ACM, 2019.
7. Johannes Lengler, Dirk Sudholt, and Carsten Witt. The complex parameter landscape of the compact genetic algorithm. *Algorithmica*, 83(4):1096–1137, 2021.
8. Manuel López-Ibáñez, Jérémie Dubois-Lacoste, Leslie Pérez Cáceres, Thomas Stützle, and Mauro Birattari. The *irace* package: Iterated racing for automatic algorithm configuration. *Operations Research Perspectives*, 3:43–58, 2016.
9. Yasha Pushak and Holger H. Hoos. Algorithm configuration landscapes: – more benign than expected? In Anne Auger, Carlos M. Fonseca, Nuno Lourenço, Penousal Machado, Luís Paquete, and L. Darrell Whitley, editors, *Parallel Problem Solving from Nature – PPSN XV – 15th International Conference, Coimbra, Portugal, September 8-12, 2018, Proceedings, Part II*, volume 11102 of *Lecture Notes in Computer Science*, pages 271–283. Springer, 2018.

Participants

- Carlos Andreu Vilarroig
Technical University of Valencia, ES
- Mayowa Ayodele
Fujitsu Research of Europe – Slough, GB
- Alexander Brownlee
University of Stirling, GB
- Josu Ceberio Uribe
University of the Basque Country – Donostia, ES
- Benjamin Doerr
Ecole Polytechnique – Palaiseau, FR
- Nikolaus Hansen
INRIA Saclay – Palaiseau, FR
- Ekhine Irurozki
Telecom Paris, FR
- Ata Kaban
University of Birmingham, GB
- Martin S. Krejca
Sorbonne University – Paris, FR
- Per Kristian Lehre
University of Birmingham, GB
- Johannes Lengler
ETH Zürich, CH
- Manuel López-Ibáñez
University of Málaga, ES
- John McCall
The Robert Gordon University – Aberdeen, GB
- Vicente Perez Soloviev
Polytechnic University of Madrid, ES
- Jonathan L. Shapiro
University of Manchester, GB
- Thomas Stützle
Free University of Brussels, BE
- Dirk Sudholt
Universität Passau, DE
- Andrew M. Sutton
University of Minnesota – Duluth, US
- Carsten Witt
Technical University of Denmark – Lyngby, DK

Report from Dagstuhl Seminar 22191

Visual Text Analytics

Christopher Collins*[1], Antske Fokkens*[2], Andreas Kerren*[3], Chris Weaver*[4], and Angelos Chatzimparmpas†[5]

1 **Ontario Tech University – Oshawa, C.** christopher.collins@ontariotechu.ca
2 **Free University – Amsterdam, NL.** antske.fokkens@vu.nl
3 **Linköping University – Norrköping, SE.** kerren@acm.org
4 **University of Oklahoma – Norman, US.** cweaver@ou.edu
5 **Linnaeus University – Växjö, SE.** angelos.chatzimparmpas@lnu.se

Abstract

Text data is one of the most abundant types of data available, produced every day across all domains of society. Understanding the contents of this data can support important policy decisions, help us understand society and culture, and improve business processes. While machine learning techniques are growing in their power for analyzing text data, there is still a clear role for human analysis and decision-making. This seminar explored the use of visual analytics applied to text data as a means to bridge the complementary strengths of people and computers. The field of visual text analytics applies visualization and interaction approaches which are tightly coupled to natural language processing systems to create analysis processes and systems for examining text and multimedia data. During the seminar, interdisciplinary working groups of experts from visualization, natural language processing, and machine learning examined seven topic areas to reflect on the state of the field, identify gaps in knowledge, and create an agenda for future cross-disciplinary research. This report documents the program and the outcomes of Dagstuhl Seminar 22191 "Visual Text Analytics".

Seminar 08.– 13. May, 2022 – https://www.dagstuhl.de/22191

2012 ACM Subject Classification Human-centered computing → Visualization techniques; Human-centered computing → Visual analytics; Human-centered computing → Information visualization; Computing methodologies → Natural language processing; Computing methodologies → Machine learning; Information systems → Information systems applications; Applied computing → Document management and text processing

Keywords and phrases *Information visualization, visual text analytics, visual analytics, text visualization, explainable ML for text analytics, language models, text mining, natural language processing*

Digital Object Identifier 10.4230/DagRep.12.5.37

* Editor / Organizer

† Editorial Assistant / Collector

Visual Text Analytics, *Dagstuhl Reports*, Vol. 12, Issue 5, pp. 37–91
Editors: Christopher Collins, Antske Fokkens, Andreas Kerren, and Chris Weaver
Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Executive Summary

Christopher Collins (Ontario Tech – Oshawa, CA)
Antske Fokkens (Free University Amsterdam, NL)
Andreas Kerren (Linköping University – Norrköping, SE)
Chris Weaver (University of Oklahoma – Norman, US)

Introduction

Visualizing textual information is a particularly challenging area of information visualization and visual analytics research. The types of data processing and analytic algorithms differ greatly from tabular or geospatial data, and the visualization techniques have additional constraints to consider, including the provision of context for text fragments of similar or different size and structure, depicting embeddings and high dimensional representations, and ensuring legibility of text incorporated into visualizations. The wide variation in the data is accompanied by the difficulties in inferring the semantic meaning of ambiguous terms, or determining the referencing between subsequent statements.

This Dagstuhl Seminar succeeded in bringing together researchers from the visualization, natural language processing (NLP), and machine learning communities, with domain experts from several text-related research areas, to identify the most pressing and promising open problems for collaborative research. This truly interdisciplinary approach offered new opportunities to capitalize on existing knowledge and recent developments across all involved disciplines. Discussions in the seminar were comprehensive, focusing on visual text analytics with the goal to provide an application-oriented research agenda.

The seminar coalesced an international community of experts from different disciplines around a research roadmap for the next 5–10 years, as documented through working group reports. The seminar generated a series of research questions which serve as a call to action to the wider community. The unique and contained setting of Schloss Dagstuhl facilitated new cross-disciplinary collaborations and allowed us to lay the groundwork for productive future collaborations, including a planned special issue of the Information Visualization journal.

Seminar Themes

The following high-level themes were discussed during the seminar. The seminar allowed attendees to critically reflect on current research efforts, the state of field, and key research challenges today. Participants also were encouraged to demonstrate their system prototypes and tools relevant to the seminar topics. As a result of the first working groups, as well as impromptu demonstrations and discussions, the actual seminar discussion topics evolved and we established a second set of working groups halfway through the week, cf. Sect. 6.

- **Data Sources and Diversity** What is the current landscape of the application fields and data domains? What are the data gaps? Can existing approaches be generalized?
- **Model Explainability and Interpretability** Can we provide more sophisticated visualizations to study how language models learn or what information they represent?
- **Evaluation and Experimental Designs** Which experimental methods best support the evaluation of techniques and processes for visualizing text information?

- **Interaction Design** What design opportunities are unique to, or more pressing, for text data? How can interaction principles be applied to any underlying NLP as well?
- **Toolkits and Standards** What success stories regarding existing text visualization approaches and systems can we learn from? What is needed?
- **TextVis Literacy** Visual text analytics can be applied across a wide variety of domains. How do we make techniques easy to learn and to interpret correctly?

Outcomes

The Dagstuhl team performed an evaluation at the end of the seminar week. The results of this survey (scientific quality, inspiration to new ideas/projects/research/papers, insights from neighboring fields, ...) were universally very good to excellent. Only a few single improvements were proposed by participants, for example, having longer breaks and mixing up the demo presentations with the other parts of the schedule. Another suggestion was to skip the intermediate group report session because it interrupted the group work.

At the end of the week the organizers agreed to proceed to arrange for a special issue of the journal Information Visualization, which will have an open call but with the intent to include any extended works resulting from the seminar. In addition, several working groups with more "position paper" style reports planed to submit these to well-read venues accepting of editorial works which motivate the research community.

Remaining Challenges in Visual Text Analytics

Not all topics identified during the seminar could be addressed in the working groups and might be left for a future Dagstuhl seminar on a similar subject area. In the following, we briefly list those topics and open problems (more are surely existing that are not mentioned here):

- *Interaction Design:* Interaction methodologies as part of any visual text analytics approach were in the focus of several working groups. A more systematic classification and evaluation of interaction techniques that are unique for text data would be useful for future developments.
- *Toolkits and Standards:* Even if many toolkits and existing standards were discussed in the seminar, a proper and comprehensive analysis of those is still missing that would be beneficial for users and developers of visual text analytics systems.
- *TextVis Literacy:* This topic is important to broaden the use of visual text analytics techniques in general and should be studied deeper in the future.
- *Focus on Text Data Aspects:* The consideration of data diversity, data fusion, and data organization in context of visual text analytics might be an interesting topic for further discussion.
- *Focus on Specific NLP and ML Methods:* The increasing number of specific/novel analytical methods (such as transfer learning or others) raise the need for specific answers from the visual text analytics community.

Acknowledgments

We would like to thank all participants of the seminar for the lively discussions and contributions during the seminar as well as the scientific directorate of Dagstuhl Castle for giving us the possibility of organizing this event. Angelos Chatzimparmpas gathered the abstracts for the overview of the invited talks, the tool demos, and the working groups in Sect. 4, Sect. 5, and Sect. 6, respectively. Once more, we are thankful to all the attendees for agreeing to compose the abstract texts and timely provide them to us in order to write this executive summary. Last but not least, the seminar would not have been possible without the great help of the staff at Dagstuhl Castle. We acknowledge all of them and their assistance.

2 Table of Contents

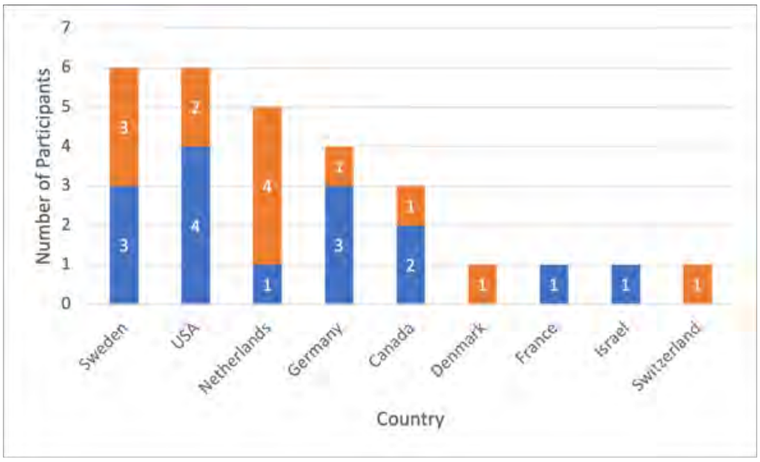

Figure 1 Attendee Statistics of Seminar #22191. Orange colored bars female participants and blue colored bars represent male.

3 Seminar Program and Activities

Christopher Collins (Ontario Tech – Oshawa, CA), Antske Fokkens (Free University Amsterdam, NL), Andreas Kerren (Linköping University – Norrköping, SE), and Chris Weaver (University of Oklahoma – Norman, US)

Participation and Program

This seminar had 28 participants from 9 different countries. Most attendees came from Sweden, USA, and the Netherlands; more attendees came from Germany, Canada, and other European countries as shown in Figure 1. Eight participants have a primary background in linguistics or NLP/ML, and the rest are information visualization and visual analytics experts.

The agenda was focused on providing time for open discussion. Before the seminar, a survey was conducted to collect ideas for discussion topics and open questions from all participants, as well as to solicit initial volunteers for project demonstrations. To engage the two main groups of participants with the richness of the interdisciplinary field, we invited two introductory talks: one on (text) visualization given by Kostiantyn Kucher and one on NLP given by Antske Fokkens. These were intended to contextualize the two fields for the benefit of the attendees from the other field, to give everyone the same general understanding of the combined research space. Following this, two participants, Narges Mahyar and Shimei Pan, gave a joint talk summarizing and reflecting on the survey results from the visualization and NLP perspectives, respectively. The introductory talks and survey summary and reflection set the groundwork for a collaborative brainstorming activity about working group topics. This discussion finalized the working groups for the week, each of which contained at least one member from visualization and NLP.

Working group discussions through the week were interspersed with report back sessions, tool demos, and mini talks in brief plenary sessions twice daily. These opportunities brought the group together to discuss progress and gave diversity to the agenda to keep the event interesting.

Table 1 Final structure of the seminar.

Monday	Tuesday	Wednesday	Thursday	Friday
Opening Remarks (organizers) Self-Introductions	Meeting (logistics) Breakout Groups (first groups)	Meeting (logistics) Breakout Groups (first groups)	Meeting (logistics) Breakout Groups (second groups)	Meeting (logistics) Group Reporting (second groups)
Introductory Talks (InfoVis & NLP, 2 talks)	Breakout Groups (first groups)	Breakout Groups (first groups) Group Reporting (first groups)	Breakout Groups (second groups)	Continued Publication & Closing Remarks
Review of Survey Discussion of Breakout Groups	Breakout Groups (first groups)	Social Event (Saar river tour, brewery visit)	Initial Group Reporting (second groups) Breakout Groups (second groups)	
Discussion of Breakout Groups (cont.) Demo Session (4 talks)	Initial Group Reporting (first groups) Demo Session (4 talks)		Breakout Groups (second groups) Demo Session (4 talks)	

Activities

Introductory Talks

The titles and presenters of the introductory talks for each application domain are listed in the following. Abstracts for the individual talks can be found in Sect. 4.

- Information Visualization
  - *Kostiantyn Kucher and Andreas Kerren:* Introduction to (Text) Visualization
- Natural Language Processing
  - *Antske Fokkens:* NLP: A Very Brief, Free-Style and Improvised Introduction

Tool Demos

In addition, a number of speakers gave a tool demo on a theme related to the research questions of the seminar. In sum, 12 demos were given during the seminar (cf. Sect. 5 for details):

- Hendrik Strobelt: *LMdiff: A Visual Diff Tool to Compare Language Models*
- Jean-Daniel Fekete: *Cartolabe: Visualization of Large Scale Publications Data*
- Nicole Sultanum: *Text Visualization and Close Reading for Journalism with Storifier*
- Yoav Goldberg: *The SPIKE Extractive Search Tool*
- Johannes Knittel: *Real-Time Visual Analysis of High-Volume Social Media Posts*
- Mennatallah El-Assady: *LingVis.io*
- Kostiantyn Kucher: *ALVA: Active Learning and Visual Analytics for Stance Classification*
- Richard Brath: *Visualizing with Text*
- Angelos Chatzimparmpas: *t-viSNE: Interactive Assessment and Interpretation of t-SNE Projections*
- Narges Mahyar: *Supporting Serendipitous Discovery and Balanced Analysis of Online Product Reviews with Interaction-Driven Metrics and Bias-Mitigating Suggestions & CommunityPulse: Facilitating Community Input Analysis by Surfacing Hidden Insights, Reflections, and Priorities*

- Pantea Haghighatkhah: *Story Trees: Representing Documents using Topological Persistence*
- Tatiana von Landesberger: *Network of Names: Visual-Interactive Exploration and Labeling of Entity Relationships*

The content of these talks, given for all seminar attendees, raised further key issues and helped the groups to discuss their individual theme from various perspectives.

Breakout Groups

As already mentioned above, the program included breakout sessions on seven specific topics, i.e., seven working groups discussed one topic at a time in parallel sessions. The themes were based on topics discussed in the original seminar proposal as well as topics that emerged in the first session on Monday afternoon. The detailed working group reports are presented in Sect. 6. In the following, we list the different groups:

1. A Critical Reflection on Uncertainty Localization and Propagation in Text Visualization
2. Annotators and their Data
3. Visual Representations of Text
4. Model Explainability and Interpretability
5. Bias and Bias Mitigation
6. Embedding Representation
7. Evaluation and Experimental Designs

4 Overview of Invited Talks

4.1 Introduction to (Text) Visualization

Kostiantyn Kucher (Linnaeus University – Växjö, SE) and Andreas Kerren (Linnaeus University – Växjö, SE)

Researchers, practitioners, and the general public interested in making sense of text data face issues as the scale of the respective data and the complexity of the respective tasks grow, especially when relying on close reading techniques only. The state-of-the-art computational methods for text analysis demonstrate very impressive results, but they are not always available or self-sufficient for particular tasks and applications, and making sense of the outputs of such methods is often an issue on its own. The methods and solutions offered within the fields of information visualization and visual analytics are thus highly relevant for numerous scenarios involving text data [23].

This talk provides a brief introduction to the respective fields with the intended audience of experts in linguistics, computational linguistics, and machine learning with the intention of establishing the common ground with visualization experts. The more general concepts and approaches are supplemented with the discussion of design spaces and particular examples in text visualization and visual text analytics, including the categorization used by the TextVis Browser online survey (cf. Figure 2).

[1] https://textvis.lnu.se

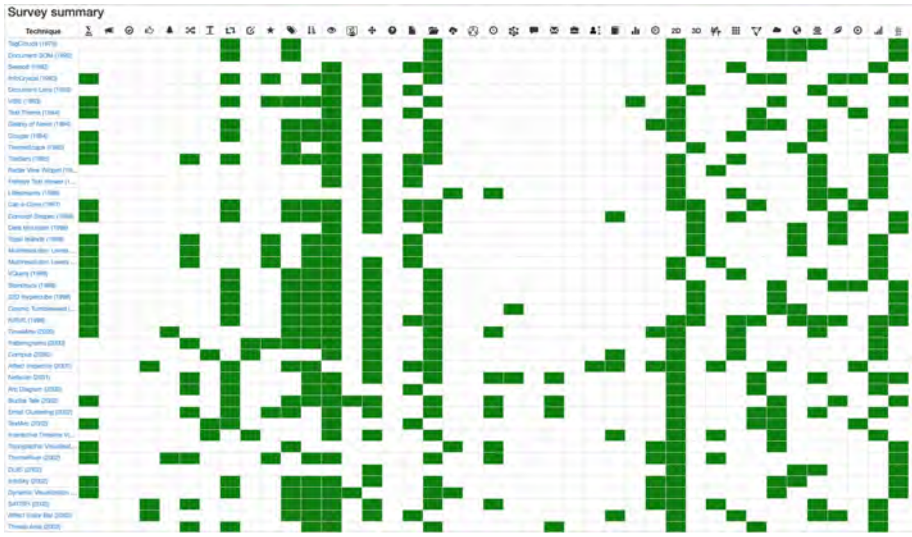

Figure 2 The summary view from the TextVis Browser, an online tool providing an interactive interface to a survey of 400+ text visualization techniques described in peer-reviewed publications and categorized according to the underlying data, high- and low-level tasks, etc. The tool is available online[1].

References

1 Kostiantyn Kucher and Andreas Kerren. Text visualization techniques: Taxonomy, visual survey, and community insights. In *Proceedings of the 2015 IEEE Pacific Visualization Symposium*, PacificVis '15, pages 117–121. IEEE, 2015.

4.2 NLP: A Very Brief, Free-Style and Improvised Introduction

Antske Fokkens (Free University Amsterdam, NL)

This presentation provided an introduction to the latest development of NLP for the members of the visual analytics community. Most invitees had worked with text and language technology before, but not everyone had been exposed to the latest developments around large language models and how they are used.

The first part of the talk mainly served as a reminder for everyone: what is NLP trying to do, in general, regardless of how it is doing this? Why is language difficult (and interesting)? How is the field organized and what are examples of tasks? What approaches were possible in the field until relatively recently (before deep learning became dominant)? With a side note that outside of academia, there are many situations where these "old" methods are preferred. I call this part a "reminder" since I assumed this part of the presentation should sound at least familiar to everyone.

The second part provided a brief introduction to neural networks, starting with those architectures that have been popular in the field: CNNs, BiLSTMs, and typical input representations, such as word embeddings and why these can be so successful compared to more traditional feature representations. After these basics, contextualized language models

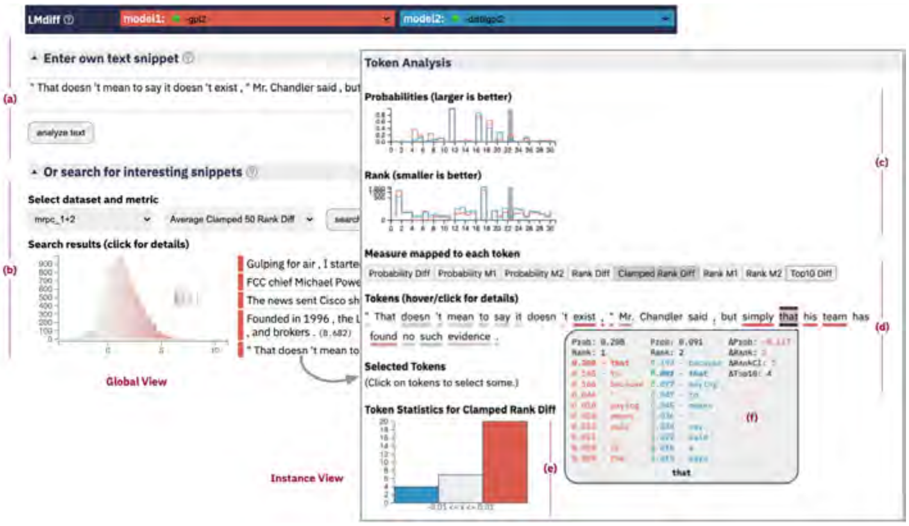

Figure 3 LMdiff interface. The Global View (a,b) allows finding interesting examples which are then selected for in-depth investigation in the Instance View (c-f). More details available online[2].

were introduced in the form of a brief explanation of the idea behind BERT's architecture and how this contextualized language model is trained. I also provided an overview of ways in which language models are trained for specific task, in terms of the input and output representations that are used. Language models raised some questions, in particular in terms of how they deal with language change (which can go fast) and other idiosyncrasies in language use. The answer is that there are ways of dealing with variations in language use, but that it remains the case that machine learning systems, no matter how fancy they get, cannot learn what they have not seen in some way in the data. Though this may seem a trivial thought, both developers and users of such systems sometimes seem to forget things. We ended the presentation with an overview of current research challenges and interests where visual analytics could be useful.

5 Tool Demos

5.1 LMdiff: A Visual Diff Tool to Compare Language Models

Hendrik Strobelt (MIT-IBM Watson AI Lab – Cambridge, US), Benjamin Hoover (MIT-IBM Watson AI Lab – Cambridge, US), Arvind Satyanarayan (MIT CSAIL – Massachusetts Institute of Technology, US), and Sebastian Gehrmann (Google Research – Harvard University, US)

Main reference Hendrik Strobelt, Benjamin Hoover, Arvind Satyanarayan, Sebastian Gehrmann: "LMdiff: A Visual Diff Tool to Compare Language Models", CoRR, Vol. abs/2111.01582, 2021.
URL https://arxiv.org/abs/2111.01582

While different language models are ubiquitous in NLP, it is hard to contrast their outputs and identify which contexts one can handle better than the other. To address this question, we introduce LMdiff [1] (cf. Figure 3), a tool that visually compares probability distributions of two models that differ, e.g., through fine-tuning, distillation, or simply training with different parameter sizes. LMdiff allows the generation of hypotheses about model behavior by investigating text instances token by token and further assists in choosing these interesting text instances by identifying the most interesting phrases from large corpora. We showcase the applicability of LMdiff for hypothesis generation across multiple case studies. A demo is publicly available [3].

References

1 Hendrik Strobelt, Benjamin Hoover, Arvind Satyanaryan, and Sebastian Gehrmann. LMdiff: A visual diff tool to compare language models. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, EMNLP '21, pages 96–105. Association for Computational Linguistics, 2021.

5.2 Cartolabe: Visualization of Large Scale Publications Data

Philippe Caillou (INRIA Saclay – Orsay, FR), Jonas Renault (INRIA Saclay – Orsay, FR), Jean-Daniel Fekete (INRIA Saclay – Orsay, FR), Anne-Catherine Letournel (INRIA Saclay – Orsay, FR), and Michèle Sebag (INRIA Saclay – Orsay, FR)

Cartolabe [4] is an online visualization tool designed to visualize large collections of publication data as maps, such as arXiv [5] (2.5 million documents) and HAL [6] (1 million documents), the French scientific articles repository. Cartolabe [1] tackles several issues related to NLP and visualization. It offers a flexible NLP pipeline to build new maps that can be used to explore the best possible transformations to turn documents into high-dimensional vectors. It offers a complete visualization pipeline that shows a multidimensional projection of millions of documents and allows exploring them through search, pan & zoom.

It is meant to be used by NLP researchers to explore their corpora, or by visualization researchers to improve the visual representations and interactions. The main challenge it faces, along with text visualization, is finding methods to measure the quality of the NLP pipeline to decide if one pipeline is better than another one. Evaluation of NLP pipelines related to human tasks remains an open problem. Cartolabe is open source and can be found at Inria's Gitlab [7].

References

1 Philippe Caillou, Jonas Renault, Jean-Daniel Fekete, Anne-Catherine Letournel, and Michèle Sebag. Cartolabe: A web-based scalable visualization of large document collections. *IEEE Computer Graphics and Applications*, 41(2):76–88, March–April 2021.

3 http://lmdiff.net
4 https://cartolabe.fr
5 https://arxiv.org/
6 https://hal.archives-ouvertes.fr/
7 https://gitlab.inria.fr/cartolabe

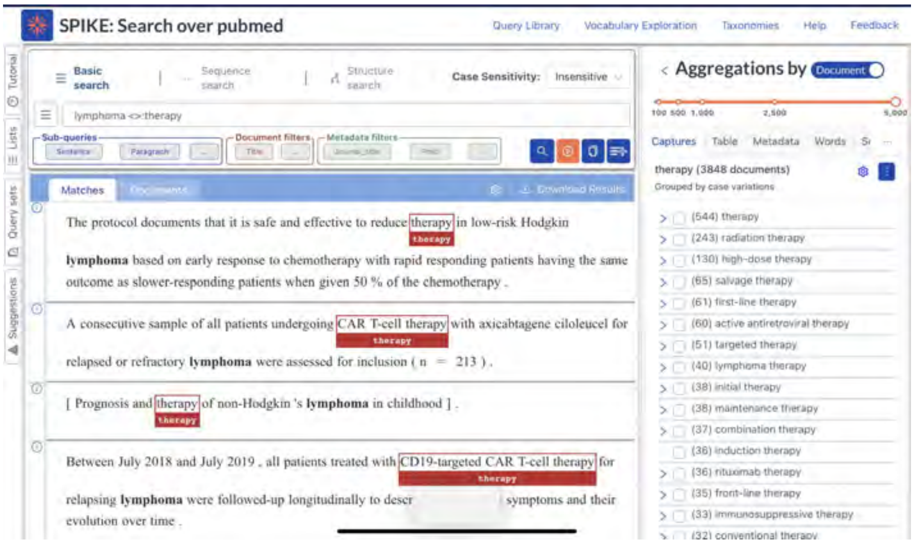

Figure 4 The visual interface of the **SPIKE** system.

5.3 The SPIKE Extractive Search Tool

Shauli Ravfogel (Bar-Ilan University – Ramat Gan, IL), Hillel Taub-Tabib (Allen Institute for AI – Sarona, IL), and Yoav Goldberg (Bar-Ilan University – Ramat Gan, IL)

I presented the SPIKE system [1] (see Figure 4), which is an implementation of a paradigm that we call "extractive search". It combines the traditional search mechanism with rich syntactic and semantic annotations and grep-like "capture slots" over the query, thus allowing to not only locate information, but also to *extract* and *aggregate* focused pieces of information, thus creating knowledge. The extractive search paradigm enables domain experts to perform various kinds of text-mining operations over a corpus using a query language, without the need to program, and to perform corpus exploration. For example, a user may search for a CHEMICAL that is mentioned in the same paragraph as a given disease name (for example, COVID-19), and with the same sentence as the word forms *treat, treatment, or treated.* By designating the CHEMICAL entity as a capture slot, the result of the query will be a list of mentioned chemicals, ranked by their frequencies. These chemicals correspond to COVID-19 treatments from the literature. Similarly, a user my search for the word "lymphoma" together with the word "therapy" and ask to *expand* the word therapy to its linguistic context, and then to *capture* the result, resulting in a ranked list of lymphoma therapies from the literature. Clicking on one of the extracted items focuses the list of results to only mentions of that item, allowing the user to verify the evidence for each result of interest.

References

1 Shauli Ravfogel, Hillel Taub-Tabib, and Yoav Goldberg. Neural extractive search. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: System Demonstrations*, pages 210–217. Association for Computational Linguistics, 2021.

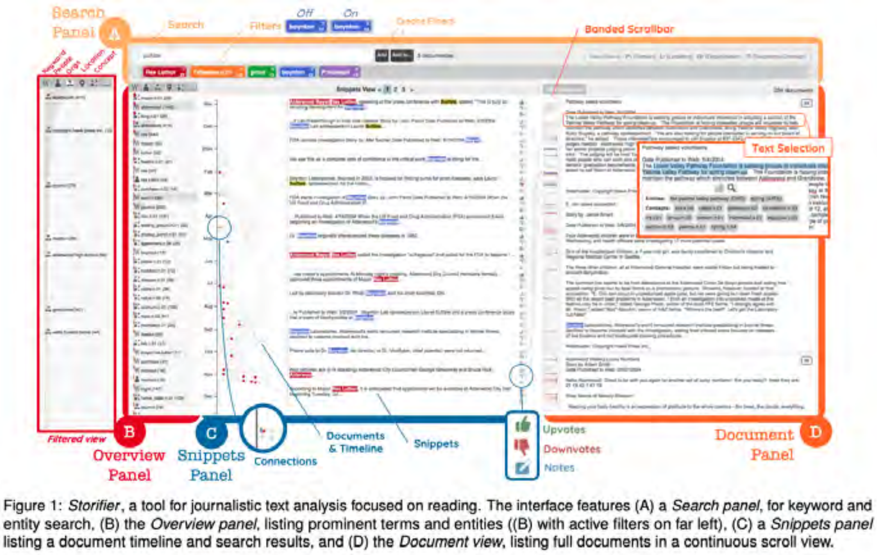

Figure 5 The **Storifier** interface. It features (A) a Search panel, for keyword and entity search, (B) the Overview panel, listing prominent terms and entities ((B) with active filters on far left), (C) a Snippets panel listing a document timeline and search results, and (D) the Document view, listing full documents in a continuous scroll view.

5.4 Text Visualization and Close Reading for Journalism with Storifier

Nicole Sultanum (University of Toronto, CA), Anastasia Bezerianos (INRIA Saclay – Orsay, FR), and Fanny Chevalier (University of Toronto, CA)

At times, interesting datasets come in the form of unstructured text. Data journalists go over these datasets to provide digestible takeaways and to bring important framings to light. Analysis strategies can vary a lot from story to story, but they all take up significant time; and good timing is critical in journalism. While text visualization tools have been helpful at expediting these analyses, journalistic analysis is very broad, and therefore, challenging, for any single tool to support. As such, we proposed a more generalizable approach that makes reading easier and faster to do, instead of making assumptions on desired journalist insights.

In this demo, we presented Storifier [1] (cf. Figure 5), a text visualization tool created in close collaboration with *Ouest France*, a large francophone news office. It is designed to span multiple levels of detail, from a list of frequent and relevant terms, to semantically grouped snippets organized by relevance and time, to full documents, for a complete contextual overview of snippets and evidence tracking. This tool was used in practice by one of our journalist collaborators and led to the publication of a news story.

References

1 Nicole Sultanum, Anastasia Bezerianos, and Fanny Chevalier. Text visualization and close reading for journalism with storifier. In *Proceedings of the 2021 IEEE Visualization Conference*, VIS '21, pages 186–190. IEEE, 2021.

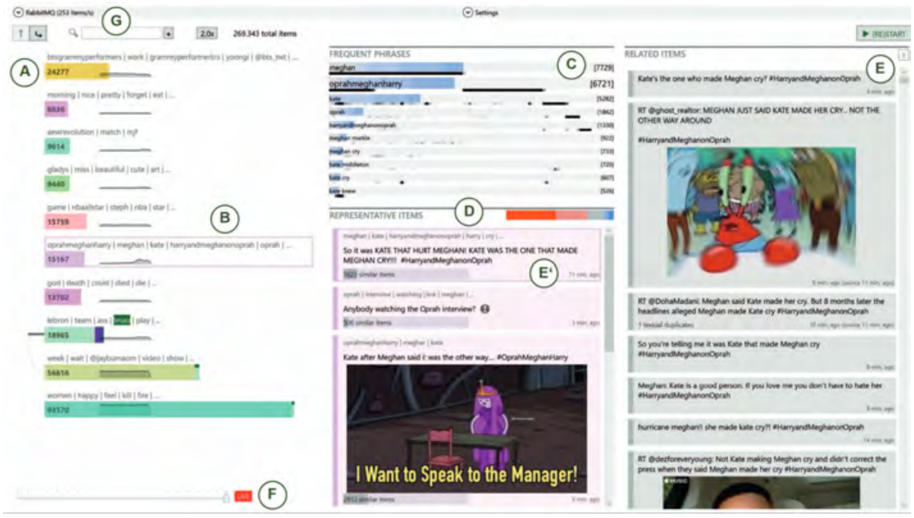

Figure 6 Overview of our system applied to a real-time stream of tweets. A: Topical overview of the 270k posts in the sliding window. B: Selected topic of interest (ToI). C: Visualization of frequent phrases in the ToI. D: Stream of representative posts in the ToI. E: List of similar posts to the representative post E'. F: History slider. G: Dive into topics based on search query or selected topics.

5.5 Real-Time Visual Analysis of High-Volume Social Media Posts

Johannes Knittel (Universität Stuttgart, DE), Steffen Koch (Universität Stuttgart, DE), Tan Tang (State Key Lab of CAD&CG – Zhejiang University, CN), Wei Chen (State Key Lab of CAD&CG – Zhejiang University, CN), Yingcai Wu (State Key Lab of CAD&CG – Zhejiang University, CN), Shixia Liu (Tsinghua University, CN), and Thomas Ertl (Universität Stuttgart, DE)

Breaking news and first-hand reports often trend on social media platforms before traditional news outlets cover them. The real-time analysis of posts on such platforms can reveal valuable and timely insights for journalists, politicians, business analysts, and first responders, but the high number and diversity of new posts pose a challenge. In this demo [1], we presented an interactive system (Figure 6) that enables the visual analysis of streaming social media data on a large scale in real-time. It is based on a new dynamic clustering algorithm that is both efficient and visually explainable. The system provides a continuously updated visualization of the current thematic landscape as well as detailed visual summaries of specific topics of interest. The parallel clustering strategy allows us to provide an adaptive stream with a digestible but diverse selection of recent posts related to relevant topics. We also integrate familiar visual metaphors that are highly interlinked for enabling both explorative and more focused monitoring tasks. Users can gradually increase the resolution to dive deeper into particular topics. In contrast to previous work, our system also works with non-geolocated posts and avoids extensive preprocessing such as detecting events.

References

1 Johannes Knittel, Steffen Koch, Tan Tang, Wei Chen, Yingcai Wu, Shixia Liu, and Thomas Ertl. Real-time visual analysis of high-volume social media posts. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):879–889, January 2022.

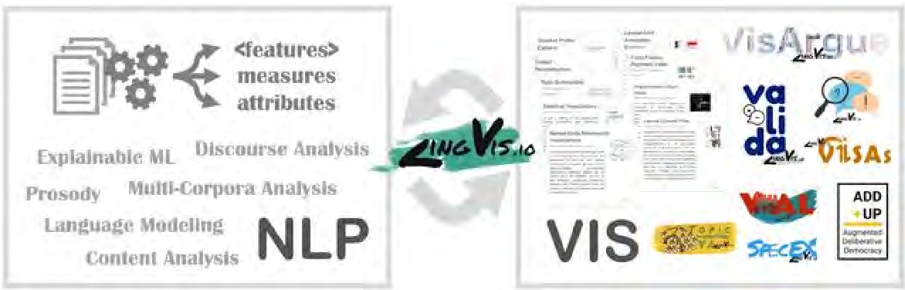

Figure 7 The LingVis.io modular framework. All the projects are available online[8].

5.6 LingVis.io

Mennatallah El-Assady (ETH Zürich, CH), Fabian Sperrle (Universität Konstanz, DE), Rita Sevastjanova (Universität Konstanz, DE), and Wolfgang Jentner (Universität Konstanz, DE)

URL https://lingvis.io/

LingVis.io [2] (see Figure 7) is a modular framework for the rapid-prototyping of linguistic, web-based, visual analytics applications. Our framework gives developers access to a rich set of machine learning and natural language processing steps, through encapsulating them into microservices and combining them into a computational pipeline. This processing pipeline is autoconfigured based on the requirements of the visualization front-end, making the linguistic processing and visualization design detached, independent development tasks. I presented the framework, which continues to support the efficient development of various human-in-the-loop, linguistic visual analytics research techniques and applications. **Concrete demos** can be found below:

- LMFingerprints: Visual Explanations of Language Model Embedding Spaces through Layerwise Contextualization Scores [1] (Figure 8). Demo available online [9].
- Explaining Contextualization through Word Self-Similarity [3] (Figure 9). Demo publicly available [10].
- Visual Comparison of Language Model Adaptation (Figure 10). Demo can be found online [11].

References

1 LMFingerprints: Visual explanations of language model embedding spaces through layerwise contextualization scores. *Computer Graphics Forum*, 41(3):295–307, June 2022.

2 Mennatallah El-Assady, Wolfgang Jentner, Fabian Sperrle, Rita Sevastjanova, Annette Hautli-Janisz, Miriam Butt, and Daniel Keim. lingvis.io – a linguistic visual analytics framework. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, ACL '19, pages 13–18. Association for Computational Linguistics, 2019.

3 Rita Sevastjanova, Aikaterini-Lida Kalouli, Christin Beck, Hanna Schäfer, and Mennatallah El-Assady. Explaining contextualization in language models using visual analytics. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, ACL-IJCNLP '21, pages 464–476. Association for Computational Linguistics, 2021.

9 https://lmfingerprints.lingvis.io/
10 https://embeddings-explained.lingvis.io/
11 https://adapters.demo.lingvis.io/

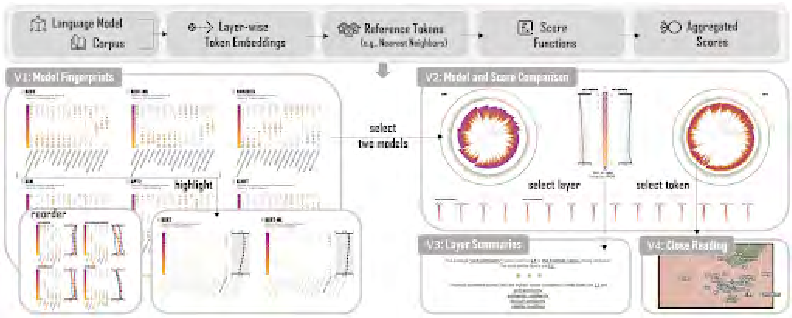

Figure 8 The visual interface of LMFingerprints.

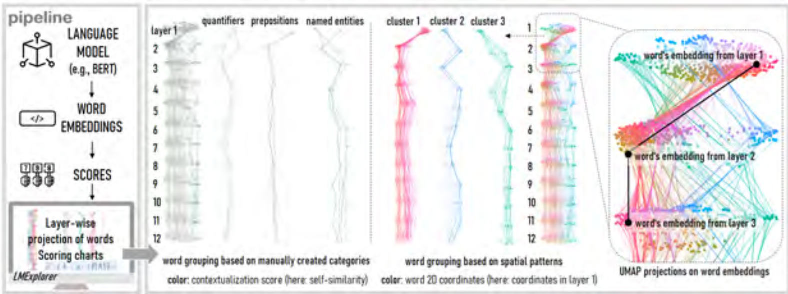

Figure 9 Embeddings explained through visualization.

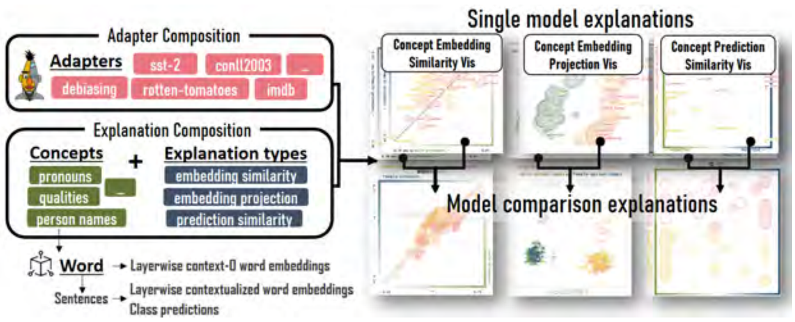

Figure 10 A flexible visual analytics workspace that enables the comparison of adapter properties.

5.7 ALVA: Active Learning and Visual Analytics for Stance Classification

Kostiantyn Kucher (Linnaeus University – Växjö, SE), Carita Paradis (Lund University, SE), Magnus Sahlgren (Swedish Institute of Computer Science and Gavagai AB, SE), and Andreas Kerren (Linnaeus University – Växjö, SE)

This talk provides a brief overview of ALVA [1] (see Figure 11), a visual analytics approach designed to facilitate the entire process of training a text classifier for the problem of multi-label stance classification. ALVA implements the functionality of annotation process management, annotation user interface, and integration with an active learning approach for selecting batches of yet unlabeled utterances (sentences) to include in the next annotation

Figure 11 One of the interactive visual interfaces in ALVA represents the current contents of the annotated dataset with a focus on combinations of labels, the data about the annotated process, and the data about the performance of the respective machine learning classifier trained over time with additional labeled data according to the active learning approach.

round. In order to allow the experts in linguistics and computational linguistics to explore the current contents of the annotated data, ALVA includes an interactive visual interface consisting of multiple views and controls. The overview of individual annotations is provided with a novel visual representation titled CatCombos, which groups the annotations with the same combinations (sets) of categories in separate blocks. Furthermore, ALVA supports visual inspection of the annotation process data (including intra- and inter-annotator agreement scores, for instance) and the performance of the classifier over the course of the active learning process.

References

1 Kostiantyn Kucher, Carita Paradis, Magnus Sahlgren, and Andreas Kerren. Active learning and visual analytics for stance classification with ALVA. *ACM Transactions on Interactive Intelligent Systems*, 7(3), October 2017.

5.8 Visualizing with Text

Richard Brath (Uncharted Software – Toronto, CA)

Direct depiction of text is often very limited in visualizations. The foundations of visualization were built on statistics, cartography, and perceptual psychology, with no consideration of text. Often text is just transformed into statistical data and plotted as a visualization. On the other hand, many examples of text embedded into graphics, art and visualizations can be seen. For example in the artworks at Dagstuhl, such as subway maps, heatmaps, medieval baneroles, and a number quilt by Jill Knuth – which use visualization-like techniques to encode information. Extending the visualization design space to use more text attributes and

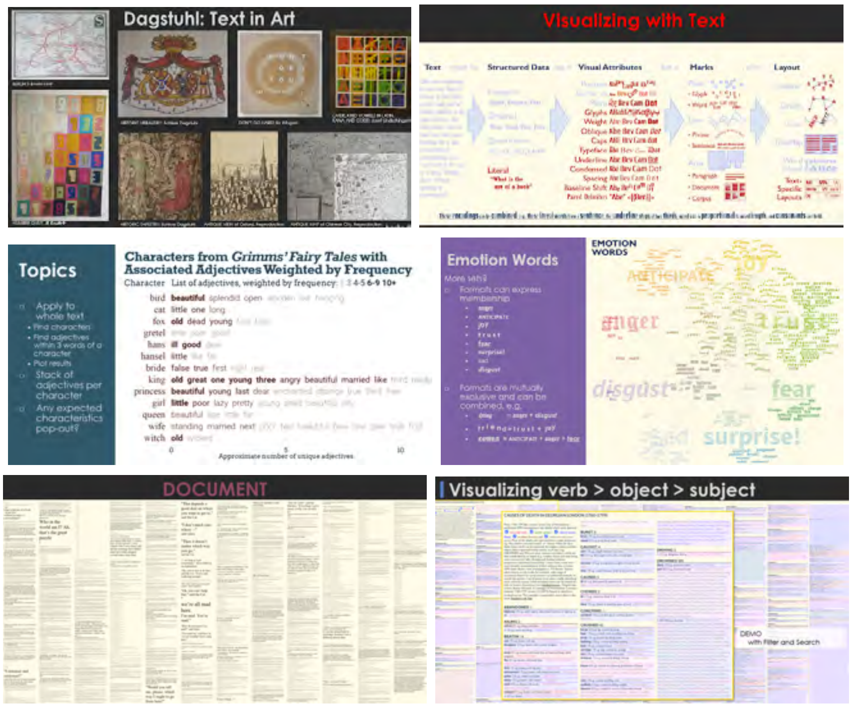

Figure 12 Clockwise from top left: a) Dagstuhl artwork using techniques such as font size, weight, color, layout to indicate data; b) a design space for data visualization extended for text (indicated in red); c) a set of words belonging to one or more emotions, indicated via color and font-style; d) a hierarchy of verbs, objects and subjects indicating causes of death from coroner inquests; e) the full text of Alice in Wonderland, with the most cited text on the Internet sized the largest; and f) adjectives most frequently associated with characters in fairy tales.

text elements allows for a wider range of encoding and expressing literal data in visualizations. Examples (and demos) shown include characterizations instead of word clouds; microtext line charts, set membership indicated by typographic attributes, rhyme patterns, and dictionary-like hierarchies instead of visualization hierarchies such as treemaps (see Figure 12). Live demos are available online [12].

5.9 t-viSNE: Interactive Assessment and Interpretation of t-SNE Projections

Angelos Chatzimparmpas (Linnaeus University – Växjö, SE), Rafael Messias Martins (Linnaeus University – Växjö, SE), and Andreas Kerren (Linnaeus University – Växjö, SE)

Several dimensionality reduction (DR) techniques exist with the goal of identifying similarities in multi-dimensional data and conveying them to users. These methods intend to produce a low-dimensional representation of high-dimensional data that preserves as much of its

[12] https://richardbrath.wordpress.com/books-and-chapters-by-richard-brath/visualizing-with-text-book-companion-web-site/#Demos

Figure 13 The visual interface of **t-viSNE**. It consists of: (a) a panel for changing datasets, choosing between grid search and a specific set of parameters, and saving (or loading earlier) executions; (b) an overview projection with color-encoded class labels (if there are any); (c) the Shepard Heatmap for inspecting the overall quality of the embedding; (d) the Density and Remaining Cost distributions; (e) a list of projections worthy of further exploration, sorted based on quality metrics; (f) the main visualization displaying the Density of neighborhoods in the multi-dimensional space and the Remaining Cost of every point; (g) the Neighborhood Preservation plot for examining the local quality of the selected group of points; (h) the three available interaction modes and a reset button; (i) the visual mapping panel with various functionalities such as the annotator; (j) the Dimension Correlation view revealing the correlations between the dimensions; and (k) the Adaptive Parallel Coordinates Plot for ranking the dimensions from the most to the least important, depending on users' selection.

local and/or global structure as feasible. Consequently, a distinct group of instances could be visualized as a well-separated cluster of points in either two or three dimensions. A well-known DR algorithm is t-Distributed Stochastic Neighbor Embedding (t-SNE), which became popular because of its usefulness in creating low-dimensional representations that accurately capture complex patterns from the high-dimensional space. However, the intrinsic complexity of t-SNE has generated questions about the reliability of the results and the high level of difficulty in understanding them. Indeed, t-SNE projections might be challenging to comprehend or even deceptive, thus undermining the credibility of the extracted insights. Furthermore, understanding the specifics of t-SNE and the rationale for certain patterns in its output can be demanding, especially for those unfamiliar with DR.

In this demo, we presented **t-viSNE** [1] (see Figure 13), a web-based tool for the visual investigation of t-SNE projections that allows analysts to examine their quality and meaning from various perspectives, such as the effects of hyperparameters, distance and neighborhood preservation, densities and costs of particular neighborhoods, and correlations between dimensions and visual patterns. We deliver an open-source tool with multiple coordinated views for exploring t-SNE projections interactively. The utility and applicability of t-viSNE are illustrated via usage scenarios using real-world datasets. We also conducted a comparative user study to evaluate our tool's effectiveness.

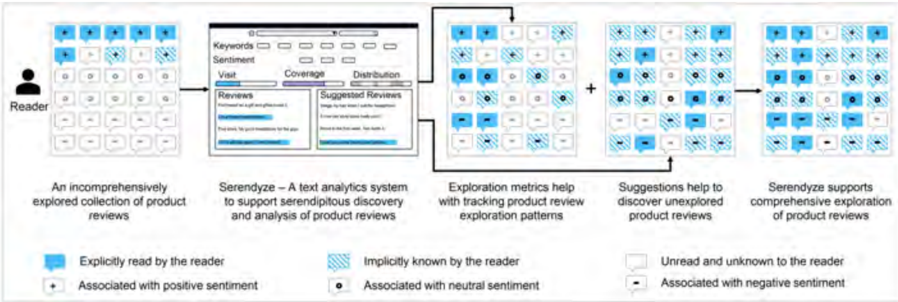

Figure 14 Serendyze is a text analytics system that uses two novel interventions – exploration metrics and a bias mitigation model – to enable readers to explore product reviews more comprehensively. The exploration metrics help readers track their data exploration across different facets, such as sentiments. The bias mitigation model suggests reviews that are semantically and sentiment-wise dissimilar to what the readers have been exploring so that they can discover a broader range of reviews. Integrated within an interactive interface, these features can enable readers to gain comprehensive knowledge about the data prior to decision-making.

References

1 Angelos Chatzimparmpas, Rafael M. Martins, and Andreas Kerren. t-viSNE: Interactive assessment and interpretation of t-SNE projections. *IEEE Transactions on Visualization and Computer Graphics*, 26(8):2696–2714, August 2020.

5.10 Supporting Serendipitous Discovery and Balanced Analysis of Online Product Reviews with Interaction-Driven Metrics and Bias-Mitigating Suggestions

Mahmood Jasim (University of Massachusetts – Amherst, US), Christopher Collins (Ontario Tech – Oshawa, CA), Ali Sarvghad (University of Massachusetts – Amherst, US), and Narges Mahyar (University of Massachusetts – Amherst, US)

Customers of online products often depend on product reviews to make data-driven purchase decisions. These product reviews – free-form text comments from previous customers that highlight their opinions and evaluations of online products – are often considered the most influential factor behind sales and attitudes towards a product. While customers might have different strategies to navigate reviews to make their decisions, those who prefer to comprehensively explore and analyze product reviews often struggle to do so due to the abundance of reviews available and the limited amount of time to accrue insights from them. As such, these customers are often unable to evaluate all available alternatives in-depth, which often results in incomplete exploration and understanding of the underlying product reviews prior to making purchase decisions.

In this demo, I presented **Serendyze** [1] (Figure 14), a text analytics system for supporting serendipitous discovery and analysis of online product reviews. The system includes two interventions – **Exploration Metrics** that can help readers understand and track their exploration patterns through visual indicators and a **Bias Mitigation Model** that intends to maximize knowledge discovery by suggesting sentiment and semantically diverse reviews.

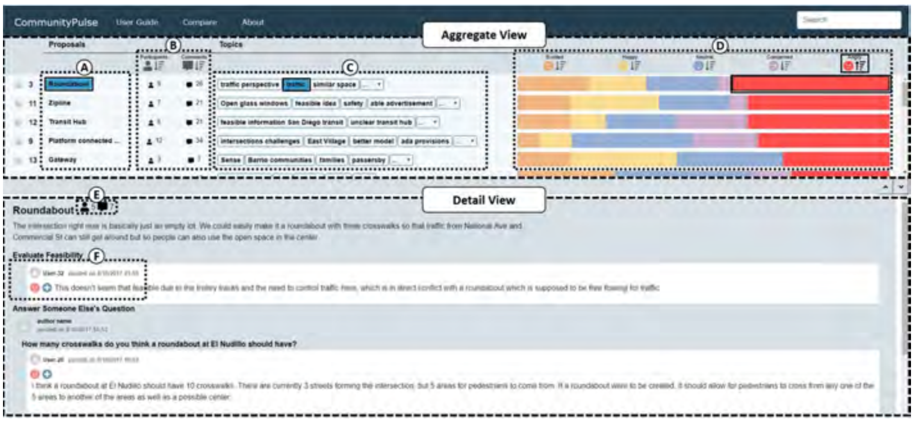

Figure 15 A snapshot of CommunityPulse. The Aggregate View shows: (A) a list of Proposals (Roundabout is selected), (B) the number of people and comments for each proposal, (C) a list of topics for each proposal (Traffic is selected), and (D) emoticons to sort the proposals based on emotions and stacked bar charts to present people's emotion distribution and drill down to actual comments (In this view, the proposals are sorted by Angry emotions and angry comments from Roundabout are selected). The Detail View is rendered and updated based on the filters used in the Aggregate View. This example shows (E) Meta-information based on the user-selected angry comments, and (F) user information for each comment, with icons to represent associated emotion and option to save the comment as a note.

These interventions are intended to support serendipitous discovery and analysis to help readers cover the reviews more comprehensively and tease apart valuable insights from reviews in a balanced way.

To evaluate our approach, we asked 100 crowd workers to use Serendyze to make purchase decisions based on product reviews. Our evaluation suggests that exploration metrics enabled readers to efficiently cover more reviews in a balanced way, and suggestions from the bias mitigation model influenced readers to make confident data-driven decisions. In the paper, we discuss the role of user agency and trust in text-level analysis systems and their applicability in domains beyond review exploration.

References

1 Mahmood Jasim, Christopher Collins, Ali Sarvghad, and Narges Mahyar. Supporting serendipitous discovery and balanced analysis of online product reviews with interaction-driven metrics and bias-mitigating suggestions. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22. Association for Computing Machinery, 2022.

5.11 CommunityPulse: Facilitating Community Input Analysis by Surfacing Hidden Insights, Reflections, and Priorities

Mahmood Jasim (University of Massachusetts – Amherst, US), Enamul Hoque (York University – Toronto, CA), Ali Sarvghad (University of Massachusetts – Amherst, US), and Narges Mahyar (University of Massachusetts – Amherst, US)

Increased access to online engagement platforms has created a shift in civic practice, enabling civic leaders to broaden their outreach to collect a larger number of community input, such as comments and ideas. However, sensemaking of such input remains a challenge due to the unstructured nature of text comments and ambiguity of human language. Hence, community input is often left unanalyzed and unutilized in policymaking.

In this demo, I presented **CommunityPulse** [1] (Figure 15), an interactive system that combines text analysis and visualization to scaffold different facets of community input. To design the system, we conducted a formative study where we interviewed 14 civic leaders to understand their practices and requirements. We identified challenges around organizing the unstructured community input and surfacing community's reflections beyond binary sentiments. Our evaluation with another 15 experts suggests CommunityPulse's efficacy in surfacing multiple facets such as reflections, priorities, and hidden insights while reducing the required time, effort, and expertise for community input analysis.

References

1 Mahmood Jasim, Enamul Hoque, Ali Sarvghad, and Narges Mahyar. CommunityPulse: Facilitating community input analysis by surfacing hidden insights, reflections, and priorities. In *Proceedings of the Designing Interactive Systems Conference 2021*, DIS '21, pages 846–863. Association for Computing Machinery, 2021.

5.12 Story Trees: Representing Documents using Topological Persistence

Pantea Haghighatkhah (TU Eindhoven, NL), Antske Fokkens (Free University Amsterdam, NL), Pia Sommerauer (Free University Amsterdam, NL), Bettina Speckmann (TU Eindhoven, NL), and Kevin Verbeek (TU Eindhoven, NL)

The primary emphasis of topological data analysis (TDA) is the inherent shape of (spatial) data. As a result, it could be considered an effective way to investigate spatial representations of linguistic data (embeddings), which have recently become a prominent component of NLP. In this demo, we presented TDA as a means to express document structure – which is a method so-called story trees [1]. Story trees are hierarchical representations produced via persistent homology using semantic vector representations of sentences (see Figure 16). They may be used to recognize and vividly picture key elements of a storyline. Finally, we also demonstrated their capabilities by using story trees to generate extractive summaries for news stories.

References

1 Pantea Haghighatkhah, Antske Fokkens, Pia Sommerauer, Bettina Speckmann, and Kevin Verbeek. Story trees: Representing documents using topological persistence. In *Proceedings of the Language Resources and Evaluation Conference*, LREC '22, pages 2413–2429. European Language Resources Association, 2022.

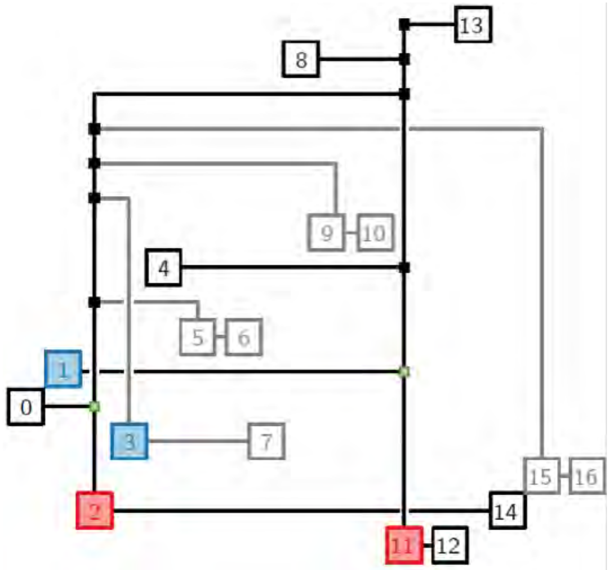

Figure 16 Story trees in action. The example with ID 60 is from the CNN/Daily Mail validation set. Salient STL sentences are red, k-center sentences are blue, grey sentences are irrelevant side-stories which are pruned from the tree.

5.13 Network of Names: Visual-Interactive Exploration and Labeling of Entity Relationships

Artjom Kochtchi (Technische Universität Darmstadt, DE), Martin Müller (Technische Universität Darmstadt, DE), Kathrin Ballweg (Technische Universität Darmstadt, DE), Tatiana von Landesberger (Technische Universität Darmstadt, DE), Seid M. Yimam (University of Hamburg, DE), Uli Fahrer (University of Hamburg, DE), Chris Biemann (University of Hamburg, DE), Marcel Rosenbach (Spiegel-Verlag, DE), Michaela Regneri (Spiegel-Verlag, DE), and Heiner Ulrich (Spiegel-Verlag, DE)

We present a visual analytics system [3] that enables to explore and annotate relationships between named entities extracted from large document collections. The relationships are visualized in a node-link diagram. Search, show important and expand on demand strategy is used to explore the relationships of interest to the user. Novel degree of interest (DOI) function enables to explore user-specified types of relationships (as in our more recent publication [2], see Figure 17). Visual and textual user-edited annotations are provided that are supported by algorithmic completion.

References

1 Artjom Kochtchi, Tatiana von Landesberger, and Chris Biemann. Networks of names: Visual exploration and semi-automatic tagging of social networks from newspaper articles. In *Proceedings of the 16th Eurographics Conference on Visualization*, EuroVis '14, pages 211–220. Eurographics Association, 2014.

[13] https://www.newsleak.io/

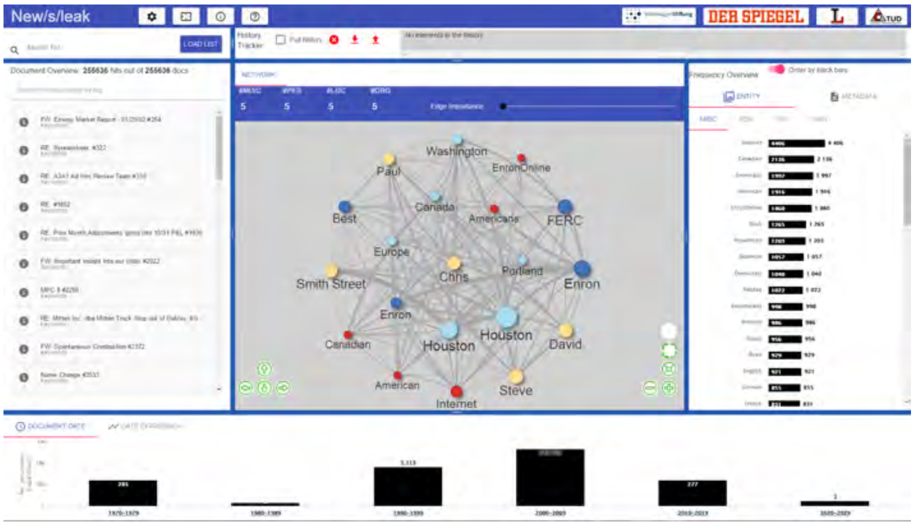

Figure 17 The user interface of the proposed software prototype. Demo available online[13].

2 Martin Müller, Kathrin Ballweg, Tatiana von Landesberger, Seid M. Yimam, Uli Fahrer, Chris Biemann, Marcel Rosenbach, Michaela Regneri, and Heiner Ulrich. Guidance for multi-type entity graphs from text collections. In *Proceedings of the EuroVis Workshop on Visual Analytics*, EuroVA '17, pages 1–6. Eurographics Association, 2017.

6 Working Groups

This section describes results from each of the seven working groups and identifies the attendees contributing to each group. The names of those people who reported for the working groups are underlined.

6.1 WG: A Critical Reflection on Uncertainty Localization and Propagation in Text Visualization

Bettina Speckmann, Carita Paradis, Jean-Daniel Fekete, Mennatallah El-Assady, Narges Mahyar, Pantea Haghighatkhah, and Vasiliki Simaki

Our group discussed the uncertainty propagation pipeline for visual text analytics. In our first attempt to clarify what we identify as **uncertainty** and what we consider **artifacts** and **errors**, we came up with the following definition:

> "**Uncertainty** refers to epistemic[14] situations involving imperfect or unknown information[15], which may come from different sources. If it comes from data, we have no control over it; if it comes from the process, we may/may not know the imperfection."

[14] https://en.wikipedia.org/wiki/Epistemology
[15] https://en.wikipedia.org/wiki/Information

Our discussion resulted in distinguishing **artifacts** as a consequence of the realm and something that can be corrected. **Errors** as deliberate and systematic problems that can not be corrected.

Then, we used dimensionality reduction as a use case to depict two levels of uncertainty that can be propagated in the pipeline; (1) *the internal pipeline uncertainty*; and (2) *the external semantic uncertainty.* Both levels directly impact the perceived and interpreted uncertainty at the end of the pipeline. In the following, we describe the different types of uncertainty we identified based on their occurrence in the pipeline (Figure 18).

(1) **Semantic uncertainty:** As Figure 18 shows, the first type of uncertainty that we identified in our discussion is located at the text production level. We named it *semantic uncertainty* on the part of the producer, with the speaker expressing uncertainty about his/her sayings, opinions, facts, or ideas, and it is usually communicated with the use of markers like *may*, *not sure*, *might*, *could*, e.g., *I am not sure how to get there* [19].

(2) **Comprehension uncertainty:** The second type of uncertainty that we identified is located at the data capturing and annotation level, and we named it *comprehension uncertainty.* Uncertainty issues at this level can be caused during the data collection process, with representativeness, balance, noise, and bias being the main factors causing uncertainty. But uncertainty is caused during the data annotation process, as text ambiguity, vague or generic annotation guidelines and the annotators' different perceptual systems can lead to different annotation decisions and thus more uncertainty about the reliability of the annotated data. For instance, in *The fruit is too soft for me* there is uncertainty on the part of the annotator with respect to the meaning of *soft* whether it is about the texture, the touch, or the smell of the fruit. There is uncertainty in the example of *What is your position?* With respect to the interpretation of *position* whether it is about a job, a posture, or an opinion [12, 18, 17].

(3) **Encoding uncertainty:** The labeled data is then mapped to a data structure, which could be a lossy representation of the input, resulting in *encoding uncertainty.*

(4) **Transformation uncertainty:** The encoding typically represents the data as embedding vectors in a high-dimensional space. These get transformed through NLP models that are either exclusively considering the internal data from the pipeline, or additionally rely on external resources, such as in language modeling or externalizing expert knowledge and feedback. The NLP models introduce *transformation uncertainty* into the pipeline.

(5) **Representation uncertainty:** The output of NLP models is another set of high-dimensional vectors that are usually represented on a visual interface through dimensionality reduction. Generating the visualizations introduces *representation uncertainty.*

(6) **Interpretation uncertainty:** Finally, at the last stage of the pipeline the receiver or analyst inspecting the visualization is interpreting the data through the lens of the visual design, their *interpretation uncertainty* is due to the mindset of the receiver (user).

6.1.1 Related work

How uncertainty in text vis is different that uncertainty visualization in general?

Prior work has addressed the importance of visualizing uncertainty (E.g., [7, 9, 10, 13, 16]). However, we argue that text visualization needs more careful consideration for uncertainty, its sources, and potential ways to visualize them [3, 3, 11].

The uncertainty in text visualization comes from many origins. First, we need to consider that text is an imperfect representation of human thoughts, therefore encoding thoughts in a text by nature produces artifacts. Another issue is that people come with different

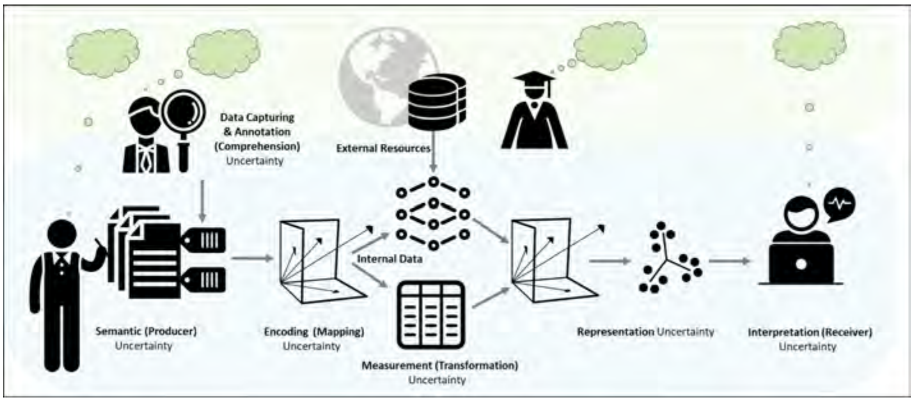

Figure 18 This figure shows a dimensionality reduction pipeline as one potential use case to show six types of uncertainties that can be propagated through the text visualization pipeline.

understanding and interpretations of the text. Hence, one single text input can result in multiple interpretations which not only affect the interpretation of the receiver but also annotators, because they might have various interpretations. The Text Encoding Initiative (TEI) Guidelines [5], designed to define best practices to encode textual sources with a rich vocabulary of annotations, mention several mechanisms for encoding uncertainty, such as "levels of certainty" and "precision" in the chapter "Certainty, Precision, and Responsibility", and encoding for text segments such as "unclear", "gap" for the transcription of text or speech. All of these textual or linguistic uncertainties are idiosyncratic and intrinsic to our languages, texts, and speech structures. The TEI also allows encoding alternative interpretations for the same text segment (using the <choice> element), as well as marking visible errors (<sic>) and possible corrections (<corr>). These annotations can become very rich and a currently not supported in a consistent way by visualization systems; they are mostly ignored.

Multidimensional Projection

Visualizing large document corpora is often done using multidimensional project techniques (also called dimensionality reduction techniques) such as t-SNE [20] or UMAP [15]. These techniques start by computing a distance between documents, such that two related documents are closer than two less related documents. There are many methods to compute the distances, from the older "bag of words" to the more recent ones using deep learning such as doc2vec, Bert, and GPT-3. From these distances, the projection methods represent a document as a point that should be placed in a 2D position such that the distances between the documents are proportional to the distances between the points. The projection methods are very effective at computing an overview but they also introduce geometric distortions and topological artifacts, and this is unavoidable. Therefore, visualizations of high-dimensional data through projections should provide mechanisms to inform users about the artifacts and, if possible, overcome them. In general, when two points are close, the documents they represent are similar, but sometimes, two points are "false neighbors". Conversely, sometimes, two points should be close by because the documents they represent are similar, but they end up being far away, they are "missing neighbors". These two artifacts cause misleading interpretations without visual warnings [1].

More generally, visualizations based on multidimensional projections are the last step of a longer analytical pipeline that can inject errors, uncertainty, distortions, and various kinds of artifacts that will end up in the final visualization. If they are not explicitly managed by the visualization technique, they lead to errors and a lack of trust [7]. Visualizing artifacts and uncertainty lead to different techniques that can sometimes be combined but always complexify the visual representation and the interaction.

Visualizing Artifacts

There have been a few articles on techniques for visualizing topological artifacts created by multidimensional projections. Overall, projections maintain local and global geometry and topology but always produce local errors (artifact). These artifacts, when not noticed, lead to errors or uncertainty. For example, a point close to a dense group could be part of the dense group (the cluster) if faithfully located, or can be erroneously located too close and misleads the user.

Therefore, a visualization should at least inform users of these possible errors, preferable indicate where (areas) where they do not happen and, if possible, allow resolving them. Currently, few visualization systems inform users of possible artifacts, and almost none provide techniques to overcome them, especially for a large number of points. Addressing that problem is very important for text visualization in particular, but also for multidimensional visualization in general. Aupetit [1] and Heulot et al. [8] have proposed a few techniques for small amounts of data. Martins et al. [14] for larger amounts but they are not understandable by large audiences.

In addition to artifacts due to the projections, uncertainty can be also intrinsic to data coming out of the NLP pipeline. The simplest form would be a scalar value associated with each point expressing its degree of certainty. In that case, simple visualization methods can be used, such as "Value Suppressing Uncertainty Palettes" by Correl et al. [6].

References

1 Michaël Aupetit. Visualizing distortions and recovering topology in continuous projection techniques. *Neurocomputing*, 70(7):1304–1330, 2007. Advances in Computational Intelligence and Learning.
2 Eric P. S. Baumer, Mahmood Jasim, Ali Sarvghad, and Narges Mahyar. Of course it's political! A critical inquiry into underemphasized dimensions in civic text visualization. *Computer Graphics Forum*, 2022. To appear.
3 Eric P. S. Baumer and Micki McGee. Speaking on behalf of: Representation, delegation, and authority in computational text analysis. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, AIES '19, pages 163–169. Association for Computing Machinery, 2019.
4 Jason Chuang, Daniel Ramage, Christopher Manning, and Jeffrey Heer. Interpretation and trust: Designing model-driven visualizations for text analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 443–452. Association for Computing Machinery, 2012.
5 TEI Consortium. TEI P5: Guidelines for electronic text encoding and interchange. http://www.tei-c.org/Guidelines/P5/, 2022.
6 Michael Correll, Dominik Moritz, and Jeffrey Heer. Value-suppressing uncertainty palettes. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 1–11. Association for Computing Machinery, 2018.
7 Miriam Greis, Jessica Hullman, Michael Correll, Matthew Kay, and Orit Shaer. Designing for uncertainty in HCI: When does uncertainty help? In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '17, pages 593–600. Association for Computing Machinery, 2017.

8 Nicolas Heulot, Michaël Aupetit, and Jean-Daniel Fekete. ProxiLens: Interactive exploration of high-dimensional data using projections. In *Proceedings of the EuroVis Workshop on Visual Analytics using Multidimensional Projections*, VAMP '13. The Eurographics Association, 2013.
9 Jake M. Hofman, Daniel G. Goldstein, and Jessica Hullman. How visualizing inferential uncertainty can mislead readers about treatment effects in scientific results. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, pages 1–12. Association for Computing Machinery, 2020.
10 Jessica Hullman. Why authors don't visualize uncertainty. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):130–139, January 2020.
11 Mahmood Jasim, Enamul Hoque, Ali Sarvghad, and Narges Mahyar. CommunityPulse: Facilitating community input analysis by surfacing hidden insights, reflections, and priorities. In *Proceedings of the Designing Interactive Systems Conference*, DIS '21, pages 846–863. Association for Computing Machinery, 2021.
12 Steven Jones, M. Lynne Murphy, Carita Paradis, and Caroline Willners. *Antonyms in English: Construals, Constructions and Canonicity*. Studies in English Language. Cambridge University Press, 2012.
13 Matthew Kay, Tara Kola, Jessica R. Hullman, and Sean A. Munson. When (ish) is my bus? User-centered visualizations of uncertainty in everyday, mobile predictive systems. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 5092–5103. Association for Computing Machinery, 2016.
14 Rafael M. Martins, Danilo Barbosa Coimbra, Rosane Minghim, and Alexandru C. Telea. Visual analysis of dimensionality reduction quality for parameterized projections. *Computers & Graphics*, 41:26–42, June 2014.
15 Leland McInnes, John Healy, and James Melville. UMAP: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.
16 Lace Padilla, Matthew Kay, and Jessica Hullman. *Uncertainty Visualization*, pages 1–18. John Wiley & Sons, Ltd, 2021.
17 Carita Paradis. Conceptual spaces at work in sensory cognition: Domains, dimensions and distances. In *Applications of Conceptual Spaces*, pages 33–55. Springer, 2015.
18 Carita Paradis. *Meanings of words: Theory and application*, pages 274–294. De Gruyter, 2015.
19 Vasiliki Simaki, Carita Paradis, Maria Skeppstedt, Magnus Sahlgren, Kostiantyn Kucher, and Andreas Kerren. Annotating speaker stance in discourse: The Brexit Blog Corpus. *Corpus Linguistics and Linguistic Theory*, 16(2):215–248, 2020.
20 Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9(86):2579–2605, 2008.

6.2 WG: Annotators and their Data

Alex Endert, Angelos Chatzimparmpas, Christofer Meinecke, Christopher Collins, José Angel Daza Arévalo, Maria Skeppstedt, Ross Maciejewski, and Tatiana von Landesberger

Textual datasets come in varying sizes and are found across a large number of application domains. In many instances, there are no corpora of sufficient sizes for state-of-the-art machine learning algorithms to be effectively applied. This problem is even more acute for

Figure 19 A picture of our group taken moments after the discussion started.

low-resource and under-resourced languages, and one major bottleneck is often the need for annotating and hand-curating a training dataset. A solution for these areas is to use pre-trained language models and then to fine-tune a large pre-trained model or apply transfer learning to the specific task or language; however, this still requires some data annotation. For specialized domains, it would be important to leverage the expert knowledge and incorporate it into the models instead of trying to only treat the annotator as a person that provides labels to the data. Given these challenges, there is an overarching need for interactive tools that can support domain experts in annotation, model development, model comparison and model transfer. Despite a large number of possibilities available to visualize and explore a text collection on which you aim to apply NLP models, it is still common to blindly use the available annotation resources (which sometimes are scarce) to label the data according to the specific task you have in mind.

Some questions we (Figure 19) had in this space were how to support the injection of expert knowledge into the labeling and annotation process. For example, labeling functions could be created to guide the first round annotation process using simple rules, but these labels would need to be refined and confirmed by someone familiar with the data. If the expert creates a rule that labels documents containing a specific word as "A", but another rule would label this document "B", how do we support the resolution of this conflict? What is the return on investment of work and time for the annotator to resolve these conflicts, in terms of improvement in trained model quality or the achievement of the ultimate task?

A more efficient approach could be to first explore the data. For instance:

- In order to know which pre-trained language model would be most suitable to fine-tune/apply transfer learning to, in order to achieve the task aimed for.
- In order to know if existing methods can be applied for generating pre-annotations of your data that the annotators can correct/adjust, speeding up the annotation process. E.g., to apply an existing machine learning model to the data, or to apply heuristic rules (stemming from the expert knowledge of the annotator) to the data, or to use a taxonomy/vocabulary for performing the aimed NLP task.
- Also in the annotation process, visualization can help. E.g., if active learning is used for actively selecting data to annotate from a pool of unlabeled data, the state of the pool of unlabeled data could be visualized.

6.2.1 Improving the experience of annotation

Given the underlying annotation needs, we wanted to consider how to support data exploration and improve the process of data annotation, making it more enjoyable for the annotator. Considerations included gamification techniques, or showing the impact of the time and effort through visualizing the improvement of model quality and annotation coherence during the work. Ideally, we want to create a system that could perhaps visualize the expected impact of completing a particular annotation, in order to incentivize the human annotator to participate in the interactive active learning loop. This could be translated into three steps, when it might be possible to use visualization to support the annotator in exploring and annotating the text that is to be annotated.

Step 1: How well can an existing model be applied to your data? Try different models and visualize the areas in the texts (or the data points in the data) that are surprising to the model.

Step 2: How well can heuristics rules, stemming from the expert knowledge of the annotator be applied to your data? Will there be conflicting points in your labeling function that stem from these heuristic rules? How can the labeling, and the conflicting labeling decisions be visualized? Can weak supervision help you? We discussed a few concrete ideas of how sets and weak supervision [5] could be used for visualizing when different heuristic rules, annotator decisions, and the output of machine learning models conflict (see Sect. 6.2.2).

Step 3: How can the annotator be given control of the active learning process. For example, how can the classification certainty of an active learning model be visualized in order to let the annotator use their expert knowledge to be in charge of actively selecting what data to annotate? Visualize the data in different ways in order to guide the annotator in how to choose data points to annotate.

6.2.2 Project idea: Expert annotation tool

Even though general commercial labeling tools already exist, such as LabelStudio and Prodigy, we discussed the need for a tool that allows people who have particular expertise in their data to annotate and label meaningful words, phrases, syntactic patterns, and other items in their texts. Such a tool would differ from the traditional labeling tools in that experts often have knowledge about their data and the domain which they would like to communicate to the system and model. Doing so by repetitively adding and correcting categorical labels on words misses out on the opportunity to capture this information directly. For instance, people may have heuristics about the syntactic structure of specific sentences that imply meaning, paragraph lengths that help them detect authorship or intentional misspellings of words that reveal hidden meanings in the texts. A tool that allows people to translate such domain knowledge into heuristics, rules, or otherwise learned model parameters may be particularly helpful in cases where the data is small. For example, Mehta et al. [4] showed how close reading can be supported by a tablet interface that allows experts to annotate their text while reading. The system uses these annotations as a method for generating recommendations of relationships to additional unseen parts of the text. Kochtchi et al. [3] proposed a visual system for users to annotate relationships between named entities extracted from the text. The user-defined annotations are processed and annotation rules are automatically extracted and applied to unexplored parts of the text. In general, human-based annotation enrichments can be assisted by a visual analytics tool aiming to reduce the effort of generating label functions from scratch. This also requires high-quality annotations and the models. As models may be uncertain about labels or lead to conflicts in labels.

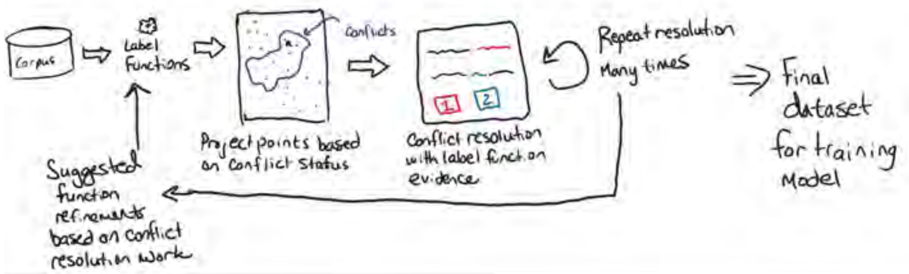

Figure 20 A pipeline for an expert annotation tool for refining labeling functions which are used for labeling training data for future machine learning. Starting with a set of rule-based labeling functions, conflicting function labels are identified and grouped in a visualization. Instances of conflicting labels are manually resolved. Suggestions for label function refinements are learned from the corrections. The iterative process continues until label conflicts are satisfactorily resolved, leaving a labeled dataset for training future models.

This uncertainty and the conflicts need to be communicated and resolved. Ambiguities or uncertainties in labels (multi-labels) can be seen as overlapping sets [1] that could show which labels are problematic and whether there are any patterns in the problematic labels that would lead to new annotation rules [2]. On the other side, models could also be updated and improved based on human-injected knowledge using the visually-supported label function creation (cf. Figure 20). The comparison of the human vs. machine-produced models and the correction of the misalignment between them may be another open research opportunity.

References

1. Bilal Alsallakh, Luana Micallef, Wolfgang Aigner, Helwig Hauser, Silvia Miksch, and Peter Rodgers. The state-of-the-art of set visualization. *Computer Graphics Forum*, 35(1):234–260, February 2016.
2. Jürgen Bernard, Matthias Zeppelzauer, Michael Sedlmair, and Wolfgang Aigner. VIAL: A unified process for visual interactive labeling. *The Visual Computer: International Journal of Computer Graphics*, 34(9):1189–1207, September 2018.
3. Artjom Kochtchi, Tatiana von Landesberger, and Chris Biemman. Networks of names: Visual exploration and semi-automatic tagging of social networks from newspaper articles. In *Proceedings of the 16th Eurographics Conference on Visualization*, EuroVis '14, pages 211–220. Eurographics Association, 2014.
4. Hrim Mehta, Adam Bradley, Mark Hancock, and Christopher Collins. Metatation: Annotation as implicit interaction to bridge close and distant reading. *ACM Transactions on Computer-Human Interaction*, 24(5), November 2017.
5. Alexander Ratner, Stephen H. Bach, Henry Ehrenberg, Jason Fries, Sen Wu, and Christopher Ré. Snorkel: Rapid training data creation with weak supervision. *The VLDB Journal*, 29(2):709–730, May 2020.

6.3 WG: Visual Representations of Text

Andreas Kerren, Antske Fokkens, Barbara Plank, Chris Weaver, Kostiantyn Kucher, Nicole Sultanum, Tatiana von Landesberger, and Yoav Goldberg

Language is complex. NLP provides ways to extract information from different levels of text (paragraph level, sentence level, phrase or entity level, etc.). We are interested in displaying related pieces of information (related paragraphs, related sentences, or related phrases) in a way which exposes their similarities and allows aggregations over similar content – in this regard, besides the more general work on text visualization or information visualization, we can consider the approaches for representing text alignment in particular [12], for instance. However, similarity phenomena are complex [2], and textual items can relate to each other in multiple different aspects, and in multiple layers of similarity. In addition, the order of similarity may not be linear. We distinguish two main forms of similarity. The first is *linguistic similarity*, which refers to two expressions exhibiting the same linguistic phenomenon (e.g., containing negation or a specific syntactic structure). The second is referential similarity (which can vary from the same referent, e.g., a specific department in a university, to referents of the same type, e.g., an educational institution). We first outline the challenges involved in visualizing this and then provide an overview of potential starting points based on prior work in information visualization and visual analytics.

6.3.1 Linguistic similarity

Computational linguistics has a tradition in creating rich evaluation sets (e.g., Lehmann et al. [7]). Recently, interest in tests that are carefully designed to cover specific linguistic phenomena has regained interest in the community, e.g., the introduction of checklists for evaluation [9]. Such datasets can provide valuable insights into which phenomena have been learned by a model and which not. The downside of looking at carefully designed sets is that it is not necessarily clear how representative they are for naturally occurring data. We have benchmark data that has been created by annotating naturally occurring samples for most tasks, but these are small and it is often hard to tell how representative they are when using models in the “real world”.

The best of both worlds would therefore be to combine the two and investigate various linguistic phenomena in real-world data. Additionally, it can be informative to explore the actual occurrence of various phenomena to gain insight into whether the success or failure of models to deal with them correctly matters for real world applications. If we can identify and show what phenomena actually occur, we can also gain insight into whether existing benchmarks are sufficiently similar to the data we intend to apply our models to and, thus, whether reported results are indicative of how the model will perform on our data. Visualizations that would support such exploration of data would therefore be highly beneficial for the field.

6.3.2 Referential similarity

As an example, consider a set of sentences describing people’s academic achievements. Items in such a set may include:

1. *Alice obtained her PhD in Computational Linguistics from Saarbrucken University.*

2. *Bob majored from MIT Business School with a Business Administration Master's degree.*
3. *Cam has a PhD from MIT.*
4. *Dan has a Master's degree in Business Administration from Saarbrucken Business School.*

One point of similarity between these sentences is their topic, and, more narrowly, the event they describe (a person obtaining some degree from some academic institution). They also share various forms of linguistic similarity such as syntax, for instance, all sentences are in active form, all have a prepositional clause starting with "from", etc.

But there are also other levels of similarity. For example, the obtained degree in items (1) and (3) is both a *"PhD"*, and in items (1) and (2) and (4) is both *"Master's"*. Moreover, both Master's degrees are degrees in *"Business Administration"*, although the syntactic structure in which this information is realized differs: *"Business Administration Master's degree"* vs. *"Master's degree in Business Administration"*.

Similarly, the *institution* slot is similar between items (1) and (4) (*"Saarbrucken University"*) and between items (2) and (3) (*"MIT"*). However, note that the institution can also be grouped differently, noting the similarity between (2) and (4), which are both *"Business School"*. If we had also a sentence about a person obtaining a degree from LMU, it would have been similar to items (1) and (4) if we were to consider the country in which the academic institution is based. We can also consider grouping the person who obtained their degree by, for example, their gender. Note that this introduces a notion of *ambiguity* or *under-specification*: the gender of *"Cam"* in sentence (3) cannot be determined based on the text alone. The *topic* of Cam's degree is also not specified. Visualizations that support exploring similarity at these levels can be of high value for non-technical end users as well as NLP experts. For non-expert users it is particularly important that the interface is intuitive and supports identifying potentially useful relations they are not a *priori* aware of.

The use cases include both corpus exploration on the individual text level (what is the structure of the individual items) as well on aggregations on the entire corpus level, and encompass either single arguments (*"which German schools are represented in the corpus"*) or collections of arguments (*"show me all PhDs graduates in Business from Saarbrucken"*).

6.3.3 Information visualization perspective

If we consider the traditional InfoVis Reference Model by Card et al. [4], the first two steps of the overall process/pipeline involve transforming raw data into *"data tables"*, which for our purposes of visual text analytics could be compared to the definition of facets/frames that should eventually be revealed by the visual representation. One issue of relying on this conceptual model in the context of dealing with text data is the richness, complexity, and ambiguity of text in comparison with the standard case of n-dimensional multivariate tabular data, which was the default case for the InfoVis Reference Model and general-purpose visualization tools.

The typical scenario of representing text data with (interactive) visual representations that we can find in the related work is driven by the particular choice of user tasks (both from the point of higher-level analytical concerns, such as revealing the main concepts, similarities, or opinions in the text data, or lower-level exploration concerns, such as providing overview or supporting navigation over a collection of text documents) and the available data. Here, we could consider an example such as *PEARL* by Zhao et al. [13] which focuses on the exploration of emotions expressed in a single user's messages on Twitter – the visual representations used by this tool rely on the particular set of computational analyses which were selected according to the task, but also the particular data formats and the respective constraints and peculiarities. PEARL could thus be viewed as a specialized solution, but cannot be considered a general-purpose text visualization approach.

Jigsaw [10] is another meaningful example to partly illustrate some of the data framing concepts (and issues) for text data representation. It was created to support intelligence analysts in investigative work to uncover hidden patterns. Its underlying data structures contain relationships between entities (people, organizations, location) and documents, which can be visualized in different ways, via a multitude of views. Each *Jigsaw* view was purposefully designed to support a specific facet of the data, and multiple views are needed to provide complementary perspectives on the data. Compared to the expectations of NLP experts, the design of *Jigsaw* lacks the support for representing the relationships between entities across facets and at *multiple levels* of hierarchies and scales present in the data (e.g., the semantic hierarchies, the entity relations, the structural progression of content, etc.)—such a general visual text representation technique would be very valuable for experts and lay users alike for a variety of use cases.

Considering the generalizability spectrum, the solutions such as *Voyant Tools* [16] lie closer to the more general side, while providing support for several common linguistic analyses and rather lightweight visual representations, including word clouds [11], which seem to be appreciated by the general public and even users from particular knowledge domains, however, this representation is rather controversial with respect to the perceptual considerations and usability concerns, as studied in the related work [1, 5, 17], and it also does not address a number of possible higher- and lower-level tasks.

6.3.4 Towards flexible visual text analytics pipelines

The considerations mentioned above led us to another iteration of our discussion, with the deficiencies of the existing visual representations of text for complex, yet generic scenarios motivating the need to reconsider the underlying data structures, analyses, and tasks besides the visual encoding or interaction technique on their own.

One of models/frameworks that we found to be relevant to such scenarios is the multilayer network model, which recognizes that the complexity of relationships between entities in real-world applications is better embraced as several interdependent layers/levels rather than a simple graph approach usually used in network visualization. A recent book publication provides an overview of the current state of the art in this emerging field [8].

For example, each word or phrase in a sentence can have a type (*location*, *person*, *organization*, ...). Each type can be seen as a layer within a multilayer network. Moreover, the types (e.g., locations) can be part of a hierarchy (e.g., *city* → *area* → *state* → *continent*) that are also layers of the hierarchical structure. The similarity between the sentences can thus be seen as a multilayer network, where each aspect of the similarity is a type of element in the sentence. Each word or phrase in the sentence would be a node of the network and similarities would be connections between elements. The order of words/phrases forms a sequence which needs to be considered in the network visualization. Thus, the exploration of similarities and their semantics can be seen as an analogous exploration process in multilayer networks. The view on words, phrases, sentences, paragraphs, and documents relates to multi-scale exploration of networks. This opens interesting ideas and challenges from a visualization perspective.

[16] https://voyant-tools.org/

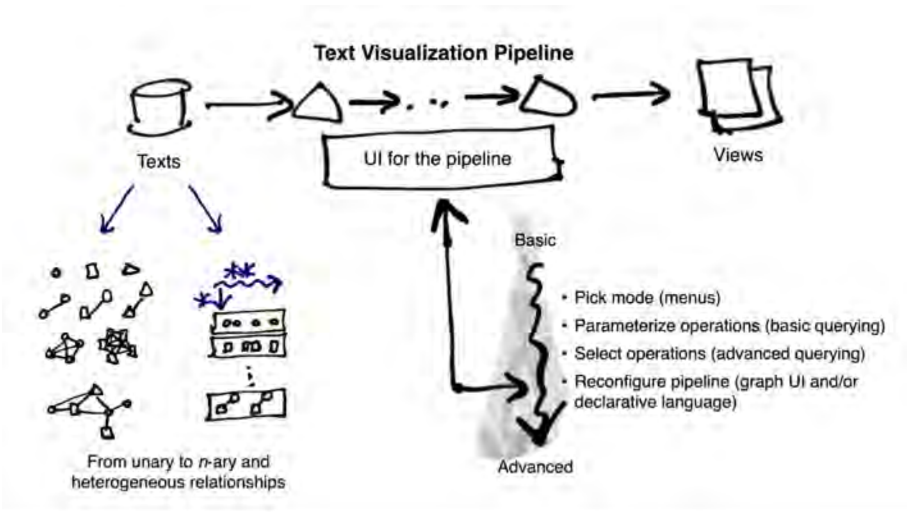

Figure 21 A sketch of a flexible text data visualization pipeline configured according to the needs of the respective visual text analytic application.

One example of an existing visual analytic approach relevant to this discussion is *FacetAtlas* [3]. *FacetAtlas* embodies some of the principles of representing rich and complex information while allowing for dynamic exploration via a compact representation of multi-layered concept networks extracted from text. Each layer represents a different facet, and relationships are computed via topic similarity; layers are superimposed and facet-level relationships are encoded as set overlaps.

Taking into account all such expectations from the experts and lay users, data models and computational approaches, and visual representation and interaction concerns, we envision one possible way forward as a flexible pipeline of data transformation and representation for text data (see Figure 21), which starts with capturing the facets and relationships on the data and domain-specific analyses side (which might include ordinal, n-ary, and heterogeneous relationships – which can also be related to the multi-layer network framework) and providing (parameterized) visual encodings that reveal the contents of these facets, but also patterns of interactions across these facets/frames.

References

1 Eric Alexander, Chih-Ching Chang, Mariana Shimabukuro, Steven Franconeri, Christopher Collins, and Michael Gleicher. Perceptual biases in font size as a data encoding. *IEEE Transactions on Visualization and Computer Graphics*, 24(8):2397–2410, August 2018.

2 Daniel Bär, Torsten Zesch, and Iryna Gurevych. A reflective view on text similarity. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing*, RANLP '11, pages 515–520. Association for Computational Linguistics, 2011.

3 Nan Cao, Jimeng Sun, Yu-Ru Lin, David Gotz, Shixia Liu, and Huamin Qu. FacetAtlas: Multifaceted visualization for rich text corpora. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1172–1181, November–December 2010.

4 Stuart K. Card, Jock D. Mackinlay, and Ben Shneiderman. *Readings in Information Visualization: Using Vision to Think.* Morgan Kaufmann Publishers Inc., 1999.

5 Cristian Felix, Steven Franconeri, and Enrico Bertini. Taking word clouds apart: An empirical investigation of the design space for keyword summaries. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):657–666, January 2018.

6 Marti A. Hearst, Emily Pedersen, Lekha Patil, Elsie Lee, Paul Laskowski, and Steven Franconeri. An evaluation of semantically grouped word cloud designs. *IEEE Transactions on Visualization and Computer Graphics*, 26(9):2748–2761, September 2020.

7 Sabine Lehmann, Stephan Oepen, Sylvie Regnier-Prost, Klaus Netter, Veronika Lux, Judith Klein, Kirsten Falkedal, Frederik Fouvry, Dominique Estival, Eva Dauphin, Hervè Compagnion, Judith Baur, Lorna Balkan, and Doug Arnold. TSNLP: Test suites for natural language processing. In *Proceedings of the 16th Conference on Computational Linguistics – Volume 2*, COLING '96, pages 711–716. Association for Computational Linguistics, 1996.

8 Fintan McGee, Benjamin Renoust, Daniel Archambault, Mohammad Ghoniem, Andreas Kerren, Bruno Pinaud, Margit Pohl, Benoît Otjacques, Guy Melançon, and Tatiana von Landesberger. Visual analysis of multilayer networks. *Synthesis Lectures on Visualization*, 8(1):1–150, 2021.

9 Marco Tulio Ribeiro, Tongshuang Wu, Carlos Guestrin, and Sameer Singh. Beyond accuracy: Behavioral testing of NLP models with CheckList. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, ACL '20, pages 4902–4912. Association for Computational Linguistics, 2020.

10 John Stasko, Carsten Gorg, Zhicheng Liu, and Kanupriya Singhal. Jigsaw: Supporting investigative analysis through interactive visualization. In *Proceedings of the 2007 IEEE Symposium on Visual Analytics Science and Technology*, VAST '07, pages 131–138. IEEE, 2007.

11 Fernanda B. Viégas and Martin Wattenberg. Tag clouds and the case for vernacular visualization. *Interactions*, 15(4):49–52, July 2008.

12 Tariq Yousef and Stefan Janicke. A survey of text alignment visualization. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1149–1159, February 2021.

13 Jian Zhao, Liang Gou, Fei Wang, and Michelle Zhou. PEARL: An interactive visual analytic tool for understanding personal emotion style derived from social media. In *Proceedings of the 2014 IEEE Conference on Visual Analytics Science and Technology*, VAST '14, pages 203–212. IEEE, 2014.

6.4 WG: Model Explainability and Interpretability

Daniel A. Keim, Hendrik Strobelt, Johannes Knittel, Pia Sommerauer, Richard Brath, and Shimei Pan

In recent years, NLP has become data-driven with very large models such as BERT, and GPT-3 providing an unprecedented performance. The latest models have billions of variables and need enormous amount of data and computing resources for training. While results in general are of high quality, there are numerous applications where explainability is of high importance [17], such as medical diagnosis or bias detection [18] [19]. We, a group of visualization researchers and NLP researchers (Figure 22), discussed when, why, and how interactive visualization has a valuable role in explaining NLP models (XAI4NLP), especially considering

[17] https://medium.com/cortico/visualizing-toxicity-in-twitter-conversations-3cd336e5db81
[18] https://www.brookings.edu/research/detecting-and-mitigating-bias-in-natural-language-processing/
[19] https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G

Figure 22 Group at work during lunch break.

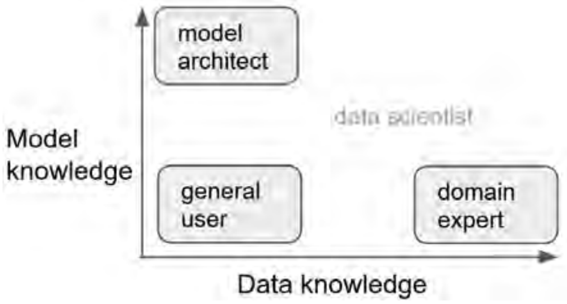

Figure 23 User groups and their knowledge with respect to models and data.

large pre-trained neural language models such as GPT-3. Explainability, interpretability, and model analysis offer slightly different perspectives on making neural models more transparent and interpretable [20].

There are many stakeholders with different levels of expertise and domain knowledge that can take advantage of XAI4NLP, but the goals and methods typically vary. We decided to consider this field from the perspective of different users who can be grouped by different goals. We identified three main target user groups (Figure 23): model architects and builders, researchers and domain experts, and general users and consumers that do not necessarily have any background in computer science and may not have deep data knowledge. Model architects and builders may not have extensive knowledge about the data their models are trained on, but they have strong expertise in which models to use when, and how they can be adapted for different target goals. Researchers, data scientists, and domain experts, on the other hand, can be placed on a broad spectrum of expertise. They range from computational linguistics and NLP experts with a detailed understanding of neural models, to experts with deep knowledge of domain-specific data, but little technical understanding. Average consumers, however, rarely know any specifics about data or models. When considering explainability using visualization, it is important to consider the goals and skills of each user group.

Different NLP approaches involve different opportunities and challenges for model in-

[20] https://pair.withgoogle.com/explorables/fill-in-the-blank/

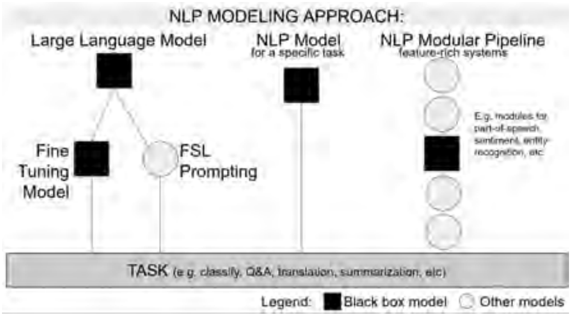

Figure 24 Different NLP modeling approaches.

terpretability (Figure 24). We decided to focus on three major approaches: The current state-of-the-art models tend to employ large-pre-trained language models and fine-tune them for particular tasks. However, other, more traditional architectures still exist. In our discussions we generally considered approaches that make use of neural models at some stage. For instance, traditional pipeline approaches can employ 'black box' neural models for particular steps in the pipeline, but are somewhat transparent in the sense that input and output can be inspected at various steps in the pipeline. Task-specific end-to-end models, however, cannot provide this transparency. The framework of pre-training and fine-tuning poses even more challenges, as the training data of pre-trained language models are extremely large and often not accessible. In some cases, the pre-trained models themselves are not available for inspection.

It is important to note that, for some tasks, there is little benefit in using interactive visualizations. For instance, visualizations typically rarely play a role in confirming hypotheses in a formal way, and for well-defined goals (e.g., text search) we may be better off using automatic methods or just basic visualizations. Interactive visualizations, though, are particularly helpful for exploring models and data, as well as for generating novel hypotheses. Beyond curiosity-driven research, NLP experts should be able to understand and predict how models behave when employed in real-world scenarios. In addition, visualizations support communicating results and insights to different audiences, to consumers as well as experts.

We discussed the importance of interacting with visualizations, particularly to scale the analysis of large NLP models (Figure 25). Analysts may either explore the data first using an overview visualizations, which allows interactive drill-downs for more specific analyses related to insights or hypotheses that they have gained thus far, or they may analyze or filter the data first and investigate a particular subset of the data and/or model. Furthermore, analysts may steer the model interactively to fit it to their needs, which allows for incorporating domain knowledge into the model.

Using text to visualize text is generally a useful technique, but we need additional means to scale this encoding (e.g., highlighting, filtering, aggregation). For large NLP models, we need multi-level scalable visualization techniques that are able to process and analyze massive amounts of training data, billions of parameters, and highly connected graphs.

We identified different points in the three different approaches at which different visualizations can be used.

Debiasing use case: Expert modelers may need to understand where bias is coming from, such as the model or training date, both of which are massive visualization challenges given the scale of data in large language models.

Labeling data use case: For constructing fine-tune models, such as classifiers, etc., data

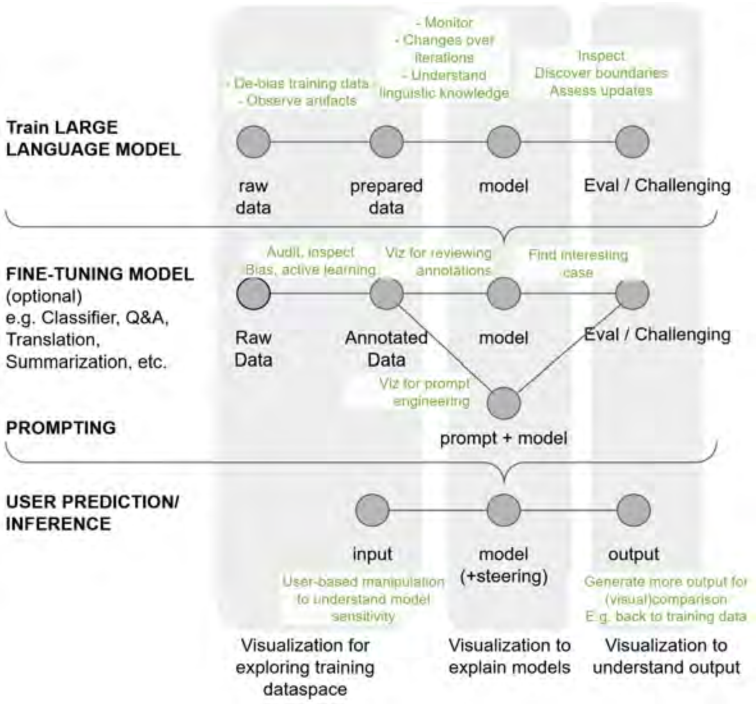

Figure 25 Areas for explainability (in green text) in relation to NLP processes in relation to the large language model, optional fine-tuning or prompting model, and use of model in inference mode.

needs to be annotated by humans. Search, navigation, filtering and marking appropriate examples benefit from visualization techniques to show relevant text in detail as well as related examples to aid increasing the performance of the annotation task. Access to these same viusalizations aids downstream explainability and interpretability tasks to assess the annotations used to derive the model.

Model structure use case: Understanding model structure aids insights into relations, such as high-dimension vector spaces of word embeddings, or layers and nodes within a language model assuming there is access to the internals of the model.

References

1 Bilal Alsallakh, Allan Hanbury, Helwig Hauser, Silvia Miksch, and Andreas Rauber. Visual methods for analyzing probabilistic classification data. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1703–1712, December 2014.

2 Rie Kubota Ando and Tong Zhang. A framework for learning predictive structures from multiple tasks and unlabeled data. *Journal of Machine Learning Research*, 6(61):1817–1853, 2005.

3 Galen Andrew and Jianfeng Gao. Scalable training of L1-regularized log-linear models. In *Proceedings of the 24th International Conference on Machine Learning*, ICML '07, pages 33–40. Association for Computing Machinery, 2007.

4 Mahzarin R. Banaji and Anthony G. Greenwald. *Blindspot: Hidden Biases of Good People.* Bantam, 2013.

5 Yonatan Belinkov, Nadir Durrani, Fahim Dalvi, Hassan Sajjad, and James Glass. What do neural machine translation models learn about morphology? *arXiv preprint arXiv:1704.03471*, 2017.
6 Yonatan Belinkov and James Glass. Analysis methods in neural language processing: A survey. *Transactions of the Association for Computational Linguistics (TACL)*, 7:49–72, 2019.
7 Jürgen Bernard, Marco Hutter, Matthias Zeppelzauer, Dieter Fellner, and Michael Sedlmair. Comparing visual-interactive labeling with active learning: An experimental study. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):298–308, January 2018.
8 Angie Boggust, Brandon Carter, and Arvind Satyanarayan. Embedding Comparator: Visualizing differences in global structure and local neighborhoods via small multiples. In *Proceeding of the 27th International Conference on Intelligent User Interfaces*, IUI '22, pages 746–766. Association for Computing Machinery, 2022.
9 Tolga Bolukbasi, Kai-Wei Chang, James Y Zou, Venkatesh Saligrama, and Adam T Kalai. Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. *Advances in Neural Information Processing Systems*, 29, 2016.
10 Richard Brath. *Visualizing with Text*. CRC Press, 2020.
11 Anamaria Crisan, Margaret Drouhard, Jesse Vig, and Nazneen Rajani. Interactive model cards: A human-centered approach to model documentation. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '22, pages 427–439. ACM, 2022.
12 Fahim Dalvi, Avery Nortonsmith, Anthony Bau, Yonatan Belinkov, Hassan Sajjad, Nadir Durrani, and James Glass. NeuroX: A toolkit for analyzing individual neurons in neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33 of *AAAI*, pages 9851–9852, 2019.
13 Marina Danilevsky, Kun Qian, Ranit Aharonov, Yannis Katsis, Ban Kawas, and Prithviraj Sen. A survey of the state of explainable AI for natural language processing. In *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*, AACL '20, pages 447–459. Association for Computational Linguistics, 2020.
14 Joseph F. DeRose, Jiayao Wang, and Matthew Berger. Attention flows: Analyzing and comparing attention mechanisms in language models. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1160–1170, February 2021.
15 Sebastian Gehrmann, Hendrik Strobelt, and Alexander Rush. GLTR: Statistical detection and visualization of generated text. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, ACL '19, pages 111–116. Association for Computational Linguistics, 2019.
16 Anthony G. Greenwald, Debbie E. McGhee, and Jordan L. K. Schwartz. Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6):1464–1480, June 1998.
17 Marti A. Hearst, Emily Pedersen, Lekha Patil, Elsie Lee, Paul Laskowski, and Steven Franconeri. An evaluation of semantically grouped word cloud designs. *IEEE Transactions on Visualization and Computer Graphics*, 26(9):2748–2761, September 2020.
18 Florian Heimerl and Michael Gleicher. Interactive analysis of word vector embeddings. *Computer Graphics Forum*, 37(3):253–265, June 2018.
19 Florian Heimerl, Steffen Koch, Harald Bosch, and Thomas Ertl. Visual classifier training for text document retrieval. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2839–2848, December 2012.
20 Evan Hernandez and Jacob Andreas. The low-dimensional linear geometry of contextualized word representations. In *Proceedings of the 25th Conference on Computational Natural*

Language Learning, CoNLL-EMNLP '21, pages 82–93. Association for Computational Linguistics, 2021.

21 Minsuk Kahng, Pierre Y. Andrews, Aditya Kalro, and Duen Horng Chau. ActiVis: Visual exploration of industry-scale deep neural network models. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):88–97, January 2018.

22 Johannes Knittel, Steffen Koch, and Thomas Ertl. PyramidTags: Context-, time- and word order-aware tag maps to explore large document collections. *IEEE Transactions on Visualization and Computer Graphics*, 27(12):4455–4468, December 2021.

23 Kostiantyn Kucher and Andreas Kerren. Text visualization techniques: Taxonomy, visual survey, and community insights. In *Proceedings of the 2015 IEEE Pacific Visualization Symposium*, PacificVis '15, pages 117–121. IEEE, 2015.

24 Johannes Lang and Miguel A. Nacenta. Perception of letter glyph parameters for infotypography. *ACM Transactions on Graphics (Proceedings of the SIGGRAPH 2022)*, 2022. To appear.

25 Jiwei Li, Xinlei Chen, Eduard Hovy, and Dan Jurafsky. Visualizing and understanding neural models in NLP. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, NAACL '16, pages 681–691. Association for Computational Linguistics, 2016.

26 Yuxin Ma, Arlen Fan, Jingrui He, Arun Reddy Nelakurthi, and Ross Maciejewski. A visual analytics framework for explaining and diagnosing transfer learning processes. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1385–1395, February 2021.

27 Burt L. Monroe, Michael P. Colaresi, and Kevin M. Quinn. Fightin' words: Lexical feature selection and evaluation for identifying the content of political conflict. *Political Analysis*, 16(4):372–403, February 2017.

28 Mohammad S. Rasooli and Joel Tetreault. Yara Parser: A fast and accurate dependency parser. *arXiv preprint arXiv:1503.06733*, 2015.

29 Clément Rebuffel, Marco Roberti, Laure Soulier, Geoffrey Scoutheeten, Rossella Cancelliere, and Patrick Gallinari. Controlling hallucinations at word level in data-to-text generation. *Data Mining and Knowledge Discovery*, 36(1):318–354, 2022.

30 Victor Sanh, Albert Webson, Colin Raffel, Stephen H. Bach, Lintang Sutawika, Zaid Alyafeai, Antoine Chaffin, Arnaud Stiegler, Teven Le Scao, Arun Raja, Manan Dey, M. Saiful Bari, Canwen Xu, Urmish Thakker, Shanya Sharma, Eliza Szczechla, Taewoon Kim, Gunjan Chhablani, Nihal V. Nayak, Debajyoti Datta, Jonathan Chang, Mike T.-J. Jiang, Han Wang, Matteo Manica, Sheng Shen, Zheng-Xin yong, Harshit Pandey, Michael Mckenna, Rachel Bawden, Thomas Wang, Trishala Neeraj, Jos Rozen, Abheesht Sharma, Andrea Santilli, Thibault Fevry, Jason A. Fries, Ryan Teehan, Tali Bers, Stella Biderman, Leo Gao, Thomas Wolf, and Alexander M. Rush. Multitask prompted training enables zero-shot task generalization. In *Proceedings of the Tenth International Conference on Learning Representations*, ICLR '22, 2022.

31 Maarten Sap, Dallas Card, Saadia Gabriel, Yejin Choi, and Noah A. Smith. The risk of racial bias in hate speech detection. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, ACL '19, pages 1668–1678. Association for Computational Linguistics, 2019.

32 Naomi Saphra and Adam Lopez. Understanding learning dynamics of language models with SVCCA. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, NAACL '19, pages 3257–3267. Association for Computational Linguistics, 2019.

33 Hendrik Strobelt, Sebastian Gehrmann, Michael Behrisch, Adam Perer, Hanspeter Pfister, and Alexander M. Rush. Seq2seq-Vis: A visual debugging tool for sequence-to-sequence

models. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):353–363, January 2019.

34 Hendrik Strobelt, Sebastian Gehrmann, Hanspeter Pfister, and Alexander M. Rush. LSTMVis: A tool for visual analysis of hidden state dynamics in recurrent neural networks. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):667–676, January 2018.

35 Tony Sun, Andrew Gaut, Shirlyn Tang, Yuxin Huang, Mai ElSherief, Jieyu Zhao, Diba Mirza, Elizabeth Belding, Kai-Wei Chang, and William Yang Wang. Mitigating gender bias in natural language processing: Literature review. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, ACL '19, pages 1630–1640. Association for Computational Linguistics, 2019.

36 Ian Tenney, Dipanjan Das, and Ellie Pavlick. BERT rediscovers the classical NLP pipeline. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, ACL '19, pages 4593–4601. Association for Computational Linguistics, 2019.

37 Ian Tenney, James Wexler, Jasmijn Bastings, Tolga Bolukbasi, Andy Coenen, Sebastian Gehrmann, Ellen Jiang, Mahima Pushkarna, Carey Radebaugh, Emily Reif, and Ann Yuan. The language interpretability tool: Extensible, interactive visualizations and analysis for NLP models. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, EMNLP '20, pages 107–118. Association for Computational Linguistics, 2020.

38 Ian Tenney, Patrick Xia, Berlin Chen, Alex Wang, Adam Poliak, R Thomas McCoy, Najoung Kim, Benjamin Van Durme, Samuel R. Bowman, Dipanjan Das, and Ellie Pavlick. What do you learn from context? Probing for sentence structure in contextualized word representations. *arXiv preprint arXiv:1905.06316*, 2019.

39 Jesse Vig. A multiscale visualization of attention in the transformer model. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, ACL '19, pages 37–42. Association for Computational Linguistics, 2019.

40 Susan Zhang, Stephen Roller, Naman Goyal, Mikel Artetxe, Moya Chen, Shuohui Chen, Christopher Dewan, Mona Diab, Xian Li, Xi V. Lin, Todor Mihaylov, Myle Ott, Sam Shleifer, Kurt Shuster, Daniel Simig, Punit S. Koura, Anjali Sridhar, Tianlu Wang, and Luke Zettlemoyer. OPT: Open pre-trained transformer language models. *arXiv preprint arXiv:2205.01068*, 2022.

6.5 WG: Bias and Bias Mitigation

Alex Endert, Angelos Chatzimparmpas, Antske Fokkens, Chris Weaver, Christopher Collins, Ross Maciejewski, Shimei Pan, and Tatiana von Landesberger

This discussion focused on definitions and categorization of bias (e.g., social bias, system bias, cognitive bias, and sample bias) and methods to identify and mitigate all in the context of text analysis and visualization. We (Figure 26) discussed the data processing pipelines from the NLP community and the data visualization community as a lens through which to discuss areas where bias can appear. A fundamental point when talking about bias is that biases can be found or introduced in every step of the pipeline. Locating bias in the pipeline can be a challenge. There can be bias in the data, this can be amplified or even introduced by the model, by choices on how to visualize the data, the transformations of the data as part of the visualization and in the eye of the beholder interpreting the results.

Figure 26 The group coming up with new and interesting ideas.

Another fundamental aspect is the question of whether bias is necessarily problematic, where the obvious answer is that it depends. It is a complex question when bias is problematic and what the ideal outcome or correction is when a problematic bias occurs [3]. This can only be addressed with input from various disciplines, from notably the humanities and social sciences as well as experts from the application domain. We agreed however that it is always good to know about bias so that we can discuss if this bias is useful or problematic, inevitable or something that can and should be fixed, or at least attempts made to mitigate it [4]. Therefore, we continued to focus on identifying and communicating bias. Several ideas for development were discussed. We mostly dived into the question of how to visualize bias in data and language models exploring prompt engineering and patterns in data.

6.5.1 Exploring social bias through prompts and completion

In large-scale NLP models (e.g., BERT, GPT-2), one mechanism to probe for biases is to utilize prompts to try and understand if the model was associating particular roles more frequently with certain protected classes (e.g., gender, ethnicity). However, this probing is often highly sequential and can be ad hoc. It does not necessarily have the capability to reveal hidden biases; instead, it may lead to only the discovery of expected biases (as opposed to unexpected biases). Current approaches for understanding if biases exist in such models include comparing completions for two prompts. Observing the visual comparison of the completions can illuminate potential biases when references are made in the text that indicate forms of social biases. However, these visual comparisons often lead to follow-up questions about specific words or phrases. Current tools do not support this iterative refinement and exploration. We discussed how visualization tools could be designed to foster this iterative exploration, including:

1. *Comparing multiple models trained on different datasets for a single prompt.*
2. *Comparing multiple prompt completions for a single model.*
3. *Visually constructing complex prompt completion templates through the use of SentenTree or WordTree visualization techniques.*
4. *Visualizing the embedding space of prompts for a given model.*

It is at this intersection where visualization could serve as an effective mechanism to support the exploration of biases.

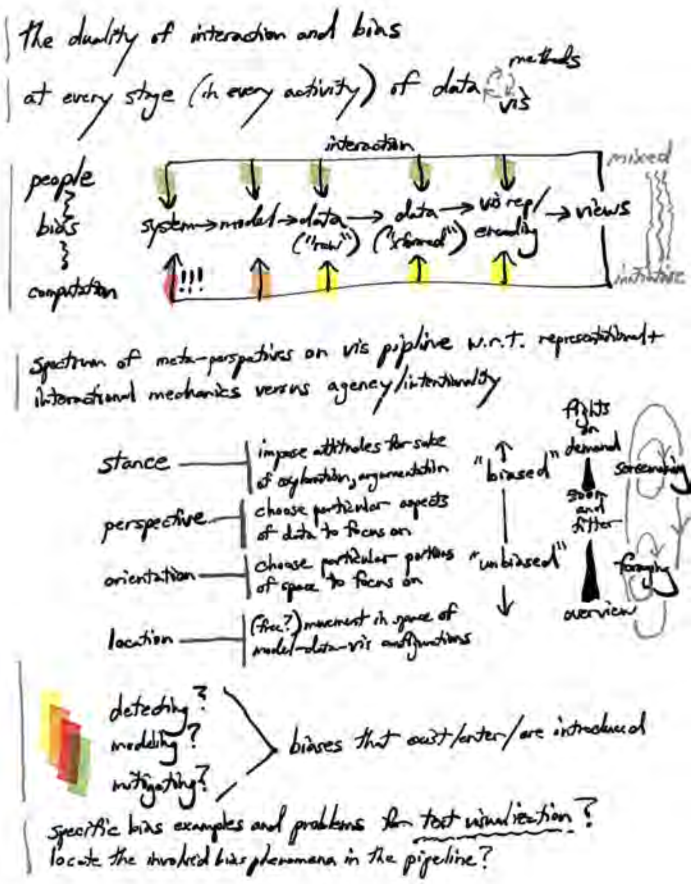

Figure 27 A sketch of a speculative model of how mixed initiative visual text analysis might introduce different kinds and severities of biases in parallel at successive stages of the visualization pipeline.

1. *Can text visualization help us to identify problems in the training data (e.g., various types of biases throughout the complete pipeline)?*
2. *Can text visualization facilitate prompt engineering?*
3. *Can interactivity support identifying bias in large datasets?*
4. *Can interactivity support framing perspectives and stances on bias?*
5. *How can we detect, warn and communicate about potential bias shifts over time?*

To that end, it would be imperative to explore which visualizations are suitable for:

- Communicating bias detected by automatic data processing.
- Identifying bias, its size, and aspects in the original dataset.
- Discovering hidden and unknown bias in large corpora.
- Comparing bias between models.
- Steering models towards less bias.
- Letting the user remove or mitigate bias (e.g., by selecting and checking datasets without bias).

- Exploring the "path" of bias from a dataset, via model to visualization – source, size, amplification, and transformation of bias through the pipeline.

Throughout the design process, considerations of the user's background (e.g., NLP experts or application domain professionals, e.g., social scientists) should be considered. Furthermore, the visual design must be careful not to suppress or amplify detected biases from the data and models. A key step to this is identifying which biases are caused by the data representations, visual representations, and cognitive biases of domain experts using the visualization [5, 1]. Another is understanding how biases are introduced, changed, removed, and even elicited in the design process and resulting pipelines (Figure 27). Finally, such designs would need to be systematically evaluated to explore the potential for unintended biases (pointer to evaluation group).

6.5.2 Project idea

Language models, such as GPT-3, can be used as a reflection upon and summary of a large corpus of text (and also the behavior of different models). We discussed using such models to compare various collections of texts, for example, training multiple instances of the model across a variety of related corpora, in order to compare the completions the variously trained models would make. Previous work [2] has explored the comparison of corpora of tweets mentioning "China virus" to those using the term "SARS Cov-2" by looking at the variety of completions generated by the models when given carefully selected prompts. One of the challenges of this type of research is to design the prompts that are used to generate the completions. This requires some domain knowledge and expertise in prompt design, as some prompts will limit the completion space significantly. For example "Dr. Fauci is a _" limits the space significantly more than "Dr. Fauci is _". How do we support the exploration of the prompt space in an interactive manner? We discussed a design in which a prompt could be used to generate a set of N (e.g., 1000) completions across multiple models. These completion sets could be turned into vectors in which each unique completion is given a cell and the cell is filled with the probability of the completion for the given model [21]. The distance between these completion vectors could represent how different (or biased) the models are on the given prompt. Given a set of possible prompts, they could be evaluated through such a pipeline to help guide an investigator toward which prompts are most likely to show a signal across models. One avenue for generating such a list of prompts could be to create a sentence with an open "slot" and a rule for its completion. For example "[NAME] is a _". The NAME slot here could be filled by mining the input training data for the top N names in the collection. This would generate N prompts which could be fed through each language model, generating the completion vectors in order to compare. In an interactive loop, one could choose a prompt, view the completions across models, then use those completions as starting points to generate new prompt rules as shown in Figure 28.

References

1 Yongsu Ahn and Yu-Ru Lin. FairSight: Visual analytics for fairness in decision making. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):1086–1095, January 2020.

2 Philip Feldman, Sim Tiwari, Charissa S. L. Cheah, James R. Foulds, and Shimei Pan. Analyzing COVID-19 tweets with transformer-based language models. *arXiv preprint arXiv:2104.10259*, 2021.

[21] https://pair.withgoogle.com/explorables/fill-in-the-blank/

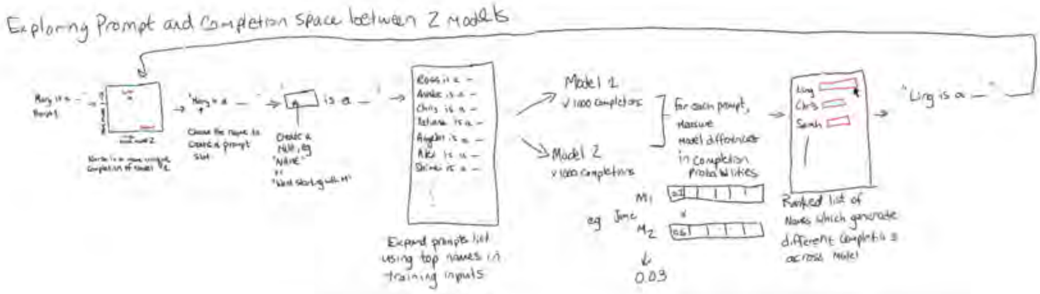

Figure 28 Bias in generative language models such as GPT-3 can be explored by examining suggested completions for given prompts. In this proposed system, a prompt is given and run through two models and the bias or preference (frequency) of completion occurrences can be compared. We expand this to explore the *prompt space*. A given word in the initial prompt can be substituted with a wildcard slot (as in “Mary” above). A corpus is used to fill this slot with likely terms from the training data. Each of these new prompts is run through both models for N (e.g., 1000) completions. The differences in model completions for each prompt is measured and visualized, to show the analyst which prompts result in the most model differences. Therefore, model bias could be revealed, for example, if a racialized name shows that the models produce highly different outputs.

3 Sorelle A. Friedler, Carlos Scheidegger, and Suresh Venkatasubramanian. The (im)possibility of fairness: Different value systems require different mechanisms for fair decision making. *Communications of the ACM*, 64(4):136–143, March 2021.

4 Archit Rathore, Sunipa Dev, Jeff M. Phillips, Vivek Srikumar, Yan Zheng, Chin-Chia Michael Yeh, Junpeng Wang, Wei Zhang, and Bei Wang. VERB: Visualizing and interpreting bias mitigation techniques for word representations. *arXiv preprint arXiv:2104.02797*, 2021.

5 Emily Wall, Leslie M. Blaha, Lyndsey Franklin, and Alex Endert. Warning, bias may occur: A proposed approach to detecting cognitive bias in interactive visual analytics. In *Proceedings of the 2017 IEEE Conference on Visual Analytics Science and Technology*, VAST ’17, pages 104–115. IEEE, 2017.

6.6 WG: Embedding Representation

Andreas Kerren, Bettina Speckmann, Carita Paradis, Christofer Meinecke, Mennatallah El-Assady, Pantea Haghighatkhah, and Yoav Goldberg

Our group discussed the challenge of creating a faithful and static low-dimensional representation of embedding vectors in high dimensions (R^n). Such a representation should convey the topological and geometric artifacts of the projection in two dimensions by means of visual cues and markers. The discussion focused on creating an intuitive depiction of distortions that requires minimal interaction of the user. These condensed representations can always be extended with suitable interactions to convey other levels of insight on the embedding vectors.

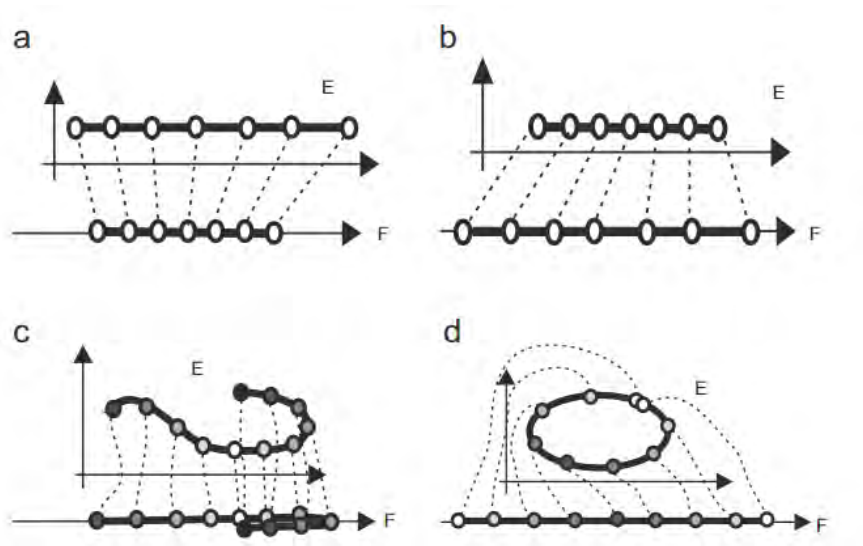

Figure 29 E is the original and F is the projected space: (a) compression, (b) stretching, (c) gluing, and (d) tearing [1].

6.6.1 Motivation

The NLP community makes extensive use of vector embeddings which relate to words, sentences, paragraphs, documents and other text elements. The vectors tend to be high dimensional (300 dimensions or more) and are often projected down to two dimensions for visualization purposes. Such visualizations then support tasks, for instance, identifying clusters in the data, inspecting the spatial structure of the space (which words/sentences relate to each other, and how are the relations organized globally), and identifying similar items or neighborhoods of similar items. Another use case is the debugging of learned representations [3].

In most related research, visualization systems use projections like PCA, t-SNE, or UMAP [5, 4]. However, projecting the high dimensional vectors to two dimensions is inherently a lossy process, and the resulting visualizations are often misleading in various ways, which in turn may lead researchers to incorrect conclusions. Interactive tools such as t-viSNE [2] can help to alleviate some of these issues, but interactivity also requires the user to know specifically what kind of distortion they are looking for. A static visualization that effectively shows additional information, which is usually lost in standard projection visualization, can go a long way in supporting more accurate interpretations of the results. In particular, a static visualization can be used to highlight the points in the visualization where an interactive exploration is required. What we aim for in our static design is to support this requirement.

6.6.2 Visualization error

There are two general types of errors that are caused by projection of embeddings to lower dimensions [1]. First is the **geometric error** that distorts the distances but keeps the neighborhood of points intact. Second is the **topological error** that disrupts both the neighborhood and the distances. Examples of geometric errors are stretching or compressing, which increase or decrease the distances between embeddings compared to the original space (see Figure 29, (a) and (b)). Examples of the topological errors are gluing or tearing, which

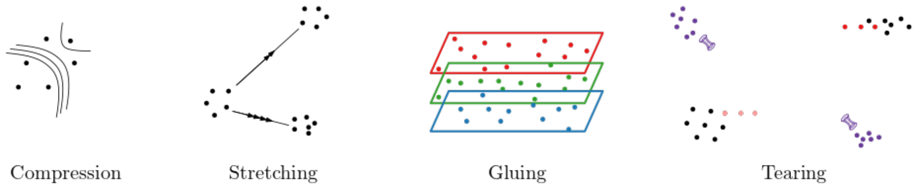

Figure 30 Proposed solution for both geometric and topological errors.

merge neighborhoods or disrupt a neighborhood into multiple neighborhoods, respectively (Figure 29, (c) and (d)).

6.6.3 Noise in the original data

Based on our discussions, we find it valuable to study the nature of noise in the data. In order to be able to communicate how much the visualization of various clusters of the data can be trusted once projected into the two-dimensional space, we first need to be able to measure to what extent the clusters are well-formed since the distortions visible in the two-dimensional visualization can be both caused by the projection or by the inherent noise in the data.

6.6.4 Solution

Once we have a way to quantify the noise relative to the general geometry and topology of the data, we aim to proceed with the visualization aspect. Our solution is to use cartographic techniques to depict both geometric and topological distortions in 2.5 dimensions. We have composed a list of visual encodings to help communicate the distortions in the visualization of data (cf. Figure 30).

Compression error: In order to communicate the geometric distortions in the visualization caused by compression, we propose to use contour lines, which are often used to communicate height differences in cartography. The areas that are compressed can contain more contour lines, indicating that in the original space there is larger distances between points. The number of contour lines can communicate the extent to which an area is compressed.

Stretching error: The distortion caused by stretching can be communicated with lines connecting the areas that are stretched out in the visualization. The extent of stretch can be communicated by means of arrow heads placed on the line. More arrow heads indicate smaller original distance between two areas.

Gluing error: The gluing error is caused by two or more neighborhoods being mapped to the same position. We distinguish between the neighborhoods by visualizing them in overlaying layers. One can think of this as a multistory parking garage where every story is representing a different neighborhood in the original space.

Tearing error: The tearing of the same neighborhood into multiple neighborhoods is visualized by means of "ghost points" where we duplicate the separated points with a more transparent visualization (ghost-like) and put them close to the points of the original neighborhood. Another way to represent this error in the visualization is to make use of color-coded "wormhole" shapes that indicate which neighborhoods are originally connected in the initial space.

References

1 Michaël Aupetit. Visualizing distortions and recovering topology in continuous projection techniques. *Neurocomputing*, 70(7):1304–1330, 2007. Advances in Computational Intelligence and Learning.

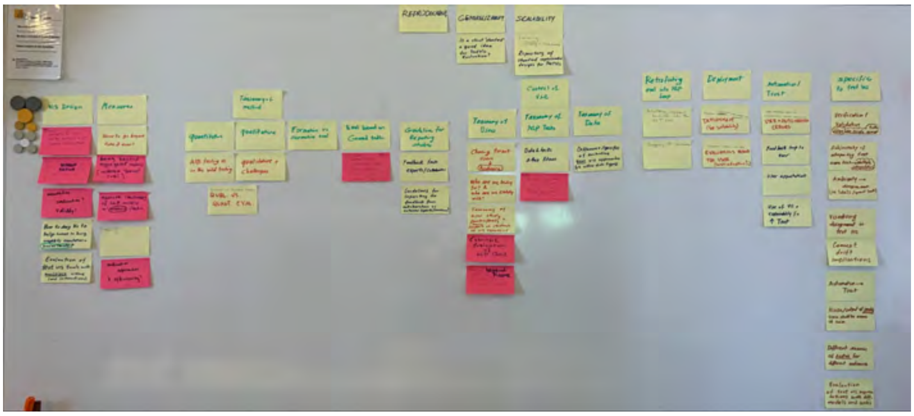

Figure 31 The notes taken at the beginning of the group discussion and reorganized according to the main groups of concerns and open challenges.

2 Angelos Chatzimparmpas, Rafael M. Martins, and Andreas Kerren. t-viSNE: Interactive assessment and interpretation of t-SNE projections. *IEEE Transactions on Visualization and Computer Graphics*, 26(8):2696–2714, August 2020.
3 Hila Gonen and Yoav Goldberg. Lipstick on a pig: Debiasing methods cover up systematic gender biases in word embeddings but do not remove them. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, NAACL '19, pages 609–614. Association for Computational Linguistics, 2019.
4 Yingfan Wang, Haiyang Huang, Cynthia Rudin, and Yaron Shaposhnik. Understanding how dimension reduction tools work: An empirical approach to deciphering t-SNE, UMAP, TriMap, and PaCMAP for data visualization. *Journal of Machine Learning Research*, 22(201):1–73, 2021.
5 Martin Wattenberg, Fernanda Viégas, and Ian Johnson. How to use t-SNE effectively. *Distill*, 2016.

6.7 WG: Evaluation and Experimental Designs

Barbara Plank, Jean-Daniel Fekete, José Angel Daza Arévalo, Kostiantyn Kucher, Maria Skeppstedt, Narges Mahyar, Nicole Sultanum, and Vasiliki Simaki

The purpose of visual text analytic approaches in most cases is to support the users in accomplishing their tasks and achieving their goals that involve text data, computational methods, and visual/interactive techniques. In order to ensure that the proposed visual text analytic tool indeed supports the users in their work and behaves in a reliable, predictable manner, the designers of such approaches must consider the respective evaluation/validation concerns [10]. The choice of appropriate evaluation techniques and experimental designs is not a trivial issue, though, and it presents a number of open challenges.

Our group focused on the issues of evaluation and experimental design and it featured an equal number of participants with a background in visualization and NLP, which helped us with the identification of concerns and concepts related to all parts of the (visual) text analytic pipeline that affect the choice of evaluation methods. Our discussion started with a short round of introductions and interests with respect to the group topics, and then we proceeded with a brainstorming session using notes from individual participants. We analyzed the contents of the collected notes and organized them into several main groups/categories (see Figure 31), which were then discussed in smaller subgroups and eventually with other group participants.

6.7.1 Verification and validation in the context of language/text data ambiguity and lack of single ground truth

Text data is rich and ambiguous, and there is no single canonical data representation or universally accepted interpretation. Ambiguity can be found in the text at several levels, from the grammatical level (e.g., run/run) to the semantic and conceptual level (relations between concepts and what the concepts mean, e.g., let's go to the bank). Ambiguity in language may lead to different interpretations of the text by the annotators, which may affect their decisions. But even in cases where annotators have a similar or even identical understanding/interpretation of the text, their annotation decisions may still differ in the end due to, for instance, their different backgrounds and world perception.

In the cases where the divergence between the annotators' decisions is high, the task can be more closely specified (and/or simplified) and the annotation guidelines can be further refined, until a desired level of inter-annotator agreement (IAA) [2] is reached. Traditionally, high IAA implies higher data quality (in terms of reproducibility of data labeling). But due to the nature of language itself and the various meanings that exist, assuming high agreement (and by consequence, a single ground truth) is at best an idealization—this neglects genuine disagreement that can provide valuable insights about language use and interpretation. For some text interpretation tasks, too restrictive guidelines might even hide the fact there is no real ground truth for this particular task.

There are many examples of visualizations that create an overview of the results of text annotations, in particular for visualizing the co-occurrence of different annotation categories. Approaches range from using standard visualization techniques (e.g. bar charts, scatter plots, Sankey and chord diagrams, and treemaps [11] to develop new visualization techniques specific to the text annotation task [12]. To the best of our knowledge, however, user interfaces for showing the difference in annotation categories between different annotators mainly focus on (i) individual annotation instances (e.g., Yimam et al. [14]), or (ii) IAA-statistics for the annotators, such as matrices showing the inter-annotator agreement for pairs of users (e.g. Grosman et al. [8]). That is, the aim of these visualizations is not to provide the user with an overview of the differences in annotation choices. This is natural, as these interfaces typically are created for curators who use them for annotation adjudication [4]. That is, curators who aim at merging each conflicting annotation into the one correct annotation and not at exploring the annotated dataset in order to gain insights into the reasons for disagreement, for instance, insights into what extent there is a single ground truth for the annotation task.

We believe that visual text analytics provides opportunities for mutual benefits of research within NLP and text visualization to provide insights into the nature of the disagreement. In addition to visualizing disagreement between annotators in terms of labeling divergences, there are additional factors that contribute to disagreement and could be visualized. For instance, the annotators could be given the opportunity to not only annotate, but also specify the level of uncertainty for the annotation, or to opt-out of annotating an instance due to it being too difficult. Also the time it takes for the annotators to make an annotation decision could be gathered as an information type showing the difficulty of the task.

6.7.2 Concerns for text vis evaluation and experimental design: Audience, tasks, scale, models

An NLP system and its visualizations should be evaluated in light of these four dimensions: target audience (who?), list of tasks to support (what for?), scales to consider in the tasks (linguistic objects of interest), and models supporting the tasks at the given scales (how?).

Audience: We start with the end-user who defines what they want to perform in the system (goals). We can define different User groups (from larger to smaller in terms of the system knowledge and expected functionality): Debugger or Developer of the system, Linguist or NLP Expert in charge of the modeling, Domain Specialist using the system with no particular knowledge of NLP or CS, and General User, where each member of a group has the ability to perform the tasks of their own group and the smaller groups contained within. Considering the use of visual text analytic approaches for applications with real-world data and real-world implications [3], the importance of carefully considering the intended audience and their goals for the design and evaluation of visual text analytics cannot be overstated.

Tasks: We then define the tasks that will achieve each of the users' goals. A task is the combination of a User+Verb and attributes; it can lead to a side effect (change in the visualization) or to a result (list of items). For example, overview (depends on the visual representation, no parameter), query, search (return a list of items), cluster (visualize groups and return a list of groups), etc. In addition to such general prototypical tasks, specialized tasks can be domain-specific, but also specific to the linguistic model (aimed at the NLP specialist) or to the programmer, allowing the respective user to understand the internal data structure of the NLP and visualizations. These tasks are not useful for the domain expert or to a lay user, for instance. Further potential tasks can be identified in the previous works on the more general task taxonomies in information visualization and visual analytics (e.g., [1] and [5]) as well as the sources focusing on text visualization and visual text analytics, such as the work by Tofiloski et al. (2015), for instance.

Scale: Each task can be applied or affect the text at a different level of granularity: word, phrase, entities, tree structure (syntax), proposition (semantics), sentence, section, paragraph, document, and corpus.

Models: Layers of linguistic information are encoded and structured differently: parse tree, part-of-speech tags, XML-encoded text (metadata), stand-off annotation (e.g., relevant spans in text), topics, etc. The model influences the tasks in the sense that it allows some tasks to be done more or less accurately, and with diverse levels of uncertainty, quality, errors, and artifacts. The models also constrain the scale of possible tasks.

6.7.3 Trust in automation in the context of visual text analysis evaluation

Given the complex, multifaceted, and potentially ambiguous nature of text, there is a wide error range in NLP outcomes, as we cannot expect NLP systems to work perfectly in every circumstance, e.g., for each task and for each dataset. This is something users must take into account when leveraging text visual analytics systems for a particular problem, and is reflected in the extent users trust automated outcomes and take them into account. User trust is an evolving process [9], and it reflects perceived system performance over use and the risk associated with decision-making. For example, spelling and grammar correction in documents is a relatively mature problem with low impact for errors, and something users are generally comfortable delegating to automation. On the other hand, automatic summarization of patient records for clinical decision-making is an open research problem with significant consequences for wrong medical choices.

Apart from design considerations to foster adequate trust, from an evaluation perspective, it is also important to consider how trust can be measured, and whether it matches NLP capabilities. There are many challenging components to measuring adequate trust. It entails (a) an understanding of the limitation areas of NLP algorithms within a particular application, (b) assessing user perception of the limits of automated outcomes, (c) assessing how well user perception matches real capabilities of automation, and (d) how trust levels evolve over time. While there is research in understanding and designing for trust [13], including visual (text) analytic systems [7, 6], strategies for measuring trust are still not widely adopted in the evaluation of visual text analytics systems. Moving forward, we posit that the community should consider developing and fostering the use of standardized and easily applicable metrics and instruments (e.g. questionnaires) for measuring trust, which should be applicable at different stages of use (i.e., measurable over a period of time).

6.7.4 Systematic analysis and guidelines for evaluation of computational and human-centered aspects of visual text analytics

One of the challenges of evaluating visual text analytic approaches is related to the breadth and variety of concerns regarding computational and interactive, human-centered aspects of such approaches and tools. The efforts focusing on systematic analysis, categorization, and formulation of guidelines for evaluation (cf. the work by Isenberg et al. [10]) would thus be beneficial for researchers, practitioners, and end-users of such visual text analytic tools. Here, the topics and dimensions would include an interdisciplinary analysis/view and include both NLP/ML-focused aspects (e.g., the methods and measurements of intra- and inter-annotator agreement, or the measurements of NLP model performance for a given task beyond standard ML metrics such as accuracy or F1-score) as well as visualization and interaction-related aspects (e.g., the methods of usability measurement and long-term adoption observation). The concerns and dimensions of data annotation, computational models, text visualization approaches, and users' trust discussed above are all highly relevant to such systematic analyses. While the complexity and variety of visual text analytic problems would most likely not allow us to recommend a single prescribed method for validation, the outcomes of the work on systematic analysis would result in a collection of guidelines that would benefit the research efforts and applications.

References

1. Robert A. Amar and John T. Stasko. Knowledge precepts for design and evaluation of information visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 11(4):432–442, July–August 2005.
2. Ron Artstein and Massimo Poesio. Inter-coder agreement for computational linguistics. *Computational Linguistics*, 34(4):555–596, December 2008.
3. Eric P. S. Baumer, Mahmood Jasim, Ali Sarvghad, and Narges Mahyar. Of course it's political! A critical inquiry into underemphasized dimensions in civic text visualization. *Computer Graphics Forum*, 2022. To appear.
4. Chris Biemann, Kalina Bontcheva, Richard Eckart de Castilho, Iryna Gurevych, and Seid Muhie Yimam. Collaborative web-based tools for multi-layer text annotation. In *Handbook of Linguistic Annotation*, pages 229–256. Springer, 2017.
5. Matthew Brehmer and Tamara Munzner. A multi-level typology of abstract visualization tasks. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2376–2385, December 2013.

6 Angelos Chatzimparmpas, Rafael M. Martins, Ilir Jusufi, Kostiantyn Kucher, Fabrice Rossi, and Andreas Kerren. The state of the art in enhancing trust in machine learning models with the use of visualizations. *Computer Graphics Forum*, 39(3):713–756, June 2020.

7 Jason Chuang, Daniel Ramage, Christopher Manning, and Jeffrey Heer. Interpretation and trust: Designing model-driven visualizations for text analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 443–452. Association for Computing Machinery, 2012.

8 Jonatas S. Grosman, Pedro H.T. Furtado, Ariane M.B. Rodrigues, Guilherme G. Schardong, Simone D.J. Barbosa, and Hélio C.V. Lopes. ERAS: Improving the quality control in the annotation process for natural language processing tasks. *Information Systems*, 93:101553, 2020.

9 Robert R. Hoffman, Matthew Johnson, Jeffrey M. Bradshaw, and Al Underbrink. Trust in automation. *IEEE Intelligent Systems*, 28(1):84–88, January–February 2013.

10 Tobias Isenberg, Petra Isenberg, Jian Chen, Michael Sedlmair, and Torsten Möller. A systematic review on the practice of evaluating visualization. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2818–2827, December 2013.

11 Ecem Kavaz, Anna Puig, Inmaculada Rodriguez, Mariona Taule, and Montserrat Nofre. Data visualization for supporting linguists in the analysis of toxic messages. *Computer Science Research Notes*, 3101:59–70, May 2021.

12 Kostiantyn Kucher, Carita Paradis, Magnus Sahlgren, and Andreas Kerren. Active learning and visual analytics for stance classification with ALVA. *ACM Transactions on Interactive Intelligent Systems*, 7(3), October 2017.

13 John D. Lee and Katrina A. See. Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1):50–80, March 2004.

14 Seid Muhie Yimam, Iryna Gurevych, Richard Eckart de Castilho, and Chris Biemann. WebAnno: A flexible, web-based and visually supported system for distributed annotations. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 1–6. Association for Computational Linguistics, 2013.

Participants

- Richard Brath
Uncharted Software – Toronto, CA
- Angelos Chatzimparmpas
Linnaeus University – Växjö, SE
- Christopher Collins
Ontario Tech – Oshawa, CA
- José Angel Daza Arévalo
Free University Amsterdam, NL
- Mennatallah El-Assady
ETH Zürich, CH
- Alex Endert
Georgia Institute of Technology – Atlanta, US
- Jean-Daniel Fekete
INRIA Saclay – Orsay, FR
- Antske Fokkens
Free University Amsterdam, NL
- Yoav Goldberg
Bar-Ilan University – Ramat Gan, IL
- Pantea Haghighatkhah
TU Eindhoven, NL
- Daniel A. Keim
Universität Konstanz, DE
- Andreas Kerren
Linköping University, SE
- Johannes Knittel
Universität Stuttgart, DE
- Kostiantyn Kucher
Linnaeus University – Växjö, SE
- Ross Maciejewski
Arizona State University – Tempe, US
- Narges Mahyar
University of Massachusetts – Amherst, US
- Christofer Meinecke
Universität Leipzig, DE
- Shimei Pan
University of Maryland – Baltimore County, US
- Carita Paradis
Lund University, SE
- Barbara Plank
IT University of Copenhagen, DK
- Vasiliki Simaki
Lund University, SE
- Maria Skeppstedt
SE
- Pia Sommerauer
Free University Amsterdam, NL
- Bettina Speckmann
TU Eindhoven, NL
- Hendrik Strobelt
MIT-IBM Watson AI Lab – Cambridge, US
- Nicole Sultanum
University of Toronto, CA
- Tatiana von Landesberger
Universität Köln, DE
- Chris Weaver
University of Oklahoma – Norman, US

Report from Dagstuhl Seminar 22192

Dynamic Traffic Models in Transportation Science

Martin Gairing*[1], Carolina Osorio*[2], Britta Peis*[3], David Watling*[4], and Katharina Eickhoff†[5]

1 **University of Liverpool, GB.** gairing@liverpool.ac.uk
2 **HEC Montréal, CA & Google – Mountain View, US.** carolina.osorio@hec.ca
3 **RWTH Aachen, DE.** britta.peis@oms.rwth-aachen.de
4 **University of Leeds, GB.** d.p.watling@its.leeds.ac.uk
5 **RWTH Aachen, DE.** katharina.eickhoff@oms.rwth-aachen.de

Abstract

Traffic assignment models are crucial for transport planners to be able to predict the congestion, environmental and social impacts of transport policies, for example in the light of possible changes to the infrastructure, to the transport services offered, or to the prices charged to travellers. The motivation for this series of seminars – of which this seminar was the third – is the prevalence in the transportation community of basing such predictions on complex computer-based simulations that are capable of resolving many elements of a real systems, while on the other hand, the theory of dynamic traffic assignments (in terms of equilibrium existence, computability and efficiency) had not matured to the point matching the model complexity inherent in simulations.

Progress has been made on this issue in the first two seminars (Dagstuhl Seminar 15412 and 18102), by bringing together leading scientists in the areas of *traffic simulation*, *algorithmic game theory* and *dynamic traffic assignment*. We continued this process this seminar. Moreover, we started to address the growing real-life challenge of new kinds of 'mobility service' emerging, before the tools are available to incorporate them in such planning models. These services include intelligent/dynamic ride-sharing and car-sharing, through to fully autonomous vehicles, provided potentially by a variety of competing operators.

Seminar May 8–13, 2022 – http://www.dagstuhl.de/22192

2012 ACM Subject Classification Networks → Control path algorithms; Networks → Network performance evaluation; Theory of computation → Design and analysis of algorithms

Keywords and phrases Algorithms and Complexity of traffic equilibrium computations, Dynamic traffic assignment models, Simulation and network optimization

Digital Object Identifier 10.4230/DagRep.12.5.92

1 Executive Summary

Martin Gairing (University of Liverpool, GB)
Carolina Osorio (HEC Montréal, CA & Google – Mountain View, US)
Britta Peis (RWTH Aachen, DE)
David Watling (University of Leeds, GB)

Traffic assignment models are crucial for transport planners to be able to predict the traffic distributions – and thereby the congestion, environmental and social impacts – of transport policies, for example in the light of possible changes to the infrastructure (e.g. traffic light

* Editor / Organizer
† Editorial Assistant / Collector

controls), to the transport services offered (e.g. the provision of public transport fleets and frequencies), or to the prices imposed on travellers (e.g. road tolls, public transport fares). The prevailing approaches used in the transportation science literature to predict such distributions can be roughly classified into mathematical approaches based on *dynamic traffic assignment models* (DTA) (using the methodology of flows over time) and *simulation-based approaches* (using large-scale microsimulations). The striking advantage of microscopic simulations over DTA models is that the latter usually ignore practically relevant side-constraints such as horizontal queueing of vehicles, the feedback of changing network conditions on user behavior, flexible departure time choice, mode choice, activity schedule choice, and such. Current simulation tools integrate all these dimensions and many more. The increase in model complexity inherent in simulations, however, is not matched by the existing DTA theory. For most simulation software, there is no "mathematical proof of concept" such as formal proofs for the following fundamental properties:

- a (dynamic) equilibrium always exists,
- an equilibrium is unique,
- an equilibrium is efficiently computable,
- there is a smooth function of equilibria with respect to parameters (such as street capacities, tolls, etc.)

In this seminar we brought together, for the third time, leading researchers from three different communities (Simulations, Dynamic Traffic Assignment and Algorithmic Game Theory) in order to bridge the gap between complex simulation based models and the existing theory. In doing so, we build on the progress made in the first two seminars, and at the same time addressed the emerging real-life challenge of modeling the new kinds of 'mobility service' that are increasingly envisaged for the future, such as intelligent/dynamic ride-sharing and car-sharing, through to fully autonomous vehicles, provided potentially by a variety of competing operators.

The purpose of this third seminar was to build on the momentum, collaborations and new directions identified at the previous two seminars, following on from the impacts identified earlier. Similarly to the first two seminars, one objective was to make new advances on identifying, formulating and solving existing open problems by combining insights from the different academic fields. Thus, we continued working on themes, initiated at the previous Dagstuhl seminars, related to the representation of queuing in dynamic models and the identification of new requirements and new open problems arising from future forms of mobility and travellers' behaviour, for example, systems of autonomous vehicles, and coordinated/shared routing.

Again, the seminar was a big success. Beside forster collaboration of the last seminars of this series, the seminar stimulates new and very fruitful ones. We got laudatory feedback from many participants which is also reflected in the survey conducted by Dagstuhl.

2 Table of Contents

3 Overview of Talks

3.1 Towards day-to-day assignment game models for mobility ecosystems

Joseph Chow (New York University, US)

While there are many methods in within-day dynamics of on-demand mobility services, the literature on day-to-day dynamics is scarce. In this open problem session, I will present a steady state assignment game model for analyzing and designing platforms of mobility ecosystems that output passenger flows, stable pricing ranges, and platform subsidies. I will discuss challenges and opportunities to identifying an equivalent day-to-day adjustment model corresponding to that model: stable outcome space becoming an attractor, learning functions for user and operators that include cost allocations, using the path-dependency of the dynamic system to tackle the nonuniqueness of stable outcome spaces corresponding to a matching solution, deterministic and stochastic variants, and introducing policy interventions to guide ecosystems toward desired attractors.

3.2 Tutorial Talk: Integrated travel behavior/network modeling

Gunnar Flötteröd (Linkp̈ing University, SE & Swedish National Road and Transport Research Institute (VTI) – Stockholm SE)

Strategic transport planning models that predict travel demand and behavior in consideration of a congested transport system have been used for many decades. Historically, these models are aggregate and continuum flow based, meaning that human behavior and physical vehicle propagation are abstracted into a limited number of representative traveler and vehicle flows. The last few decades have seen the convergence of network flow meso/microsimulations and activity-based travel demand models into "agent-based" model systems that simulate individual synthetic travelers and vehicles instead of largely anonymous flows. Greater realism may be expected from agent-based models, given their potentially higher level of detail. However, careful structural modeling is needed to leverage these capabilities. The talk presents the possible advantages that arise from an agent-based approach and emphasizes the accompanying modeling challenges.

3.3 Existence and Complexity of Approximate Equilibria in Weighted Congestion Games

Yiannis Giannakopoulos (Universität Erlangen-Nürnberg, DE)

Joint work of Giorgos Christodoulou, Martin Gairing, Yiannis Giannakopoulos, Dogo Poças, Clara Waldmann

We study the existence of approximate pure Nash equilibria (α-PNE) in weighted atomic congestion games with polynomial cost functions of maximum degree d. Previously it was known that d-approximate equilibria always exist, while nonexistence was established only for small constants, namely for 1.153-PNE. We improve significantly upon this gap, proving that such games in general do not have $\tilde{\Theta}(\sqrt{d})$-approximate PNE, which provides the first super-constant lower bound.

Furthermore, we provide a black-box gap-introducing method of combining such nonexistence results with a specific circuit gadget, in order to derive NP-completeness of the decision version of the problem. In particular, deploying this technique we are able to show that deciding whether a weighted congestion game has an $\tilde{O}(\sqrt{d})$-PNE is NP-complete. Previous hardness results were known only for the special case of exact equilibria and arbitrary cost functions.

The circuit gadget is of independent interest and it allows us to also prove hardness for a variety of problems related to the complexity of PNE in congestion games. For example, we demonstrate that the question of existence of α-PNE in which a certain set of players plays a specific strategy profile is NP-hard for any $\alpha < 3^{d/2}$, even for unweighted congestion games.

Finally, we study the existence of approximate equilibria in weighted congestion games with general (nondecreasing) costs, as a function of the number of players n. We show that n-PNE always exist, matched by an almost tight nonexistence bound of $\tilde{\Theta}(n)$, which we can again transform into an NP-completeness proof for the decision problem.

3.4 Dynamic Traffic Assignment for Electric Vehicles

Lukas Graf (Universität Augsburg, DE)

Joint work of Lukas Graf, Tobias Harks, Prashant Palkar

Electric vehicles (EVs) are a great promise for the coming decades in order to allow for mobility but at the same time take measures against the climate change by reducing the emissions of classical combustion engines.

However, those new types of vehicles also come with new challenges like the limited range and the more complex and time consuming recharging process. Modelling traffic involving EVs, thus, requires a more involved model then the one used for traditional vehicles. To accommodate this, we augment the deterministic queuing model of Vickrey with energy constraints and study capacitated dynamic equilibria in this model. Here, in contrast to the traditional dynamic equilibria, we can neither assume that particles only take simple paths (leading to an infinite dimensional strategy space) nor synchronized intermediate arrival times (i.e. particles starting at the same time from the same source towards the same destination might still arrive at intermediate nodes at different times).

We will explore these difficulties and show how we can overcome them to show existence of equilibria for EVs as well as an even more general traffic models. We complement our theoretical results by a computational study in which we apply a fixed-point algorithm to compute energy- dynamic equilibria.

3.5 Tutorial Talk: The role of Information in DTA Models

Tobias Harks (Universität Augsburg, DE)

I will first give a brief overview on some existing DTA models assuming i) perfect and full information on travel times, flows etc. and the other extreme (ii) assuming only instantaneous (incomplete) information. After this overview I introduce a DTA model, where agents base their instantaneous routing decisions on real-time delay predictions. I formulate a mathematically concise model and derive properties of the predictors that ensure a dynamic prediction equilibrium exists.I will show that the framework subsumes the well-known full information and instantaneous information models, in addition to admitting further realistic predictors as special cases.

3.6 Signaling in Network Congestion Games

Martin Hoefer (Goethe-Universität – Frankfurt am Main, DE)

Joint work of Svenja Griesbach, Martin Hoefer, Max Klimm, Tim Koglin

A largely untapped potential for improving traffic flows is rooted in the inherent uncertainty of travel times. Large mobility services have an informational advantage over single network users as they are able to learn traffic conditions from data. A benevolent provider may use this informational advantage in order to steer a traffic equilibrium into a favorable direction. The resulting optimization problem is a task commonly referred to as signaling or Bayesian persuasion.

In this talk, we discuss our recent progress on this signaling problem. We characterize cases in which optimal signaling schemes are structurally simple and/or can be computed in polynomial time. Our results reveal interesting connections between optimal signaling, network structure, and support enumeration for equilibria.

3.7 Stability and Instability in DTA

Takamasa Iryo (Tohoku University – Sendai, JP)

Stability and instability issues in DTA have been studied actively in recent years but still many things are not well understood. This talk will give a brief introduction of existing studies dealing with stability issues in DTA, especially those on for DUE (dynamic user-equilibrium) solutions, and raise possible future research directions to tackle with this problem.

3.8 Equilibria in Multiclass and Multidimensional Atomic Congestion Games

Max Klimm (TU Berlin, DE)

We study the existence of pure Nash equilibria in atomic congestion games with different user classes where the cost of each resource depends on the aggregated demand of each class. A set of cost functions is called consistent for this class if all games with cost functions from the set have a pure Nash equilibrium. We give a complete characterization of consistent sets of cost functions showing that the only consistent sets of cost functions are sets of certain affine functions and sets of certain exponential functions. This characterization is also extended to a larger class of games where each atomic player may control flow that belongs to different classes.

3.9 Something about traffic lights and bicycles

Ekkehard Köhler (BTU Cottbus-Senftenberg, DE) and Martin Strehler (Westsächsische Hochschule Zwickau, DE)

Joint work of Ekkehard Köhler, Markus Rogge, Robert Scheffler, Martin Strehler

We report on work in progress on the question of finding good paths for bicycles in a traffic network with (possibly coordinated) traffic lights. While the standard dynamic shortest path problem in time-dependent networks with cyclic time windows is known to be easy, we study the related problem of finding routes that keep a small number of stops. We show that the problem is strongly NP-hard for paths and trails, whereas it is weakly NP-hard for walks. We give a pseudo-polynomial algorithm for this third option. Further we report about ongoing work on designing algorithms for the case of variable speeds and the employment of appropriate power consumption and recovery models for bicycles for the above formulated problem.

3.10 Hitting paths with random sets in abstract networks

Jannik Matuschke (KU Leuven, BE)

Consider a set family $\mathcal{P}$ on a common ground set E, together with a requirement $r(P) \in [0,1]$ for each $P \in \mathcal{P}$. Each element $e \in E$ is equipped with a marginal probability $p(e)$ such that $\sum_{e \in P} p(e) \geq r(P)$ for all $P \in \mathcal{P}$. In this talk, we investigate the question under which conditions these marginal probabilities can be turned into a probability distribution on the subsets of E such that each $P \in \mathcal{P}$ is intersected by the resulting random set with probability at least $r(P)$.

Motivated by a network security game, Dahan et al. (2001) showed the existence of such a distribution for the case that $\mathcal{P}$ is the family of maximal chains in a partial order (equivalently: the set of s-t-paths in a directed acyclic graph) and the requirements r fulfill a natural conservation property. We extend this result to the case that $\mathcal{P}$ is an abstract network fulfilling a certain ordering assumption. We also give a significantly simplified argument for the original result, using a classic idea from the theory of flows over time. In particular, our results imply that equilibria for the aforementioned security game can be efficiently computed even when the underlying digraph contains cycles.

3.11 Tutorial Talk: Continuity, Uniqueness and Long-Term Behaviour of Nash Flows Over Time

Neil Olver (London School of Economics, GB)

Joint work of Neil Olver, Leon Sering, Laura Vargas Koch

We consider a dynamic model of traffic that has received a lot of attention in the past few years. Users control infinitesimal flow particles aiming to travel from a source to destination as quickly as possible. Flow patterns vary over time, and congestion effects are modelled via queues, which form whenever the inflow into a link exceeds its capacity. We answer some rather basic questions about equilibria in this model: in particular uniqueness (in an appropriate sense), and continuity: small perturbations to the instance or to the traffic situation at some moment cannot lead to wildly different equilibrium evolutions.

To prove these results, we make a surprising connection to another question: whether, assuming constant inflow into the network at the source, do equilibria always eventually settle into a "steady state" where all queue delays change linearly forever more? Cominetti et al. proved this under an assumption that the inflow rate is not larger than the capacity of the network – eventually, queues remain constant forever. We resolve the more general question positively.

3.12 Combining Bayesian optimization with analytical dynamic traffic models to tackle dynamic simulation-based transportation optimization problems

Carolina Osorio (HEC Montréal, CA & Google – Mountain View, US)

Cities and companies alike often resort to the use of detailed, dynamic, and stochastic traffic simulators to tackle their transportation optimization problems. Methods from the field of simulation-based optimization are often used. However, their use to tackle problems with time-dependent decision variables is limited. In this talk, we present recent work that formulates an analytical,, differentiable, and compute efficient dynamic traffic model, and integrates it within a Bayesian optimization (BO) framework to tackle high-dimensional dynamic transportation problems. We present validation results on small network problems, as well as a large-scale traffic signal control case study for Midtown Manhattan, New York City. We discuss how the proposed approach enhances the scalability of BO methods.

3.13 In Congestion Games, Taxes Achieve Optimal Approximation

Dario Paccagnan (Imperial College London, GB) and Martin Gairing (University of Liverpool, GB)

We consider the problem of minimising social cost in atomic congestion games and show, perhaps surprisingly, that efficiently computed taxation mechanisms yield the *same* performance achievable by the best polynomial time algorithm, even when the latter has full control over the players' actions. It follows that no other tractable approach geared at incentivising desirable system behaviour can improve upon this result, regardless of whether it is based on taxations, coordination mechanisms, information provision, or any other principle. In short: Judiciously chosen taxes achieve optimal approximation.

Three technical contributions underpin this conclusion. First, we show that computing the minimum social cost is NP-hard to approximate within a given factor depending solely on the admissible resource costs. Second, we design a tractable taxation mechanism whose efficiency (price of anarchy) matches this hardness factor, and thus is optimal. As these results extend to coarse correlated equilibria, any no-regret algorithm inherits the same performances, allowing us to devise polynomial time algorithms with optimal approximation.

3.14 Stackelberg Max Closure

Britta Peis (RWTH Aachen, DE)

Joint work of Karsten Jungnitsch, Marc Schröder, Britta Peis

In a Stackelberg max closure game, we are given a digraph whose vertices correspond to projects from which firms can choose and whose arcs represent precedence constraints. Some projects are under the control of a leader who sets prices in the first stage of the game, while in the second stage, the firms choose a feasible subset of projects of maximum value. For a single follower, the leader's problem of finding revenue-maximizing prices can be solved in strongly polynomial time. In this paper, we focus on the setting with multiple followers and distinguish two situations. In the case in which only one copy of each project is available (limited supply), we show that the two-follower problem is solvable in strongly polynomial time, whereas the problem with three or more followers is NP-hard. In the case of unlimited supply, that is, when sufficient copies of each project are available, we show that the two-follower problem is already APX-hard. As a side result, we prove that Stackelberg min vertex cover on bipartite graphs with a single follower is APX-hard.

3.15 The Complexity of Gradient Descent

Rahul Savani (RWTH Aachen, DE)

Joint work of John Fearnley, Paul Goldberg, Alexandros Hollender, Rahul Savani

PPAD and PLS are successful classes that each capture the complexity of important game-theoretic problems: finding a mixed Nash equilibrium in a bimatrix game is PPAD-complete; and finding a pure Nash equilibrium in a congestion game is PLS-complete. Many important problems, such as solving a Simple Stochastic Game or finding a mixed Nash equilibrium of a congestion game, lie in both classes. However, it was strongly believed that their intersection does not have natural complete problems. We show that it does: any problem that lies in both classes can be reduced in polynomial time to the problem of finding a stationary point of a function. Our result has been used to show that computing a mixed equilibrium of a congestion game is also complete for the intersection of PPAD and PLS.

3.16 Learning Augmented Shortest Paths: Dijkstra with Predictions

Guido Schäfer (CWI – Amsterdam, NL)

We study the use of machine learning techniques to solve a fundamental shortest path problem, known as the *single-source many-targets shortest path problem (SSMTSP)*. Basically, our idea is to equip an adapted version of Dijkstra's algorithm with machine learning predictions to solve this problem: Based on the trace of Dijkstra's algorithm, we design a neural network that predicts the shortest path distance after a few iterations. The prediction is then used to prune the search space explored by Dijkstra's algorithm, which significantly reduces the number of operations on the underlying priority queue. Crucially, our algorithms always compute the exact shortest path distances (even if the prediction is inaccurate) and never use more queue operations than the standard algorithm. We provide a probabilistic analysis of our new algorithm which proves a lower bound on the number of saved queue operations. Our bound quantifies these savings depending on the accuracy of the prediction and applies to arbitrary graphs with (partial) random weights. We also present extensive experimental results on random instances showing that the actual savings are oftentimes significantly larger.

3.17 Bicriteria Nash Flows over Time

Daniel Schmand (Universität Bremen, DE)

Joint work of Tim Oosterwijk, Daniel Schmand, Marc Schröder

Flows over time are a natural way to incorporate flow dynamics that arise in various applications such as traffic networks. In this work we introduce a natural variant of the deterministic fluid queuing model in which users aim to minimize their costs subject to arrival at their destination before a pre-specified deadline. We determine the existence and the structure of Nash flows over time and fully characterize the price of anarchy for this model. The price of anarchy measures the ratio of the quality of the equilibrium and the quality of the optimum flow, where we evaluate the quality using two different natural performance measures: the throughput for a given deadline and the makespan for a given amount of flow. While it turns out that both prices of anarchy can be unbounded in general, we provide tight bounds for the important subclass of parallel path networks. Surprisingly, the two performance measures yield different results here.

3.18 Something about Complexity of Equilibria or Bicycles

Alexander Skopalik (University of Twente – Enschede, NL)

In this talk we revisit the complexity of computing a(n) (approximate) pure equilibrium in atomic congestion games. While our understanding of the complexity of pure equilibria is reasonably well, i.e, it is PLS-hard in general but we have a good picture of the landscape of tractable spacial cases, the same is not true of approximate pure equilibria. We discuss the gap between general inapproximability result of Vöcking and S. and the approximation algorithm of Caragiannis et al. for certain cost function. We briefly outline a possible approach towards closing this gap.

3.19 Tutorial Talk: A Faster Algorithm for Quickest Transshipments via an Extended Discrete Newton Method

Martin Skutella (TU Berlin, DE)

Joint work of Miriam Schlöter, Martin Skutella, Khai Van Tran

The Quickest Transshipment Problem is to route flow as quickly as possible from sources with supplies to sinks with demands in a network with capacities and transit times on the arcs. It is of fundamental importance for numerous applications in areas such as logistics, production, traffic, evacuation, and finance. More than 25 years ago, Hoppe and Tardos presented the first (strongly) polynomial-time algorithm for this problem. Their approach, as well as subsequently derived algorithms with strongly polynomial running time, are hardly practical as they rely on parametric submodular function minimization via Megiddo's method of parametric search. The main contribution of this paper is a considerably faster algorithm for the Quickest Transshipment Problem that instead employs a subtle extension of the Discrete Newton Method. This improves the previously best known running time of $\tilde{O}(m^4k^{14})$ to $\tilde{O}(m^2k^5 + m^3k^3 + m^3n)$, where n is the number of nodes, m the number of arcs, and k the number of sources and sinks.

3.20 Stackelberg Max Closure

Bernhard von Stengel (London School of Economics, GB)

LP duality (the strong duality theorem of linear programming) and the minimax theorem for zero-sum games are considered "equivalent" in the sense that one can easily be proved from the other. However, the classic proof by Dantzig (1951) of LP duality from the minimax theorem is flawed. It needs an additional assumption of strict complementarity. We show that this assumption amounts to assuming the Lemma of Farkas, which implies LP duality directly. We fix this with a new, different proof. We also mention a beautiful existing direct proof of the Lemma of Farkas based on minimally infeasible systems of inequalities.

4 Open problems

4.1 Controllability of Traffic Networks

Richard Connors (University of Luxembourg, LU)

Let $G = (N, E)$ be a graph, where N is the set of nodes, and A is the set of arcs. This graph can be taken to represent a transportation network where the nodes are intersections and the arcs are links between intersections. On this network, there is a set of origin-destination (OD) pairs W and a set of paths P which connect OD pairs. The cost, c_a, of travelling on a particular link, a, is given as a sum of the travel time $t_a(v_a)$ and a monetary cost or toll τ_a. The travel time is a function of the flow v_a on the link a, and is assumed to be strictly monotonically increasing. The demand for OD pair w is q^w. The following constraints apply over the system:

$$f_p^w \geq 0 \quad \text{for all } p \in P, w \in W, \qquad q^w = \sum_p f_p^w \quad \text{for all } w \in W, \qquad v_a = \sum_w \sum_p f_p^w \delta_{ap}^w$$

Here f_p^w is the path flow on path p between OD pair w and δ_{ap}^w is an indicator equal to 1 if link a is on path p between OD pair w, and zero otherwise. We will write $\Omega(q)$ as the feasible space of link flows defined in this way.

The link costs are given by strictly increasing functions $c_a(v_a) = t_a(v_a) + \tau_a$. User Equilibrium (UE) wherein the OD flows satisfy Wardrop's first principle, is a solution to $\min_v \int_0^{v_a} c_a(z)\, dz$ for demand-feasible link flow vectors $v \in \Omega(q)$.

It is well-known that that UE is typically inefficient, in that it does not minimise the total network travel time. The flow pattern corresponding to the minimum network travel time satisfies $\min_v \sum_a t_a(v_a)v_a$ again constrained to demand-feasible flows.

The UE path flows equilibrate the total OD cost $\sum_a c_a \delta_{ap}^w$ on each path p connecting OD w, including the cost arising from any link tolls. Applying network tolls therefore enables us to achieve the system optimum (minimum travel time) flow pattern, while satisfying UE with respect to the total costs.

We can apply tolls to achieve many other demand-feasible flow patterns e.g. for a single OD network we could ensure non-zero flow on only one path by applying sufficiently high tolls on the links we wish to remain unused. In this case the toll levels are not unique, and the set of tolled links may be non-unique. This is particularly the case if we allow "negative tolls".

For a given network we say that we have **full controllability** if we can push the UE solution to be ANY feasible link flow solution $v \in \Omega(q)$.

Problem 1: Minimum toll set for full controllability under fixed demand
For a given network with fixed demand, using positive or negative tolls, determine the minimum number of links to toll in order to achieve full controllability (i.e. the toll levels can be varied to achieve the different flow patterns, but the set of tolled links is fixed).

Problem 2: Minimum toll set to achieve system optimal under any demand
For a given network, determine the minimum number of links to toll (positively or negatively) so that UE gives the system optimum flow pattern at ANY level of OD demand(s).

Notes: In the simplest case there is only one OD pair. Certain network topologies may be trivial e.g. spanning tree. For which network topologies can bounds be established?

4.2 Embedding multiple trains in a given network

Katharina Eickhoff (RWTH Aachen, DE)

Let $G = (V, E)$ be a directed graph which corresponds to the infrastructure of a train network. We like to embed two (or multiple) trains with given origins and destinations and given departure and arrival times. Some vertices/edges may be blocked in a specific time window, for example since another fixed train uses the track at this moment.

The easier case is to embed a single train in the network which takes the times the infrastructure can be used into account. One approach to solve the problem is to construct a time expanded network, delete all vertices/edges which cannot be used and find a path in the time expanded network.

For two (or multiple) trains we have to find disjoint path so that the trains do not collide. Shiloach and Perl [1] provide an algorithm which solves this problem in directed acyclic graphs. But disjoint path in general are not enough: Due to safety reasons, the trains will block the nodes/edges longer than the time they are on them. That raises the question how to find "very disjoint" paths in a directed graph.

The discussion during the seminar leads to promising ideas, e.g. how to adapt the algorithm by Shiloach and Perl [1]. However, constructing the time expanded network blows up the instance. Thus, this idea is very time consuming and the question arise how to speed up.

References

1 Shiloach, Y., and Y. Perl. *Finding two disjoint paths between two pairs of vertices in a graph*. Journal of the ACM (JACM) 25(1), pp. 1-9, 1978.

4.3 Concepts for driver's route choice behaviour

Takamasa Iryo (Tohoku University – Sendai, JP)

Dynamic traffic assignment problems considering atomic vehicles can be described as strategic games. On the other hand, finding a pure Nash equilibrium in DTA games is not always easy. Some of them can be NP-hard. Even the existence of a pure Nash equilibrium solution is not a general property. Considering a mixed Nash equilibrium solution is a way to resolve the issue of existence at least, but it may not be natural to assume that drivers stochastically choose routes day-to-day.

Therefore it would be an open question to investigate which concept is more suitable to describe drivers' route choice behaviour other than pure/mixed Nash equilibria. I have asked participants to say some candidates, and the following have been raised:

- Approximate pure nash equilibrium
- Evolutionary dynamics
- CURB (Closed Under Rational Behaviour)
- Preparation set (Voorneveld, 2004)
- Sink equilibria
- Level-K

4.4 Hitting Sets, Randomly

Jannik Matuschke (KU Leuven, BE)

Let $\mathcal{P}$ be a set system on a ground set E. Let $r : \mathcal{P} \to [0, 1]$ be a requirement function on the sets and let $p : E \to [0, 1]$ be a vector of marginals. We are interested in characterizing marginals for which there is a probability distribution on the subsets of E such that the resulting random set intersects each $P \in \mathcal{P}$ with probability at least $r(P)$. We call such a distribution a *feasible decomposition* of p.

It is easy to see that the following condition is necessary for the existence of a feasible decomposition:

$$\sum_{e \in P} p(e) \geq r(P) \qquad \forall P \in \mathcal{P}. \tag{$\star$}$$

The above setting was first studied by Dahan et al. [1], motivated by an application in a network security game. They showed that ($\star$) is also sufficient for the existence of a feasible decomposition, if $\mathcal{P}$ is the set of maximal chains of a partial order $\preceq$ on E (or, equivalently, the set of s-t-paths in a directed acyclic graph) and the requirements r fulfill the following *conservation law*:

$$r(P) + r(Q) = r(P \times_e Q) + r(Q \times_e P) \qquad \forall P, Q \in \mathcal{P}, e \in P \cap Q,$$

where $P \times_e Q = \{f \in P : f \preceq e\} \cup \{f \in Q : e \preceq q\}$. In addition, they provide an efficient algorithm for computing a feasible decomposition in the *affine setting*, where the the requirements are of the form $r(P) = 1 - -\sum_{e \in P} \mu(e)$ for some $\mu(e) \in [0, 1]^E$.

For this affine setting, both the existence result and the algorithm by Dahan et al. [1] can be extended to the case where $\mathcal{P}$ is an abstract network [3]. An abstract network, introduced by [2], is a system $(E, \mathcal{P})$, where the elements of $\mathcal{P} \subseteq 2^E$ are called *paths*. Each path $P \in \mathcal{P}$ is equipped with a complete order $\preceq_P$ of the elements in P, and the following switching axiom is fulfilled: For each $P, Q \in \mathcal{P}$ and each $e \in P \cap Q$, there is a path $P \times_e Q \in \mathcal{P}$ with $P \times_e Q \subseteq \{f \in P : f \preceq_P e\} \cup \{f \in Q : e \preceq_Q q\}$. In particular, the set of s-t-paths in a digraph is an abstract network. As a consequence of the aforementioned generalization, equilibria for the security game studied by Dahan et al. [1] can be efficiently computed even when the underlying digraph contains cycles.

It is also possible to show that when $\mathcal{P}$ is the set of s-t-paths in a directed acyclic graph, the conservation law is actually equivalent to the affine case, and a corresponding weight vector μ can be found efficiently [3]. As a result, a feasible distribution consistent with the marginals can also be computed efficiently in this case.

The conservation law can be extended to general digraphs (with cycles) and even abstract networks. In fact, it resembles Hoffman's [2] supermodularity condition for abstract networks:

$$r(P) + r(Q) \leq r(P \times_e Q) + r(Q \times_e P) \qquad \forall P, Q \in \mathcal{P}, e \in P \cap Q. \tag{1}$$

This leads to the following open question.

Open Problem 1: Assuming that $\mathcal{P}$ is the set of s-t-path in a digraph and that r fulfills (1), is $(\star)$ a sufficient condition for the existence of a feasible decomposition?

One can also investigate the existence of feasible decompositions other set systems. However, it seems that, when going beyond paths (in abstract networks), $(\star)$ is no longer a sufficient condition – simple counter-examples exist, e.g., when $\mathcal{P}$ is the set of bases of a matroid or the set of perfect matchings in a graph [3]. These counter-examples all have a similar underlying structure. So it appears that stronger assumptions on r and $\mathcal{P}$ are necessary here, leading to the second open question.

Open Problem 2: Is there a way to characterize the sufficiencey of $(\star)$ by means of forbidden substructures in r and $\mathcal{P}$?

References

1 Mathieu Dahan, Saurabh Amin, and Patrick Jaillet. *Probability distributions on partially ordered sets and network interdiction games*. Mathematics of Operations Research, 47:458–484, 2022.

2 A. J. Hoffman. *A generalization of max flow-min cut*. Mathematical Programming, 6(1):352–359, 1974.

3 Jannik Matuschke. *Hitting paths with random sets in abstract networks*. Unplushied note. Presented at Dagstuhl Seminar "Dynamic Traffic Models in Transportation Science".

4.5 The Expected Price of Anachy in Wardrop Routing

Daniel Schmand (Universität Bremen, DE)

The price of anarchy (PoA) is a standard concept for analyzing the inefficiency of equilibria in strategic games. Researchers have mainly focused on worst-case bounds for a specific class of games, e.g. Wardrop routing games. It is well known that the PoA is at most 4/3 in Wardrop routing games with affine-linear cost functions and that this bound is tight [1]. If cost functions are polynomials of maximal degree k, this bound increases in the order of $\Theta(k/\log k)$ [2].

However, the known worst case examples no not only need a bad network structure but also a demand that is tailored to the given network. In real-life traffic networks the demand is typically not fixed, but varying over time and due to uncertainties. These facts motivate a study of the Price of Anarchy dependent on the demand. Numerical studies suggest that in many networks the price of anarchy is only large for a small range of demands [3]. Recently, it was shown that the price of anarchy tends to 1 if the demand goes to 0 or infinity [5]. Additionally, Cominetti, Dose, and Scarsini analyzed the behavior and properties of the PoA-function in dependence of the demand [4].

The question of establishing a bound on the expected price of anachy for a demand drawn from some given distribution is still open. Specifically, for a fixed network let W_d be the cost of a Wardrop equilibrium, and O_d the cost of an optimal flow, both with demand d. We are interested in establishing a bound on $\mathbb{E}\left[\frac{W_d}{O_d}\right]$, where the expectation is taken over the demand drawn from a given probability distribution. During and after the 2015 Dagstuhl seminar, we have analyzed the Pigou network with costs 1 and x^k. We have shown that the expected price of anarchy for a demand drawn from a uniform distribution U[0,D] for $D \in \mathbb{R}$ is in the order of $\Theta(\log k)$. It is still open to extend this result to other network structures.

References

1 Roughgarden, Tardos: *How bad is selfish routing*
2 Roughgarden: *The price of anarchy is independent of the network topology*
3 O'Hare, Connors, Watling: *Mechanisms that Govern how the Price of Anarchy varies with Travel Demand*
4 Cominetti, Dose, Scarsini: *The price of anarchy in routing games as a function of the demand*
5 Colini-Baldeschi, Cominetti, Mertikopoulos, Scarsini: *When is selfish routing bad? The price of anarchy in light and heavy traffic*

4.6 Complexity of network loading in the fluid-queue model

Martin Skutella (TU Berlin, DE)

Our open problem is to prove results on the complexity of the network loading problem in the fluid-queue model. The most basic yet interesting setting is as follows: We consider a directed graph $D = (V, A)$ with capacities $u \in \mathbb{Z}^A_{>0}$ and transit times $\tau \in \mathbb{Z}^A_{\geq 0}$ on the arcs. Moreover, we are given directed paths $P_1, \dots, P_k$ along with fixed flow rates $r_1, \dots, r_k$. Assuming that we start to send flow at rates r_i into every path P_i at time 0, the network loading problem is to compute the resulting flow rate functions on the arcs of the network. Notice that the time at which a flow particle traveling along path P_i arrives on some arc $a \in P_i$ not only depends on its departure time at the start node of P_i but also on the latencies experienced on the predecessor edges on path P_i. This fact induces involved interdependencies among the flow rate functions on the arcs of the network. Notice that we can assume without loss of generality that all paths share the same start- and end-node.

We are interested in the computational complexity of decision problems such as, for example: For a given arc $a \in A$, a given point in time θ, and a given threshold value $s \in \mathbb{Q}_{>0}$,

- is the flow rate on arc a at time θ at least s?
- is there a point in time $\theta' \geq \theta$ such that the flow rate on arc a at time θ' is at most s?

The described problem setting can also be modified, for example by considering more general inflow patterns on the given paths (e.g., piecewise constant with a limited number of break points), or by allowing for time-dependent arc capacities etc.

The open problem is to come up with NP-hardness, PSPACE-hardness, or maybe even undecidability results in this context.

4.7 Towards defining a set of open problems for transportation analysis

David Watling (University of Leeds, GB)

Many branches of mathematics have benefited from having well-defined, unsolved problems, yet to my knowledge there has never been a clear set of such problems defined in transportation analysis. Such problems could serve, for example, as a focus for theoretical work to prove some property, or to extend the sufficient conditions under which a property is known to hold. Alternatively they could be a focus for numerical simulation work, developing a body of evidence that might lead to a more well-defined conjecture, or finding counter-examples that we then know must be ruled out by any sufficient conditions of a theoretical proof. If an unproven conjecture amasses sufficient numerical evidence in support of it, then it might be used as a starting point from which to prove other results.

The purpose of the presentation is to develop this suggestion further, with the intention that it could lead to the start of a technical discussion group that would aim to develop an initial specification of such a set of open problems. The motivation for selecting problems could be diverse. For example, it could be in order to bring together a wide body of knowledge in order to make a clear statement of what is currently known and not known; for example, in dynamic traffic assignment, for which kinds of component sub-models, network structure, etc. are we able to guarantee existence, uniqueness or stability of equilibrium, and in which cases do we know of counter-examples? Alternatively, the selection of such problems could be also be based on what the community believe would be useful results, even if extremely difficult to prove theoretically, as a basis for developing further work. A third possibility is to consider the way in which real transportation systems are changing, and are expected to change in the future, and the new demands they place on the modelling of strategic agents; for example, automated vehicles and more shared transport options give rise to very different kinds of problem that will need pressing attention. The talk will not actually aim to precisely define any such problems, but is intended as a stimulus for more discussion at Dagstuhl and beyond.

4.8 Identifying paths within a given bound

David Watling (University of Leeds, GB)

A kind of "behavioural" route choice model has been defined in the transportation literature where drivers choose routes based on random utility theory, but where the random error terms are bounded in a special way, such that the difference in random utility between any given route and a minimum cost route for that movement is a random variable on a closed interval. The consequence of this is that travellers use all routes with a cost no greater than $x\%$ more than minimum cost. In order to implement such models in an equilibrium framework it would be desirable to be able to solve sub-problems which, for fixed (i.e. flow-independent) arc costs, identify the set of all routes that are within such a bound. The discussions discussed the behavioural aspects of this model and the complexity of the problem. As part of the discussions ensuing, an outline of a potential solution approach was also proposed (by Martin Strehler):

1. Do shortest path from origin to all vertices.
2. Do shortest path from destination to all vertices with edges reversed (so gives shortest paths from all vertices to destination).
3. Check for vertices which can be deleted by seeing which routes via the vertex violates bound.
4. Identify edges by the same kind of way.
5. What remains is a minimal subnetwork containing all paths (but not all paths in the subnetwork will satisfy bound).

Participants

- Joseph Chow
New York University, US
- Richard Connors
University of Luxembourg, LU
- Katharina Eickhoff
RWTH Aachen, DE
- Miriam Fischer
Imperial College London, GB
- Gunnar Flötteröd
Linköping University, SE & Swedish National Road and Transport Research Institute (VTI) – Stockholm SE
- Martin Gairing
University of Liverpool, GB
- Yiannis Giannakopoulos
Universität Erlangen-Nürnberg, DE
- Lukas Graf
Universität Augsburg, DE
- Tobias Harks
Universität Augsburg, DE
- Martin Hoefer
Goethe-Universität – Frankfurt am Main, DE
- Takamasa Iryo
Tohoku University – Sendai, JP
- Max Klimm
TU Berlin, DE
- Ekkehard Köhler
BTU Cottbus-Senftenberg, DE
- Jannik Matuschke
KU Leuven, BE
- Haruko Nakao
University of Luxembourg, LU
- Neil Olver
London School of Economics, GB
- Carolina Osorio
HEC Montréal, CA & Google – Mountain View, US
- Dario Paccagnan
Imperial College London, GB
- Britta Peis
RWTH Aachen, DE
- Rahul Savani
University of Liverpool, GB
- Guido Schäfer
CWI – Amsterdam, NL
- Daniel Schmand
Universität Bremen, DE
- Alexander Skopalik
University of Twente – Enschede, NL
- Martin Skutella
TU Berlin, DE
- Martin Strehler
Westsächsische Hochschule Zwickau, DE
- Bernhard von Stengel
London School of Economics, GB
- David Watling
University of Leeds, GB

Report from Dagstuhl Seminar 22201

The Constraint Satisfaction Problem: Complexity and Approximability

Martin Grohe*[1], Venkatesan Guruswami*[2], Dániel Marx*[3], and Stanislav Živný*[4]

1 **RWTH Aachen, DE.** grohe@informatik.rwth-aachen.de
2 **UC Berkeley, US.** venkatg@berkeley.edu
3 **CISPA – Saarbrücken, DE.** marx@cispa.de
4 **University of Oxford, GB.** standa.zivny@cs.ox.ac.uk

Abstract

Constraint satisfaction has always played a central role in computational complexity theory; appropriate versions of CSPs are classical complete problems for most standard complexity classes. CSPs constitute a very rich and yet sufficiently manageable class of problems to give a good perspective on general computational phenomena. For instance, they help to understand which mathematical properties make a computational problem tractable (in a wide sense, e.g., polynomial-time solvable, non-trivially approximable, fixed-parameter tractable, or definable in a weak logic). In the last 15 years, research activity in this area has significantly intensified and hugely impressive progress was made. The Dagstuhl Seminar 22201 "The Constraint Satisfaction Problem: Complexity and Approximability" was aimed at bringing together researchers using all the different techniques in the study of the CSP so that they can share their insights obtained during the past four years. This report documents the material presented during the course of the seminar.

Seminar May 15–20, 2022 – http://www.dagstuhl.de/22201

2012 ACM Subject Classification Theory of computation → Problems, reductions and completeness

Keywords and phrases Constraint satisfaction problem (CSP); Computational complexity; Hardness of approximation; Universal algebra; Semidefinite programming

Digital Object Identifier 10.4230/DagRep.12.5.112

1 Executive Summary

Martin Grohe
Venkatesan Guruswami
Dániel Marx
Stanislav Živný

The *constraint satisfaction problem*, or CSP in short, provides a unifying framework in which it is possible to express, in a natural way, a wide variety of computational problems dealing with mappings and assignments, including satisfiability, graph/hypergraph colorability, and systems of equations. The CSP framework originated around 40 years ago independently in artificial intelligence, database theory, and graph theory, under three different guises, and it was realised only in the late 1990s that these are in fact different faces of the same fundamental

* Editor / Organizer

problem. Nowadays, the CSP is extensively used in theoretical computer science, being a mathematical object offering a good balance between generality and structure that provides an excellent laboratory both for classification methods and for algorithmic techniques, while in AI and more applied areas of computer science this framework is widely regarded as a versatile and efficient way of modelling and solving a variety of real-world problems, such as planning and scheduling, software verification and natural language comprehension, to name just a few. An instance of CSP consists of a set of variables, a set of values for the variables, and a set of constraints that restrict the combinations of values that certain subsets of variables may take. Given such an instance, the possible questions include (a) deciding whether there is an assignment of values to the variables so that every constraint is satisfied, or optimising such assignments in various ways, or (b) finding an assignment satisfying as many constraints as possible. There are many important modifications and extensions of this basic framework, e.g., those that deal with counting assignments or involve soft or global constraints.

Constraint satisfaction has always played a central role in computational complexity theory; appropriate versions of CSPs are classical complete problems for most standard complexity classes. CSPs constitute a rich and yet sufficiently manageable class of problems to give a good perspective on general computational phenomena. For instance, they help to understand which mathematical properties make a problem tractable (in a wide sense, e.g., polynomial-time solvable or non-trivially approximable, fixed-parameter tractable or definable in a weak logic). One of the most striking features of this study is the variety of different branches of mathematics (including universal algebra and logic, combinatorics and graph theory, probability theory and mathematical programming) that are used to achieve deep insights into CSP, and this Dagstuhl Seminar aims to contribute towards further synergy in the area.

After about 15 years of intense research activity and hugely impressive progress, the culmination of the algebraic-approach to fixed-template CSPs was the resolution of the Feder-Vardi conjecture independently by Bulatov and Zhuk in 2017. While some fundamental questions (such as a fine-grained understanding of tractable CSPs) remain open, new research directions on generalizations of CSPs started emerging. The fixed-template promise CSP (PCSP) is among the most promising new directions of research motivated by better understanding computational hardness or tractability. PCSPs are a vast generalization of CSPs where each predicate has a strong and a weak form and given a CSP instance, the objective is to distinguish if the strong form can be satisfied vs. even the weak form cannot be satisfied. A prime and well-known example is the approximate graph coloring problem: distinguish k-colorable graphs from graphs that are not even ℓ-colorable, for some fixed $k < \ell$. The main topic of this seminar is PCSPs, a highly ambitious research direction with intriguing connections to both old open problems (such as the approximate graph coloring problem) and new research directions (such as generalizations of submodularity, a key concept in optimization).

The recent flurry of activity on the topic of the seminar is witnessed by five previous Dagstuhl seminars, titled "Complexity of constraints" (06401) and "The CSP: complexity and approximability" (09441, 12541, 15301, 18231), that were held in 2006, 2009, 2012, 2015, and 2018 respectively. This seminar was a follow-up to the 2009, 2012, 2015, and 2018 seminars. Indeed, the exchange of ideas at the 2009, 2012, 2015, and 2018 seminars has led to ambitious new research projects and to establishing regular communication channels. There is clearly the potential for further systematic interaction that will keep on cross-fertilising the areas and opening new research directions. The 2022 seminar brought together 46 researchers from different highly advanced areas of constraint satisfaction and involved many specialists who use universal-algebraic, combinatorial, geometric, and probabilistic techniques to study CSP-related algorithmic problems. 10 of the participants attended remotely owing to the global COVID pandemic. The participants presented, in 24 talks, their recent results on a number of important questions concerning the topic of the seminar. One particular feature of this seminar is a significant increase in the number of talks involving multiple subareas and approaches within its research direction – a definite sign of the growing synergy, which is one of the main goals of this series of seminars.

Concluding remarks and future plans. The seminar was well received as witnessed by the high rate of accepted invitations and the great degree of involvement by the participants. Because of a multitude of impressive results reported during the seminar and active discussions between researchers with different expertise areas, the organisers regard this seminar as a great success. With steadily increasing interactions between such researchers, we foresee another seminar focusing on the interplay between different approaches to studying the complexity and approximability of the CSP. Finally, the organisers wish to express their gratitude to the Scientific Directors of the Dagstuhl Centre for their support of the seminar.

Description of the Topics of the Seminar

With the resolution of the Feder-Vardi conjecture on finite-domain CSPs by Bulatov and Zhuk in 2017, the field has moved on to more generalizations of finite-domain decision CSPs. One of the main emerging areas is that of Promise CSPs (PCSPs), which are a huge generalization of CSPs.

Promise CSPs

The study of PCSPs is about approximability of perfectly satisfiable instances. This exciting area been the main theme of this workshop, with about half of the talks on PCSPs.

The workshop started with a 2-hour tutorial on the basics of PCSPs.

- In the first part, Opršal talked about basic tools used in the complexity analysis of PCSPs, namely minions and free structures.
- In the second part, Opršal talked about an application of topology in the study of complexity of graph colorings.

There were three topologists attending the seminar and one of them gave a talk.

- Kozlov gave a talk on uses of topology in theoretical computer science and homomorphism questions in particular.

There were 9 contributed talks on recent results on PCSPs.

- Barto talked about his work on "baby PCP", which is a combinatorial variant of a (weaker version of) the PCP theorem.
- Brakensiek gave two talks: one on minions related to SDP relaxations for PCSPs, and one on very recent results on robust solvability of certain PCSPs.
- Butti gave a talk about Sherali-Adams relaxations for valued PCSPs.
- Ciardo gave a talk about Sherali-Adams relaxations for PCSPs, and more generally about studying hierarchies of convex relaxations via tensorisation.
- Dalmau gave a talk about his work on a condition that PCSPs solvable by local consistency algorithms have to satisfy.

- Kompatscher gave an overview of existing results on finite tractability of PCSPs.
- Mottet gave a talk about his very recent work on first-order definable PCSPs and a relationship to PCSPs solved by local consistency algorithms.
- Nagajima gave a talk on the complexity of linearly-ordered colorings.

Approximability of CSPs

The second main theme of the seminar was recent progress in the area of approximability of CSPs. This part of the programme started with a 2-hour tutorial:
- Tulsiani gave a tutorial on how to approximate CSPs via the Sum of Squares hierarchy of SDP relaxations, including related work on high-dimensional expanders.

Three participants gave talks on approximability of CSPs.
- Bhangale gave a talk on approximability of solvable linear equations over non-Abelian groups.
- Kothari presented a refutation algorithm for unsatisfiable instances of SAT and CSP.
- Potechin talked about an algorithm for Max NAE-SAT with clauses of lengths 3 and 5.

Infinite-domain CSPs

The area of infinite-domain CSPs has seen some exciting progress in last few years with new techniques being developed.
- Bodirsky talked about methods for transferring complexity classification result between different classes of infinite-domain CSPs.
- Nagy talked about the complexity of CSPs over random hypergraphs.
- Pinsker gave an overview of recent developments around the dichotomy conjecture for infinite-domain CSPs.

Generalisations of CSPs

There are several interesting research directions related to CSPs, including the number of solutions, CSPs with a bound on the number of occurrences for each variable, quantified CSPs, a generalisation to the ideal membership problem, etc. The seminar featured 5 talks on these topics.
- Austrin presented an algorithmic lowerbound for perfect matching on random graphs, a canonical example of an edge CSP.
- Bulatov gave a talk the ideal membership problem.
- Chen presented his results on property testing of homomorphism inadmissibility.
- Kazeminia presented a classification of counting graph homomorphism modulo 2.
- Zhuk presented a complexity classification of quantified CSPs.

2 Table of Contents

3 Overview of Talks

3.1 Perfect Matching in Random Graphs is as Hard as Tseitin

Per Austrin (KTH Royal Institute of Technology – Stockholm, SE)

Joint work of Per Austrin, Kilian Risse
Main reference Per Austrin, Kilian Risse: "Perfect Matching in Random Graphs is as Hard as Tseitin", in Proc. of the 2022 ACM-SIAM Symposium on Discrete Algorithms, SODA 2022, Virtual Conference, pp. 979–1012, SIAM, 2022.
URL https://doi.org/10.1137/1.9781611977073.43

We study the complexity of proving that a sparse random regular graph on an odd number of vertices does not have a perfect matching, and related problems involving each vertex being matched some pre-specified number of times. We show that this requires proofs of degree $\Omega(n/\log n)$ in the Polynomial Calculus (over fields of characteristic $\neq 2$) and Sum-of-Squares proof systems, and exponential size in the bounded-depth Frege proof system. This resolves a question by Razborov asking whether the Lovász-Schrijver proof system requires n^δ rounds to refute these formulas for some $\delta > 0$. The results are obtained by a worst-case to average-case reduction of these formulas relying on a topological embedding theorem which may be of independent interest.

3.2 Combinatorial Value and Gap Amplification

Libor Barto (Charles University – Prague, CZ)

Joint work of Libor Barto, Filip Bialas, Marcin Kozik
Main reference Libor Barto, Marcin Kozik: "Combinatorial Gap Theorem and Reductions between Promise CSPs", in Proc. of the 2022 ACM-SIAM Symposium on Discrete Algorithms, SODA 2022, Virtual Conference, pp. 1204–1220, SIAM, 2022.
URL https://doi.org/10.1137/1.9781611977073.50

The combinatorial value of a CSP instance is a certain positive integer, which is equal to one iff the instance is satisfiable. In a joint work with Marcin Kozik we proved that a natural reduction amplifies the gap between value one and value strictly greater than one. I will talk about this result, its motivations, applications, recent developments (with Filip Bialas), and future directions.

3.3 On Characterizing the Inapproximability of Satisfiable CSPs

Amey Bhangale (University of California – Riverside, US)

Joint work of Amey Bhangale, Subhash Khot, Dor Minzer
Main reference Amey Bhangale, Subhash Khot: "Optimal inapproximability of satisfiable k-LIN over non-abelian groups", in Proc. of the STOC '21: 53rd Annual ACM SIGACT Symposium on Theory of Computing, Virtual Event, Italy, June 21-25, 2021, pp. 1615–1628, ACM, 2021.
URL https://doi.org/10.1145/3406325.3451003

In his beautiful result, Raghavendra gave a complete characterization of the inapproximability for every constraint satisfaction problem, assuming the Unique Games conjecture. This result however loses perfect completeness. In this talk, I will discuss the inapproximability of satisfiable linear equations over non-Abelian groups. Unlike Abelian groups, it is NP-complete to decide if a given system of linear equations over non-Abelian groups is satisfiable or not. We show tight NP-hardness results for approximately solving a satisfiable system of linear equations over non-Abelian groups. I will also discuss new techniques for characterizing the inapproximability of satisfiable CSPs.

3.4 CSP Complexity Classification Transfer

Manuel Bodirsky (TU Dresden, DE)

Joint work of Manuel Bodirsky, Peter Jonsson, Barnaby Martin, Antoine Mottet, Žaneta Semanišsinová

Many natural classes of computational problems can be formulated as CSPs for first-order expansions of some base structure, such as $(\mathbb{Q}, <)$. In this talk I will present methods how to transfer classification results for such classes from one base structure to another. This is particularly interesting if the base structure is a product structure, in which case one may hope to obtain the classification from the classifications of the factors.

Our results imply new classification results for CSPs that concern Allen's Interval Algebra, the n-dimensional Block Algebra, and the Cardinal Direction Calculus and solve some open problems about these formalisms from 1999 and 2002.

3.5 Robust Algorithms for Promise CSPs

Joshua Brakensiek (Stanford University, US)

Joint work of Joshua Brakensiek, Venkatesan Guruswami, Sai Sandeep

For a constraint satisfaction problem (CSP), a robust satisfaction algorithm is one that outputs an assignment satisfying most of the constraints on instances that are near-satisfiable. It is known that the CSPs that admit efficient robust satisfaction algorithms are precisely those of bounded width.

In this talk, I will discuss our recent work on robust satisfaction algorithms for Promise CSPs, which are a vast generalization of CSPs. I will present robust SDP rounding algorithms under some general conditions, namely the existence of majority or alternating threshold polymorphisms. On the hardness front, I will prove that the lack of such polymorphisms makes the PCSP hard for a large class of symmetric Boolean predicates. Our method involves a novel method to argue SDP gaps via the absence of certain colorings of the sphere, with connections to sphere Ramsey theory.

3.6 Some Interesting Minions

Joshua Brakensiek (Stanford University, US)

Joint work of Joshua Brakensiek, Venkatesan Guruswami, Sai Sandeep

In the development of the theory of Promise CSPs, it has become increasingly common to capture the structure of tractable classes of PCSPs through the existence of a minion homomorphism, where one maps a chosen minion into the polymorphisms of the particular PCSP. As the field of PCSPs is still quite young, only a handful of interesting minions are currently understood. In this talk, I will present some recently-discovered minions with interesting properties, including one that captures tractability according to the "exact" basic semi-definite program.

3.7 On the Complexity of CSP-Based Ideal Membership Problems

Andrei A. Bulatov (Simon Fraser University – Burnaby, CA)

Joint work of Andrei A. Bulatov, Akbar Rafiey
Main reference Andrei A. Bulatov, Akbar Rafiey: "On the Complexity of CSP-based Ideal Membership Problems", CoRR, Vol. abs/2011.03700, 2020.
URL https://arxiv.org/abs/2011.03700

In this talk we consider the Ideal Membership Problem (IMP for short), in which we are given polynomials $f_0, f_1, \ldots, f_k$ and the question is to decide whether f_0 belongs to the ideal generated by $f_1, \ldots, f_k$. In the more stringent version the task is also to find a proof of this fact. The IMP underlies many proof systems based on polynomials such as Nullstellensatz, Polynomial Calculus, and Sum-of-Squares (SOS). In such applications the IMP usually involves so called combinatorial ideals that arise from a variety of discrete combinatorial problems. This restriction makes the IMP significantly easier and in some cases allows for an efficient solution algorithm.

The first part of this paper follows the work of Mastrolilli [SODA 2019] who initiated a systematic study of IMPs arising from Constraint Satisfaction Problems (CSP) of the form CSP(G), that is, CSPs in which the type of constraints is limited to relations from a set G.

We show that many CSP techniques can be translated to IMPs thus allowing us to significantly improve the methods of studying the complexity of the IMP. We also develop universal algebraic techniques for the IMP that have been so useful in the study of the CSP. This allows us to prove a general necessary condition for the tractability of the IMP, and three sufficient ones. The sufficient conditions include IMPs arising from systems of linear equations over GF(p), p prime, and also some conditions defined through special kinds of polymorphisms.

Our work has several consequences and applications. First, we introduce a variation of the IMP and based on this propose a unified framework, different from the celebrated Buchberger's algorithm, to construct a bounded degree Groebner Basis. Our algorithm, combined with the universal algebraic techniques, leads to polynomial-time construction of Groebner Basis for many combinatorial problems.

3.8 Sherali-Adams Meets Weisfeiler-Leman on (Promise Valued) CSPs

Silvia Butti (UPF – Barcelona, ES)

Joint work of Libor Barto, Silvia Butti, Víctor Dalmau

In this talk I will give an overview of recent work concerning the relationship between levels of the Sherali-Adams hierarchy for CSP and the equivalence relation induced by the 1-dimensional Weisfeiler-Leman algorithm on the set of CSP instances. In particular, I will discuss how the feasibility of a linear program obtained using the Sherali-Adams method can be decomposed into three separate components, and how this fact can be used to infer results about solvability of CSPs by distributed algorithms. I will explain how these results can be extended to the more general framework of Promise Valued CSP, and discuss how the first level of the Sherali-Adams hierarchy is no longer equivalent to the Basic Linear Programming relaxation in this broader framework.

3.9 Testability of Homomorphism Inadmissibility

Hubie Chen (Birkbeck, University of London, GB)

Joint work of Hubie Chen, Yuichi Yoshida
Main reference Hubie Chen, Yuichi Yoshida: "Testability of Homomorphism Inadmissibility: Property Testing Meets Database Theory", in Proc. of the 38th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems, PODS 2019, Amsterdam, The Netherlands, June 30 – July 5, 2019, pp. 365–382, ACM, 2019.
URL https://doi.org/10.1145/3294052.3319679

Traditional notions of algorithmic efficiency, such as that of polynomial time, permit each input to be read in its entirety. However, the modern necessity of performing computations on large bodies of data poses situations where inputs are so large that it is prohibitively expensive to fully read them. These situations motivate the study of algorithms that only read a part of each input. Property testing provides a framework wherein it is possible to present such algorithms.

In property testing, the input is typically read via random access queries, and an algorithm's desideratum is to distinguish between inputs that satisfy a property and inputs that are ϵ-far from satisfying the property. Intuitively speaking, an object is called ϵ-far from satisfying a property if one needs to modify more than an ϵ-fraction of the object in order for it satisfy the property; the exact definition depends on the context. The restriction on the allowed inputs leads to the possibility of algorithms that do not read the entire input.

In this work, we utilize the perspective of property testing to consider the testability of relational database queries. A primary motivation is the desire to avoid reading an entire database to decide a property hereof. We focus on conjunctive queries, which are the most basic and heavily studied database queries. Each conjunctive query can be represented as a relational structure $\mathbf{A}$ such that deciding if the conjunctive query is satisfied by a relational structure $\mathbf{B}$ is equivalent to deciding if there exists a homomorphism from $\mathbf{A}$ to $\mathbf{B}$. We phrase our results in terms of homomorphisms. Precisely, we study, for each relational structure $\mathbf{A}$, the testability of homomorphism inadmissibility from $\mathbf{A}$. We consider algorithms that have oracle access to an input relational structure $\mathbf{B}$ and that distinguish, with high probability, the case where there is no homomorphism from $\mathbf{A}$ to $\mathbf{B}$, from the case where one needs to remove a constant fraction of tuples from $\mathbf{B}$ in order to suppress all such homomorphisms.

We provide a complete characterization of the structures **A** from which one can test homomorphism inadmissibility with one-sided error by making a constant number of queries to **B**. Our characterization shows that homomorphism inadmissibility from **A** is constant-query testable with one-sided error if and only if the core of **A** is α-acyclic. We also show that the injective version of the problem is constant-query testable with one-sided error if **A** is α-acyclic; this result generalizes existing results for testing subgraph-freeness in the general graph model.

3.10 Relaxing with Tensors

Lorenzo Ciardo (University of Oxford, GB)

Joint work of Lorenzo Ciardo, Stanislav Živný
Main reference Albert Atserias, Víctor Dalmau: "Promise Constraint Satisfaction and Width", in Proc. of the 2022 ACM-SIAM Symposium on Discrete Algorithms, SODA 2022, Virtual Conference, pp. 1129–1153, SIAM, 2022.
URL https://doi.org/10.1137/1.9781611977073.48

The Sherali-Adams linear programming hierarchy is a powerful algorithmic framework that is naturally applicable to the context of promise constraint satisfaction problems. This talk, based on recent joint work with Standa Živný, aims to show that the structure lying at the core of the hierarchy has a multilinear nature, as it essentially consists in a space of tensors enjoying certain increasingly tight symmetries. The geometry of this space allows then to establish non-solvability of the approximate graph coloring problem via constantly many rounds of Sherali-Adams. Besides this primary application, our tensorisation approach introduces a new tool to the study of various hierarchies of algorithmic relaxations for computational problems within (and, possibly, beyond) the context of constraint satisfaction. For example, both the algorithmic frameworks known as local consistency checking and "sum of squares" Lasserre hierarchy for the SDP relaxation can be naturally described geometrically through the tensorisation technique, by considering different spaces of tensors.

3.11 Promise Constraint Satisfaction and Width

Víctor Dalmau (UPF – Barcelona, ES)

Joint work of Albert Atserias, Víctor Dalmau
Main reference Albert Atserias, Víctor Dalmau: "Promise Constraint Satisfaction and Width", in Proc. of the 2022 ACM-SIAM Symposium on Discrete Algorithms, SODA 2022, Virtual Conference, pp. 1129–1153, SIAM, 2022.
URL https://doi.org/10.1137/1.9781611977073.48

We study the power of the bounded-width consistency algorithm in the context of the fixed-template Promise Constraint Satisfaction Problem (PCSP). Our main technical finding is that the template of every PCSP that is solvable in bounded width satisfies a certain structural condition implying that its algebraic closure-properties include weak near unanimity polymorphisms of all large arities. While this parallels the standard (non-promise) CSP theory, the method of proof is quite different and applies even to the regime of sublinear width. We also show that, in contrast with the CSP world, the presence of weak near unanimity polymorphisms of all large arities does not guarantee solvability in bounded width. The separating example is even solvable in the second level of the Sherali-Adams (SA) hierarchy of linear programming relaxations. This shows that, unlike for CSPs, linear programming can be stronger than bounded width. A direct application of these methods also show that the problem of q-coloring p-colorable graphs is not solvable in bounded or even sublinear width, for any two constants p and q such that $3 \leq p \leq q$. Turning to algorithms, we note that Wigderson's algorithm for coloring 3-colorable graphs with n vertices is implementable in width 4. Indeed, by generalizing the method we see that, for any $\epsilon > 0$ smaller than 1/2, the optimal width for solving the problem of $O(n^{\epsilon})$-coloring 3-colorable graphs with n vertices lies between $n^{1-3\epsilon}$ and $n^{1-2\epsilon}$. The upper bound gives a simple $\exp(\Theta(n^{1-2\epsilon}\log(n)))$-time algorithm that, asymptotically, beats the straightforward $\exp(\Theta(n^{1-\epsilon}))$ bound that follows from partitioning the graph into $O(n^{\epsilon})$ many independent parts each of size $O(n^{1-\epsilon})$.

3.12 Applied Topology and CSPs

Dmitry Feichtner-Kozlov (Universität Bremen, DE)

In this talk we give a brief introduction to the subject of Combinatorial Algebraic Topology, which lies at the crossroads of combinatorics, topology, and computation. A special emphasis is given to the family of the so-called Hom-complexes.

Given any two graphs G and H, it is possible to construct a combinatorial cell complex (more precisely a prodsimplicial complex) whose set of vertices is the set of all graph homomorphisms from G to H, and whose cellular structure reflects the commutativity relations within various groups of homomorphisms. We illuminate various structural properties of this family of cell complexes, especially with a view towards their recent applications to the complexity theory of CSPs.

3.13 Modular Counting of Graph Homomorphisms

Amirhossein Kazeminia (Simon Fraser University – Burnaby, CA)

Joint work of Andrei A. Bulatov, Amirhossein Kazeminia
Main reference Andrei A. Bulatov, Amirhossein Kazeminia: "Complexity classification of counting graph homomorphisms modulo a prime number", CoRR, Vol. abs/2106.04086, 2021.
URL https://arxiv.org/abs/2106.04086

Counting graph homomorphisms and its generalizations such as the Counting Constraint Satisfaction Problem (CSP), its variations, and counting problems in general have been intensively studied since the pioneering work of Valiant. While the complexity of exact counting of graph homomorphisms (Dyer and Greenhill, 2000) and the counting CSP (Bulatov, 2013, and Dyer and Richerby, 2013) is well understood, counting modulo some natural number has attracted considerable interest as well. In their 2015 paper Faben and Jerrum suggested a conjecture stating that counting homomorphisms to a fixed graph H modulo a prime number is hard whenever it is hard to count exactly, unless H has automorphisms of certain kind. In this talk we talk about this conjecture and how to prove it. As a part of this investigation we develop techniques that widen the spectrum of reductions available for modular counting and apply to the general CSP rather than being limited to graph homomorphisms.

3.14 Finitely Tractable PCSPs

Michael Kompatscher (Charles University – Prague, CZ)

Besides linear programming relaxations, one of the main strategies in proving the tractability of promise constraint satisfaction problems is to reduce them to tractable CSPs. In particular, PCSP($\mathbf{A}, \mathbf{B}$) is called finitely tractable, if there is a finite structure $\mathbf{C}$ such that there are homomorphisms $\mathbf{A} \to \mathbf{C} \to \mathbf{B}$ and CSP($\mathbf{C}$) is tractable. In this case every algorithm for CSP($\mathbf{C}$) also solves PCSP($\mathbf{A}, \mathbf{B}$). In this talk I will give an overview of known results on finitely tractable PCSPs and discuss some open questions.

3.15 Refuting Smoothed k-SAT Formulas and a Proof of Feige's Conjecture

Pravesh K. Kothari (Carnegie Mellon University – Pittsburgh, US)

Joint work of Venkatesan Guruswami, Pravesh K. Kothari, Peter Manohar
Main reference Venkatesan Guruswami, Pravesh K. Kothari, Peter Manohar: "Algorithms and Certificates for Boolean CSP Refutation: "Smoothed is no harder than Random"", CoRR, Vol. abs/2109.04415, 2021.
URL https://arxiv.org/abs/2109.04415

I'll present a new algorithm to refute, that is, efficiently find certificates of unsatisfiability, for smoothed instances of k-SAT and other CSPs. Smoothed instances are produced by starting with a worst-case instance and flipping each literal in every clause independently with a small constant probability. The "clause" structure in such instances is arbitrary and the only randomness is in the literal patterns. The trade-off between the density, i.e., the number of constraints in the instance, and the running time required by the algorithm matches (up to polylogarithmic factors density) the best known (and possibly sharp) trade-off for random k-SAT. As a corollary, we also obtain a simpler analysis for refuting random k-SAT formulas.

Our analysis inspires a new method that we call spectral double counting that we use to prove Feige's 2008 conjecture on the extremal trade-off between the size of the hypergraph and its girth that generalizes the Moore bound on the girth of irregular graphs proved by Alon, Hoory, and Linial (2002). As a corollary, we obtain that smoothed instances of 3-SAT with arbitrary clause structure and $n^{1.4}$ ($\ll n^{1.5}$, the threshold for known efficient refutation) constraints admit polynomial-size certificates of unsatisfiability with appropriate generalization to all k-CSPs. This extends the celebrated work of Feige, Kim, and Ofek (2006) who proved a similar result for random 3-SAT. FKO's proof uses a 2nd-moment-method based argument that strongly exploits the randomness of the clause structure. Our proof gives a simpler argument that extends to "worst-case" clause structures.

Taken together, our results show that smoothed instances with arbitrary clause structure and significantly less randomness enjoy the same thresholds for both refutation algorithms and certificates of unsatisfiability as random instances.

3.16 Sandwiches for PCSPs

Antoine Mottet (TU Hamburg, DE)

A PCSP is said to be first-order definable if there exists a first-order sentence that holds on all Yes-instances and does not hold on No-instances. I will prove that PCSP($\mathbf{A}, \mathbf{B}$) is first-order definable exactly when there exists a finite sandwich $\mathbf{A} \to \mathbf{C} \to \mathbf{B}$ where $\mathbf{C}$ has finite duality. Using similar tools, I will show how to give a sandwich-characterisation of PCSPs solvable by local consistency.

3.17 Hypergraphs in the Post-Proof Era

Tomáš Nagy (TU Wien, AT)

We consider CSPs over random k-uniform hypergraphs. Using the theory of smooth approximations, we can reduce the problem of finding an injective solution to such CSPs to a finite-domain CSP. On the other hand, surprisingly the problem of finding an arbitrary solution to such CSPs does not naturally reduce to a finite-domain CSP. We will present algorithmic techniques based on polymorphisms that allow us to reduce the latter problem to the aforementioned problem of finding an injective solution. This way, we confirm the Bodirsky-Pinsker conjecture for the random k-uniform hypergraphs for any k.

3.18 Linearly Ordered Colourings of Hypergraphs

Tamio-Vesa Nakajima (University of Oxford, GB)

Joint work of Tamio-Vesa Nakajima, Stanislav Živný
Main reference Tamio-Vesa Nakajima, Stanislav Zivný: "Linearly ordered colourings of hypergraphs", CoRR, Vol. abs/2204.05628, 2022.
URL https://doi.org/10.48550/arXiv.2204.05628

A linearly ordered (LO) k-colouring of an r-uniform hyper graph assigns an integer colour from 1 to k to each vertex of the hyper graph, such that every edge contains a unique maximum colour. We will discuss two results relating to LO colouring promise problems. First, we will show the relationships between the polymorhpism minions of LO 2 vs LO k colouring for all $k \geq 3$, and we use this to prove that LO k vs LO ℓ colouring is NP-hard for uniformity $r \geq l - k + 4$. We will also discuss an algorithm that, if given an LO 2-colourable 3-uniform hypergraph, will find an LO $O(\sqrt{(n \log\log n}/\log n)$ colouring.

3.19 An Introduction to Promise CSP, Free Structures, and Minions

Jakub Opršal (University of Oxford, GB)

Joint work of Libor Barto, Jakub Bulín, Andrei A. Krokhin, Jakub Opršal
Main reference Libor Barto, Jakub Bulín, Andrei A. Krokhin, Jakub Oprsal: "Algebraic Approach to Promise Constraint Satisfaction", J. ACM, Vol. 68(4), pp. 28:1–28:66, 2021.
URL https://doi.org/10.1145/3457606

The promise CSP is a certain "structural approximation" variant of the CSP. The key difference from standard CSP is that the template consists of two structures $\mathbf{A}$ and $\mathbf{B}$, s.t., $\mathbf{A}$ maps homomorphically to $\mathbf{B}$. The search version of promise CSP can be formulated as: given an instance $\mathbf{X}$ of CSP($\mathbf{A}$) that is promised to have a solution, i.e., there is a homomorphism from $\mathbf{X}$ to $\mathbf{A}$, find a homomorphism from $\mathbf{X}$ to $\mathbf{B}$. For example, we could ask to find a 6 colouring of a graph that is promised to be 3 colourable. This promise CSP gained a lot of attention in recent years for a few reasons: it has substantial interactions with other variants of approximation, and it gives many interesting insights into the theory of classical CSPs.

In the tutorial I will give a brief overview of the basic abstract theory of promise CSPs with focus on several key concepts: gadget reductions, polymorphisms, minions, free structures, and their interactions with standard CSPs.

This is the first part of the tutorial on promise CSPs.

3.20 Topology in Approximate Graph Colouring

Jakub Opršal (University of Oxford, GB)

Joint work of Andrei Krokhin, Jakub Opršal, Marcin Wrochna, Stanislav Živný
Main reference Andrei A. Krokhin, Jakub Oprsal, Marcin Wrochna, Stanislav Zivný: "Topology and adjunction in promise constraint satisfaction", CoRR, Vol. abs/2003.11351, 2020.
URL https://arxiv.org/abs/2003.11351

Approximate graph colouring is a prominent promise CSP, it asks to find a colouring using k colours of a graph that is 3-colourable where $k > 3$. It is a notorious open problem in computational complexity. We could also ask whether the problem gets easier if we strengthen the promise but insist on finding a 3-colouring. In the talk, we will discuss a method, based on topological ideas first used in graph colouring by Lovász, that shows that strengthening the promise does not make the problem easier (unless the promise implies that the graph is 2-colourable), and where these topological ideas might lead in classification of computational complexity of similar problems in the near future.

This is the second part of the tutorial on promise CSPs.

3.21 Surprises in Infinite-Domain Constraint Satisfaction, or: An Order Out of Nothing

Michael Pinsker (TU Wien, AT)

We give an overview of recent developments around the dichotomy conjecture for CSPs of infinite-domain structures such as the order of the rationals or the random graph. It was believed until recently that in the context of CSPs, the mathematical theory for those structures with an order (such as the order of the rationals) was quite different from those without an order (such as the random graph), and that in particular the latter were always solvable by a certain natural reduction to finite-domain CSPs, whereas the former can never be solved that way. We show that this dividing line between ordered and non-ordered structures does not provide this expected dichotomy of approaches, and that orders can appear out of nothing in unordered structures. As a consequence, this forces us to give up the project of always reducing "non-ordered" CSPs to finite-domain CSPs, and to take on the challenge of directly lifting Zhuk's finite-domain algorithmic techniques to infinite-domain CSPs.

3.22 On the Mysteries of MAX NAE-SAT

Aaron Potechin (University of Chicago, US)

Joint work of Joshua Brakensiek, Neng Huang, Aaron Potechin, Uri Zwick
Main reference Joshua Brakensiek, Neng Huang, Aaron Potechin, Uri Zwick: "On the Mysteries of MAX NAE-SAT", in Proc. of the 2021 ACM-SIAM Symposium on Discrete Algorithms, SODA 2021, Virtual Conference, January 10 – 13, 2021, pp. 484–503, SIAM, 2021.
URL https://doi.org/10.1137/1.9781611976465.30

The NAE-SAT predicate (where NAE stands for not all equal) is a simple predicate which is 0 if all of its inputs are equal and 1 otherwise. For any fixed k, there is a $7/8$ or better approximation algorithm for instances of MAX NAE-SAT where each clause has length exactly k. However, the approximability of MAX NAE-SAT with a mixture of different clause sizes is much less well understood. Before our work, it was an open problem whether or not there is a $7/8$ approximation algorithm for MAX NAE-SAT where all clause lengths are allowed.

In this talk, I will prove that the approximation ratio for MAX NAE-SAT with clauses of lengths 3 and 5 is at most 0.8739 (assuming unique games is hard). I will then describe an approximation algorithm for almost satisfiable instances of MAX NAE-SAT with clauses of lengths 3 and 5 which we conjecture gives a 0.8728-approximation. After describing this algorithm, I will describe how we found this algorithm, why we conjecture that it gives a 0.8728-approximation, and why it may well be the optimal approximation algorithm for almost satisfiable instances of MAX NAE-SAT with clauses of lengths 3 and 5.

3.23 Approximating CSPs via the Sum of Squares Hierarchy (and the Role of Expansion)

Madhur Tulsiani (TTIC – Chicago, US)

In this tutorial, we will give an overview of existing and some recent results on approximating CSPs via the Sums of Squares hierarchy. We will also examine the role played by various notions of expansion, for proving both upper and lower bounds.

3.24 PSpace-hard vs Π_2^p Dichotomy for the QCSP

Dmitriy Zhuk (Moscow State University, RU)

The Quantified Constraint Satisfaction Problem (QCSP) is the generalization of the Constraint Satisfaction problem (CSP) where we allow both existen- tial and universal quantifiers. Formally, the QCSP over a constraint language Γ is the problem to evaluate a sentence of the form

$$\forall x_1 \exists y_1 \forall x_2 \exists y_2 \ldots \forall x_n \exists y_n (R_1(\ldots) \wedge \cdots \wedge R_s(\ldots)),$$

where $R_1, \ldots, R_s$ are relations from Γ. While CSP remains in NP for any Γ, QCSP can be PSpace-hard, as witnessed by Quantified 3-Satisfiability or Quantified Graph 3-Colouring.

It turned out that there are no constraint languages whose QCSP complexity is in between PSpace and Π_2^p. The proof is based on a natural reduction of a QCSP instance to an exponential size CSP instance. We show that, unless the problem is PSpace-hard, for any false instance there exists a polynomial part of the CSP instance without a solution, which immediately brings us to the Π_2^p complexity class. In the talk we will discuss the proof as well as a concrete constraint language whose QCSP is Π_2^p-complete.

Participants

- Kristina Asimi
Charles University – Prague, CZ
- Per Austrin
KTH Royal Institute of Technology – Stockholm, SE
- Libor Barto
Charles University – Prague, CZ
- Manuel Bodirsky
TU Dresden, DE
- Bertalan Bodor
Charles University – Prague, CZ
- Andrei A. Bulatov
Simon Fraser University – Burnaby, CA
- Silvia Butti
UPF – Barcelona, ES
- Lorenzo Ciardo
University of Oxford, GB
- Victor Dalmau
UPF – Barcelona, ES
- Dmitry Feichtner-Kozlov
Universität Bremen, DE
- Martin Grohe
RWTH Aachen University, DE
- Venkatesan Guruswami
Carnegie Mellon University – Pittsburgh, US
- Johan Hastad
KTH Royal Institute of Technology – Stockholm, SE
- Mark R. Jerrum
Queen Mary University of London, GB
- Amirhossein Kazeminia
Simon Fraser University – Burnaby, CA
- Michael Kompatscher
Charles University – Prague, CZ
- Pravesh K. Kothari
Carnegie Mellon University – Pittsburgh, US
- Marcin Kozik
Jagiellonian University – Kraków, PL
- Andrei Krokhin
Durham University, GB
- Barnaby Martin
Durham University, GB
- Dániel Marx
CISPA – Saarbrücken, DE
- Antoine Mottet
TU Hamburg, DE
- TomášNagy
TU Wien, AT
- Tamio-Vesa Nakajima
University of Oxford, GB
- Daniel Neuen
Simon Fraser University – Burnaby, CA
- Joanna Ochremiak
University of Bordeaux, FR
- Jakub Oprsal
University of Oxford, GB
- Michael Pinsker
TU Wien, AT
- Aaron Potechin
University of Chicago, US
- Jakub Rydval
TU Dresden, DE
- Madhur Tulsiani
TTIC – Chicago, US
- Caterina Viola
University of Oxford, GB
- Albert Vucaj
TU Wien, AT
- Uli Wagner
IST Austria – Klosterneuburg, AT
- Dmitriy Zhuk
Moscow State University, RU
- Stanislav Živný
University of Oxford, GB

Remote Participants

- Albert Atserias
UPC Barcelona Tech, ES
- Amey Bhangale
University of California – Riverside, US
- Joshua Brakensiek
Stanford University, US
- Hubie Chen
Birkbeck, University of London, GB
- Florian Frick
Carnegie Mellon University – Pittsburgh, US
- Pasin Manurangsi
Google – Mountain View, US
- Dana Moshkovitz
University of Texas – Austin, US
- Aditya Potukuchi
University of Illinois – Chicago, US
- Akbar Rafiey
Simon Fraser University – Burnaby, CA
- Sai Sandeep
Carnegie Mellon University – Pittsburgh, US

Report from Dagstuhl Seminar 22202

Anticipatory Human-Machine Interaction

Jelmer Borst*[1], Andreas Bulling*[2], Cleotilde Gonzalez*[3], and Nele Russwinkel*[4]

1 **University of Groningen, NL.** j.p.borst@rug.nl
2 **University of Stuttgart, DE.** andreas.bulling@vis.uni-stuttgart.de
3 **Carnegie Mellon University, US.** coty@cmu.edu
4 **TU Berlin, DE.** nele.russwinkel@tu-berlin.de

Abstract

Even after three decades of research on human-machine interaction (HMI), current systems still lack the ability to predict mental states of their users, i.e., they fail to understand users' intentions, goals, and needs and therefore cannot anticipate their actions. This lack of anticipation drastically restricts their capabilities to interact and collaborate effectively with humans. The goal of this Dagstuhl Seminar was to discuss the scientific foundations of a new generation of human-machine systems that anticipate, and proactively adapt to, human actions by monitoring their attention, behavior, and predicting their mental states. Anticipation might be realized by using mental models of tasks, specific situations and systems to build up expectations about intentions, goals, and mental states that gathered evidence can be tested against.

The seminar provided an inter-disciplinary forum to discuss this emerging topic by bringing together – for the first time – researchers from a range of fields that are directly relevant but hitherto haven't met on this topic so far. This includes human-computer interaction, cognitive-inspired AI, machine learning, computational cognitive science, and social and decision sciences. We discussed theoretical foundations, key research challenges and opportunities, new computational methods, and future applications of anticipatory human-machine interaction.

Seminar May 16–20, 2022 – http://www.dagstuhl.de/22202

2012 ACM Subject Classification Human-centered computing → Collaborative and social computing; Human-centered computing → Human computer interaction (HCI); Human-centered computing → Collaborative interaction; Computing methodologies → Artificial intelligence; Computing methodologies → Machine learning

Keywords and phrases Human-Computer Interaction, Anticipation, Collaboration, Collaborative Intelligence, Human-AI Teaming, Multi-Agent Simulation, Artificial Intelligence

Digital Object Identifier 10.4230/DagRep.12.5.131

1 Executive Summary

Jelmer Borst (University of Groningen, NL, j.p.borst@rug.nl)
Andreas Bulling (University of Stuttgart, DE, andreas.bulling@vis.uni-stuttgart.de)
Cleotilde Gonzalez (Carnegie Mellon University, United States, coty@cmu.edu)
Nele Russwinkel (TU Berlin, DE, nele.russwinkel@tu-berlin.de)

There is growing interest in interdisciplinary understanding of the anticipatory processes regarding other people's actions and also one's own actions. This surge in interest is especially important given that anticipatory interaction is needed in the growing area of intelligent

* Editor / Organizer

systems that work closely together with human partners. Decades of research on human-machine interaction (HMI) have resulted in significant advances in theories, tools, and methods to facilitate, support, and enhance interactions between humans and computing systems. Despite the fundamental importance of HMI for our information society and numerous advances towards making interactions with machines more human-like, current systems still fall short in one core human ability – Theory of Mind (ToM). ToM allows us to attribute mental states to others and anticipate their actions; and is thus essential for us to interact naturally, effortlessly, and seamlessly.

ToM shapes how we interact with each other and is most easily observable in physical tasks, such as moving a table together. In this scenario, we rely on ToM abilities to attribute intentions to others and in turn, continuously adapt our own behaviour to accommodate the intentions of others, resulting in seamless collaboration. ToM begins to develop in early childhood. Even small children who are not able to develop the full cognitive ability of ToM can anticipate others' intentions using simple mental models of tasks and following familiar sequences of actions (e.g. Fiebich, 2018).

Deficits in ToM are closely linked to developmental disorders, such as autism, and current machines are similarly mind-blind. That is, they fail to sense users' attention and predict their intentions, and therefore lack the ability to anticipate and pro-actively adapt to users' actions. This limits machines to operating after the fact, i.e. to merely react to user input. This fundamentally limits the naturalness, efficiency, and user experience of current interactions. Imagine how difficult it would be to move a table with a robot: as the robot cannot anticipate easy and hard actions for the user (e.g., the table is heavy and needs to be rested for a moment, a corner has to be navigated for which the user has to switch her grip, etc.), it would not be a very helpful teammate. We believe anticipatory HMI has significant potential to bring systems that possess ToM to a new, exciting level of development. However, the building blocks required for anticipatory HMI are still at a very early research stage. This is in stark contrast to the large body of work on (Computational) Theory of Mind in the cognitive and neurosciences and the variety of potential applications in which artificial anticipatory behaviour could have a revolutionary and paradigm-changing influence.

With this seminar we have made a first step in bridging this gap and have discussed theoretical foundations, key research challenges and opportunities, new computational methods, as well as applications and use cases of anticipatory HMI. Research on anticipatory HMI is inherently interdisciplinary and draws from a number of fields, most notably human-computer interaction, machine learning, computer vision, computational neuroscience, computational cognitive science, privacy and security, as well as context-aware computing. Consequently, in this seminar, we have brought together junior and established researchers within these communities for the first time to explore the scientific foundations of anticipatory HMI.

Three perspectives on human-machine anticipation

In the seminar we have discussed three perspectives on human-machine anticipation.

Anticipating the intentions, beliefs, mental states and next actions of the user. This perspective focuses on the challenges, as well as possible solutions, for machine anticipation of user intentions, goals, motivations, and behaviors. This includes the underlying question of which mental states need to be modeled and anticipated in the first place. This might be different depending on the purpose of the collaboration or on the specific circumstances.

Anticipating the outcome of the machine's own actions. The ability of a machine to anticipate its own actions and outcomes is important, especially for interaction in teams and for learning purposes. Artificial agents need to have a concept of the outcome of their own actions, as that will naturally affect the collaboration with the human or other agents. In addition, in case the expected outcome of an action is not reached, they have to learn how they can achieve the expected outcome in an alternative manner.

Anticipating the outcome of collaborative work and human-machine teaming. The third perspective builds on the previous two. In human-machine teaming, a machine needs to anticipate the human partners' actions, identify problems, and develop ideas on how the partners in a team can be supported. We will discuss how artificial agents can engage in effective teamwork with humans and which specific abilities are required. We will also discuss whether we can construct artificial agents that behave indistinguishable from human agents – and whether this is even desirable.

Underlying concepts and methods

In addition to the three different perspectives on anticipatory behaviour, we have discussed the underlying concepts and computational methods.

On the conceptual side, we have discussed definitions and developed a common language. This included, for example, concepts such as cognitive state vs. mental states (which are not always concisely defined across disciplines). Furthermore, to support our cross-disciplined, multi-context approach, we have collected different forms and definitions of "anticipation". A related but separate concept that we have discussed is what different disciplines understand by ToM and what related terms that should be differentiated?. We found that there are considerable differences of opinion on this that we formulated and exchanged.

On the methodological side, we have discussed the ways in which different domains and disciplines have their own methods, approaches, and challenges. Subsequently, groups have exchanged and discussed these collections to gain a deeper understanding of available methodological approaches. We have also talked about current approaches in modelling and cognitive science, and exchange ideas with those of us from other disciplines. Which approaches support the development of better applications? More general is the question how user behaviour can be measured. What data can or should be collected (CV, natural language, neurobiological measures, behavior,..), how can we process it most effectively, and how can we incorporate it into cognitive or computational models of the other users? Finally, we have discussed what the best way is to integrate and fuse different sources of information about the user, and how this can be integrated with the current context.

2 Table of Contents

3 Overview of Talks

3.1 Anticipation for Assistive and Collaborative Human-Machine Interaction

Henny Admoni (Carnegie Mellon University, USA, henny@cmu.edu)

Joint work of Reuben Aronson, Abhijat Biswas, Maggie Collier, Kris Kitani, Ben Newman, Ada Taylor
Main reference Benjamin A. Newman, Reuben M. Aronson, Kris Kitani, Henny Admoni: "Helping People Through Space and Time: Assistance as a Perspective on Human-Robot Interaction", Frontiers Robotics AI, Vol. 8, p. 720319, 2021.
URL https://doi.org/10.3389/frobt.2021.720319

Robots can help people live better lives by assisting on complex tasks involved in everyday activities. For example, robot arms can help people with motor impairments eat food or prepare a meal; robot servers can support waitstaff in a busy restaurant; and social robot sidekicks can improve the conversational agency of users with severe speech impairments. To be fully capable of providing fluent assistance, robots must understand *what* their human partner wants to do, *how* they would like to do it, and *when* the robot should step in.

This capability of providing *proactive assistance* relies on the robot's understanding of implicit situational context, including their human partner's intentions, preferences, and knowledge. A key insight that drives this research is that mental states, while not directly observable, are often linked to observable behaviors. For example, where someone gazes is tied to what task they're trying to achieve, and hence their goal and likely future actions [1]. Thus, understanding the link between human behavior and mental states enables robots to anticipate human needs. Conversely, robots can use implicit signals as part of providing assistance to enable the human to anticipate the robot's actions [2].

In this research, we develop algorithms that model assistance-relevant human mental states (e.g., needs, desires, and preferences) from observable behaviors, and produce robot assistance that is effective, fluent, and interpretable. The long-term research goal is to bridge the gap between human behavior and robot algorithms, which will enable researchers to develop more effective assistive and collaborative robots. In this multi-disciplinary work, we apply knowledge, tools, and methods from fields such as robotics, machine learning, artificial intelligence, computer perception, and cognitive science. We demonstrate the utility of proactive assistance across many levels of robot autonomy and in diverse domains, such as physical support on activities of daily living, driving support in autonomous vehicles, and social support during conversation

References

1 Aronson, Reuben M.; Admoni, Henny, *Gaze Complements Control Input for Goal Prediction During Assisted Teleoperation*, in Proceedings of Robotics: Science and Systems, 2022.
2 Taylor, Ada V.; Mamantov, Ellie; Admoni, Henny, *Wait Wait, Nonverbally Tell Me: Legibility for Use in Restaurant Navigation.*, in Workshop on Social Robot Navigation at Robotics: Science and Systems, 2021.

3.2 From Anticipatory to Socially-Aware User Interfaces

Elisabeth André (Augsburg University, DE, elisabeth.andre@uni-a.de)

The development of anticipatory user interfaces is a key issue in human-centred computing. A major challenge of any anticipatory interface is to identify relevant information about the users including their environment that may be used by the system to prepare or adjust for future interactions.

In our very early work, we followed the paradigm of the so-called *anticipation feedback loop* [1] to improve the comprehensibility of multimodal presentation. The overall idea was to simulate how a user would perceive the output generated by a system and to adjust if necessary. For example, by parsing an utterance planned by the system, the generator made sure that the utterance did not contain unintended structural ambiguities. The implicit assumption was that we can predict the user's understanding processes from the system's understanding processes.

Our more recent work focuses on the paradigm of *Socially-Aware Interfaces* that adjust their behaviours based on an analysis of implicitly provided user feedback, such as social and affective signals, as an integral component of an anticipatory mechanism. For example, we investigated to what extent user engagement may be automatically detected from the user's gestures, postures, facial expressions and speech to adapt a robot's communicative style if necessary [2]. This approach is based on a direct observation of the user as opposed to a simulation. It bears the challenge that human expressions do not follow the precise mechanism of a machine, but are tainted with a high amount of variability, uncertainty and ambiguity.

References

1 Wahlster, W., André, E., Graf, W. & Rist, T. Designing Illustrated Texts: How Language Production Is Influenced By Graphics Generation, *5th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pp. 8-14, 1991, `https://aclanthology.org/E91-1003/`

2 Ritschel, H., Baur, T. & André, E., Adapting a Robot's linguistic style based on socially-aware reinforcement learning, *26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* 2017, pp. 378-384, `https://doi.org/10.1109/ROMAN.2017.8172330`

3.3 Sensing and Modelling Human Attention towards Anticipatory Human-Computer Interaction

Mihai Bâce (University of Stuttgart, DE, mihai.bace@vis.uni-stuttgart.de)

The vision for the next generation of intelligent user interfaces (UIs) is to seamlessly assist users during everyday tasks, similar to how humans interact with one another. This is in stark contrast to current UIs, which are primarily reactive and lack the ability to understand or anticipate user actions, intentions, or goals. One fundamental requirement for developing systems that can anticipate human behaviour is the ability to sense and model human attention. While human attention is key to many tasks in human-computer interaction (HCI), such as predicting interruptibility, boredom, or confusion, prior works have mostly modelled attention through proxies. Proxies such as user interactions or self-reports cannot fully capture the temporal, fine-grained dynamics of human attention.

One goal in my research is to leverage either sensed [1] or computational models of human attention for tasks such as predicting user interruptibility, modelling task interleaving, or in the context of work and productivity. Prior research has shown that, e.g. users can be in mental states that make them more likely to be interrupted [2]. My goal is to anticipate or predict when such mental states are likely to occur either using models of attention only, or in a multimodal setting (e.g. leveraging user interactions as well).

References

1 Bâce, M., Staal, S. & Bulling, A. Quantification of Users' Visual Attention During Everyday Mobile Device Interactions. *Proceedings Of The 2020 CHI Conference On Human Factors In Computing Systems.* (2020)

2 Mark, G., Iqbal, S., Czerwinski, M. & Johns, P. Focused, Aroused, but so Distractible: Temporal Perspectives on Multitasking and Communications. *Proceedings Of The 18th ACM Conference On Computer Supported Cooperative Work & Social Computing.* (2015)

3.4 Anticipatory Human-Machine Interaction with Cognitive Models

Jelmer Borst (University of Groningen, NL, j.p.borst@rug.nl)

Effective anticipatory human-machine interaction (aHMI) requires an accurate model of the human agent. To develop such models, cognitive architectures provide a promising starting point, as such architectures have long been used to develop detailed models of the human cognitive system at a level of abstraction that matches well to aHMI.

A direct example of such an approach is the Slimstampen system for fact learning ([3]; `http://www.slimstampen.nl`). The core of this system is a model of human memory taken from the cognitive architecture ACT-R (Anderson, 2007), which estimates the memory activation level of each fact that needs to be learned. Based on this information, the system anticipates the next fact that should be presented, thereby optimizing the scheduling of facts (e.g., present a fact again just before it will be forgotten). As a result, users can learn 10% more facts in the same amount of time as compared to standard methods such as flash-card learning.

As a second example, we used ACT-R to develop a model that explains under what circumstances interruptions are more or less disruptive [1]. Based on the detailed account of this model, we then built an interruption management system that measures pupil size to anticipate the best moments for interruptions [2]. By scheduling interruptions at opportune moments in the task, users finished the main task in 10% less time – while being interrupted just as often.

Together these examples show how the use of detailed cognitive models of the human mind can lead to very effective aHMI.

References

1 Borst, J., Taatgen, N. & Rijn, H. What makes interruptions disruptive? A process-model account of the effects of the problem state bottleneck on task interruption and resumption. *Proceedings Of The 33rd Annual ACM Conference On Human Factors In Computing Systems*. pp. 2971-2980 (2015)

2 Katidioti, I., Borst, J., Haan, D., Pepping, T., Vugt, M. & Taatgen, N. Interrupted by your pupil: An interruption management system based on pupil dilation. *International Journal Of Human–Computer Interaction*. **32**, 791-801 (2016)

3 Sense, F., Behrens, F., Meijer, R. & Rijn, H. An individual's rate of forgetting is stable over time but differs across materials. *Topics In Cognitive Science*. **8**, 305-321 (2016)

3.5 Learning to Cooperate by Anticipating Others' Mental States

Matteo Bortoletto (University of Stuttgart, DE, matteo.bortoletto@vis.uni-stuttgart.de)

Theory of Mind refers to humans' ability to infer and represent mental states of others. Such mental states include, for example, beliefs, desires and intentions. Computational Theory of Mind tries to build a model capable of inferring such states [1, 2]. As we want to infer mental states of other agents, Theory of Mind naturally deals with multi-agents systems, in which we have more than one agent observing or acting in the same environment. Multi-agents systems are quite difficult to deal with because one agent's action can influence other agents' actions. In this sense, anticipation plays a central role. By inferring beliefs, intentions and desires we can have cooperative agents capable of learning to assist humans or other agents [3]. Agents capable of anticipating others' behaviour can plan and act accordingly, avoiding mistakes or coordination errors. In principle, this strategy can be applied in both cooperative and competitive tasks or games. Multi-agent AI research has seen most success in competitive two-player zero-sum settings, like chess or Go. However, most real world interactions are cooperative, therefore applying Theory of Mind to improve human-AI cooperation, AI-AI cooperation or even human-human cooperation certainly represents one of the next big challenges in the field.

References

1 Rabinowitz, N., Perbet, F., Song, F., Zhang, C., Eslami, S. A., & Botvinick, M. (2018). Machine theory of mind. International Conference on Machine Learning, 4218–4227.

2 Chuang, Y.-S., Hung, H.-Y., Gamborino, E., Goh, J. O. S., Huang, T.-R., Chang, Y.-L., Yeh, S.-L., & Fu, L.-C. (2020). Using Machine Theory of Mind to Learn Agent Social Network Structures from Observed Interactive Behaviors with Targets. 2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), 1013–1019.

3 Bard, N., Foerster, J. N., Chandar, S., Burch, N., Lanctot, M., Song, H. F., Parisotto, E., Dumoulin, V., Moitra, S., Hughes, E., & others. (2020). The hanabi challenge: A new frontier for ai research. Artificial Intelligence, 280, 103216.

3.6 The role of feedback in anticipatory human-computer-interaction

André Brechmann (Leibniz Institute for Neurobiology Magdeburg, DE, brechmann@lin-magdeburg.de)

In human-computer-interaction the action of the technical system can be viewed as feedback to the prior action of the user. Such feedback can fulfil a number of different functions, e.g. signalling that a user's action has been registered by the technical system, execution of a predefined function, actively provide task-related information, providing motivational feedback, or evaluate a user's decision. In the neurosciences feedback is inextricably linked with reward rooted in traditional reinforcement-based theories that were developed based on animal studies that applied reward and/or punishment for the animal to learn. This research has even coined the term reward system of the brain consisting of midbrain dopaminergic areas (substantia nigra/VTA) and main projection areas, e.g. the striatum and prefrontal cortex. Following this tradition, the vast majority of studies on the neural correlates of feedback processing in humans have been designed as reinforcement experiments using (monetary) rewards. Consequently, any findings of activation of the brain's "reward" system has almost exclusively been attributed to mechanisms of reward processing (see Wolff et al., 2020 for references). When we entered the human-computer-interaction domain, we started with the most basic function of feedback, i.e. to signal that a users decision (button press) has been registered by the technical system to study effects of delayed and omitted system response and adaptive processes in the user's brain towards delayed system responses (Kohrs et al. 2012, 2016). To our surprise, we showed the reward system of the brain to be specifically involved in the processing of such registering feedback. A consistent interpretation of these findings relates to the user making predictions about the feedback of the technical system independent of any reward. For me, this opened up a novel neuroscience perspective on feedback processing in human-computer-interaction because humans anticipate reactions of interactions partners and probably assume such skills to be present in intelligent technical systems, too. We are currently interested in understanding how a technical system can support a user in the course of learning. As use case, we designed a multidimensional category learning paradigm where humans learn the conjunction of two rules by trial-and-error within a session of 180 trials. We observed that many participants persistently stick to a strategy that only considers one to the two rules. In order to prevent the consolidation of such an incorrect strategy, we implemented a heuristic that anticipates such behaviour based on the decision history of the learners. This heuristic then triggers a feedback intervention providing helpful information towards the solution of the learning problem. The results of this intervention in terms of subsequent performance and subjective ratings by the learner how helpful they perceived the intervention was a key motivation to pursue the topic of anticipatory human-computer-interaction: Even though the intervention was objectively helpful, many of the subjects did not rate the support very highly even if they improved in performance. We are now working towards understanding the reasons for this adverse effects of feedback intervention. A simple explanation could be that learners did not like to be interrupted at all or at a later moment or that the mode of neural processing was not compatible with the need to reflect on the previous performance. Whatsoever, my future goal is to include the patterns of brain activity during the course of learning into the process of anticipating the need and the best moment for supportive feedback intervention and to learn from the impact of an intervention on behavioural, psycho- and neuro-physiological data for anticipating the consequences of a proactive intervention of a technical system.

References

1. Kohrs C, Angenstein N, Scheich H, Brechmann A. 2012. Human striatum is differentially activated by delayed, omitted, and immediate registering feedback. Frontiers in Human Neuroscience. 6(AUGUST). https://doi.org/10.3389/fnhum.2012.00243
2. Wolff S, Kohrs C, Angenstein N, Brechmann A. 2020. Dorsal posterior cingulate cortex encodes the informational value of feedback in human-computer interaction. Scientific Reports. 10(1):Article 13030. https://doi.org/10.1038/s41598-020-68300-y

3.7 Anticipatory Human-Computer Interaction

Andreas Bulling (University of Stuttgart, DE, andreas.bulling@vis.uni-stuttgart.de)

Main reference Julian Steil, Philipp Müller, Yusuke Sugano, Andreas Bulling: "Forecasting user attention during everyday mobile interactions using device-integrated and wearable sensors", in Proc. of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI 2018, Barcelona, Spain, September 03-06, 2018, pp. 1:1–1:13, ACM, 2018.
URL https://doi.org/10.1145/3229434.3229439

Even after three decades of research on human-computer interaction (HCI), current general-purpose user interfaces (UI) still lack the ability to attribute mental states to their users, i.e. they fail to understand users' intentions and needs and to anticipate their actions. This drastically restricts their interactive capabilities. We aim to establish the scientific foundations for a new generation of user interfaces that pro-actively adapt to users' future input actions by monitoring their attention and predicting their interaction intentions – thereby significantly improving the naturalness, efficiency, and user experience of the interactions. Realising this vision of *anticipatory human-computer interaction* requires groundbreaking advances in everyday sensing of user attention from eye and brain activity. We further require methods to predict entangled user intentions and forecast interactive behaviour with fine temporal granularity during interactions in everyday stationary and mobile settings. Finally, we require fundamental interaction paradigms that enable anticipatory UIs to pro-actively adapt to users' attention and intentions in a mindful way. We have identified four challenging application areas where these new capabilities have particular relevance: 1) mobile information retrieval, 2) intelligent notification management, 3) Autism diagnosis and monitoring, and 4) computer-based training. Anticipatory human-computer interaction offers a strong complement to existing UI paradigms that only react to user input post-hoc. If successful, work on this topic will deliver the first important building blocks for implementing Theory of Mind in general-purpose UIs. As such, anticipatory HCI has the potential to drastically improve the billions of interactions we perform with computers every day, to trigger a wide range of follow-up research in HCI as well as adjacent areas within and outside computer science, and to act as a key technical enabler for new applications, e.g. in healthcare and education.

References

1. Müller, P., Sood, E. & Bulling, A. Anticipating Averted Gaze in Dyadic Interactions. *Proc. ACM International Symposium On Eye Tracking Research And Applications (ETRA)*. pp. 1-10 (2020)
2. Hu, Z., Bulling, A., Li, S. & Wang, G. FixationNet: Forecasting Eye Fixations in Task-Oriented Virtual Environments. *IEEE Transactions On Visualization And Computer Graphics (TVCG)*. 27, 2681-2690 (2021)

3 Strauch, C., Hirzle, T., Stigchel, S. & Bulling, A. Decoding binary decisions under differential target probabilities from pupil dilation: A random forest approach. *Journal Of Vision (JOV)*. 21, 1-13 (2021)

3.8 AI-driven Personalization and Anticipatory HMI

Cristina Conati (Department of Computer Science, University of British Columbia, CA, conati@cs.ubc.ca)

My research focuses on AI-Driven personalization, an interdisciplinary field at the intersection of Artificial Intelligence (AI), Human Computer Interaction (HCI) and Cognitive Science. The general objective of this field is to support effective AI-human collaboration by enabling AI systems to personalise their actions and behaviours to the specific needs, states, preferences and abilities of their individual user. Examples of active research in this field include recommender systems, intelligent-tutoring systems, conversational agents, and affect-aware systems.

To provide personalised interaction, an AI system needs to have a *personalization loop* in which it acquires a model of its user by inferring relevant user properties (see User Model in Figure 1, right) from available interaction behaviors (see Input Sources in Figure 1, bottom) and decides how to personalise the interaction accordingly (see Forms of Personalization in Figure 1, top). In my research I have been investigating AI-driven personalization in the context of Intelligent Tutoring Systems (ITS), intelligent help agents for complex interfaces, and personalized support to visualization processing.

My work in ITS revolves around the objective of moving beyond systems designed to provide support for problem solving, by modeling and adapting to students meta-cognitive abilities and by targeting educational activities that are more open-ended in nature, such as learning from interactive simulations or playing educational games, e.g., [2, 3]. I have done extensive work on using eye-tracking data to inform user modelling, showing that information from eye-tracking can help predict both user *short-term states* (e.g. confusion, affective valence, learning) as well as *long-term* cognitive abilities (e.g. perceptual skills, reading proficiency) and traits (e.g. personality). A key application of eye-tracking for user modelling is in driving personalization for tasks that are mainly perceptual in nature, such as processing data visualisations, which has been the focus of my research on user-adaptive visualisations [1]. The key idea underlying this research is to complement traditional approaches to visualisation design based on properties of the target data and tasks by providing real-time support to visualisation processing driven by relevant users states (e.g. attention, confusion) and abilities (e.g. levels of visual literacy, perceptual speed, visual memory etc.).

My most recent research proposes the concept of personalised Explainable AI. This work extends the recent interest in AI systems that can explain their inner workings to their end-users with the idea that these explanations should be part of the personalisation loop (see Figure 1, top), namely they should be dynamically personalised to factors including context, task criticality and user differences (e.g., cognitive abilities and transient states like confusion or cognitive load).

A distinction that needs to be made to understand how AI-driven personalization relates to Anticipatory Human-Machine Interaction (AHMI) is what triggers personalization (often referred to as adaptation in the literature), namely whether the AI agent is personalising

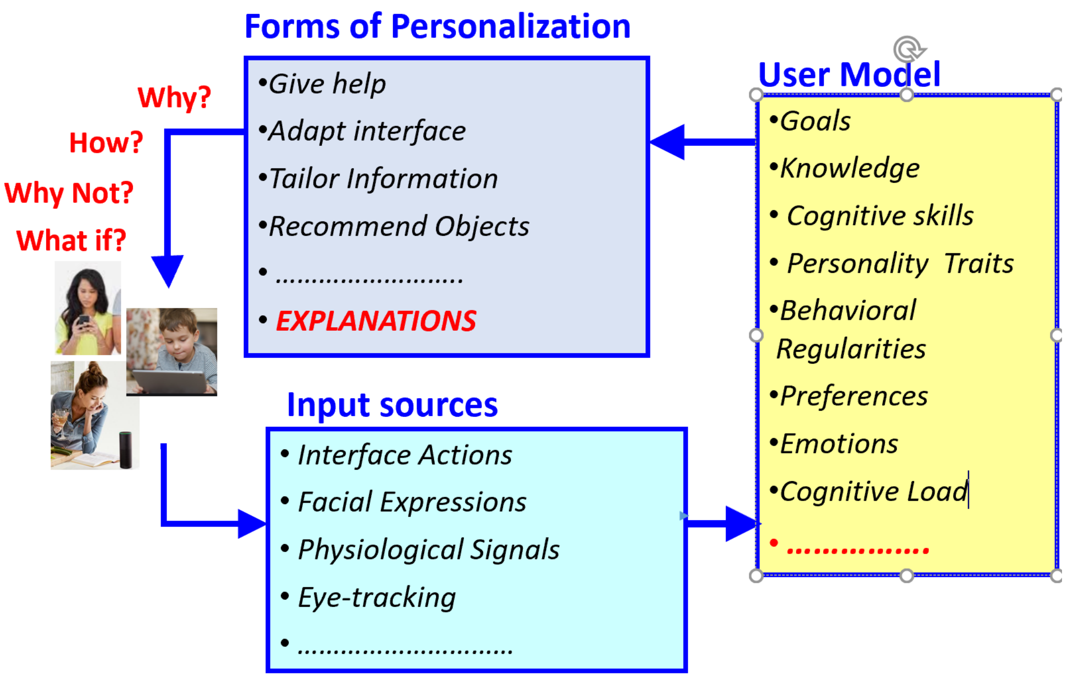

Figure 1 Adaptive Loop for AI-Driven personalization, with explanations of the AI actions added as a form of personalization.

to a *current* user state that it has detected/assessed/perceived or to a user state that is *predicted* to be happening at some point in the future. When the personalization is driven by the current user state, it is *reactive* in nature[1]. Personalization driven by a *predicted* future state of the user can be viewed as *preventive/proactive/anticipatory*. I see the latter as being the key aspect of AHMI. I will illustrate this distinction with examples based on some of my work.

An example of *reactive personalization* in my ITS research comes from my work on ITS for supporting students in acquiring meta-cognitive skills relevant for learning. Specifically, in [2] we focused on the meta-cognitive skill known as self-explanation, i.e. one's tendency to explain and elaborate instructional material to themselves when studying. There is ample evidence in Cognitive Psychology that this skill is conducive to better learning, and that students who lack the skill can acquire it if they undergo instructional sessions during which a tutor prompts them to self-explain if they do not do it. During these sessions, students are asked to self-explain aloud, just so that the tutor can detect if they are self-explaining or not. However, self-explanation usually happens in one's head. In [2], we worked on an ITS that assesses if a student self-explains in their head while studying given material, based on their gaze patterns and interface actions. The ITS then reacts if it detects that self-explanation did not happen for a specific aspect of the instructional material, by generating prompts to encourage and guide the student to self-explain that material.

An example of *anticipatory personalization* in my ITS research comes from my work on ITS for supporting students learning through explorations in open-ended-learning environments.

[1] Note that this reactive personalization is different, and more sophisticated, that the reactivity embedded in traditional interfaces, where each interface action (e.g. a button press) is associate by design to a specific system's response, which does not depend on the user who performed the action.

The capability to explore effectively is relevant to many tasks involving interactive systems, but not all users possess this ability. I have been investigating ITS that can provide personalized real-time support to users who are having difficulties with an exploratory task, while interfering as little as possible with the unconstrained nature of the interaction. Building such ITS is challenging because it requires having a user model that can predict the effectiveness of open-ended behaviours for which there is often no formal definition of correctness. Our approach relies on clustering and association rule mining to discover from data the user exploratory behaviours conducive or detrimental to learning. Supervised machine learning is applied to the resulting clusters to derive classifiers that predict in real-time whether a user will learn from the interaction and, if not, it anticipates how the user behavior should change to increase their learning. Based on this aanticipation, the ITS then generates hints designed to foster the desired change in behavior [3].

The distinction between reactive and anticipatory personalization is one that has not yet been made in the literature, but it is important to think about the unique advantages and challenges that anticipatory personalization brings to AI-Human collaboration.

References

1 Conati, C., Carenini, G., Toker, Dereck., Lallé, S., *Towards User-Adaptive Information Visualization*, Twenty-Ninth AAAI Conference on Artificial Intelligence, 2015.
2 Conati, C., Merten, C., *Eye-tracking for user modeling in exploratory learning environments: An empirical evaluation*, Knowledge Based Systems 20(6): 557-574, 2007.
3 Kardan, S., Conati, C., *Providing Adaptive Support in an Interactive Simulation for Learning: An Experimental Evaluation*, Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 2015.

3.9 Cognitive Argumentation and Sensemaking

Emmanuelle Dietz (Airbus Central R&T – Hamburg, DE, emmanuelle.dietz@airbus.com)

Systems built with the purpose of human-machine interaction do not necessarily need to behave optimally in each circumstance when interacting with humans. However, they should have the ability to provide socially acceptable explanations for their decisions. They might generate such explanations by having an internal representation of the current state of the human mind (theory of mind). This representation might allow the system to predict future events and possibly prepare by intervening with some action (e.g. to avoid catastrophic outcomes or to improve the human's situation).

The internal representation of predicting and deciding on which action to choose can happen on various levels. An important requirement however, is that the system should have the ability to provide socially acceptable explanations for their decisions. More specifically, their explanations should have some basic reasons of support (attributive), explain why a conclusion is supported in contrast to opposing conclusions (contrastive) and provide information that guides on how to act following the conclusion (actionable). Cognitive argumentation fulfills these requirements. In cognitive argumentation, having its roots in computational argumentation, the dialectic argumentation process is assumed to be heavily guided by biases or heuristics [1, 2]. Initially, a root argument for a certain position is built (e.g. based on the awareness in the current situation), (strong) counterarguments are searched (e.g. factual information contracts the argument), which the root argument defends, and is either replaced by stronger arguments or extended with the defense. This process seems similar to sensemaking, in which humans generate mental models to explain events, question these models, and reframe or replace these models by more adequate ones.

In the context of human-machine interaction anticipatory thinking can be implemented through future oriented cognitive argumentation by generating future arguments for or against carrying out a possible future action (of the human or machine) that leads to the most desired state.

References

1 Dietz, E.; Kakas, A., *Cognitive Argumentation and the Selection Task*. Proceedings of the Annual Meeting of the Cognitive Science Society, pp. 1588–1594, 2021.
2 Dietz, E., *Argumentation-based Reasoning guided by Chunk Activation in ACT-R*. In Proceedings of the 20th International Conference on Cognitive Modelling, 2022.

3.10 Anticipation of User Performance For Adapting Visualization Systems

Steffen Frey (University of Groningen, NL, s.d.frey@rug.nl)

My research concerns the development of visualization methods to gain insights from large quantities of scientific data (typically acquired from extensive experiments and simulations). Interactive approaches enable explorative user analysis of rich and complex data, but respective visualization systems generally exhibit numerous and degrees of freedom regarding how information is presented and explored.

In many of my previous works, techniques and systems are configured in a data-driven way, e.g., which time steps [1] or spatio-temporal subsets [2] to convey and how to layout large numbers of graphical tiles depicting members of an ensemble [3] (Figure 2). For this, these approaches generally optimize application-specific metrics or employ self-supervised learning. For example, metrics quantify how well dynamics in a full time series are captured by a subset [1] or the similarity of ensemble members positioned nearby in a grid layout [3]), and learning can be used identify unusual behavior [2].

Conceptually, these approaches optimize the transformation of the data into an (interactive) visualization, but largely neglect – or only indirectly consider – the human decision-making process operating on this basis. Explicitly incorporating this aspect, the anticipation of user performance for a given (configuration of) visualization system allows to explicitly optimize the visual representation and interface for a user to gain insights more quickly and/or more accurately.

References

1 Frey, S.; Ertl, T., *Flow-based Temporal Selection for Interactive Volume Visualization*, Computer Graphics Forum, 2017.
2 Tkachev, G., Frey, S., Ertl, T., *Local Prediction Models for Spatiotemporal Volume Visualization*, IEEE Transactions on Visualization and Computer Graphics, 2021.
3 Frey, S., *Optimizing Grid Layouts for Level-of-Detail Exploration of Large Data Collections*, Computer Graphics Forum, 2022.

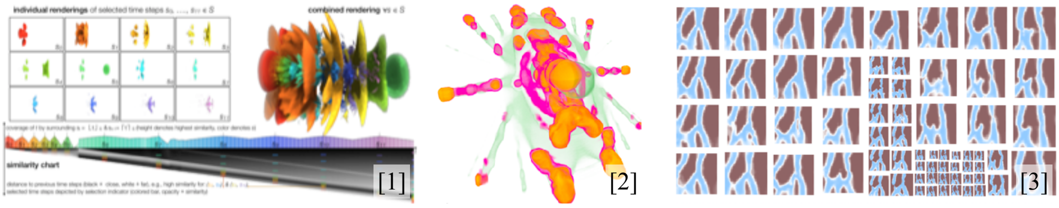

Figure 2 Examples of data-driven visualization via the optimization of metrics [1,3] and unsupervised learning [2].

3.11 Decision Making, Learning, and Adaptability in Human-Machine Collaborations

Cleotilde Gonzalez (Carnegie Mellon University, USA, coty@cmu.edu)

A major challenge for research in artificial intelligence is to develop systems that can infer the goals, beliefs, and intentions of others (i.e., systems that have theory of mind, ToM) (Nguyen & Gonzalez, 2021). To generate relevant Human-Machine collaborations, one needs to create computational representations of human behavior. In the past decades cognitive models have advanced towards more accurately representing and replicating human cognitive processes. In particular, we have been involved in generating models that can explain how humans make decisions, learn, and adapt to improve their choices in dynamic tasks. We rely on Instance-Based Learning Theory (Gonzalez, Lerch & Lebiere, 2003; Nguyen, Phan, & Gonzalez, 2022) to construct models of human dynamic decisions in a large number of situations. Our applications for anticipatory Human-Machine Collaborations include but are not limited to: Applications in cybersecurity, replicating attacker's behaviors that can be used to inform defenders; and replicating end-users' phishing behaviors that can help prevent successful attacks. We have also used these models in dynamic control tasks and scenarios for search and rescue tasks, where predicting the state of the mind of teammates' decisions is essential.

References

1 Gonzalez, C.; Lerch, F.J.; Lebiere, C., *Instance-based learning in dynamic decision making.*. Cognitive Science 27, pp. 591–635, 2003.
2 Nguyen, T.N.; Gonzalez, C., *Theory of Mind from Observation in Cognitive Models and Humans.*. Topics in Cognitive Science, 2021.
3 Nguyen, T.N.; Phan, N.D.; Gonzalez, C., *SpeedyIBL: A Comprehensive, Precise, and Fast Implementation of Instance-Based Learning Theory.*. Behavior Research Methods, 2022.

3.12 Integrating Models of Cognitive and Physical Human-Robot Interaction

Chenxu Hao (Friedrich-Alexander-Universität Erlangen-Nürnberg, DE, chenxu.hao@fau.de)

Joint work of Chenxu Hao, Nele Russwinkel, Daniel F.B. Haeufle, Philipp Beckerle

Research in human-robot interactions (HRI) usually focuses on either cognitive interactions or physical interactions. However, characteristics of HRI in real-life scenarios often involve both information exchanges on the cognitive level and force exchanges on the physical level. Specifically, the human agent and the robotic agent need to infer each other's intentions, predict each other's actions, and act adaptively in the given environment [1]. In such dynamic interactions, the robotic agent also needs to anticipate in order to support the human agent with flexibility. Therefore, it is crucial to create unified models of HRI that can capture the dynamic exchanges on both cognitive and physical levels while taking the environment into account (see [2], for an example of such a unified model of a single agent). Our work aims to provide a conceptual framework with possibilities to connect models of physical and cognitive HRI. We also hope to potentially apply our framework to an anticipatory robotic agent that provides the human agent with different levels of assistance.

References

1 Ho, M. K., & Griffiths, T. L. *Cognitive science as a source of forward and inverse models of human decisions for robotics and control.* arXiv:2109.00127 [cs, eess], 2021.

2 Kahl, S., Wiese, S., Russwinkel, N., & Kopp, S. *Towards autonomous artificial agents with an active self: modeling sense of control in situated action.* Cognitive Systems Research, 72, 50–62, 2022.

3.13 Anticipatory Human-Machine Interaction to Explain AI Systems

Susanne Hindennach (University of Stuttgart, DE, susanne.hindennach@vis.uni-stuttgart.de)

Humans have the astounding ability to attribute mental states to their interaction partners and reason about their behaviour – so-called Theory of Mind (ToM). Similarly, humans also attribute a mind and human abilities to AI systems (like intelligence and learning). The attribution of a human-like mind to the AI system puts the focus on the AI systems, thereby loosing sight of the human decisions about who and what gets counted when building them. The thought processes, intents and decisions can be surfaced by redirecting Theory of Mind away from the AI system and towards the AI experts who built the system. The redirection necessitate anticipatory human-machine interaction in order to (1) infer the mental states of the AI experts and (2) predict the explanations needed by the user of the AI system. In previous work, we have shown that intents – a mental state with explanatory power – can be predicted from mouse and keyboard action in a text formatting task[1].

References

1 Zhang, G., Hindennach, S., Leusmann, J., Bühler, F., Steuerlein, B., Mayer, S., Bâce, M. & Bulling, A. Predicting Next Actions and Latent Intents during Text Formatting. *Proceedings Of The CHI Workshop Computational Approaches For Understanding, Generating, And Adapting User Interfaces.* pp. 1-6 (2022)

3.14 The Transportation Domain can Benefit From Anticipatory Models of Human-machine Interaction

Christian P. Janssen (Utrecht University, NL, c.p.janssen@uu.nl)

Main reference Christian P. Janssen, Shamsi T. Iqbal, Andrew L. Kun, Stella F. Donker: "Interrupted by my car? Implications of interruption and interleaving research for automated vehicles", Int. J. Hum. Comput. Stud., Vol. 130, pp. 221–233, 2019.
URL https://doi.org/10.1016/j.ijhcs.2019.07.004

At this Dagstuhl seminar, I have taken the position that the transportation domain can benefit from anticipatory models for human-machine interaction. One specific domain that I have investigated over the last few years is semi-automated driving. Here, the vehicle takes over some components of the driving task (such as lateral or longitudinal control) under specific conditions (or operational design domains; such as highway roads). However, typically, the vehicle cannot drive under all circumstances, and some human intervention is occasionally needed. The default model in the literature is to look at what the minimum time is that humans need between a warning and the time where they need to act. There seems to be consensus that somewhere between 5 to 8 seconds is a minimum time, and indeed the large majority of the literature has looked at response time situations of around 2 seconds. However, as I have argued elsewhere in more detail [2], although it is beneficial for research to understand what the minimum response interval needs to be, there can be many situations where it is useful to think more ahead and maybe have "pre-alerts" or forewarnings. One reason might be that as automated vehicles gain more capacity, human drivers might not always pay as much attention to the as they would under non-automated conditions, and start performing other tasks. Working on non-driving related activity is something that some drivers already do under regular (non-automated) driving, but that they want to do under higher levels of automation. In these situations, where attention has been to another task or activity than the driving task, one cannot assume that the attention of the driver has been on the road and that they had sufficient situation awareness to act in future (emergency) situations. Moreover, they might not even process the alert itself instantly. So, instead of thinking about fast task switches within, say, 5-8 seconds, one can instead think of how the process of returning attention from a non-driving activity to driving unfolds. I have proposed a conceptual model to capture this [2], of which the first empirical tests have taken place [2]. Now, the next step is to develop models that try to analyze human state and act accordingly, to interrupt at opportune moments. Such models need to take into account the user state, but also the environment and the driving scenario (e.g., is it a situation where there are many alternative actions, or is the intended action quickly clear from the context?). Personally, I would hope that techniques such as Hidden Markov Models can be effective in capturing human states and its dynamics [3]. Without anticipatory models, interruptions and alerts might come at less opportune moments, which can lead to increased stress or workload, and leave the driver in no capacity to act quickly and safely. Although I have specified these principles for driving, I anticipate that they also apply to other traffic domain such as railroads, ships, and airplanes.

References

1 C. P. Janssen, S. T. Iqbal, A. L. Kun, and S. F. Donker, "Interrupted by my car? Implications of interruption and interleaving research for automated vehicles," International Journal of Human Computer Studies, vol. 130, 2019, doi: 10.1016/j.ijhcs.2019.07.004.

2 D. Nagaraju, A. Ansah, A., N.A.N. Ch, C. Mills, C.P. Janssen, O. Shaer, & A.L. Kun, "How Will Drivers Take Back Control in Automated Vehicles? A Driving Simulator Test of an Interleaving Framework," in 13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, 2021, pp. 20–27. doi: 10.1145/3409118.3475128.
3 C. P. Janssen, L. N. Boyle, A. L. Kun, W. Ju, and L. L. Chuang, "A Hidden Markov Framework to Capture Human–Machine Interaction in Automated Vehicles," International Journal of Human-Computer Interaction, vol. 35, no. 11, pp. 947 – 955, 2019, doi: 10.1080/10447318.2018.1561789.

3.15 Probabilistic modelling for anticipating AI-assistants: Towards human-AI teams for experimental design and decision making

Samuel Kaski (Aalto University, FI & University of Manchester, UK, samuel.kaski@aalto.fi)

Current AI-assistants are only able to help us when we can precisely specify what we want – that is, the goal and the problem setting. However, that arguably is not when we most need assistance, and effectively the current assistants require us to program them, just with a different kind of an interface such as giving precise spoken commands in restricted language. Instead, we need assistance the most we do not *yet* know precisely what we want, because the goal may be tacit, unclear or evolving. I am developing AI-assistance principles for situations such as drug design or scientific research, where the goal initially is unclear and evolves during the process.

Such assistants need to anticipate their users to be helpful to them: both their actions, some of which they can assume to arise from the user's planning towards their goals, but also their goals and the development of the goals, and furthermore how the users would understand the assistant's actions. This is particularly difficult because of the nature of the design-type tasks, where each task is different and laborious, resulting in scarcity of data for learning the goals. In particular, this culminates in "zero-shot assistance" for design and decision making, which I believe requires (i) probabilistic modelling to handle the uncertainty arising from scarcity of data and a large number of potential explanations of behaviour, (ii) nested POMDP-type modelling to capture a "theory of mind" of the user, to both infer the goal by reverse-engineering the behaviour and to anticipate how the user would understand the assistant's action, and (iii) novel computational approximations to actually do the necessary inferences.

As a result, we hope to have better tools for solving the grand challenge problems humanity currently is facing.

References

1 Mustafa Mert Çelikok, Frans A. Oliehoek, and Samuel Kaski. Best-response Bayesian reinforcement learning with BA-POMDPs for centaurs. In *Proc. AAMAS 2022*, accepted for publication.
2 Sebastiaan De Peuter and Samuel Kaski. Zero-Shot Assistance in Novel Decision Problems. arXiv:2202.07364, 2022.
3 Antti Kangasrääsiö, Jussi P.P. Jokinen, Antti Oulasvirta, Andrew Howes, and Samuel Kaski. Parameter inference for computational cognitive models with approximate Bayesian computation. *Cognitive Science*, 43:e12738, 2019.

3.16 Surprise and Alignment: Their Role in Anticipation

Frank Keller (University of Edinburgh, UK, keller@inf.ed.ac.uk)

Main reference Moreno I. Coco, Rick Dale, Frank Keller: "Performance in a Collaborative Search Task: The Role of Feedback and Alignment", Top. Cogn. Sci., Vol. 10(1), pp. 55–79, 2018.
URL https://doi.org/10.1111/tops.12300

In psycholinguistics, surprise is a key concept. It measures how predictable the next word in a sentence is, given the prior context of the word. Surprise has been shown to correlate with behavioral measures such as cloze probabilities, reading times [1], or EEG signatures such as the N400. Mathematically, surprise is the negative log of the conditions probability of word $n+$ given words $1 \ldots n$, where a language model is typically used to estimate this probability. Recent work in computational cognitive modeling has used surprise as a component in a more comprehensive cognitive model. For example, the Neural Attention Tradeoff model of [2] models the allocation of human attending during reading: it predicts which words are skipped by the reader, and which ones are fixated. Skipping in the model depends on surprise, but also on aspects of the reading task (e.g., reading for pleasure, proofreading, reading to answer a question).

There is an intuitive connection between anticipation and surprise. Assume an agent predicts the next word, and based on that decides whether to skip it or not. If that decision turns out to be incorrect, then the agent will be surprised, which is something we can measure behaviorally. Surprise can therefore be conceptualized as unsuccessful anticipation. It is worth exploring this idea in the present seminar, and in particular discuss to what extend surprise generalizes to domains other than language processing.

Cognitive scientists have not only studied language processing in single readers and speakers, but have also investigated pairs of speakers (dyads) that communicate and interact to solve a shared task. In such a situation, dialogue participants are known to align with each other. This means that their linguistic productions become more similar across a whole range of dimensions: they use similar words, similar syntactic structures, adapt their speech rate and even their pronunciation. At the same time, they also align non-verbal behavior such as gaze, posture, and gesture. The degree of alignment typically increases over the course of the dialogue, and alignment is correlated with task success. A classical conceptualization of this process in the Interactive Alignment Model [3]. More recent work has refined our understanding of alignment: for instance, using a collaborate search task, Coco et al. (2018) show that the relationship between alignment and task success is not absolute, but depends on how interactive the task is.

Alignment is relevant in the context of anticipation. A participant that aligns with their collaborator predicts the future behavior of that collaborator and decides to make their own behavior more similar. Alignment is therefore a form of long-term anticipation and, like anticipation more generally, it makes collaborating on a task more efficient. This is another observation that could feed into the discussion at this seminar, where questions to be explored include: Does the relationship between alignment and anticipation generalize to non-verbal tasks? Does alignment also make human-machine interaction more efficient? Should machines align with humans? Are humans able to align with machine?

References

1 Demberg, Vera and Frank Keller. 2008. Data from Eye-tracking Corpora as Evidence for Theories of Syntactic Processing Complexity. Cognition 109:2, 193–210.

2 Hahn, Michael, and Frank Keller. 2018. Modeling Task Effects in Human Reading with Neural Attention. https://arxiv.org/abs/1808.00054
3 Pickering, Martin J. and Simon Garrod. 2004. *Toward a mechanistic psychology of dialogue.* Behavioral and Brain Sciences 27:2, 169–190.

3.17 Anticipatory Human-machine Interaction in Single-pilot Operations

Oliver W. Klaproth (Airbus – Hamburg DE, oliver.klaproth@airbus.com)

Main reference Sebastian Blum, Oliver Klaproth, Nele Russwinkel: "Cognitive Modeling of Anticipation: Unsupervised Learning and Symbolic Modeling of Pilots' Mental Representations", Topics in Cognitive Science, Vol. 14, pp. 718-738, 2022.
URL https://doi.org/10.1111/tops.12594

The interaction of two pilots in the cockpit benefits from their ability to anticipate each other's behaviour. Single pilot operations will be characterised by increased interaction with higher levels of automation instead of with another human pilot. By modelling the pilot and their mental representation of the automation, the automation can anticipate the pilot's behaviour and adapt its own behaviour accordingly to stay in line with the pilot's anticipations [1]. Cognitive architectures like ACT-R have been used to model pilots [2, 3]. In particular, instance-based learning in ACT-R has been shown to be effective in learning mental representations through observations of behaviour sequences and in modelling intuitive associative processes underlying anticipation. We hypothesise that instance-based ACT-R models of pilots, for example based on training on flight recorder data, can enable anticipatory human machine interaction in the cockpit during single pilot operations.

References

1 Blum, S., Klaproth, O. & Russwinkel, N. Cognitive Modeling of Anticipation: Unsupervised Learning and Symbolic Modeling of Pilots' Mental Representations. *Topics In Cognitive Science.* (2022)
2 Klaproth, O., Halbrügge, M., Krol, L., Vernaleken, C., Zander, T. & Russwinkel, N. A neuroadaptive cognitive model for dealing with uncertainty in tracing pilots' cognitive state. *Topics In Cognitive Science.* **12**, 1012-1029 (2020)
3 Klaproth, O., Vernaleken, C., Krol, L., Halbruegge, M., Zander, T. & Russwinkel, N. Tracing pilots' situation assessment by neuroadaptive cognitive modeling. *Frontiers In Neuroscience.* **14** pp. 795 (2020)

3.18 Proactive Machine Assistance in Visual Analytics

Alvitta Ottley (Washington University in St. Louis – St. Louis, USA, alvitta@wustl.edu)

The visual analytics community has made significant strides in developing systems that facilitate the interplay between humans and machines in exploratory data analysis and sensemaking [1]. These advances have enabled us to learn from user interactions and uncover their analytic goals (e.g., [2, 3]). Moreover, they have set up the foundation for creating visual analytic systems that anticipate and assist or guide users during data exploration. Providing such guidance will likely become more critical as datasets grow in size and complexity, overwhelming the limited screen real estate and human cognitive power. My team and I use interdisciplinary approaches to identify optimal data representations for effective decision-making and design human-in-the-loop visual analytics interfaces more attuned to domain scientists' exploration goals. Still, some open anticipatory challenges and considerations may be unique to the visual analytics setting. In particular, the role of the AI may depend on the task and include but is not limited to: (1) the AI can be a teammate, working with the human to collaboratively solve a problem by either performing the same task concurrently or separate complementary sub-tasks. (2) the AI may be an assistant, observing actions and taking preparatory steps such as prefetching and precomputation. (3) The AI may serve as a learner, proactively observing actions to create a knowledge graph or process script. Moreover, the user's tasks will likely evolve. Thus, anticipation should happen at a global level (e.g., deciding the role the AI needs to assume based on the observed state) and at a local level (e.g., deciding which data points to prefetch). Finally, the community should address ethical considerations for how AI interventions might influence the user's analysis. For example, we need to preemptively consider how to track biases in the data shown to the user.

References

1 Crouser, R., Ottley, A. & Chang, R. Balancing human and machine contributions in human computation systems. *Handbook Of Human Computation*. pp. 615-623 (2013)

2 Ottley, A., Garnett, R. & Wan, R. Follow the clicks: Learning and anticipating mouse interactions during exploratory data analysis. *Computer Graphics Forum*. **38**, 41-52 (2019)

3 Monadjemi, S., Garnett, R. & Ottley, A. Competing models: Inferring exploration patterns and information relevance via bayesian model selection. *IEEE Transactions On Visualization And Computer Graphics*. **27**, 412-421 (2020)

3.19 Cognitive models of anticipating physical interactions

Nele Russwinkel (Technische Universität Berlin, DE, nele.russwinkel@tu-berlin.de)

For gaining mutual understanding between a human and an intelligent agent the intelligent agent needs to develop a model of the human partner and the environment as well as of the task. Depending on the type of interaction and purpose of the interaction different requirements arise for a fluent and efficient collaboration. For more complex interactions the assumption of a shared representation about the individual goals, the task and situation understanding is essential. The assumed shared representation of technical system and human might differ. Direct and indirect communication is necessary to provide information about the individual shared representations. Especially for physical Interactions with technical Interactions a lot of different cues make sure that e.g. a handing over procedure of an object will be well coordinated in time, location and type. What types of cognitive models are needed to address these issues? Since anticipation is the ability to prepare in time for problems and opportunities –it is functional in form of preparing for future even unlikely but critical events and not simply predicting what might happen. So context and dynamical representations would be relevant. It could be useful to focus the (cognitive) model rather on the relevant components of e.g. (dynamic) situation understanding and decision making for being able to trace behavior and cognitive state updating rather than capturing all aspects of cognition (Klaproth et al. 2021, Scharfe et al 2019). For modelling anticipatory ability in a physical HRI further fine-grained social cues must be detected and also produced for realizing joint action. One of the big challenges I see here is to find good ways for information integration and hybrid modelling approaches that are able to meet all the different requirements the individual tasks pose.

Details are given in [1] and [2].

References

1 Klaproth, O. W., Halbruegge, M., Krol, L. R., Vernaleken, C., Zander, T. O. and Russwinkel, N. (2020). A Neuroadaptive Cognitive Model for Dealing With Uncertainty in Tracing Pilots' Cognitive State. Topics in Cognitive Science, 12(3), p. 1012-1029.

2 Scharfe, M. and Russwinkel, N. (2019). Towards a Cognitive Model of the Takeover in Highly Automated Driving for the Improvement of Human Machine Interaction. In Proceedings of the 17th International Conference on Cognitive Modelling.

3.20 Anticipatory Brain-Computer Interfaces

Andreea Ioana Sburlea (University of Groningen, Groningen, NL, a.i.sburlea@rug.nl)

Intention-related neural signals may be exploited by different human-machine interaction systems. In particular, brain-computer interfaces (BCIs) could be used to restore communication functions and facilitate motor control in paralysed users. BCIs record the brain activity of the user, detect the user's intended action and send the corresponding execution command to an artificial effector system, e.g. computer cursor, prosthetic device, etc. Non-invasive brain recording modalities such as electroencephalography benefit from a fine temporal resolution for the signal acquisition which facilities the anticipation of different cognitive or motor states in a timely manner.

In my talk, I presented an approach that involves the simultaneous acquisition of multimodal data coming from both neural and behavioral information [1], with the purpose of understanding the underlying mechanisms that link perception to action within the framework of brain-computer interfaces. Specifically, in the context of grasping movements, I showed the importance of disentangling at a neural and behavioral level between the properties of the objects (such as shape and size) and the properties of the movement (such as type of grasp, number of fingers involved) [2]. Moreover, I showed that this disentanglement is similarly observed across modalities in separate stages going from perception, observation to action. Furthermore, I opened the conversation about the influence of the goal and feedback on the perception of action in anticipatory human-machine interaction cases involving different levels of attention to the action.

References

1 Sburlea, Andreea I., and Gernot R. Müller-Putz. *Exploring representations of human grasping in neural, muscle and kinematic signals*. Scientific reports 8:1 (2018): 1–14

2 Sburlea, A. I., Wilding, M., & Müller-Putz, G. R. (2021). *Disentangling human grasping type from the object's intrinsic properties using low-frequency EEG signals*. Neuroimage: Reports, 1(2), 100012.

3.21 Anticipation of User Goals from Human Behavioural Data for Human-AI Collaboration

Florian Strohm (University of Stuttgart, DE, florian.strohm@vis.uni-stuttgart.de)

In many machine learning areas, the goal for a model is to solve a specific task faster or better than humans can. However, there are many areas where humans cannot be replaced by an artificial intelligence (AI) system for different reasons, e.g., because humans might not want to be replaced but only seek assistance from an AI, or if the task to solve necessarily includes a human or at least human information. As a consequence, there is an increasing need for AI systems that collaborate with humans, thereby enable them to solve a certain task more efficiently or to achieve a certain goal faster or more easily. For an AI system to collaborate with humans, it is important that the AI understands which information the user is perceiving and processing, as it then can use this information to anticipate user intents. Once the AI system identified user intents it can, depending on the environment and task, anticipate the user's intent, generate results, or perform actions to assist the user. In order to understand which information a human is perceiving and processing it needs to receive information about the user. We specifically focus on using behavioural data as a source of information about the user. The advantage of behavioural data is that the user does not have to explicitly provide it but the AI can automatically obtain user information through their interaction with an application or through external devices like an eye-tracker or using electroencephalography (EEG). [1] One of the main challenges of using implicit behavioural data is the understanding of such data and the subsequent extraction of task relevant information. We plan to develop methods to extract the most relevant information, especially concerning gaze data, e.g., how task relevant a specific fixation is. Since collecting large datasets of such attentive behavioural data is time-consuming and expensive, we will propose methods that allow to generate synthetic datasets and model human behaviour to train intelligent systems.

References

1 Strohm, F., Sood, E., Mayer, S., Müller, P., Bâce, M., & Bulling, A. (2021). Neural Photofit: Gaze-based Mental Image Reconstruction. *In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 245-254).*

3.22 The Role of Skills in Anticipatory Human-Computer Interaction

Niels Taatgen (University of Groningen, NL, n.a.taatgen@rug.nl)

My specialization is Cognitive Modeling using Cognitive Architectures. I have been involved in ACT-R for a long time, and have contributed mechanisms for production compilation, multitasking and time perception. More recently, I have developed my own variant on ACT-R called PRIMs [1]. The motivation for this was the fact that ACT-R and many cognitive architectures are focused on individual tasks, whereas humans have a history of prior experience that they can use to perform new tasks.

To support models that can perform multiple tasks, I defined an additional level of abstraction between production rules and (task) goals: the level of skills. The idea is that if people have to perform a novel task, they try to mobilize the skills necessary for that task from their available repertoire [2]. Only when they miss particular skills they need to do more extensive learning in order to acquire the missing pieces.

An application area of the skill idea is education, in particular cognitive tutors. Using unsupervised learning methods, we can extract the skill set for a particular domain. Then, give a student's performance on assignments, we can determine which skills the student has mastered, and which not. Based on that, the optimal material or assignment can be selected for the student. In addition, it can be used to provide feedback to both the learner and the teacher.

References

1 Taatgen, N.A. (2013). The nature and transfer of cognitive skills. *Psychological Review, 120*(3), 439-471.

2 Hoekstra, C., Martens, S., & Taatgen, N. A. (2020). A Skill-Based Approach to Modeling the Attentional Blink. *Topics in Cognitive Science, 12*(3), 1030-1045.

4 Breakout Groups (Monday)

4.1 Shared Representations and Communication

Alvitta Ottley (Washington University in St. Louis, USA, alvitta@wustl.edu)
Matteo Bortoletto (University of Stuttgart, DE, matteo.bortoletto@vis.uni-stuttgart.de)
Nele Russwinkel (TU Berlin, DE, nele.russwinkel@tu-berlin.de)
Oliver Klaproth (Airbus – Hamburg, DE, oliver.klaproth@airbus.com)
Susanne Hindennach (University of Stuttgart, DE, susanne.hindennach@vis.uni-stuttgart.de)

This breakout group discussed shared representations and communication in anticipatory human-machine interaction.

4.1.1 Shared Representations

Shared representations are often established through discussions/negotiations and can also be reflected in tangible artefacts like sketches/visualisations/mind maps. They comprise representations of task and environment or goals, and can comprise representations of (human/machine) agents and their internal states. The individual representation/understanding of what is shared might differ and such mismatches can be detected and resolved in communication that might require meta-cognition and/or meta-communication.

Conflicts can arise when machines are able to adapt their actions to confirm human anticipation, but know there is a better (optimal) way to achieve the goal. In some situations, the best policy might be to concede, observe and try to learn from human behaviour. In other situations negotiating tasks and goals and providing explanations will be more appropriate. Finding out when to apply what policy might be critical. From an ethical/legal perspective, authority and responsibility should remain with the human where possible.

4.1.2 Communication

Communication is the best way to detect mismatches in representations. We discussed three ways to classify communication.

- **indirect/direct**: Direct communication is about the means of communication and describes interaction without intervening factors to a designated communication partner-/audience, indirect communication is interaction with intervening factors.
- **explicit/implicit**: Explicit communication is about the clarity of communication and describes unambiguous interaction, implicit interaction describes ambiguous communication.
- **symmetric/asymmetric**: Symmetric communication describes balance in the use of knowledge, means or tools in communications between two communication partners, asymmetric implies imbalance.

Explicit communication of representations requires effort; implicit detection of mismatches might be more suitable in some situations but challenging. Physiological indicators of representations, intent, etc. can facilitate communication.

4.2 Timescales at which Anticipatory models for HMI can operate

Christian P. Janssen (Utrecht University, NL, c.p.janssen@uu.nl)
Cristina Conati (University of British Columbia – Vancouver, CA, conati@cs.ubc.ca)
Niels A. Taatgen(University of Groningen, NL, n.a.taatgen@rug.nl)

This group thought about the time scales at which anticipatory human-machine interaction could operate. Starting point was Newel's time scales that distinguish at which time interval behavior (and a model) occurs: milliseconds (biological band), seconds (cognitive band), minutes (rational band), or hours or longer (social band) ([1], see also chapter 1 of [2]).

After some discussion, we thought it might be interesting to identify two dimensions of time: (1) the time dimension at which a phenomenon (such as boredom) occurs, and (2) the time dimension at which a system might measure something about the human to infer user state. These might sometimes be at the same time interval scale, or sometimes differ. In some cases it might be ambiguous what the exact scale is. There might also be other ways to categorise time that our classification does not capture (yet).

Below we have five pictures of four example domains where different scenarios might take up different proportions of these time dimensions. In each figure there are "clouds" that indicate the rough position of the phenomenon and its measurement time. Within that cloud we also added text: the top text describes what the rough phenomenon is, and the bottom text indicates our thoughts about what the phenomenon can be. What is interesting to note about these figures, is that there are opportunities to consider anticipatory systems at many time scales, beyond systems that require immediate action. Although many clouds are positioned close to the diagonal, this should not be interpreted as that these are in principle more common. It might rather show our limitation in thinking of more diverse examples.

References

1 Newell, A. (1990). Unified theories of cognition. Cambridge, MA: Harvard University Press.
2 Salvucci, D. D., & Taatgen, N. A. (2010). The multitasking mind. Oxford University Press.

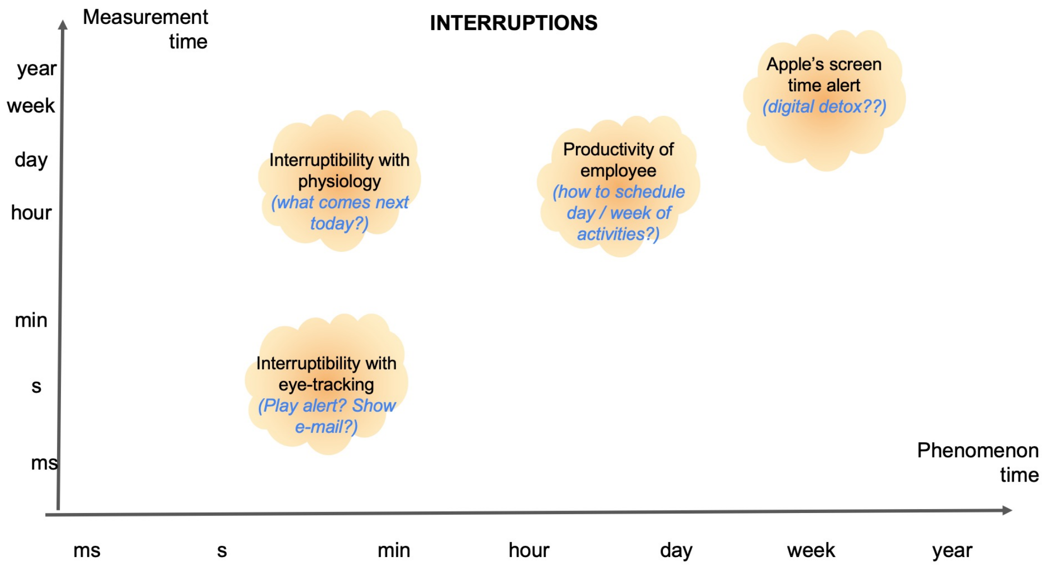

Figure 3 Example scenarios related to the domain of interruptions.

4.3 Data-driven vs. rule-based models for anticipation

Jelmer Borst (University of Groningen, NL, j.p.borst@rug.nl)
Andreas Bulling (University of Stuttgart, DE, andreas.bulling@vis.uni-stuttgart.de)
Emmanuelle Dietz (Airbus Central R&T, Hamburg, DE, emmanuelle.dietz@airbus.com)
Chenxu Hao (Friedrich-Alexander-Universität Erlangen-Nürnberg, DE, chenxu.hao@fau.de)
Frank Keller (University of Edinburgh, UK, keller@inf.ed.ac.uk)
Florian Strohm (University of Stuttgart, DE, florian.strohm@vis.uni-stuttgat.de)
Andreea Ioana Sburlea (University of Groningen, NL, a.i.sburlea@rug.nl)
Samuel Kaski (Aalto University, FI & University of Manchester, UK, samuel.kaski@aalto.fi)
Elisabeth André (Augsburg University, DE, elisabeth.andre@uni-a.de)
Henny Admoni (Carnegie Mellon University, USA henny@cmu.edu)

The purpose of this group was to discuss the suitability of rule-based models of anticipation compared to data-driven models. We start with a definition of these terms. Traditional examples of rule-based models include the ACT-R cognitive architecture. However, even ACT-R contains a learning component, e.g., the activation levels are trainable. Other rule-based models includes most models used for planning, and more generally logic-based models. The group felt that some tasks are specified symbolically (e.g., in robotics, where the starting state and the end state of an action is given), and this might bias any modeling effort towards rule-based models. Also, it is important to bear in mind that even rule-driven models are built based on data, except that the model designer comes up with the abstractions (e.g., based on experimental results), and formulates these abstractions in the form of rules.

The prototypical data-driven models are deep-learning models, which are black-box and contain very little internal structure, and are trained on massive datasets. Models with a simple structure are also included, e.g., models that learn state transitions (HMMs,

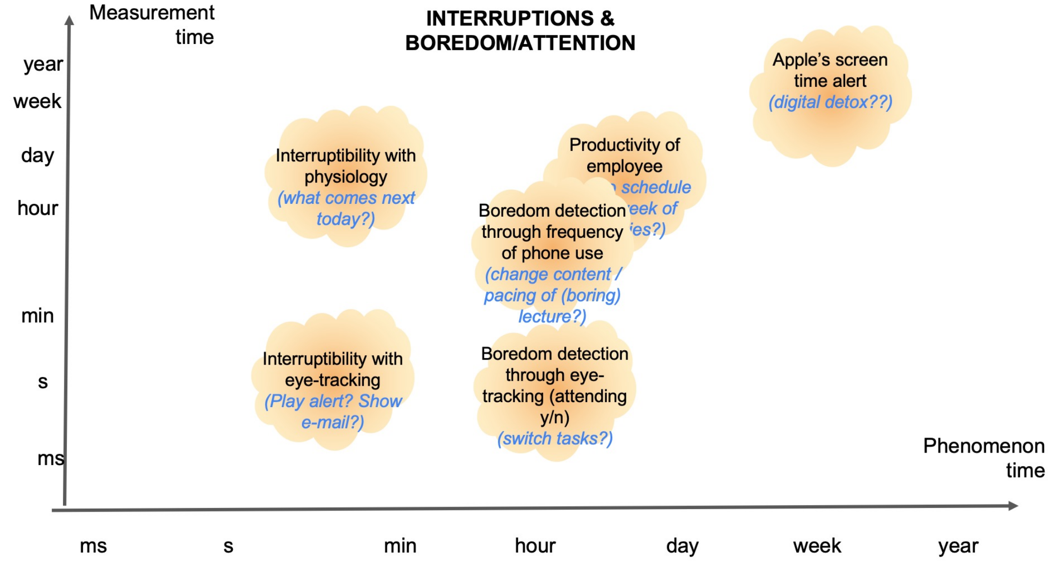

Figure 4 Example scenarios related to the domain of interruptions and boredom.

POMDPs). Instance-based learning is another example. It needs to be taken into account, though, that some data-driven models have internal structure (trees, graphs), for example graph convolutional neural networks.

The group also considered hybrid models and a third model type that combines both data-drive and rule-based models. Hybrid approaches are prima facie plausibility, for example because humans probably learn in a hybrid way, combining pre-specified rules (instincts) with data obtained from the environment. Language acquisition is a potential example: the learning mechanism seems to be innate, but external data is required to trigger the acquisition of an individual language. So it seems clear that you can't learn without data, but also not without (possibly pre-defined) structure. A classic example of a hybrid model in machine learning are Bayes Nets, which combine pre-specified (graph) structure with probabilities learned from data.

The group then switched to the discussion of Theory of Mind (TOM) and its relevance to modeling anticipation. As to the definition of TOM, we can think of it as a nested model, i.e„ a model within a model. More specifically, having a TOM means having a model of another agent that accounts for their knowledge, beliefs, and intentions. This raises the question of what TOM is useful for and which behaviors we can't do without TOM. Traditionally, non-literal behavior is thought to require TOM: lying, irony, sarcasm, empathy. The group then examined the question whether TOM can be learned in a data-driven way. This is probably possible in specific instances, but generalization would require a rule-based approach. The group also believed that TOM would not be needed for anticipation in the general case; limited, task-specific anticipation behavior could be learned from data. The group also noted that a higher level of abstraction is required to generalize anticipation across tasks, and TOM may be required for that. The link between (impaired) TOM and autism was noted.

There were other issues that seemed important to the group, but were not discussed in detail because of lack of time. This includes the question of how a simulation is different from a model. The group felt that a simulation required models of both the user and their environment. A simulation could also be used for reasoning, e.g., to simulate what a user

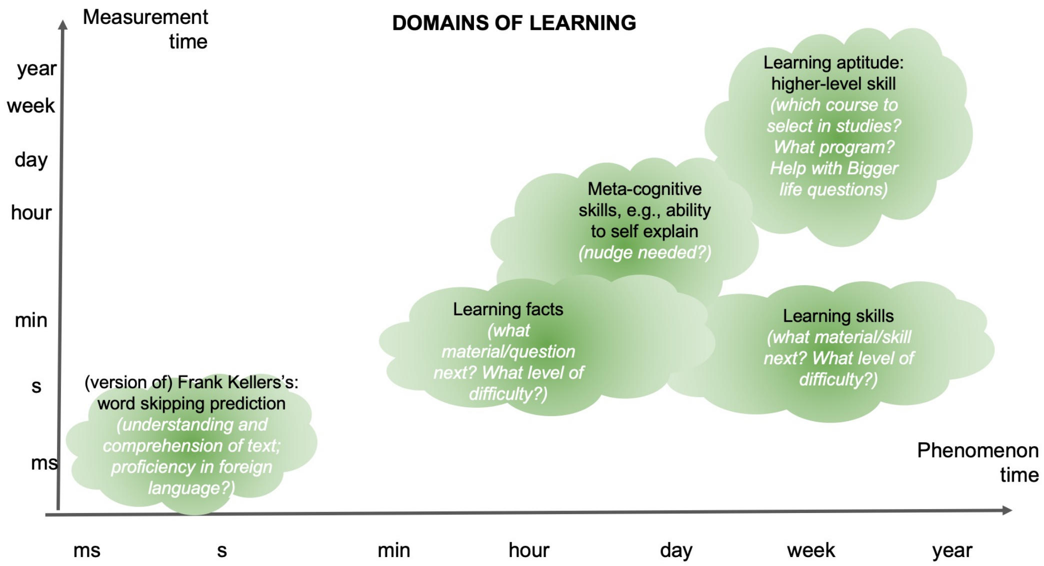

Figure 5 Example scenarios related to the domain of learning.

does in a given environment and therefore reason about what the user would end up doing. Another important topic that the group touched upon was explainability. Humans constantly generate explanations for their behavior and their actions, they tell stories that help them to make sense of the world and to interpret other people and their actions. Another issue the group touched upon was how TOM relates to sense making. Sense making is a way of model criticism, of challenging a model, adapting and modifying it. It was also pointed out that Inference is possible without probabilities, e.g., in Approximate Bayesian Computation (ABC). Another important issue seems to be causality. Can data-driven models be causal? The group felt that this was possible in certain cases, a good example are Bayes Nets.

Finally, the group discussed evaluation. This is potentially an issue with respect to theory of mind: how can you tell that a system has TOM? It may be possible to use tests designed in developmental psychology to test TOM in human participants, e.g., the false belief task (Sally-Anne task). These tests would have to be adapted for a human-machine interaction setting. This raises the wider question of how to evaluate anticipation. How do we tell if a system anticipates successfully? One approach would be to rely on task-based evaluation, i.e., just measure task success. This gives us an indirect way of evaluating anticipation. However, a direct evaluation measure ("anticipation accuracy") would be desirable, but this remains an issue for future work. Furthermore, the group noted that data-driven models cannot be trained without an evaluation measure that informs the design of the loss function.

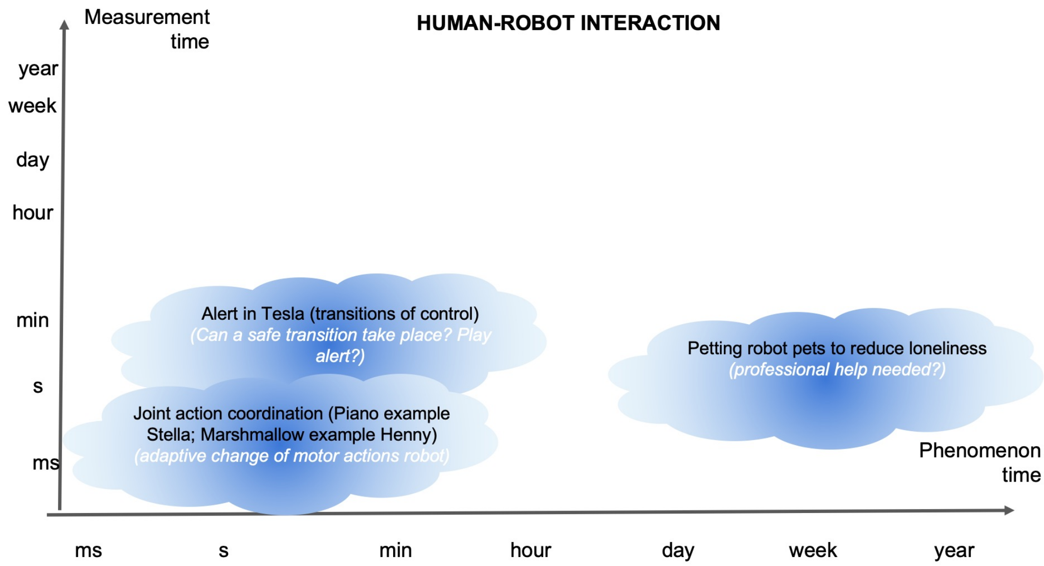

Figure 6 Example scenarios related to the domain of human-robot interaction

5 Breakout Groups (Tuesday and Wednesday)

5.1 Formalisation of Anticipation and Use Cases

Cristina Conati (University of British Columbia – Vancouver, CA, conati@cs.ubc.ca)
Emmanuelle Dietz (Airbus Central R&T, Hamburg, DE, emmanuelle.dietz@airbus.com)
Steffen Frey (University of Groningen, NL, s.d.frey@rug.nl)
Susanne Hindennach (University of Stuttgart, DE, susanne.hindennach@vis.uni-stuttgart.de)
Alvitta Ottley (Washington University in St. Louis, USA, alvitta@wustl.edu)
Niels A. Taatgen (University of Groningen, NL, n.a.taatgen@rug.nl)
Matteo Bortoletto (University of Stuttgart, DE, matteo.bortoletto@vis.uni-stuttgart.de)
Henny Admoni (Carnegie Mellon University – Pittsburgh, USA, henny@cmu.edu)

Anticipation is about the future. A system that anticipates can possibly perceive (or recognise) the current (and potentially past) target's state (e.g. human's cognitive state), which determines the prediction of a future state with respect to a set of available actions. Given some objective or utility function (e.g. to prevent a negative outcome or facilitate a positive outcome), the system decides which action to choose (e.g. prepare to act). The system might evaluate the outcome (the perceived future state) in order to evaluate and adapt accordingly.

5.1.1 Notation

The top of Figure 8 (above the green dotted line) represents the perceived environment and the bottom (below the green dotted line) represents the actual (target) environment, i.e. what the system intends to perceive.

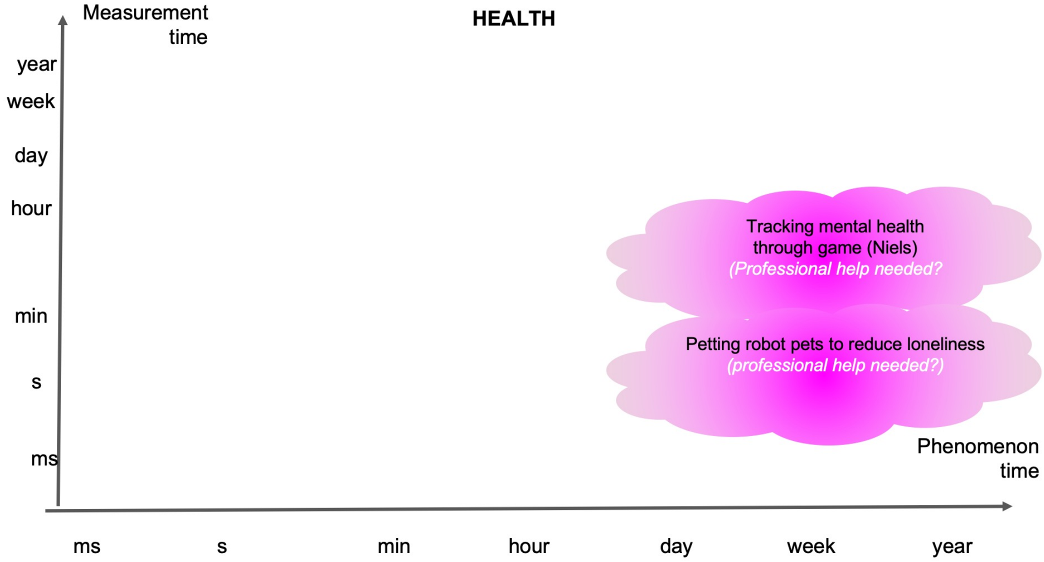

Figure 7 Example scenarios related to the domain of health

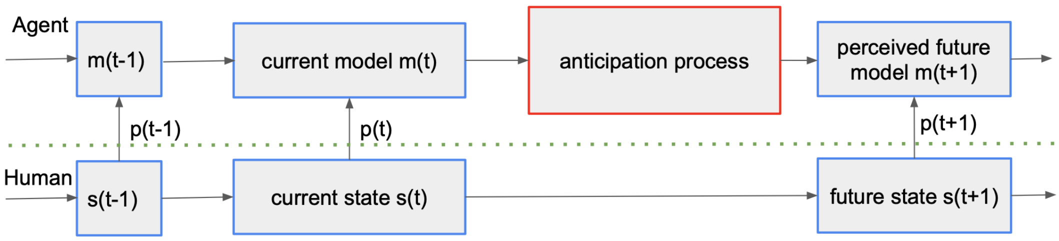

Figure 8 Overview of the Formalisation of Anticipation.

The models in the perceived environment are from the view of the system (e.g. artificial agent or even a human). The perceived environment is a sequence of models that can change over time ...$m(t-1), m(t), m(t+1)$, ... where $t > 1$. The states within the actual environment are actual (target) states (i.e. states that the system intends to perceive), which can range from very simple states (e.g. button on/ off) to very complex states (e.g. the cognitive state of a human). The actual environment is a sequence of states that can change over time, ...$s(t-1), s(t), s(t+1)$, ... where $t > 1$. The perceived and the actual environment can differ, depending on the specification of the actual states and the accuracy of the perception function p.

We further introduce the following notation:

- S is a set of states of the world, where each state $s(t) \in S$ is specified with respect to time t;
- M is a set of models (representations) of the world, where each model $m(t) \in M$ is specified with respect to time t;
- A is a set of possible actions in the world, where each action $a(t) \in A$ is specified with respect to time t;
- P is a set of perceptions of the world, where each perception $p(t) \in P$ is specified with respect to time t;

- $\text{Obj} : M \to \mathbb{R}$, the objective function, mapping a model m to a value;
- $\rho : S \to P$, the perception function, mapping a state $s(t)$ to a perception p;
- $\mathit{update} : M \times P \to M$ is the (expectation) update function. It updates a model $m(t)$ given the current perception $p(t)$;
- $\mathit{prediction} : M \times A \to M$, the prediction function. It predicts the future model $m'(t+1)$ given the current chosen action $a(t)$ and current model $m(t)$.

The action that give the highest predicted value from model m at time t m$m(t)$, is computed as follows:

$$a(t) = \operatorname*{argmax}_{a_i \in A} [\text{Obj}(\mathit{prediction}(m(t), a_i))] \tag{1}$$

where $\mathit{prediction}(m(t), a_i) \equiv m'(t+1)$ is the predicted effect of action a_i in $m(t)$ and $\text{Obj}(\mathit{prediction}(m(t), a_i))$ is the predicted value of $m'(t+1)$.

The predicted model according to the selected action producing an *anticipation* is computed by the prediction function *prediction* follows:

$$m'(t+1) = \mathit{prediction}(m(t), a(t)) \tag{2}$$

The model $m(t+1)$ is computed according to the predicted model $m'(t+1)$ and the perception of state $s(t+1)$ by the update function *update*:

$$m(t+1) = \mathit{update}(m'(t+1), \rho(s(t+1))) \tag{3}$$

5.1.2 Instantiating the formalism for a particular case

To apply the formalism to a particular problem, the components need to be specified, so we need to provide $\langle M, A, P, \text{Obj}, \rho, \mathit{update}, \mathit{prediction}\rangle$. Note that we do not specify S, because the system has no direct access to it.

Each of the individual components can be very complex and the focus of the research, or can be very simple. For example, ρ is simple if all we need to do is check whether a student answered a multiple choice question correctly, or complex when mapping EEG measurements onto workload level. In particular, a lot of variation can be expected in the model M. A model can be not much more than the outcome of perception, but can also be a complete cognitive model with theories about learning, and a record of all past events. Moreover, the model can represent the goal of the system, and a partial plan that is update with experience. An elaborate model will also, in most cases, have more elaborate *update* and *prediction* functions.

5.1.3 Algorithm

Algorithm 1 shows an algorithm applied on a use case about the interaction with a tutoring system. In this task the student interacts with a tutoring system and the systems' goal is that the student gets better at the assignment. Algorithm 2 shows one possible way how the prediction function could be defined for this use case.

Note: The objective function Obj can return how likely an action is to lead to a desired future state or action, that is, how well action a prepares for some goal.

Algorithm 1 Tutoring example.

S ▷ what student knows about algebra
A ▷ set of assignments that address specific skills
$M = \{0, 1\}^k$ ▷ ability on k algebraic skills
$P = \{0, 1\}$ ▷ correct or incorrect
$\text{Obj} = \sum m_i$
$m \leftarrow 0^k$ ▷ student has no knowledge to start
while $\text{Obj}(m) < k$ **do**
  $a \leftarrow \underset{a_i \in A}{\text{argmax}}(\text{Obj}(\textit{prediction}(m, a_i))$ ▷ select an action
  $m' \leftarrow \textit{prediction}(m, a)$ ▷ make a prediction
  $s \leftarrow$ execute a
  $\textit{observation} \leftarrow \textit{update}(m', \rho(s))$ ▷ perceive the new states
  if *observation* is correct **then** $m \leftarrow m'$
  else $m \leftarrow m$ with some deduction ▷ student not as far as we thought
  end if
end while

Algorithm 2 Prediction function: predict learning gain given an assignment.

Given $a \in A$ ▷ set of actions (assignments)
Given $m \in M$ ▷ model (skill levels of the student)
$\textit{skill} \leftarrow [0, 1]^k$ ▷ which skills are trained by a
δ ▷ learning rate
τ ▷ learning threshold
$\textit{student\_improvement} \leftarrow \sum \textit{relu}\{\textit{skill}(a) - -m\}$ ▷ student skill improvement
if $\textit{student\_improvement} < \tau$ **then**
  $m \leftarrow m + (\delta \cdot \textit{skill}(a))$ ▷ update student's knowledge
else
  $m \leftarrow m$ ▷ expectation is that assignment is too hard
end if

5.1.4 Partial/ Uncertain Observations

There will always be some uncertainty as to what is perceived from the system and what is happening in the user (cognitive states). This uncertainty depends on the complexity of what is intended to be perceived. If, for instance, the only requirement is to perceive whether a button is on or off, then the perceived state and the actual state are likely to be the same most of the time (except e.g. if some sensor is broken). However, if the system intends to perceive a human's cognitive state based on the human's behavior, then the perceived state can only provide an approximation.

5.2 Transportation-driven Anticipatory Human-Machine Interaction

Elisabeth André (Augsburg University, DE, elisabeth.andre@uni-a.de)
Mihai Bâce (University of Stuttgart, DE, mihai.bace@vis.uni-stuttgart.de)
Jelmer Borst (University of Groningen, NL, j.p.borst@rug.nl)
André Brechmann (Leibniz Institute for Neurobiology Magdeburg, DE, brechmann@lin-magdeburg.de)
Andreas Bulling (University of Stuttgart, DE, andreas.bulling@vis.uni-stuttgart.de)
Christian P. Janssen (Utrecht University, NL, C.P.Janssen@uu.nl)

5.2.1 A formalism for anticipatory human-machine interaction

We defined anticipatory human-machine interaction as predicting future states, which can include information about human behaviour, the environment, or the system, to prepare for a future machine action.

$$s_{h,t+i} = f(s_{h,t}, a_{m,t+j})$$

where $s_{h,t+i} \in \mathit{DesiredStates}$, $s_{h,t+i} \notin \mathit{BlacklistedStates}$, and $t + j < t + i$.

5.2.2 Domain

In our group, the discussions were mainly driven by one key application domain, namely safety-critical human-machine interaction. In particular, our formalism was defined with an aviation use case in mind. Despite a focus on this domain, our formalism can also be applied to a tutoring scenario where a positive intervention could be used to improve the learning process. Finally, we also thought about an application scenario in which the machine schedules interruptions at opportune moments in a task.

5.2.3 Methods

An anticipatory agent has to perceive, reason and react.

Prediction or state tracking is needed to keep track of whether the user has achieved the desired state. Measurement of human / user sensing (perceive) is needed to know the human state, because we want to achieve some desired human state or behaviour.

Learning might be needed to be adaptive to different contexts / settings AND to adjust underlying models in case it (almost) failed in the past (if anticipation was wrong).

We identified a number of methods, according to the different function that they have to fulfil:

- Recurrent networks
- Classifiers
- Anything using Bayesian statistics
- Some form of user / cognitive model

Learning: (can be model-free or model-based) (reason)

- Multi-agent RL
- Symbolic models
- Measurements of human / user sensing (Perceive)
- Physiological + brain

- Question/self-declaration/self-reports
- Observation of actions (taken yes / no? Correct? speed?)

In addition to the above methods that can be used for different abilities of the agent, we discussed possible interruption or intervention modalities.

- Giving a subtle nudge
- Giving a more prominent intervention (such as an alert)
- Brute force taking over behaviour by the system (e.g., transition of control in car)
- Notifying/informing something or someone outside of the system and user (e.g., notifying air traffic control that the pilot seems drowsy, or is in difficult conditions)

5.2.4 Case Study

We discussed the following case study relevant to the aviation domain. Let us assume an unexpected situation in the plane's cockpit, such as a thunderstorm coming up that the pilot may not be aware of. A challenge for anticipatory human-machine interaction: Can the machine (e.g. the plane or a system in the plane) anticipate that the pilot is unaware of the upcoming storm? If yes, how can the system prepare the pilot? The anticipation task in this case it to anticipate the pilot's state of whether they are aware of the storm or not. The machine could suggest alternative routes or request remote support (e.g., air traffic control, etc.).

5.2.5 Challenges

In the breakout session, our group identified a number of key challenges:

- What aspects of user behaviour or state does one need to track, measure, or sense. The state has to be complete (have all the information you need), it should be privacy sensitive (no information you should not need), and feasible to enable reasoning within limited time. Finally, the system should avoid being "over sensitive". Some of the above also hold for systems / machines.
- What are the blacklisted and desired states?
- Act in time before the state is reached, without negatively affecting the user
- Avoid recursive (?) anticipation / be "dynamic" in adjusting actions real-time
- Multi-agent ToM, e.g., system estimates what experienced pilot assumes about state of novice pilot.
- Within aviation domain: there are consequences and systems might not be changed overnight and only once they have been heavily tested. E.g., system needs all kinds of certifications. Transparent, traceable
- Have unambiguous system actions (to avoid that the human/pilot gets distracted by an intervention or misinterpret)

5.3 Planning in Anticipation

Leslie Blaha (AFRL – Wright Patterson, US, lblaha@andrew.cmu.edu)
Matteo Bortoletto (University of Stuttgart, DE, matteo.bortoletto@vis.uni-stuttgart.de)
Samuel Kaski (Aalto University, FI & University of Manchester, UK, samuel.kaski@aalto.fi)
Andreea Ioana Sburlea (Univerity of Groningen, NL, a.i.sburlea@rug.nl)

5.3.1 Definition of planning

Planning is the process of thinking regarding the actions to take in order to maximise a form of utility, given a context. Such context can be defined in terms of the following quantities:

- an action space, defining the set of possible actions
- a conditional state space, defining the set of possible states of the environment
- an observation space, defining the set of observations an agent can experience of its environment
- a transition function, giving for each environment state and agent action a probability distribution over environment states
- an observation function, which gives, for each action and resulting state, a probability distribution over possible observations
- a reward function

5.3.2 Methods for planning in Anticipatory Human-Machine Interaction

Two methods are required for planning in Anticipatory Human-Machine Interaction (AHMI). First, we need a method to compute the actual plan. Such method heavily depends on the specific task we are dealing with. Classic approaches involve game theory, probabilistic models, constraint optimisation, causal inference and simulator based inference. Development in deep learning provided other valuable tools, such as active learning and self-play. In particular, combining deep learning methods like self-play and heuristic search algorithms like Monte Carlo tree search led to incredibly powerful models, e.g. AlphaGo. More complex scenarios could require Theory of Mind, e.g. for user modelling. In these cases, inferring humans' mental states could help the AI-assistant to efficiently cooperate with users.

5.3.3 Use cases

The variety of use cases that require (strategic) planning is potentially unlimited. Most decision making scenarios require planning. Some examples are teamwork, model design, education, games or sports, social events, emergency response, crime prevention, rehabilitation, surgery, transportation and logistics.

5.3.4 A formal definition of Anticipatory Human-Machine Interaction in terms of planning

We propose a definition of AHMI based on the concept of planning. *AHMI is helping humans thinking ahead to place resources to maximise eventual utility*, over the space of potential plans. Anticipation is naturally linked to thinking ahead and as we are dealing with human-machine interaction the goal is to help humans. Resources represent some assets that will result on actions taken in order to maximise a form of utility, which is defined by the context.

5.4 Physical Interaction

Nele Russwinkel (TU Berlin, DE, nele.russwinkel@tu-berlin.de)
Frank Keller (University of Edinburgh, UK, keller@inf.ed.ac.uk)
Florian Strohm (University of Stuttgart, DE, florian.strohm@vis.uni-stuttgart.de)
Chenxu Hao (Friedrich-Alexander-Universität Erlangen-Nürnberg, DE, chenxu.hao@fau.de)
Henny Admoni (Carnegie Mellon University, Pittsburgh, USA, henny@cmu.edu)

Physical human-agent (or human-robot) interaction involves interaction between one or more people and one or more emobodied agents in some kind of physical space. This kind of interaction often, but not always, includes the agent having some kind of physical effect in the world. One domain in which physical human-agent interaction has become very salient in recent years is autonomous vehicles (AVs), where drivers, pedestrians, and vehicles must collaborate to navigate safely. For example, a driver may anticipate that the AV's autopilot is likely to fail, and take over control before a critical state is reached. In contrast, a vehicle might anticipate that the driver is available for a control handoff, and yield control of the vehicle. An AV may avoid collisions by anticipating pedestrian actions (e.g., crossing the street), or anticipating the behaviour of other vehicles when merging lanes or performing other negotiations with drivers.

5.4.1 Discussed Problems

Physical human-agent interaction provides a special case of anticipation. Because it involves an embodied agent interacting in the real world, the challenges of the domain can be divided into sensing, planning, and acting. Specifically, some challenges include:

- Sensing
  1. Perception of environment
  2. Perception of user (including social cues)
- Planning
  1. Knowledge of task (e.g., transition function)
  2. Maintaining safety
  3. In collaborative scenarios, also:
     a. Shared representation
     b. Bi-directional communication
- Acting
  1. Proactive action
  2. Transparent and explainable action
  3. Adhering to social norms / cues

There are also other challenges to anticipation in physical interactions. These include how to integrate information from multiple input sources, and how to create enough model flexibility to enable generalisation to new situations.

5.4.2 Possible Approaches

Methods for enabling anticipation in physical human-agent interactions centre around enabling bi-directional communication between human and AI agents. This communication can be explicit or implicit, verbal or nonverbal, and direct or indirect.

Direct and explicit communication mechanisms include speech and text. These are generally easy to understand but can be difficult to produce because of their specificity. Implicit behaviours can integrate more seamlessly into ongoing interactions and can add information that supports and augments explicit communication.

For the robot to anticipate the human implicitly, it can sense aspects of human behaviour such as gaze using eye tracking, gestures using motion tracking, and brain signals using brain-computer interfaces or EEG. Other elements that can be sensed from humans include physiological cues and social cues.

Other methods focus on allowing the human to anticipate the robot. To make it easier for humans to do so, robots can generate implicit, interpretable behaviours of their own. These include eye gaze (or head direction), gestures (including pointing), and legible motion. Research has investigated how to automatically generate these behaviours and their effect on human understanding.

To model human and robot in a way that enables physical interaction, robots must be able to represent certain aspects of their environment and partner. These include: a value function to represent reward, a state representation that captures relevant aspects of the world including representations of the human partner(s) and environment, a task representation that allows the robot to plan, and an action space and motion planner that enables the robot to act. Other components of models that are less critical but still useful include an ability to learn or adapt the model, flexibility, and a notion of uncertainty over states and actions.

There are, of course, a number of modeling methods that have been used for anticipation in physical interactions. These include now-standard approaches such as MDP/POMPDs, cognitive models like ACT-R and SOAR, deep neural networks, and Bayes Nets.

5.4.3 Conclusions

Physical human-agent interaction has need of anticipation to ensure safety, task efficiency, and user comfort. Domains like autonomous vehicles or collaborative building require interactions that would benefit from an agent being able to anticipate and respond to human needs, goals, and expectations. The challenges of this domain exist throughout the sense-plan-act cycle. Approaches to address anticipation in physical human-agent interaction must involve representations that enable robot behavior (e.g., state and value functions, motion planners) as well as representations that model the human to be anticipated.

Participants

- Henny Admoni
Carnegie Mellon University – Pittsburgh, USA
- Elisabeth André
Augsburg University, DE
- Mihai Bâce
University of Stuttgart, DE
- Leslie Blaha
Air Force Research Laboratory – Wright Patterson, USA
- Jelmer Borst
University of Groningen, NL
- Matteo Bortoletto
University of Stuttgart, DE
- André Brechmann
Leibniz Institute for Neurobiology Magdeburg, DE
- Andreas Bulling
University of Stuttgart, DE
- Cristina Conati
University of British Columbia – Vancouver, CA
- Emmanuelle Dietz
Airbus Central R&T – Hamburg, DE
- Steffen Frey
University of Groningen, NL
- Cleotilde Gonzalez
Carnegie Mellon University – Pittsburgh, USA
- Chenxu Hao
Friedrich-Alexander-Universität Erlangen-Nürnberg, DE
- Susanne Hindennach
University of Stuttgart, DE
- Christian P. Janssen
Utrecht University, NL
- Samuel Kaski
Aalto University, FI & University of Manchester, UK
- Frank Keller
University of Edinburgh, UK
- Oliver W. Klaproth
Airbus – Hamburg, DE
- Alvitta Ottley
Washington University in St. Louis, USA
- Nele Russwinkel
TU Berlin, DE
- Andreea Ioana Sburlea
University of Groningen, NL
- Florian Strohm
University of Stuttgart, DE
- Niels Taatgen
University of Groningen, NL

Report from Dagstuhl Seminar 22221

Exponential Analysis: Theoretical Progress and Technological Innovation

Annie Cuyt*[1], Wen-shin Lee*[2], Gerlind Plonka-Hoch*[3], and Ferre Knaepkens†[4]

1 **University of Antwerp, BE.** annie.cuyt@uantwerpen.be
2 **University of Stirling, GB.** wen-shin.lee@stir.ac.uk
3 **Universität Göttingen, DE.** plonka@math.uni-goettingen.de
4 **University of Antwerp, BE.** Ferre.Knaepkens@uantwerpen.be

Abstract

Multi-exponential analysis might sound remote, but it touches our daily lives in many surprising ways, even if most people are unaware of how important it is. For example, a substantial amount of effort in signal processing and time series analysis is essentially dedicated to the analysis of multi-exponential functions. Multi- exponential analysis is also fundamental to several research fields and application domains that have been the subject of this Dagstuhl seminar: remote sensing, antenna design, digital imaging, all impacting some major societal or industrial challenges such as energy, transportation, space research, health and telecommunications. This Seminar connected stakeholders from seemingly separately developed fields: computational harmonic analysis, numerical linear algebra, computer algebra, nonlinear approximation theory, digital signal processing and their applications, in one and more variables.

Seminar May 29 – June 3, 2022 – http://www.dagstuhl.de/22221
2012 ACM Subject Classification Mathematics of computing → Computations on polynomials; Mathematics of computing → Interpolation; Mathematics of computing → Numerical analysis
Keywords and phrases inverse problem, remote sensing, sparse interpolation, spectral analysis, structured matrices
Digital Object Identifier 10.4230/DagRep.12.5.170

1 Executive Summary

Annie Cuyt (University of Antwerp, BE)
Wen-shin Lee (University of Stirling, GB)
Gerlind Plonka-Hoch (Universität Göttingen, DE)

For the analysis and representation of stationary signals and images the conventional Fourier- and wavelet-based methods are particularly appropriate. However, in many areas in science and engineering we are faced with the problem to interpret digital signals and images which are not band-limited and have a non-stationary behaviour. Frequently, there are even further obstacles. The acquisition of signal or image measurements may be very expensive and therefore limited. In other applications, measurement sets are huge but contaminated by noise. Examples of the above are encountered in magnetic resonance imaging, infrared microscopy, fluorescence-lifetime imaging microscopy (FLIM), the analysis of seismic signals

* Editor / Organizer
† Editorial Assistant / Collector

in geophysics, radar imaging (SAR/ISAR), tissue ageing models, vibration analysis, direction of arrival (DOA) detection, texture classification, radio frequency identification (RFID), non-destructive testing, satellite navigation, time series analysis, echolocation, induction motor diagnostics (MCSA), to name just a few.

Within the last few years, research on Prony-based methods has been intensified, as they offer an alternative to the compressed sensing approach. One essential advantage of the Prony method is that it does not need randomly collected measurements but works with deterministic sampling based on a sampling scheme which is adapted to the nonlinear signal model. At the same time, the Prony approach does not suffer the well-known curse of dimensionality in the multivariate case.

This Dagstuhl Seminar brought together a number of researchers from different areas in mathematics, engineering and industry. Topics included new mathematical insights and efficient numerical algorithms for problems based on exponential analysis as well as applications of exponential analysis models in engineering and life sciences. During this Seminar, the participants presented their newest results and discussed several open problems and applications from different perspectives.

The talks in the workshop partially had the character of survey talks and were arranged in different main research topics, as new mathematical theory (Day 1 and Day 5), new computational approaches (Day 2) and new results in applications in engineering (Day 3) and life sciences (Day 4).

On the first day, the talks focussed on the **mathematical theory of exponential analysis**. B. Beckermann and A. Matos emphasized the close connection between exponential analysis and rational approximation, which leads to new application areas as density reconstruction in an equilibrium problem in logarithmic potential theory and improved rational approximation of Markov functions. As shown in the talk by T. Sauer, there is also a close connection between multi-exponential analysis and continued fractions.

In the afternoon, the participants started four smaller **thematic discussion groups** to discuss new challenges and open problems throughout the week. These discussion groups particularly focussed on theory, more efficient computational algorithms and applications in engineering and biosciences. The discussions included general observations on the further development of exponential analysis and its ties to other subjects as well as further specific approaches for application of the theory to application problems in radar imaging or MRI. Still, we are far away from a full understanding of the relations between the different methods for the stable reconstruction of a parametric signal model, as well as from a systematic construction of Fourier analytic methods for the improved analysis of non-stationary signals.

In practice, Prony-based methods often suffer from a bad conditioning of the involved structured matrices and some extra effort is required to reliably execute the corresponding algorithms. Among the more successful implementations, the ESPRIT method, the Matrix-Pencil method, the approximate Prony method, and validated exponential analysis are established. The problem statement is also closely related to rational approximation theory and the structured low-rank approximation of structured matrices. These connections are still not completely understood and may lead to strongly improved reconstruction algorithms, able to treat more general sampling sets and deliver super-resolution results.

The second day was devoted to the problem of **efficient computations** for reconstruction or approximation of functions using multi-exponential analysis. First numerical methods like MUSIC are already implemented in the computer algebra system Maplesoft, as presented in the talk by J. Gerhard. The advantage of the application of Maplesoft is that it can process the data within a desired very high precision and therefore successfully handle these reconstruction problems which are known to be ill-conditioned. Another way to improve the numerical stability is the application of more sophisticated numerical approaches to structured matrices and the connection to rational approximation, such that the existing stable algorithms for rational approximation can be used, see the talk of G. Plonka. The survey presentation by H. Mhaskar turned the attention to the application of neural networks to approximation problems, and D. Potts presented new efficient Fourier methods to compute the so-called ANOVA decomposition for approximation of multivariate functions.

In the afternoon a mini-course on newest features of Maplesoft was presented by J. Gerhard, with a special focus on the problem of how to connect MAPLESOFT with other Software as MATLAB etc. The remaining time was used for further scientific discussions in smaller groups.

On the third day, the talks surveyed different new **applications of exponential analysis in engineering**. The presentation of F. Knaepkens showed new approaches for direction of arrival (DOA) estimation, image denoising and inverse synthetic aperture radar. Chromatic aberration in large antenna systems have been studied by D. de Villiers. R.-M Weideman showed her results on antenna position estimation through sub-sampled exponential analysis of signals in the near-field. The talks by R. Beinert and J. Prestin showed applications concerning phase retrieval in optical diffraction tomography and detection of directional jumps in images.

The survey presentations on the fourth day focussed on **exponential models in life sciences**. J. Gielis showed in his talk, how generalized Möbius-Listing bodies (GML) can be employed for better modelling and understanding of certain dynamical processes in the natural sciences. D. Li explained recent progress of advanced time-resolved imaging techniques based on exponential analysis models and their applications in life sciences, for example to reveal biological processes at the molecular level. In a further talk, R.G. Spencer reported on newest developments in Magnetic Resonance Relaxometry and Macromolecular Mapping to achieve more accurate myelin quantification in the brain that permits the establishment of physiological correlations. The underlying reconstruction problem is a seriously ill-posed inverse problem.

The last day of the workshop was again devoted to further results in **mathematical theory of exponential analysis** and connections to other areas of mathematics. In the talk by D. Batenkow, the degree of ill-posedness of the parameter reconstruction problem based on the exponential sum model was studied in more detail. The degree of condition of the problem essentially depends on the distribution of the frequency parameters. H. Vesolovska discussed the problem of recovering an atomic measure on the unit 2-sphere $\mathbb{S}^2$ given finitely many moments with respect to spherical harmonics. A connection of exponential analysis to computer algebra problems was brought to our attention by M. Ishteva. She showed how the joint decomposition of a set of non-homogeneous polynomials can be computed using the canonical polyadic decomposition, and how this decomposition can be applied in nonlinear system identification.

This Dagstuhl meeting has been an important milestone for improved understanding of the large impact of exponential analysis tools in both theory and practice. During the thematic discussions in small groups in this meeting, several new aspects have been considered and several collaborations have been initiated or continued. Examples include new approaches for an improved modelling of antenna system frequency responses in radio frequency (Cuyt, De Villiers, Weideman) and for stabilised parameter estimation in exponential models using iterative factorizations of matrix pencils of Loewner matrices (Beckermann, Plonka-Hoch).

We mention that this seminar is related to Dagstuhl seminar 15251 on "Sparse modelling and multi-exponential analysis" that took place in 2015. The discussions at the latter have led to many interesting collaborative projects, among which a funded Horizon-2020 RISE project (Research and Innovation Staff Exchange) with the acronym EXPOWER, standing for "Exponential analysis Empowering innovation" (grant agreement No 101008231).

It is our experience that these Dagstuhl seminars are timely and seminal. Through the meetings new collaborations and new potential are unlocked. There is a clear need to further connect stakeholders from the new theoretical developments and the identified industrial applications, as is our objective here and in the future.

2 Table of Contents

3 Overview of Talks

3.1 Limits of sparse super-resolution

Dmitry Batenkov (Tel Aviv University, IL)

We consider the recovery of parameters $\{c_j, x_j\}$ in exponential sums

$$f(\omega) = \sum_{j=1}^{s} c_j \exp(\imath x_j \omega)$$

from bandlimited and noisy samples

$$f(\omega_i) + \epsilon_i, \qquad \omega_i \in [-\Omega, \Omega], \quad |\epsilon_i| \leq \varepsilon.$$

We discuss the conditioning of the problem when some of the exponents $\{x_i\}$ become close to each other, and show that all model parameters can be stably recovered provided that $\varepsilon \leq c(\Omega\Delta)^{2\ell-1}$, where ℓ is the maximal number of exponents which can be within an interval of size $\approx 1/\Omega$, and Δ is the a-priori minimal separation between the $\{x_j\}$.

We also discuss extensions of the analysis to generalized exponential sums $f(\omega) = \sum_{j=1}^{s}(a_j + b_j\omega)\exp(\imath x_j \omega)$, and connections to spectral properties of (confluent) Vandermonde matrices with nodes on the unit circle.

References

1 D. Batenkov, L. Demanet, G. Goldman, and Y. Yomdin, "Conditioning of Partial Nonuniform Fourier Matrices with Clustered Nodes," *SIAM J. Matrix Anal. Appl.*, **44**(1), 199–220, 2020.
2 D. Batenkov, B. Diederichs, G. Goldman, and Y. Yomdin, "The spectral properties of Vandermonde matrices with clustered nodes," *Linear Algebra and its Applications*, **609**, 37–72, 2021.
3 D. Batenkov and G. Goldman, "Single-exponential bounds for the smallest singular value of Vandermonde matrices in the sub-Rayleigh regime," *Applied and Computational Harmonic Analysis*, **55**, 426–439, 2021.
4 D. Batenkov, G. Goldman, and Y. Yomdin, "Super-resolution of near-colliding point sources," *Inf Inference*, **10**(2), 515–572, 2021.
5 D. Batenkov and N. Diab, "Super-resolution of generalized spikes and spectra of confluent Vandermonde matrices," *arXiv:2203.11923*.

3.2 Best rational approximants of Markov functions

Bernhard Beckermann (University of Lille, FR)

Main reference Bernhard Beckermann, Joanna Bisch, Robert Luce: "On the rational approximation of Markov functions, with applications to the computation of Markov functions of Toeplitz matrices", Numer Algor 91, 109–144, (2022).
URL https://doi.org/10.1007/s11075-022-01256-4

The study of the error of rational approximants of Markov functions

$$f^{[\mu]}(z) = \int \frac{d\mu(x)}{z - x}, \quad \text{supp}(\mu) \subset [a, b],$$

on some $E \subset R \setminus [\alpha, \beta]$ has a long history, with a well-established link to orthogonal polynomials. For example, Zolotarev more than 100 years ago described best rational approximants and their error for the particular Markov function

$$f^{[\nu]}(z) = \frac{\sqrt{|a|}}{\sqrt{(z-a)(z-b)}} = \int \frac{d\nu(x)}{z-x}, \quad \frac{d\nu}{dx}(x) = \frac{\sqrt{|a|}}{\pi\sqrt{(x-a)(b-x)}},$$

for closed intervals E. The aim of this talk is to show that

$$\min_{r \in \mathcal{R}_{m-1,m}} \|1 - r/f^{[\mu]}\|_{L^\infty(E)} \leq 3 \min_{r \in \mathcal{R}_{m-1,m}} \|1 - r/f^{[\nu]}\|_{L^\infty(E)},$$

that is, up to some modest factor, the particular Markov function $f^{[\nu]}$ gives the worst relative error among all Markov functions $f^{[\mu]}$. In our proof we show similar inequalities for rational interpolants and Padé approximants.

3.3 Total-Variation-Based Phase Retrieval in Optical Diffraction Tomography

Robert Beinert (TU Berlin, DE)

Joint work of Robert Reinert, Michael Quellmalz
Main reference Robert Beinert, Michael Quellmalz: "Total Variation-Based Reconstruction and Phase Retrieval for Diffraction Tomography", SIAM Journal on Imaging Sciences, Vol. 15(3), pp. 1373–1399, 2022.
URL http://dx.doi.org/10.1137/22M1474382

In optical diffraction tomography (ODT), the three-dimensional scattering potential of a microscopic object rotating around its center is recovered by a series of illuminations with coherent light. Reconstruction algorithms such as the filtered backpropagation require knowledge of the complex-valued wave at the measurement plane, whereas often only intensities, i.e., phaseless measurements, are available in practice.

In this talk, we propose a new reconstruction approach for ODT with unknown phase information based on three key ingredients. First, the light propagation is modeled using Born's approximation enabling us to use the Fourier diffraction theorem. Second, we stabilize the inversion of the non-uniform discrete Fourier transform via total variation regularization utilizing a primal-dual iteration, which also yields a novel numerical inversion formula for ODT with known phase. The third ingredient is a hybrid input-output scheme. We achieved convincing numerical results, which indicate that ODT with phaseless data is possible. The so-obtained 2D and 3D reconstructions are even comparable to the ones with known phase.

3.4 Chromatic Aberration in Large Antenna Systems

Dirk de Villiers (University of Stellenbosch, ZA)

It is well understood in electromagnetic field theory that harmonic plane waves, traveling in the same direction but emanating from different locations, interfere to cause a total field with a complex exponential frequency dependence. This effect often manifests in large (many wavelengths in size) antennas as a frequency ripple in the port reflection coefficient and radiation pattern magnitudes. The mechanism generating the ripple, or chromatic aberration [1], can normally be attributed to either multiple reflections from different structures in the antenna, or due to interference between direct and diffracted waves in the structure [2].

Often the resulting effect is small by virtue of the fundamental design of the structures at hand, but for certain applications even such small effects can influence the efficacy of the total system. Many radio astronomy science cases are examples of such sensitive applications, where the experiment searches for extremely feint signals (often deep in the noise) that need to be accurately characterized as a function of frequency. Here, we need to model our antenna system frequency responses very accurately in order to de-embed them from the data, and the chromatic aberration effects must be accounted for. Antennas typically used for radio astronomy, namely large reflectors or smaller antennas suspended over an artificial or natural ground plane, all suffer from this aberration to some degree. Previous efforts to model this effect are rather crude, and opens the door for more careful consideration and application of exponential analysis to reconstruct the required frequency responses without resorting to extremely time consuming high frequency resolution numerical analyses.

References

1. D. Morris, Chromatism in radio telescopes due to blocking and feed scattering. Astronomy and Astrophysics 67, 221-–228, 1978.
2. D.I.L. de Villiers, Prediction of Aperture Efficiency Ripple in Clear Aperture Offset Gregorian Antennas. IEEE Transactions on Antennas and Propagation 61(5), 2457-–2465, 2013.
3. D.I.L. de Villiers, Wideband Reflector Antenna Beam Pattern Frequency Variation Modeling through the Characteristic Basis Function Pattern Method. In Int. Conference on Electromagnetics in Advanced Applications (ICEAA), 220–223., Cartagena de Indias, 2018.

3.5 New Features in Maple 2022

Jürgen Gerhard (Maplesoft– Waterloo, CA)

An overview of the new features in Maple 2022 will be given, including formal power series, step-by-step solutions, intersection multiplicities, print layout mode, signal processing, and many more.

3.6 Generalized Möbius-Listing bodies– new models for the sciences

Johan Gielis (NL)

Generalized Möbius-Listing surfaces and bodies (GML) were introduced by Ilia Tavkhelidze [1, 2], building on the idea of Gaspard Monge to understand complex movements as the composition of simple movements. Another motivation is that the solutions of BVP for partial differential equations strongly depend on the topological properties of the domain in which the problem is considered.

A main focus of our joint research has been to classify all possible ways of cutting GML bodies (in analogy with the cutting of Möbius bands) [2, 3]. In general the cutting process yields intertwined and linked bodies or surfaces of complex topology.

We defined the conditions under which the cutting process results in a single surface or body only, displaying the Möbius phenomenon of one-sidedness [4]. At the crossroads of geometry, topology, algebra and number theory, this gives rise to new ways of modeling and understanding certain dynamical processes in the natural sciences.

References

1. I. Tavkhelidze, On the some properties of one class of geometrical figures. *Reports of Enlarged Sessions of the Seminar of I. Vekua Institute of Applied Mathematics* **16(1)**, 35-38, 2001
2. I. Tavkhelidze, C. Cassisa, P.E. Ricci, About connection of generalized Möbius-Listing surfaces with sets of knots and links. *Lecture notes of Seminario Interdisciplinare di Matematica* **IX**, 183-200 2010.
3. J. Gielis and I. Tavkhelidze, The general case of cutting of GML surfaces and bodies. *4OPEN* **3**(7), 2020.
4. J. Gielis, P.E. Ricci, I. Tavkhelidze, The Möbius phenomenon in Generalized Möbius-Listing surfaces and bodies, and Arnold's Cat phenomenon. *Advanced Studies: Euro-Tbilisi Mathematical Journal* **14**(4), 17-35, 2021.

3.7 A modified Waring's problem: decoupling a multivariate polynomial

Mariya Ishteva (KU Leuven– Geel, BE)

Joint work of Mariya Ishteva, Philippe Dreesen
Main reference Philippe Dreesen, Mariya Ishteva, Johan Schoukens: "Decoupling Multivariate Polynomials Using First-Order Information and Tensor Decompositions", SIAM J. Matrix Anal. Appl., Vol. 36(2), pp. 864–879, 2015.
URL http://dx.doi.org/10.1137/140991546

The Waring's problem for polynomials is a fundamental problem in mathematics, concerning the decomposition of a homogeneous multivariate polynomial $f(u_1, \ldots, u_m)$ of degree d as

$$f(u_1, \ldots, u_m) = \sum_{i=1}^{r} w_i (v_1 u_1 + \cdots + v_m u_m)^d,$$

in which r denotes the so-called Waring rank.

We consider a set of non-homogeneous polynomials and show how their joint decomposition can be computed using the canonical polyadic decomposition, which is a well-studied tensor decomposition. We also mention an application in nonlinear system identification.

3.8 Least-squares Multidimensional Exponential Analysis

Ferre Knaepkens (University of Antwerp, BE)

Joint work of Annie Cuyt, Yuan Hou, Ferre Knaepkens, Wen-shin Lee
Main reference Ferre Knaepkens, Annie Cuyt, Wen-Shin Lee, Dirk I. L. de Villiers: "Regular Sparse Array Direction of Arrival Estimation in One Dimension", IEEE Transactions on Antennas and Propagation, Vol. 68(5), pp. 3997–4006, 2020.
URL http://dx.doi.org/10.1109/TAP.2019.2963618
Main reference Annie Cuyt, Yuan Hou, Ferre Knaepkens, Wen-shin Lee: "Sparse Multidimensional Exponential Analysis with an Application to Radar Imaging", SIAM Journal on Scientific Computing, Vol. 42(3), pp. B675–B695, 2020.
URL http://dx.doi.org/10.1137/19M1278004

Exponential analysis consists in extracting the coefficients α_j, $j = 1, \dots, n$, and exponents ϕ_j, $j = 1, \dots, n$, of an exponential model

$$f(x) = \sum_{j=1}^{n} \alpha_j \exp(\phi_j x),$$

from a limited number of observations of the model's behaviour. Since there are $2n$ unkown parameters for n exponential terms, this directly leads to a non-linear square system of $2n$ equations. However, in practice the signal is often perturbed by noise, hence, additional samples are collected and accumulated in an overdetermined non-linear system. Now the question remains how such a noisy overdetermined system behaves and how we can use this information to improve the accuracy of the results. In the case of a square system it is shown in [1] that the exponential analysis problem is deeply intertwined with Padé approximation theory and symmetric tensor decomposition, for both the one-dimensional and multi-dimensional cases. In particular the connection with Padé approximations is very interesting, since it allows the use of Froissart doublets to effectively filter out the noise and correctly estimate the number of terms n. It still remains to be shown that these properties also hold for the least-squares setting of the exponential analysis problem.

Furthermore, we focus on three different application domains, ranging from only one dimension to three-dimensional problems, each with its own challenges. First up is one-dimensional direction of arrival estimation, then image denoising of structured images and finally inverse synthetic aperture radar. We combine sub-Nyquist sampling, a validation technique based on Froissart doublets and new matrix pencil methods in order to tackle these challenging engineering applications.

References

1 F. Knaepkens, A. Cuyt, and W.-s. Lee, From Exponential Analysis to Padé Approximation and Tensor Decomposition, in One and More Dimensions. In *Computer Algebra in Scientific Computing*, G. Vladimir, K. Wolfram, S. Werner, and V. Evgenii (eds.), 116–130. Lille, 2018.

3.9 Advanced time-resolved imaging techniques and analysis and their applications in life sciences

David Li (The University of Strathclyde– Glasgow, GB)

Main reference Dong Xiao, Zhenya Zang, Natakorn Sapermsap, Quan Wang, Wujun Xie, Yu Chen, David Day Uei Li: "Dynamic fluorescence lifetime sensing with CMOS single-photon avalanche diode arrays and deep learning processors", Biomed. Opt. Express, Vol. 12(6), pp. 3450–3462, Optica Publishing Group, 2021.
URL http://dx.doi.org/10.1364/BOE.425663

Advanced time-resolved imaging techniques can reveal biological processes at the molecular level. They can unravel mechanisms underlying diseases and facilitate drug development. However, we also enter the low light regime (only a few photons are acquired), and new acquisition and analysis approaches are sought after.

3.10 Solving a class of singular integral equations using rational approximation

Ana C. Matos (Lille I University, FR)

Joint work of Ana C. Matos, Bernhard Beckermann

We are interested in computing the unknown density of an equilibrium problem in logarithmic potential theory, where the support of the equilibrium measure is a finite union of distinct intervals. We will show that this problem is equivalent to solving a system of singular integral equations with Cauchy kernels. After fixing the functional spaces where we search for the solution , we obtain a theorem of existence and unicity. We then develop a general framework of a spectral method to compute an approximate solution , giving a complete error analysis. We will consider polynomial [1] and rational approximations [2, 3], showing the advantage of using rational interpolation when the intervals are close. Inspired by the third Zolotareff problem, the poles and the interpolation points are chosen in such a way that we can ensure small errors. Some numerical examples showing the good approximation results will be given.

References

1 A. Bultheel, P. Gonzalez Vera, J. Van Deun, On computing Rational Gauss-Chebyshev Quadrature Formulas, *Mathematics of Computation*, **Vol. 75**253, 307- 326, 2006.
2 A. Bultheel, P. Gonzalez Vera, E. Hendriksen and O. Njastad, *Orthogonal rational functions*, Cambridge Monographs on Applied and Computational Mathematics, vol. 5, Cambridge University Press, 1999.
3 S. Olver, Computation of equilibrium measures, *Journal of Approximation Theory* **163**, 1185-1207, 2011.

3.11 Function approximation and machine learning

Hrushikesh N. Mhaskar (Claremont Graduate University, US)

Main reference Hrushikesh N. Mhaskar: "A direct approach for function approximation on data defined manifolds", Neural Networks, Vol. 132, pp. 253–268, 2020.
URL http://dx.doi.org/10.1016/j.neunet.2020.08.018

Two of the fundamental problems of machine learning are the following: (1) Given random samples from an unknown probability distribution, estimate the probability measure, (2) Given samples $\{(y_j, z_j)\}$ from an unknown probability distribution, estimate a functional relationship between z_j and y_j. We explain why classical approximation theory as it was during our childhood is not adequate to solve the problem, and explain our efforts to use ideas of approximation theory to give a direct solution to the problem of estimating the functional relationship. We demonstrate that the problem of density estimation can be considered as a dual problem.

3.12 From ESPRIT to ESPIRA: Estimation of Signal Parameters by Iterative Rational Approximation

Gerlind Plonka-Hoch (Universität Göttingen, DE)

Joint work of Nadiia Derevianko, Gerlind Plonka, Markus Petz
Main reference Nadiia Derevianko, Gerlind Plonka, Markus Petz: "From ESPRIT to ESPIRA: estimation of signal parameters by iterative rational approximation", IMA Journal of Numerical Analysis, 2022.
URL http://dx.doi.org/10.1093/imanum/drab108

We consider exponential sums of the form

$$f(t) = \sum_{j=1}^{M} \gamma_j \, \mathrm{e}^{\phi_j t} = \sum_{j=1}^{M} \gamma_j \, z_j^t,$$

where $M \in \mathbb{N}$, $\gamma_j \in \mathbb{C} \setminus \{0\}$, and $z_j = \mathrm{e}^{\phi_j} \in \mathbb{C} \setminus \{0\}$ with $\phi_j \in \mathbb{C}$ are pairwise distinct. The recovery of such exponential sums from a finite set of possibly corrupted signal samples plays an important role in many signal processing applications, see e.g. in phase retrieval, signal approximation, sparse deconvolution in nondestructive testing, model reduction in system theory, direction of arrival estimation, exponential data fitting, or reconstruction of signals with finite rate of innovation.

Often, the exponential sums occur as Fourier transforms or higher order moments of discrete measures (or streams of Diracs) of the form $\sum_{j=1}^{M} \gamma_j \, \delta(\cdot - T_j)$ with $T_j \in \mathbb{R}$, which leads to the special case that $\phi_j = \mathrm{i}T_j$ is purely complex, i.e., $|z_j| = 1$.

We introduce a new method for **E**stimation of **S**ignal **P**arameters based on **I**terative **R**ational **A**pproximation (ESPIRA) for sparse exponential sums. Our algorithm uses the AAA algorithm for rational approximation of the discrete Fourier transform of the given equidistant signal values. We show that ESPIRA can be interpreted as a matrix pencil method applied to Loewner matrices. These Loewner matrices are closely connected with the Hankel matrices which are usually employed for recovery of sparse exponential sums. Due to the construction of the Loewner matrices via an adaptive selection of index sets, the matrix pencil method is stabilized. ESPIRA achieves similar recovery results for exact data as ESPRIT and the matrix pencil method (MPM) but with less computational effort. Moreover, ESPIRA strongly outperforms ESPRIT and MPM for noisy data and for signal approximation by short exponential sums.

3.13 Interpretable Approximation of High-Dimensional Data based on ANOVA Decomposition

Daniel Potts (TU Chemnitz, DE)

Joint work of Daniel Potts, Michael Schmischke
Main reference Daniel Potts, Michael Schmischke: "Approximation of High-Dimensional Periodic Functions with Fourier-Based Methods", SIAM J. Numer. Anal., Vol. 59(5), pp. 2393–2429, 2021.
URL http://dx.doi.org/10.1137/20M1354921
Main reference Daniel Potts, Michael Schmischke: "Interpretable Approximation of High-Dimensional Data", SIAM J. Math. Data Sci., Vol. 3(4), pp. 1301–1323, 2021.
URL http://dx.doi.org/10.1137/21M1407707

We consider fast Fourier based methods for the approximation of high-dimensional multivariate functions. Our aim is to learn the support of the Fourier coefficients in the frequency domain, where only function values of scattered data available. Based on the fast Fourier transform for nonequispaced data (NFFT) we will use the ANOVA (analysis of variance) decomposition in combination with the sensitivity analysis in order to obtain interpretable results. We will couple truncated ANOVA decompositions with the NFFT and compare the new method to other approaches on publicly available benchmark datasets.

References

1 D. Potts and M. Schmischke, Approximation of high-dimensional periodic functions with Fourier-based methods. *SIAM J. Numer. Anal.* 59, 2393-2429, arXiv:1907.11412, 2021
2 D. Potts and M. Schmischke, Interpretable Approximation of High-Dimensional Data. *SIAM J. Math. Data. Sci.* 3, 1301-1323, arXiv:2103.13787, 2021
3 F. Bartel, D. Potts, and M. Schmischke, Grouped Transformations in High-Dimensional Explainable ANOVA Approximation. SIAM J. Sci. Comput., arXiv:2010.10199 2022

3.14 Detection of directional higher order jump discontinuities with shearlets

Jürgen Prestin (Universität zu Lübeck, DE)

Joint work of Kevin Schober, Jürgen Prestin, Serhii A. Stasyuk
Main reference Kevin Schober, Jürgen Prestin, Serhii A. Stasyuk: "Edge detection with trigonometric polynomial shearlets", Adv. Comput. Math., Vol. 47(1), p. 17, 2021.
URL http://dx.doi.org/10.1007/s10444-020-09838-3

In this talk we consider trigonometric polynomial shearlets which are special cases of directional de la Vallée Poussin type wavelets to detect singularities along curves of periodic bivariate functions. This generalises the one-dimensional case discussed in [1], where singularities occur in single points only.

Here we provide sharp lower and upper estimates for the magnitude of inner products of the shearlets with the given function. The size of these shearlet coefficients depends not only on the distance to the curve singularity, but also on the direction of the singularity.

In the proofs we use orientation-dependent localization properties of trigonometric polynomial shearlets in the time and frequency domain, cf. [1]. We also discuss jump discontinuities in higher order directional derivatives along edges, see [2].

References

1 H. N. Mhaskar and J. Prestin, On the detection of singularities of a periodic function. *Adv. Comput. Math.* **12**(2-3), 95–131, 2000.

2 K. Schober, J. Prestin and S. A. Stasyuk, Edge detection with trigonometric polynomial shearlets. *Adv. Comput. Math.* **47**(1), Paper No. 17, 41 p., 2021.

3 K. Schober and J. Prestin, Analysis of directional higher order jump discontinuities with trigonometric shearlets. *Math. Found. Comput.* Online First, 2021.

3.15 Multi-exponential Functions and Continued Fractions

Tomas Sauer (Universität Passau, DE)

Main reference Tomas Sauer: "Prony's method in several variables: Symbolic solutions by universal interpolation", J. Symb. Comput., Vol. 84, pp. 95–112, 2018.
URL http://dx.doi.org/10.1016/j.jsc.2017.03.006

The recovery of multi-exponential functions of the form

$$f(x) = \sum_{k=1}^{r} f_k \, \rho_k^x,$$

from integer samples is closely related to the theory of continued fractions. The connection is, among others, made by the Hankel matrices

$$H_n := \big(f(j+k) : j,k = 0, \dots, n \big)$$

and the rational best approximants to the Laurent series

$$\lambda(x) = \sum_{k=0}^{\infty} f(k) \, x^{-k-1}, \qquad x \neq 0.$$

The talk highlights some of these connections and also discusses some possibilities for extensions to several variables.

References

1 T. Sauer, *Continued Fractions and Signal Processing.* Springer, 2021.

3.16 An Inverse Problems Framework for Magnetic Resonance Relaxometry and Macromolecular Mapping

Richard G. Spencer (US)

Joint work of Richard G. Spencer, Chuan Bi
Main reference Richard G. Spencer, Chuan Bi: "A Tutorial Introduction to Inverse Problems in Magnetic Resonance", NMR in Biomedicine, Vol. 33(12), p. e4315, 2020.
URL http://dx.doi.org/10.1002/nbm.4315

The success of conventional magnetic resonance imaging (MRI) is partly attributable to the fact that it is a Fourier technique. Data is collected in the space reciprocal to spatial coordinates, known as k-space, and then (inverse) Fourier transformed to produce an image. This reconstruction has the very attractive property of being mathematically well-conditioned, with condition number of unity, so that noise in the acquired data is necessarily transmitted to the image domain but is not magnified. Because of this, early studies performed at low magnetic field and with relatively unsophisticated radio frequency technology were able to yield useful images. Correspondingly, the quality of conventional MRI has increased roughly in proportion to improvements in acquisition SNR. However, the newer technique of MR relaxometry is not so fortunate; the reconstruction of acquired data to obtain parameter distribution functions is via a version of the inverse Laplace transform. This arises from the classically ill-posed problem of inverting the Fredholm equation of the first kind. One implication is that noise is amplified in the reconstruction process, so that brute force efforts to improve SNR rapidly reach the point of diminishing returns, and other means must be undertaken to produce useful results. For this, the inverse problems perspective has proven to be enormously fruitful. We will discuss some aspects of this formalism applicable to MR relaxometry and related experiments. Our main application is to macromolecular mapping, particularly myelin mapping in the brain, and we will show how more accurate myelin quantification permits physiological correlations to be established. Our studies have the twofold goal of improving the capacity of MR to diagnose pathology and monitor disease progression, and of developing methods of general use for inverse problems.

3.17 Super-Resolution on the Two-Dimensional Unit Sphere

Hanna Veselovska (TU München, DE)

Joint work of Frank Filbir, Kristof Schröder, Anna Veselovska
Main reference Frank Filbir, Kristof Schröder, Anna Veselovska: "Recovery of Atomic Measures on the Unit Sphere", Numerical Functional Analysis and Optimization, Vol. 43(7), pp. 755–795, Taylor & Francis, 2022.
URL http://dx.doi.org/10.1080/01630563.2022.2052319

In this talk, we discuss the problem of recovering an atomic measure on the unit 2-sphere $\mathbb{S}^2$ given finitely many moments with respect to spherical harmonics. Our analysis relies on the formulation of this problem as an optimization problem on the space of bounded Borel measures on $\mathbb{S}^2$ as it was considered by Y. de Castro & F. Gamboa [1] and E. Candés & C. Fernandez-Granda [2]. We construct a dual certificate using a kernel given in an explicit form and make a concrete analysis of the interpolation problem, which provides theoretical guarantees for the recovery of atomic measure on $\mathbb{S}^2$. Next to that, we analyze such a problem from a numerical perspective in an extensive series of experiments, using a semidefinite formulation of the optimization problem and its discretized counterpart.

This is a joint work with Frank Filbir and Kristof Schröder.

References

1 De Castro, Y., & Gamboa, F., Exact reconstruction using Beurling minimal extrapolation. *Journal of Mathematical Analysis and applications*, **395**(1), 336–354, 2012.

2 Candés E.J., Fernandez-Granda C. , Super-Resolution from Noisy Data. *J. Fourier Anal. Appl.*, **19**, 1229–1254, 2013.

3 Filbir, F., Schröder, K., & Veselovska, A., Recovery of Atomic Measures on the Unit Sphere. *Numerical Functional Analysis and Optimization*, 1–41, 2022.

3.18 Antenna Position Estimation through Sub-Sampled Exponential Analysis of Signals in the Near-Field

Rina-Mari Weideman (University of Stellenbosch, ZA)

Joint work of Ridalise Louw, Ferre Knaepkens, Annie Cuyt, Wen-shin Lee, Stefan J. Wijnholds, Dirk I. L. de Villiers, Rina-Mari Weideman
Main reference Ridalise Louw, Ferre Knaepkens, Annie Cuyt, Wen-shin Lee, Stefan J. Wijnholds, Dirk I. L. de Villiers, Rina-Mari Weideman: “Antenna Position Estimation Through Subsampled Exponential Analysis of Signals in the Near Field”, in URSI Radio Science Letters, vol. 3, 2021.
URL https://doi.org/10.46620/21-0062

Antenna position estimation is an important problem in large irregular arrays where the positions might not be known very accurately from the start. We present a method using harmonically related signals transmitted from an Unmanned Aerial Vehicle (UAV), with the added advantage that the UAV can be in the near-field of the receiving antenna array. The received signal samples at a chosen reference antenna element are compared to those at every other element in the array in order to find its position. We show that the method delivers excellent results using ideal synthetic data with added noise. Furthermore, we also simulate the problem in a full-wave solver. Although the results are less accurate than when synthetic data are used, due to the effects of mutual coupling, the method still performs well, with errors smaller than 4% of the smallest transmitted wavelength. Finally, we show how our method can detect whether the cables of two antennas were accidentally switched, and how a simple mutual coupling calibration method can improve the results even further.

Participants

- Dmitry Batenkov
Tel Aviv University, IL
- Bernhard Beckermann
University of Lille, FR
- Robert Beinert
TU Berlin, DE
- Wen-Yang Chu
Vitronomy, DE
- Annie Cuyt
University of Antwerp, BE
- Dirk de Villiers
University of Stellenbosch, ZA
- Jürgen Gerhard
Maplesoft – Waterloo, CA
- Johan Gielis
Genicap, NL
- Mark Giesbrecht
University of Waterloo, CA
- Mariya Ishteva
KU Leuven– Geel, BE
- Ferre Knaepkens
University of Antwerp, BE
- George Labahn
University of Waterloo, CA
- Wen-shin Lee
University of Stirling, GB
- David Li
The University of Strathclyde– Glasgow, GB
- Ridalise Louw
University of Stellenbosch, ZA
- Ana C. Matos
University of Lille, FR
- Hrushikesh N. Mhaskar
Claremont Graduate University, US
- Jesús Ortega Almirón
Siemens PLM Software, BE
- Gerlind Plonka-Hoch
Universität Göttingen, DE
- Daniel Potts
TU Chemnitz, DE
- Jürgen Prestin
Universität zu Lübeck, DE
- Tomas Sauer
Universität Passau, DE
- Richard G. Spencer
National Institute on Aging, US
- Hanna Veselovska
TU München, DE
- Rina-Mari Weideman
University of Stellenbosch, ZA

Report from Dagstuhl Seminar 22222

Radical Innovation and Design for Connected and Automated Vehicles

Wendy Ju*[1], Bastian Pfleging*[2], and Andreas Riener*[3]

1 **Cornell Tech – New York, US.** wendyju@wendyju.com
2 **TU Bergakademie Freiberg, DE.** bastian.pfleging@informatik.tu-freiberg.de
3 **TH Ingolstadt, DE.** andreas.riener@thi.de

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 22222 "Radical Innovation and Design for Connected and Automated Vehicles".

Automated driving will most likely be the next big change in individual mobility. While research is still primarily concerned with technical challenges of Automated Driving Systems (ADS), HCI researchers and designers already started to develop concepts on how to use privately owned ADS as a space for non-driving-related activities, going beyond what is possible today. There is, however, room to think about creative ways to use automated vehicles (AV) and connected technology towards the public interest beyond incremental changes, which is what we addressed in this seminar. We challenged the current generative/evaluative research approach for automated driving systems against a radical innovations attempt and questioned whether the current incremental research approach is appropriate for the development of future vehicles. As an integral part of the seminar we wanted participants to learn from each other and to disseminate each individual's experience for boosting subsequent research by trying out different methods that support the "out of the box" thinking (e.g., brainwriting, bodystorming, focus groups, World Café, amongst others).

Seminar May 29–June 3, 2022 – http://www.dagstuhl.de/22222
2012 ACM Subject Classification Human-centered computing → Accessibility; Human-centered computing → HCI design and evaluation methods
Keywords and phrases Racidal Innovation, Automated Driving, Future Mobility, Sustainability
Digital Object Identifier 10.4230/DagRep.12.5.188

1 Executive Summary

Andreas Riener (TH Ingolstadt, DE)
Wendy Ju (Cornell Tech – New York, US)
Bastian Pfleging (TU Bergakademie Freiberg, DE)

Automated driving will most likely be the next big change in individual mobility. Currently, research is still primarily concerned with technical challenges of Automated Driving Systems (ADS), such as sensing, data processing, communication, or steering. HCI researchers have already started to develop concepts for working, relaxing, or recreational activities (in privately owned) ADS as rather incremental innovations. There is, however, room to think about creative ways to use automated vehicles (AV) and connected technology towards the public interest beyond incremental changes. Current open questions are, amongst others:

* Editor / Organizer

Dagstuhl Seminar 22222: Draft Program

7:30-8:45 Breakfast				
Introduction	Design Methods II	Recap Tuesday	Prototyping	Plotting outcomes II
Mid-morning break				
Radical Innovation Panel	Novel UsersPanel	Prototyping	Plotting outcomes	Summary
12:15-13:45 lunch				
Design Methods I	Propose Radical Innovation Initiatives	Group photo	Drafting documents, videos, prototypes	
15:30-16:00 coffee/cake				
Walk and Talk: Mobility of the Future	World Cafe	Hiking	Dagstuhl homework	
18:00-19:00 dinner		Hiking dinner	18:00-19:00 dinner	
Movie Night			Q&D video screening	

Figure 1 Schedule of Dagstuhl seminar 22222 (May 29 to June 3, 2022).

- What are possible innovative and groundbreaking visions for future human transport concepts?
- How can we apply radical innovation and design and leverage AV technology to other applications and use cases?
- What types of research and design methods are able to contribute to radical designs, instead of incremental?

This Dagstuhl Seminar aimed to contribute to continued research that is able to challenge the current generative/evaluative research approach for automated driving systems against a radical innovations attempt. Similar to the invention of the car, which was not just an incremental improvement of a horse carriage, we want to question whether the current incremental research approach is appropriate and want to provoke novel opportunities. We further wanted to challenge whether a human-centered design approach is appropriate in the domain of connected, automated vehicles or if related innovation-centered approaches like design thinking are more constructive.

An integral part of the seminar was to reach the next level of product quality, to learn from each other, and to disseminate individual's experience for boosting subsequent research. In order to give the seminar participants something they can use directly from the seminar, we dedicated two blocks of time for teaching methodological knowledge in addition to the creative sessions, interactive panels and practical work. The goal was to give the participants the opportunity to try out new methods, such as brainwriting, bodystorming, focus groups, user enactment, WorldCafe, amongst others, that they would like to use in their own research in the future – thus a directly applicable result of the seminar. Details on the schedule of the seminar are shown below.

To promote trans- and interdisciplinarity, we invited computer scientists/engineers, interaction designers, UI/UX designers, market and consumer psychologists as well as urban planners from industry and academia to discuss and design future mobility and vehicle concepts. We expected three types of results from this seminar – and will discuss at the end of this report which contributions were created in the seminar and to which degree these results were achieved by the end of the seminar.

- **New visions for transportation and interaction**. The seminar will focus on radical and innovative visions for the future of human transport and will employ innovation-centered approaches to foster out-of-the-box thinking.
- **Novel applications for AV technology**. Perhaps the most significant contribution of the seminar should be novel applications for automated vehicle technology that will stretch the research and science community into new directions. We expect that in the coming 3 to 10 years, these new ideas should serve as inspiration for the research of the seminar attendees and, more broadly, the communities involved in designing automotive and mobility user interfaces.
- **Roadmap(s) for research**. Another goal is the elaboration of a roadmap to outline proposed research collaborations and recommend new funding mechanisms. Furthermore, the roadmap should lay out plans for disseminating results such that members of our community are well informed, and such that they can effectively interact with researchers and practitioners in related communities, such as human-computer interaction, human-factors, user experience, automotive engineering, psychology, and urban/traffic planning.

Perhaps the most tangible and easy-to-communicate outcome of the seminar for other interested parties are short videos/video prototypes of ideas created over the week in group work. All of the videos were professionalized in the aftermath of the seminar and submitted, along with one short paper each, to the ACM AutomotiveUI conference, where three of the videos were accepted in the video track and shown to the audience at the conference in Seoul (videos are available as supplementary material in the ACM DL).

References

1 Nikolas Martelaro, Debargha Dey, Gary Burnett, Helena K. Strömberg, Jonas Andersson, and Andreas Löcken. 2022. How to Manage Social Order in Shared Automated Vehicles. In Adjunct Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '22). Association for Computing Machinery, New York, NY, USA, 201 – 203. https://doi.org/10.1145/3544999.3550154

2 Kai Holländer, Lutz Morich, Natasha Merat, Gary Burnett, Virpi Roto, Wendy Ju, and David Sirkin. 2022. Radical Innovations in Future Mobility and Virtual Assistants. In Adjunct Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '22). Association for Computing Machinery, New York, NY, USA, 209 – 211. https://doi.org/10.1145/3544999.3551501

3 Regina Bernhaupt, Mark Colley, David Goedicke, Alexander Meschtscherjakov, Bastian Pfleging, Andreas Riener, and Shadan Sadeghian. 2022. A Critical Perspective on Radically Innovating Personal Mobility. In Adjunct Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '22). Association for Computing Machinery, New York, NY, USA, 215 – 218. https://doi.org/10.1145/3544999.3551689

2 Table of Contents

3 Overview of Talks

3.1 If and how AI can replace the authority figure of a human driver

Jonas Andersson (RISE Research Institutes of Sweden – Göteborg, SE)

Some of my current research interests include remote monitoring and operation of different types of vehicles and how the system-of-systems of humans, technologies and organisations evolve as the development of autonomous vehicles progress. Currently, most autonomous vehicles require a human safety operator inside the vehicle, and autonomous vehicles without a human "fallback" might be distant. At the same time, having a human safety operator inside autonomous vehicles contradicts the promises of autonomous driving transport efficiency. To bridge this gap, stakeholders are exploring remote monitoring and control. But remote operation comes with challenges, both from technical and human behaviour perspectives. The challenges entail both remote management and control per se, as well as the social impacts of removing the human driver since the human driver can also have a social function. This is true for public transport in particular. As the human authority in the vehicle is removed, there is a possibility to explore if and how AI can replace the authority figure of a human driver. At the seminar, deep discussions on human-automation interaction paradigms and conceptual solutions to the interaction design challenges we face with introducing this new technology. Social aspects of interactions were also discussed in-depth in different interactive sessions. For example, we explored possible futures people might meet in autonomous pods using video prototyping and enactment methodologies. The question explored was: What happens when drivers are removed, and may it be necessary for the intelligent vehicle to display authority, and how can that be done? Potential solutions were also developed, prototyped and played out in a characteristic setting of daily travel where we wanted to extend the perception of autonomous vehicles as they are often portrayed as always clean and having well-behaved travellers. To reflect real-life public transport, a situation where fellow passengers can be an annoyance was chosen as the use case. In this scenario, we wanted to highlight what might happen when the authority figure in the vehicle (bus or tram driver) is removed and illustrating it in a short video is an effective way to disseminate results accessible to academics, industry professionals and the public alike.

The experience I take with me from the week at Dagstuhl is that of a warm and welcoming community with an immense amount of knowledge, understanding and ideas about the future of human mobility. My motivation for joining the Dagstuhl seminar was to experience the opportunity of having in-depth discussions without the time constraints typically found in meetings and conferences and that definitely came true. Having prior experience in remote operation human factors research, the seminar has given me the inspiration to broaden my design methodological approach in this field towards a more explorative stance. The seminar has also expanded and enriched my professional network and opened many new possibilities for future collaborations.

3.2 The Metaverse of Driving

Gary Burnett (University of Nottingham, GB)

My research interests focus around the Human Factors issues for future vehicles, which has increasingly considered the complex design challenges associated with higher levels of autonomy in cars and other forms of personal and shared mobility. More recently, I have also become interested in the huge potential of social VR (and associated Metaverse technologies), primarily as a "radical" research method, but also as a means to engage with my students in Higher Education teaching. In this year's Dagstuhl seminar, conducted in the current hiatus in the Covid-19 pandemic, there are obvious reasons to consider how we can conduct human-centred research in such a way that our participants are physically remote from us (and each other) but still present and fully engaged in a space – collecting both qualitative and quantitative data. As always, I have thoroughly enjoyed my Dagstuhl experience, reuniting with old friends and forming new collaborations. A particular highlight was the video prototyping session where our team articulated a key research question for future shared autonomous vehicle mobility, and envisaged several interesting and novel scenarios and solutions all in a matter of hours and documented within a 2-minute video. This was a highly creative and social exercise – something I'd forgotten how much I missed during the last 2 years of disruption. I also was reminded of the huge importance of in-person incidental conversations that occurred outside of scheduled work (usually while eating/drinking). In a specific example, during an evening meal I realized that everyone on my table was extremely knowledgeable in the area of external human-machine interfaces for autonomous vehicles (some being world experts). This was significant to me personally, as my team in Nottingham is currently designing a field study on this particular topic. In summary, the Automotive User-Interface community will increasingly be looking to the Metaverse as inspiration for future autonomous vehicle research. Moreover, in these uncertain times, the massive value of extended in-person seminars has been reinforced to me in this last week. I am looking forward to attending more seminars on autonomous vehicles in future years.

3.3 The Role of Communication and Computational Advances Towards Accessible Automotive User Interfaces

Mark Colley (Universität Ulm, DE)

Main reference Mark Colley, Elvedin Bajrovic, Enrico Rukzio: "Effects of Pedestrian Behavior, Time Pressure, and Repeated Exposure on Crossing Decisions in Front of Automated Vehicles Equipped with External Communication", in Proc. of the CHI '22: CHI Conference on Human Factors in Computing Systems, New Orleans, LA, USA, 29 April 2022 – 5 May 2022, pp. 367:1–367:11, ACM, 2022.
URL https://doi.org/10.1145/3491102.3517571
Main reference Kai Holländer, Mark Colley, Enrico Rukzio, Andreas Butz: "A Taxonomy of Vulnerable Road Users for HCI Based On A Systematic Literature Review", in Proc. of the CHI '21: CHI Conference on Human Factors in Computing Systems, Virtual Event / Yokohama, Japan, May 8-13, 2021, pp. 158:1–158:13, ACM, 2021.
URL https://doi.org/10.1145/3411764.3445480

Mobility is an important aspect and a facilitator of industry, academia, and general life quality. Automation efforts offer possibilities toward economic growth and increased quality of life for all people. However, mobility also contributes to phenomena such as global warming due to its reliance on fossil fuels. Additionally, current mobility often treats accessibility as an afterthought. Automation, and with it the fundamental changes in mobility, are, therefore, transformative and must thus be carefully planned, designed, and evaluated. For example, automated vehicles propose novel research and design challenges in the social-technical systems due to their encounters with vulnerable road users. This has led to the formation of a novel subfield of research towards this communication (e.g., [1]) and also regarding the definition of novel taxonomies more appropriate to this context [2]. While automation is often viewed with a singular focus on improved user experience of the user, potentially enhanced via novel possibilities enabled via machine learning, the challenges currently faced could require a fundamental reshaping of mobility. One crucial aspect for equitable access is including accessibility from the start. This intersection, future mobility supported by novel computation advances to strongly support accessibility needs, is at the core of my work.

Therefore, this Dagstuhl seminar greatly helped understanding the numerous stakeholder perspectives and the various challenges – technical, political, design-related, human-factor-related. Based on the knowledge, experience, and vision of the participants, the great challenges of mobility were outlined – which will shape the research agenda for the future years.

References

1 Colley, Mark, Elvedin Bajrovic, and Enrico Rukzio. "Effects of pedestrian behavior, time pressure, and repeated exposure on crossing decisions in front of automated vehicles equipped with external communication." In *CHI Conference on Human Factors in Computing Systems*, pp. 1-11. 2022.

2 Holländer, Kai, Mark Colley, Enrico Rukzio, and Andreas Butz. "A taxonomy of vulnerable road users for hci based on a systematic literature review." In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1-13. 2021.

Contribution to Novel Users Panel

Moderated by Prof. Riener, a panel consisting of Regina Bernhaupt, Mark Colley, Andreas Löcken, and Nikolas Martelaro discussed the possibilities of future users of automated mobility. In this panel, questions regarding current challenges and possible avenues to a more equitable future were discussed. The questions were especially:

- What new user groups can automated driving open up?
- What would be prerequisites for a successful integration into future transport?
- What accessibility problems are currently present in private/public transport?
- What would be a radical innovation for blind people in future mobility?
- How to ensure all people have access to future mobility?

The panel and the attendees identified interesting avenues for future methodical, educational, and human-factors-related research areas. Especially relevant are the requisites of designing not for a "standard" user but for a user range. The according tools and evaluation protocols for such designs were proposed as a highly relevant research direction. This includes software-defined vehicles both for appropriating hard- as well as software needs.

Schemes for Collaboration

Ulm University with Human-Computer Interaction Institute – Carnegie Mellon University – Nikolas Martelaro
Based on a common research interest in supporting the analysis of driving-related data, the future development of an analysis tool for rich data and metadata in the automotive domain is envisioned. This collaboration could include requirements analysis of the tool, workshops on usability enhancement, and data sharing for the advancement of understanding user behavior both in manual and automated driving.

Ulm University with Centre for Accident Research & Road Safety, Queensland University of Technology, Brisbane, CRICOS 00213J, Australia – Ronald Schroeter
The Centre for Accident Research & Road Safety currently evaluates automated vehicles and their behavioral impact. This includes the recording of numerous variables related to the driving task as well as the driver or user state. Prior to this, a vehicle was developed that is allowed to drive on various roads. The rich information logged by these vehicles can be used for deeper analysis and to guide future research in domains such as computer vision or interaction with automated vehicles.

Ulm University with RISE Research Institutes of Sweden – Jonas Andersson
Until automation capabilities become so advanced that a truly autonomous service is supported, that is, path planning, path execution, interaction with vehicle users, moderation for varying needs, handling of all imaginable situations, and maintenance work planning, some degree of human supervision is necessary. Currently, the arrival of this future is not predictable. Therefore, to still benefit of the advancements in technology and with the pitfalls of supervising automation in mind, remote control is developed and evaluated both from a technical and a Human Factors perspective. In this potential collaboration, remote control interfaces intended for multi-vehicle and/or multi-supervisor should be developed, implemented, and evaluated.

Ulm University with Cornell Tech – Wendy Ju and David Goedicke
Based on the simulators developed by David Goedicke and Wendy Ju at Cornell Tech, the evaluation of numerous aspects regarding the interaction with an (automated) vehicle is possible. For example, a real vehicle equipped with AR glasses can be used to study braking behavior for simulated crossings of pedestrians, the interaction with real forces within the vehicle, or as a tool to use with a Wizard-of-Oz approach for the participants to take the place of the pedestrian. This enables the evaluation and assessment of numerous concepts designed and evaluated in Virtual Reality.

3.4 Radical Innovation and Design for Connected and Automated Vehicles

Regina Bernhaupt (TU Eindhoven – Eindhoven, NE)

When attending a workshop named Radical Innovation and Design for Connected and Automated Vehicles one of the key expectations is to find ways on how to design for connected an automated vehicles that will very likely be available in a not so distant future.

What was impressive in our workshop was the change towards even more critically reflecting on the assumptions we make for such a radical shift. Where would be the revolution – and not only the small evolution.

Reflecting on develpments in socio-technical systems and on their design in the past 30 years, we structured our thoughts on three dimensions: (1) users and groups of users, (2) systems and technology, and (3) the stakeholders or companies. We were able to identify a set of critical aspects and assumptions that need to be discussed and focused on in the near future: The need of negotiation between stakeholders and companies and the users, in terms of overall goals for transportation and mobility, be it monetary or based on other key characteristics of society like health, well-being or sustainability.

For the planning horizon of 15 – 20 years it became clear to me, that this lovely and inspiring group in this workshop is currently designing technology that will be key in my early retirement years. This led to a shift in my reflections and focus: while a 15 to 20 year horizon is key for a researcher mid-career, my research contribution should be more on the evolution and the more immediate improvement as well as the lay-out of a framework that can guide today's discussion.

Based on a number of brainstorming sessions and group work, I came up with a triangle depicting user, system and stakeholders and allowing to group the various discussion points and possible conflicting requirements in one central figure. It can be a basis to summarize the contributions and might lead to a better communication for researchers with varying research backgrounds (as experienced in the Dagstuhl seminar).

Regina Bernhaupt
Professor for Digital-Physical Data Interaction (Systemic Change)
Industrial Design
Eindhoven University of Technology

3.5 From interaction and interfaces to a holistic end-to-end experience

Debargha Dey (TU Eindhoven, NL)

It is interesting to observe how the core focus of our discussion is gradually shifting from designing for interactions and interfaces of the personal automobile to designing for a holistic end-to-end experience of smart and sustainable mobility. The future of mobility points increasingly in the direction of shared mobility and the gradual phasing out of the personal car. Public transport offer several benefits such as sustainability, better economic viability, and reduced environmental impact. However, this also comes at a cost of aspects such as convenience, flexibility, privacy, and speed. Here, the question arises: to what extent can the "sacrifices" to personal convenience and comfort be addressed and/or mitigated? The benefits of end-to-end, seamless mobility afforded by personal vehicles soon disappear when adapting to the currently fragmented nature of public transport that is often plagued by unpredictability, lack of flexibility, and lack of support for first/last mile conveyance. Ignoring the emotional aspects of car ownership (in terms of attachment to a car, or the thrill and feeling of freedom), I believe that these core practicality aspects of individual car ownership needs to be addressed for us to be even remotely successful in realizing this vision. I see this therefore as an interesting but uphill challenge (problem statement) for us interaction designers, HCI researcher, and policy experts to consider and address. Undoubtedly, the next decade(s) will see some unprecedented innovations in terms of mobility (combination of new powertrains/electric propulsion, connected and automated driving, and MaaS) leading to radical disruptions in the realization of sustainable mobility.

3.6 What could radical innovation for transportation look like?

David Goedicke (Cornell Tech – New York, US)

The activities in the workshop focused on understanding how transportation can be solved in this new environment. A new environment that requires is filled with companies pushing for automation, a public that demands "greener" technologies, and a global political situation that restricts access to traditional energy sources.

Through discussion with the different experts that joined the conference, it became apparent that a new radically new perspective is necessary. The future of individual transport was taken under scrutiny and modern concepts of transport were explored and envisioned (like on-call bus services, ride-sharing concepts etc.).

Even though the perspective was on finding modern solutions, a substantial amount of time was also spent on disassembling techno-optimist solutions to the current transportation problem. It became evident that car manufacturers are having difficulty changing their strategy. Modern idealized transportation concepts are often present in a simplified view and, even worse, often ignore marginalized uses of private transport (Home health workers).

The activities in the seminar, from walks to movie nights to board games, created an overall collegial environment and even a playing field where everyone contributed. The fact that multiple videos produced during this week are planned to be submitted to the AutoUI conference speaks to the progress made this week.

Personal note

As a finishing Ph.D. candidate, participation in such an event was most useful to better understand European academic systems, plot the next steps, and get exposed to higher-level planning processes that are often part of running a research lab.

3.7 Radical perspectives on the future of mobility

Kai Holländer (LMU München, DE)

In this work we show novel approaches to fully automated driving.

On the mechanical side, we propose to build modular cars which are split in multiple subsystems and allow for a multitude of different combinations to fit the user, usage context, application area and environment. In addition, we present the idea of a digital agent which helps travelers to decide on means of transportation in accordance to their individual preferences, such as using green modes of transport, their budget, supporting privacy during the journey or travelling as fast or convenient as possible. The focus of the agent lies on seamless interaction of different modes of transport where cars driven on fossil fuels get excluded as much as possible.

We believe these ideas form a base for future research in the mobility domain.

3.8 Transforming Experiences in Transportation

Wendy Ju (Cornell Tech – New York, US)

The key insight I gained from the Dagstuhl workshop on Radical Innovation in Connected and Autonomous Vehicles is that we need to re-invent the experience people have with transportation. Currently, people are solving a lot of their problems with the "silver bullet" of their personally owned car. They might appreciate that this solution comes at some cost – to the environment, to their health, to their everyday routines – but also it allows them to get to work, to explore the world, to meet with people, and to enjoy their lives.

I believe it is possible to help transition people to finding new ways to solve their problems, so that they might reduce or remove their vehicle usage, to the point that they might not need personally owned vehicles at all. To do this, though, people need information about the myriad options for transportation and confidence about their ability to adopt new modes of mobility. To this end, I believe that we should consider ways to improve self-efficacy in mobility. Perhaps, for example, people need a travel concierge to help them brainstorm new ways to travel, and help them through the "but but but"s that keep them from trying and then adopting new travel. Key aspects of this need is to provide people with specific solutions to solve their specific issues – such as needing to transport children, or needing to arrive quickly or cheaply, being able to fit a meal or a workout.

Over the course of the workshop, we learned many ways of exploring and expanding ideas – design thinking, user enactments, remote ethnography, video prototyping, collecting field data. These skills were put to use in the development of an idea for a virtual agent to help people plan travel without personal cars. The super group of Natasha Merat, David Sirkin, rOnny Schroeter, Lutz Morich, Virpi Roto, Kai Hollander, and myself made a humorous video prototype of a virtual agent that aids future mobility for a person with disabilities who is planning a bank heist. I am excited to see the way that this group will build on the skills and methods learned in this week to tackle the great challenges that lie ahead in transforming people's experience of mobility.

3.9 Are We Ready to Take Away People's Cars?

Andreas Löcken (TH Ingolstadt, DE)

Current challenges such as climate or demographic change point to a radically different future. To maintain (or even improve) our standard of living, we need to keep up. And what could be more radical (for automotive human-computer interaction researchers) than removing cars from traffic? Automated shuttle buses or personal "transportation-pods" may replace personally owned vehicles while being more sustainable and enabling more people to participate in traffic. In our own research, we focused a lot on how the automated vehicle should communicate with its passengers (e.g., [1]) or its surroundings (e.g., [2]. But will people accept "giving up" their cars? In my opinion, the most interesting challenges for a radically new future of mobility (that HCI researchers can tackle) will be to create inclusive solutions (and not "just" add accessibility options) and meet user experience needs that manually controlled vehicles will no longer provide. If we cannot give users an alternative to their cars, the industry has no incentive to provide the technology or stakeholders to update policies. The answer to the title is: "probably not". I look forward to continuing the discussion in future workshops and seminars. I thank Wendy, Bastian, Andreas and the Dagstuhl organization for allowing me to make new contacts and deepen existing connections.

References

1 Andreas Löcken, Anna-Katharina Frison, Vanessa Fahn, Dominik Kreppold, Maximilian Götz, and Andreas Riener. 2020. Increasing User Experience and Trust in Automated Vehicles via an Ambient Light Display. In MobileHCI '20: 22nd International Conference on Human-Computer Interaction with Mobile Devices and Services, Oldenburg, Germany, October 5-9, 2020. ACM, 38:1-38:10. DOI: https://doi.org/10.1145/3379503.3403567.

2 Debargha Dey, Azra Habibovic, Andreas Löcken, Philipp Wintersberger, Bastian Pfleging, Andreas Riener, Marieke Martens, and Jacques Terken. 2020. Taming the eHMI jungle: A classification taxonomy to guide, compare, and assess the design principles of automated vehicles' external human-machine interfaces. Transportation Research Interdisciplinary Perspectives 7, 100174. DOI: https://doi.org/10.1016/j.trip.2020.100174.

3.10 Radically personalized mini-mobility automotive design

Nikolas Martelaro (Carnegie Mellon University – Pittsburgh, US)

To help reduce the impact of automobiles on traffic congestion and the environment, I believe there needs to be a shift towards smaller vehicles. To make this more agreeable to people, I also believe that there needs to be ways to make sure that such vehicles meet people's needs. I suggest that this could be done through new design systems that help people to radically personalize their vehicles for their needs.

A current issue with today's automobiles is that they are physically too large. As people begin to shift towards electric vehicles, I believe that we are also starting to further understand what our driving needs actually are and are thus changing what we believe we need in a vehicle. For example, a primary issue that many people have had with electric vehicles is they have too little range for their driving behavior. While this is true for some drivers, it is not true for many drivers, especially those who live in cities, as they may only drive short distances and thus would be ok with smaller vehicles with less range. Moreover, with the convenience of charging at home (for some), range becomes a low priority issue once a certain threshold of range capacity is reached (I hypothesize that this is about 240 KM of range). While new battery technologies that allow more dense energy storage should be explored, I believe that there is not as much of a need for creating physically larger vehicles to meet consumers; transportation needs.

If we can reduce the size of vehicles then, we can also reduce their impact on congestion on the road and while parked. While smaller vehicles are much more common in Europe, in the US, vehicle design is trending larger and larger. Much of this is likely based on consumer psychology as people get a vehicle for the 5% of driving that they do rather than the 95% of driving. For so long, vehicles were affordable enough that people could own a vehicle to express their identity. However, with rising manufacturing costs, vehicles are now going to be more expensive, making it harder for people to own the vehicles that they may desire. One way to combat such rising prices though is to reduce the material costs of a vehicle, for example by making smaller vehicles. Still, such small vehicles will need to have great designers in order to fulfill people's physical needs as well as their psychological and aspirational needs. I believe that one radical further for personal automobiles is in highly personalized, ultra-compact cars that are designed specifically for every person to fit their needs. Such a future would likely require a few key technical innovations that require research and development.

First, EV drivetrains will likely need to be modularized so that the most expensive parts of a vehicle can be mass-produced and interchangeable with different body styles. Such technology is already in development with the concept of EV drivetrain "skateboard" platforms that allow different body styles to be bolted onto the skateboard frame. Second, we will need better ways of determining people's physical needs in regards to range and functionality of a vehicle. While one may assume that people determine their own physical needs when choosing a vehicle, I am not as sure about this and believe that many people consider such needs but in more aspirational or speculative ways. Often, people drive alone on daily commutes and often do not take the long road trip with many people or haul lots of gear to go camping. I believe that there is research to do on how well people vehicle choices meet their physical needs. Such research might include quantified self research around driving habits and ethnographic studies at the scale of people's use of their cars. I hypothesize that many cars are highly underutilized. Based on such research, I also believe that such ethnographic and data collection systems can be used for the design of new kinds of vehicles that do fit people needs. To accomplish such personalized design on a massive scale would require new kinds of design tools and likely computational assistance. Research into generative design systems may lead to new tools that can allow a person paired with a computational "designer" to create a car body/interior design just for them. Such systems could design vehicles for people who do transport many other people, or who go on trips to the woods to ski and camp, or who want to have a stylish vehicle to show off their personal tastes. One can even imagine that as people's needs change over time they can change out their bolt-on bodies for new ones while keeping the same skateboard frame.

Such an idea is not without its potential issues and here I try to describe some that may need to be overcome when considering such a personalized automotive design future. First, what happens when a vehicle does not meet someone's needs any more and must be resold? Will it be so personalized that it does not work for another user? Or will there need to be a service to find similar enough users who could then purchase a customized vehicle from another person? Or, are there ways to make modifications on a customized vehicle that allow someone to keep most of what is there and change out only key areas? Another issue is that of fulfilling driver's needs when they do take the 5% of trips where their vehicle is not a good fit? This might be overcome with some innovation in car sharing / rental business models where people can rent the car they need for the few times a year that they need it. Finally, there is a question as to if more investment should be made into automobile travel or into other forms of transportation. I personally believe that we will need a diverse set of transportation options well into the future, however, such technologies that can help to understand people's needs and personalize their car might also be the foundation for systems that can design personalized mobility plans for people that includes personal vehicles, public transportation, bicycling, and other forms of micromobility.

3.11 Designing the future automated vehicle: Including the excluded

Natasha Merat (University of Leeds, GB)

Background. The vision for implementing autonomous (automated) vehicles (AVs) in our cities has been around since the 1980s. These new forms of mobility are proposed to be safer, more efficient, and more economical, helping to prolong the mobility of older and impaired citizens, making transport more sustainable, and cleaner. Currently, personally-owned versions of these vehicles, which can maintain the designated speed of the road, and keep within the lane, are available for specific road conditions – such as highways, during good daylight conditions, with clear weather, etc. An alternative model to this personal model, is a shared automated vehicle (SAV), with no driver controls, but controlled by a safety driver inside the vehicle – e.g. with a joystick, or remotely, by a human operator. These shared automated vehicles typically operate in designated urban areas, such as airports, carparks or hospital/shopping grounds. Despite the plethora of effort and $$$ spent on the development of these vehicles in the past 40+ years, a number of key barriers currently exist, which prevent their mass deployment, acceptance and adoption. These barriers are partly based on inadequate technological advancements, such as imperfect cameras, sensors and radars, which are not yet 100% reliable in all road and weather conditions, and are particularly challenging in rural settings, or a mixed traffic environment containing other (manually driven) vehicles, as well as cyclists, pedestrians and users of other micromobility services (collectively termed Vulnerable Road Users, or VRUs). In addition to such technological challenges, the "everywhere, all the time" vision for AVs is currently not feasible, because they do not meet the needs of all citizens, with particular challenges for accessibility, inclusion, and affordability.

Motivation. My motivation for attending the Dagstuhl Seminar on "Radical Innovation and Design for Connected and Automated Vehicles", was to have the opportunity to meet colleagues from moderately different research backgrounds, and discuss the challenges which currently exist for the deployment of AVs with a group of experts from Europe and North America, using a diverse range of research methods and philosophies. I was also very excited and honoured to join a meeting structured around scientific debate and discussion, which my busy diary has not allowed me to experience since my PhD days, 25 years ago!

Vision and Methodology. The vision for the meeting was for each group member to outline their areas of interest in this field, defining the problems that need to be addressed before mass deployment of AVs is a possibility, and the timeline for addressing these problems. A wide range of interests, focusing on new design solutions for the user, and outlining the role of society and policy (rather than technology) were identified, and discussed. Three main themes were developed from this diverse range, for further debate amongst smaller groups. These included: 1) A policy push for removal of all personal vehicles from our roads by 2032, and the introduction of shared automated vehicles and alternative forms of transport, 2) Better understanding and coordination between high levels of automation and drivers, and 3) creation of a modular, personalised vehicle of the future. Some of these ideas were then taken further to create short demonstration videos, illustrating the solutions offered for a radical change in the mobility of the future.

Main takeaways. Interacting with an almost completely new team of researchers was enlightening, exciting, and super rewarding. I may have started on a slightly different page, but quickly learned to see the benefits of different approaches and the contribution of different disciplines, learning the value of new and different research tools and methods. Solving problems in a team with peers (rather than students) was also very special, especially after 2 years of lone working, due to COVID-19. Professor Natasha Merat, Chair in Human Factors of Transport Systems. Institute for Transport StudiesUniversity of Leeds

3.12 Radical Innovations for Future Vehicles Needs Radical Methods and a Holistic Understanding of Users, Stakeholders and Society

Alexander Meschtscherjakov (Universität Salzburg, AT)

When debating future mobility concepts there is always a tension between the doable and thinkable. One way how to look at the future is to expand it from the past and think of how the future may evolve. In the context of automated vehicles this is often done through the lens of automation levels. Coming from manually driven vehicles (level 0) lateral and longitudinal control of the vehicle, as well as the necessity to observe and react to the environment are omitted. At level 1 the vehicles accelerate and break by means of ACCs or keeps the lane by utilizing lane-keeping assistance systems. On level higher the vehicle does the driving and the driver "only" needs to observe the environment and react if necessary. On level 3 non-driving related tasks (NDTR) are possible, but the driver needs to react to take-over-requests (TORs). Higher level (i.e. 4 and 5) even relives drivers from any driving activity allowing them (maybe) to intervene in tactical driving decisions (e.g. shall the vehicle overtake or not) and strategic ones (e.g. telling the automated vehicle where to go). Increasing levels allow us to think of different kinds of transport systems. Vehicles evolve and become more and more autonomous. This is one way to do it.

Another way is to come from another point of view: thinking of and exploring solutions by asking "What if" questions. This allows us to re-think the future of mobility more radically. But what are good "What if" questions? How far do we have to go to be radical enough? Whom do we have to involve in this process? This Dagstuhl seminar was a starting point for me to think about radical changes in our premises leaving (hopefully) to novel insights on how the future of mobility could look like. Questions that have been (or will be) tackled are the following:

- What if all public transport would be free of costs for people who use it?
- What if we prohibit people of owning their own vehicle?
- What if we forbid individual cars at all?
- What if we get rid of all streets?
- What if we forbid manually driven vehicles?

By making this “what if” statements to our premise we were forced to come up with radical new solutions but also by looking at problems that may arise for individuals triggered by these radical changes.

3.13 Abstract on Seminar „Radical Innovartion and Design for Connected and Automated Vehicles”

Lutz Morich (frE3-innovations, DE)

As a novice and industry participant at a Dagstuhl seminar, all estimates and impressions are based on an absolute jumping off basis without relative comparison to comparable seminars.
My goals before the seminar were the following:

1. broaden my perspective regarding mobility (narrow due to industrial background)
2. networking with mobility research experts in academia
3. learning methods for innovation and knowledge generation

My main takeaways: My expectations were exceeded by the very diverse group from the science of HCI in combination with mobility issues. My dull glasses of industrial imprinting have wich not only cleared up but also significantly expanded. Through the dynamic composition of working groups in connection with challenging and innovative methods, both my goal of networking and methodological competence enhancement was fulfilled. Connections and perspectives have emerged that promise a closer cooperation of my private economic background with science in the field of automated and networked mobility in the future in a general and concrete way.

Lutz Morich
frE3-Innovations
Josef-Fleischmann-Str. 7
85139 Wettstetten
Tel.: +491605739144
Lutz.Morich@fre3-innovations.com

3.14 Radical Innovation and Design of Future Mobility – Where should we head?

Bastian Pfleging (TU Bergakademie Freiberg, DE)

Human mobility has massively changed throughout the last centuries. Humans had used horses (or similar animals), horse-drawn carriages, or just their own feet to get to a destination for thousands of years. This hugely limited individual mobility and only with the invention of bicycles, trains, cars, busses, and trucks humans gained mobile flexibility and this massively extended their motion range. Vehicle automation, climate change, and novel drive technology are expected to again change to a large extent how we will use and experience mobility.

There is a variety of concepts for automated cars and how mobility could look like in the future (e.g., 2030 or 2050). This includes concepts of automated individual vehicles and shared mobility using various types of vehicles. What many of these concepts have in common with current research and development practices is that they rely on rather incremental and evolutionary steps. An open question is whether this approach is suitable to use or whether radical innovation and design could be more suitable.

Looking at other fields such as communication industry and especially phones, the history of the phone shows that a similar evolutionary approach was used multiple times: operators were replaced by machine switching, wireless phones at home replaced traditional fixed phones (=the user got rid of the phone cord) and later traditional mobile phones ("feature phones") offered mobility to their users as they were able to use their phone not only at a fixed location but everywhere. For quite a while, such phones were improved incrementally to satisfy specific user needs (e.g., longer standby time, better sound quality, ...). The introduction of the smartphone instead can be seen as a more revolutionary step, since prior achievements or principles (e.g., standby time, privacy) were sacrificed for a novel type of device (large touch screen) and an ever-increasing App ecosystem. As history now tells, the gained benefits of smartphones outweighed the user need for week-long standby times and former global players, who sticked mostly to their good old feature phones, became mostly meaningless with regard to phone production. Another interesting aspect in communication industry can be found when looking at today's business models for mobile phone (operating) systems when looking at the two biggest players and ecosystems: What both have in common is that they offer large App ecosystem. One of the biggest difference between them is that one generates revenue from selling a combination of hardware (phone) and operating system, while the other one mostly offers a free operating system (leaving production of hardware to other companies) while generating revenue from collecting and using data with all its pros and cons.

The previous case of the mobile phone sector, where the business has change massively during the last decades, raises a few questions for the automotive field:

- What does it need to prevent current players in (automotive) mobility from becoming the "next Nokia" and instead contribute successfully to novel forms of mobility?
- What are the consequences of different business models and how do they affect society?
- How can we ensure to develop future mobility in an ecological way where users, society, and environment are first and business/revenue are only second?

Taking the experience from the smartphone sector into account, also related to market power, I consider it essential that we do not make the same mistakes again that happened in the phone industry when it comes to (re-)inventing future mobility. Thus, I hope that the discussion and ideation of how our mobility will look like in the next 10, 30, or 50 years happens equally in industry, society, and academia. With the outcomes of this Dagstuhl seminar, I hope that we can contribute to and stimulate this discussion.

I thoroughly enjoyed to be back at Dagstuhl in person and discuss our ideas and perspectives with old and new collaborators. What I found remarkable is that the focus of the seminar shifted very quickly into novel directions such as "what if ... the individual car will/needs to disappear" and it was a pleasure to discuss creative solutions and approaches to radical changes in the mobility sector with all attendees.

3.15 Radical Innovation and Design for Connected and Automated Vehicles

Andreas Riener (TH Ingolstadt, DE)

Joint work of Andreas Riener, Bastian Pfleging, Wendy Ju
Main reference Anna-Katharina Frison, Andreas Riener, Myounghoon Jeon, Bastian Pfleging, Ignacio J. Alvarez: "Workshop on Designing Highly Automated Driving Systems as Radical Innovation", in Proc. of the Adjunct 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2018, Toronto, ON, Canada, September 23-25, 2018, pp. 37–43, ACM, 2018.
URL https://doi.org/10.1145/3239092.3239097

We met for a week in Dagstuhl to discuss radical innovation approaches to change future transport, to identify problems, and to develop solutions. The aspects and dimensions brought in were diverse and the discussions from different perspectives very fruitful.

But we also realized how difficult it is to break out of the "normal" iterative development process and think "radically" or "innovative"; after numerous presentations, discussion rounds, panels, prototyping workshops and, of course, evenings in the wine cellar of Schloß Dagstuhl, we gave birth to great ideas to revolutionize (individual) transport in the future. At the same time, however, we have also realized that there are a number of obstacles to be cleared out on the way to more sustainable mobility. On the one hand, car manufacturers and the automotive lobby are strong advocates of individual cars and it will be difficult to initiate the rethinking process there. On the other hand, political decision-makers also need to be convinced to launch new forms of mobility.

Today, however, at the end of the seminar, we are confident that we will be able to influence future decisions and implement some of the ideas developed in the seminar. More on the short-term, some ideas for bilateral/multilateral collaborations, research exchanges and last but not least, conference papers and grant proposal have emerged.

Thank you for the opportunity to organize this seminar in Dagstuhl and to let all these creative, innovative minds work on problems of the future for one week and to foster personal interaction after a long time of remote meetings.

The Game is NOT Over – we are coming back!

Andreas Riener Co-organizer of Dagstuhl Seminar 22222

References

1 Anna-Katharina Frison, Andreas Riener, Myounghoon Jeon, Bastian Pfleging, and Ignacio Alvarez. 2018. Workshop on Designing Highly Automated Driving Systems as Radical Innovation. In *Adjunct Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (*AutomotiveUI '18*). Association for Computing Machinery, New York, NY, USA, 37 – 43. https://doi.org/10.1145/3239092.3239097

2 Anna-Katharina Frison, Philipp Wintersberger, Andreas Riener, and Clemens Schartmüller. 2017. Driving Hotzenplotz: A Hybrid Interface for Vehicle Control Aiming to Maximize Pleasure in Highway Driving. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (*AutomotiveUI '17*). Association for Computing Machinery, New York, NY, USA, 236 – 244. https://doi.org/10.1145/3122986.3123016

3 Anna-Katharina Frison, Bastian Pfleging, Andreas Riener, Myounghoon Philart Jeon, Ignacio Alvarez, and Wendy Ju. 2017. Workshop on User-Centered Design for Automated Driving Systems. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications Adjunct* (*AutomotiveUI '17*). Association for Computing Machinery, New York, NY, USA, 22 – 27. https://doi.org/10.1145/3131726.3131734

3.16 Experience goal-driven design

Virpi Hannele Roto (Aalto University, FI)

Radical innovation requires powerful innovation methods. Through some years of development and testing, we have developed the concept of experience goal and the method of experience goal -driven design. In this seminar, we experimented using the method on innovating future vehicles. Pairs of participants started by defining experience goals for a given use case, e.g., Eric, 75y, having Covid-31 and needing transportation to hospital. What kind of emotional needs does he have? An experience goal should communicate the experience or psychological need that the person has in the situation. The experience goal should then drive the idea generation: what kind of future vehicle could enable the experience goal to realize? In this case, the group found it is important to support feelings of safety and being nurtured. The social needs include contacting the closest people and informing about the status.

3.17 Individual and social experiential aspects in the design of automated vehicles

Shadan Sadeghian Borojeni (Universität Siegen, DE)

Mobility is both a desire and a necessity. Mobility practices, such as commuting to work, shopping, or, traveling to visit loved ones are an integral part of our everyday lives. Currently, a large proportion of these practices are performed through manually driven vehicles. With the rise of driving automation, however, these practices will change. Driving automation promises to relieve people from driving-related tasks and provide them with the opportunity to engage in non-driving-related activities. To fulfill this, of course, the automated vehicles should ensure safety and reliable maneuvering which requires systems that are able to sense and understand the environment and act flawlessly. This pragmatic aspect of automated driving has been researched in the last decades. It is nevertheless, important to investigate how these automated vehicles change the mobility practices and what experiences emerge through their use. This as well raises the question of whether the new practices are accepted by users and can be integrated into their everyday lives. While the decisions related to the design of technology for automated vehicles are mostly technical, they have social consequences. Creating technology that forces users to follow rules that the designers have planned for them without considering their individual needs or social norms can create a host of negative experiences. Therefore, the design of automated vehicles should consider the individual and social experiential aspects besides the pragmatic and safety-oriented factors.

3.18 Friendly metaphors for human-AV interaction

Ronald Schroeter (Queensland University of Technology – Brisbane, AU)

My interests are to make Automated Vehicles safe and fun to use, so I attended the seminar to be inspired to explore new ideas to that end. Here is what I walked away with: Let's apply a "dog" metaphor to future Automated Vehicle (AV) design! Dogs have an innate connection to human behavior. They are receptive to their owner's cues, sense their owner's feelings or moods (or "states"), adjust or tailor their behaviour accordingly, and even probe their owners for explicit cues or responses when needed (by lifting paws, making sounds, nudging, etc.). As a result, humans develop a deep connection to their dog. How can we create such deep connections in the context of human-AV interaction, and opportunities for mutual learning processes?

We propose a vehicle decision-making framework that incorporates cues obtained from the vehicle user and an adaptive multimodal HMI (human-machine-interface, considering all human senses, so including vehicle dynamics) that probes the user. The framework integrates the user's state (attention, arousal, stress, ...) and the external environment (complexity, planned path & maneuvers) while making operational decisions, and applies active interventions on user states. This bi-directional, adaptive tuning mechanism or "language" between AV and user, may then potentially be used to

- enhance safety (aligning cognitive state with environment requirements),
- improve user experience (aligning with or playfully manipulating emotional states),
- create user profiles that optimise AV behaviour, and
- improve user experience.

3.19 Prototyping Interactions with Autonomous Vehicles in the Field

David Sirkin (Stanford University, US)

Fully or partially autonomous vehicles are a rich but vexing context for interaction designers to study in real road contexts. On one hand, their design draws on established lines of research in physicality and human factors, and in psychology and perception. On the other hand, for now, real autonomous vehicles are difficult to acquire and program, and thus apply to interaction design research questions. One approach is to use Wizard of Oz prototyping, typically hiding a human driver, which can work for participants both inside and outside of the vehicle.

Monday's tutorial focused on collecting field data in the context of Wizard of Oz prototyping of human-vehicle interactions in the field. We covered several examples of prior studies, which led to both the kinds of findings we were looking for1 (about deceleration profiles, pedestrian perceptions) and those we had not expected2 (that some road users "grief" autonomous vehicles). These experiences suggested the value of many hours spent selecting and driving in various contexts to capture a breadth of behaviors. The conversation quickly moved to sharing practical tips during field work, such as the amount of time needed to plan a field study, the daily setup and breakdown of equipment, a suitable number of interviewers for each outing, alternative ways to design questionnaires, and how to increase participation from passers-by, such as offering to walk along with those who are reluctant to stop and talk with interviewers.

My personal goals for the program were to meet human factors experts, computer scientists, engineers, and other designers, share our research methods, and plan future collaborations. While the program had many organized activities, these 2 goals were those where I focused most of my effort and have already found the most value.

References

1 Moore, D., Currano, R., Strack, G. E., & Sirkin, D. 2019. The case for implicit external human-machine interfaces for autonomous vehicles. In Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, 295 – 307.
2 Moore, D., Currano, R., Shanks, M., & Sirkin, D. 2020. Defense against the dark cars: Design principles for griefing of autonomous vehicles. In Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction, 201 – 209.

3.20 Radical mobility innovations for a messy everyday life in the future

Helena K. Strömberg (Chalmers University of Technology – Göteborg, SE)

My research interest is driven by a vision of sustainable everyday mobility, where mobility is something that allows people to go where they need to be and do what they need to do, and where people have equal access to functions in society, healthy environment, and safe, appropriate, and effective ways of getting around. Partial or full automation of vehicles can play a part in facilitating such a future vision, but it is necessary to investigate how and where to best implement automation so that society can gain the most benefits from it – to potentially radically rethink the way the mobility system is built today.

Currently, there is a strong technology push in relation to automation, and a weak societal pull or engagement. There are risks associated with this dynamic, as technology becomes implemented where it is possible instead of where it is beneficial. Research and industry efforts are then spent trying to handle the resulting consequences of this sub-optimal implementation, including overtrust in automation, managing sudden and dangerous handovers and the difficulties in communicating among different road users. There is a need to take a step back and look at the interplay between humans and automated systems from a more comprehensive and multi-stakeholder perspective. Which role(s) in mobility should automation play, and which should be played by humans? During use, which actions should be done by whom? And who should make which decisions and take which responsibilities? Taking a step back will open up the design space to not only create radical technical inventions, but also truly radical innovations that are implemented in a way that benefits both society and the environment.

A radical rethink requires a combination of critical and creative approaches, and multiple perspectives. My motivation to join the seminar is to meet new people (and old acquaintances) to get new ideas for approaches and clash my perspectives against others, and in the long term create new collaborations. I am especially interested methods that allow us to explore an unknown future in such a way that we can explore the consequences for interaction, use and society at large. My previous experiences include enacted and real-world trials of mobility solutions, with the intended users in their intended contexts, as well as more future-oriented and speculative approaches using provocative enactments and prototypes to evoke a discussion on the desired future. It would be interesting to find other ways of exploring future practices, to predict, enact and evaluate that practice in advance, in order to understand if ideas will be useful and valuable in the future.

This seminar has explored multiple ways of doing so. At the same time, the activities and mix of interests represented has pushed us into a simultaneously critical and open mind to the visions of the future available in the world or that we can come up with ourselves. This fruitful exercise has allowed us to find the faults in our thinking, challenge our assumptions and realise who we have included and excluded in the future, Importantly, this leads to the realisation that people will still be people in the future, lives will be messy and any radical mobility solution must cater to many conflicting interests

I come away from the seminar with a renewed interest for research into the experiences and relationships that forms with, within and beyond the vehicle, validating the work I and my research group have investigated previously. I am also empowered by the discussions that more people (in a car dominated research area) feel the need to challenge the car or even get rid of it altogether. However, it is important not to lose track of the important role that personal cars can play – it is, and most likely will be, a part of the diversity of solutions that are a key to successful sustainable everyday mobility in the future.

4 Workshop Activities

4.1 Welcome to Dagstuhl seminar 2222: Opening and Brainstorming

Andreas Riener (TH Ingolstadt, DE), Wendy Ju (Cornell Tech – New York, US), and Bastian Pfleging (TU Bergakademie Freiberg, DE)

In order to collect uninfluenced interesting topics, open problems, and challenges that the participants would like to discuss more deeply during the week, we started the Dagstuhl seminar on Monday morning (first activity, even before the Introduction, Figure 2) with an open brainstorming session.

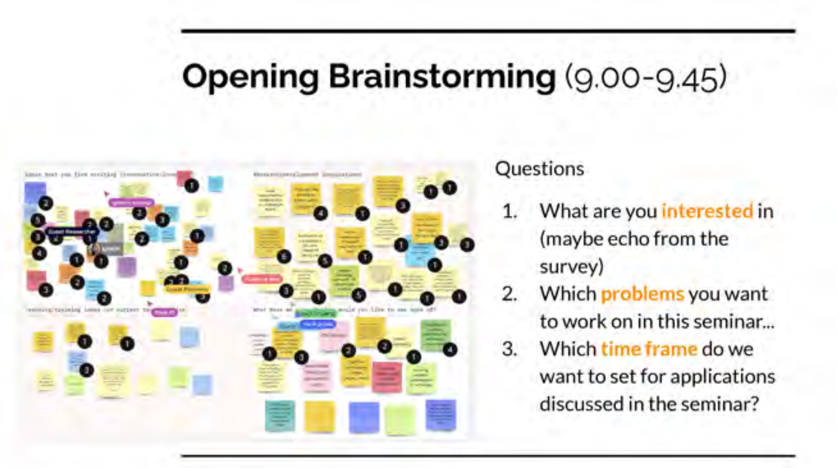

Figure 2 Starting slide for the first activity in the seminar: Open brainstorming.

For this purpose, the participants were invited to write post-its on the 3 topics of

1. What are you interested in?
2. Which problems do you want to work on in this seminar?
3. Which time frame do we want to set for applications discussed in the seminar?

collect them, present them and stick them on a brainstorming wall (see Figure 3). The seminar organizers later clustered them and derived topics for further activities.

After this warm-up activity, the co-organizers opened the seminar, walked through the agenda, and also gave a little bit of history of Schloss Dagstuhl (see Figure 6).

Next, and in order to get known to the other participants and their interests and goals for the seminar, we used the rest of Monday morning to hear short ignite talks from each participant. In order to better recognize the heterogeneity and diversity of the participants, individual slides from these presentations are shown below (Figure 4, 5). All audience were encouraged to jot down comments, questions, points of interest for the talks and add to the brainstorming wall (see above).

Figure 3 Brainstorming wall as result of the first activity: Participants' input on interesting topics, open problems, and challenges.

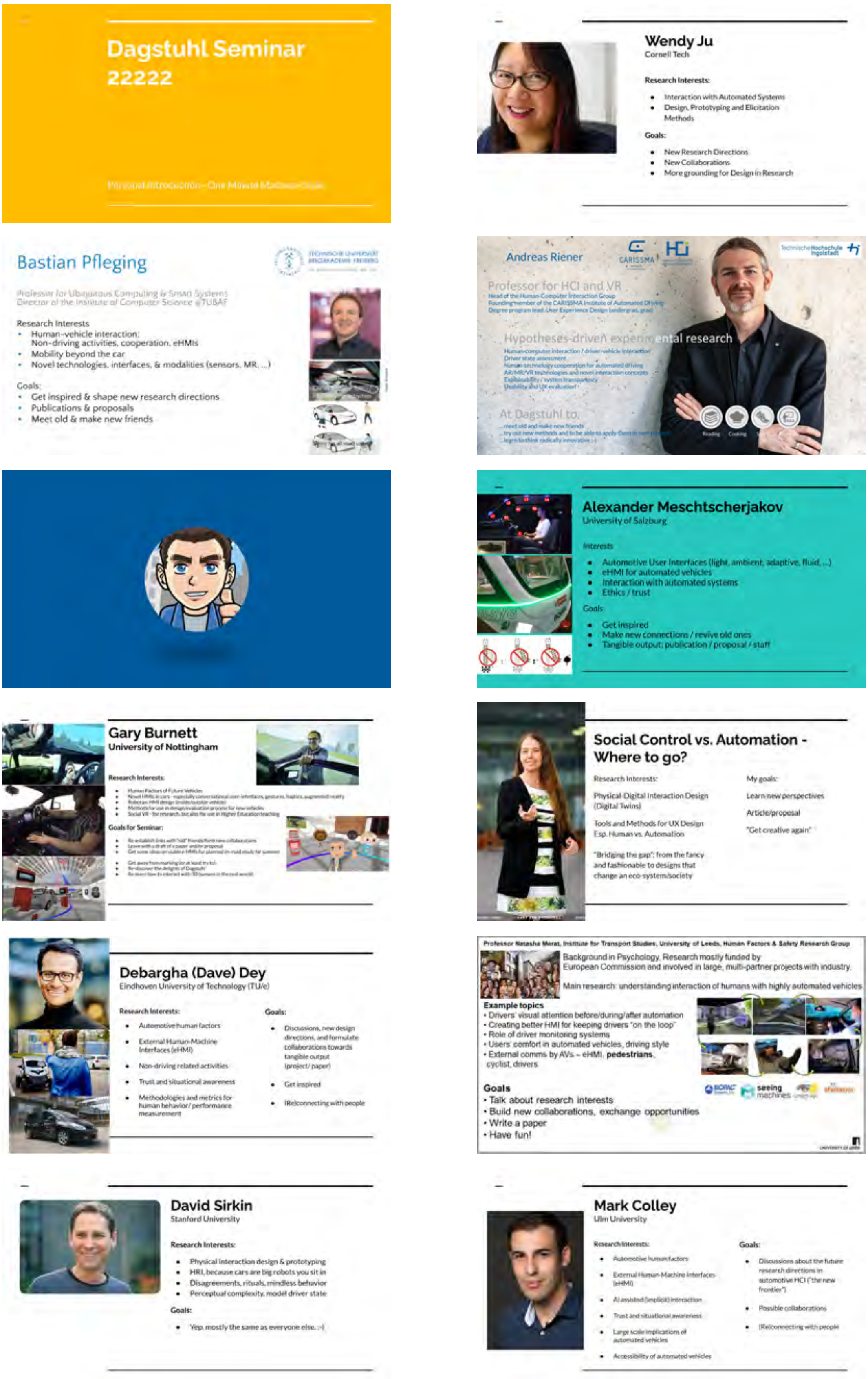

Figure 4 Selected slides from the ignite talks.

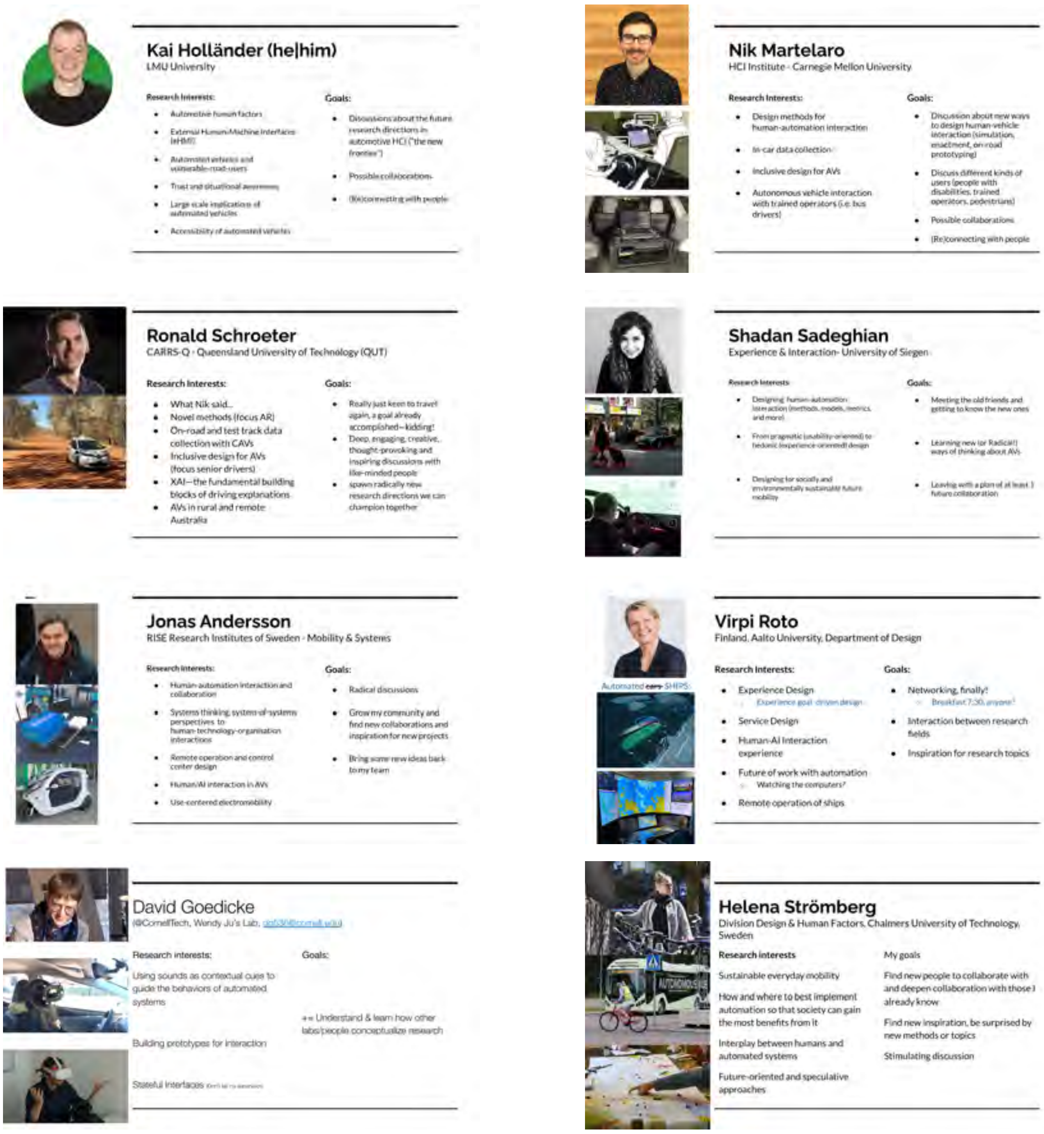

Figure 5 Selected slides from the ignite talks (continued).

Figure 6 Participants were impressed and interested in the history of Dagstuhl Castle.

4.2 Walk and Talk: Mobility of the Future

Bastian Pfleging (TU Bergakademie Freiberg, DE) and Lutz Morich (frE3-innovations, DE)

Joint work of All of the Dagstuhl 22222 participants

As one the goals of the seminar was to think of innovative and groundbreaking visions for future human transportation, the organizers invited all participant to prepare for these topics by a) reflecting on current transportation options and b) envisioning riding in a futuristic, revolutionary vehicle.

Therefore, we asked everyone to find a few minutes during/after the ride to Dagstuhl to think about and reflect about this particular (or a similar) ride and ideate about a future ride to get into the mood for the seminar. In particular, we asked to reflect on the following four questions:

- What did you enjoy / dislike during this ride?
- What did you miss during the ride?
- How would such a ride look like in a revolutionary vehicle? What would it entail?
- What is needed to overcome the limitations of incremental progress and research to come up with more revolutionary concepts and vehicles?

Taking these stimulating questions as starting point, we took all attendees on a "walk & talk" tour around the seminar venue, similar to sessions in prior Dagstuhl seminars (e.g. `https://www.dagstuhl.de/12351`). For each segment of the tour, we asked the participants to discuss certain questions in small groups, for instance picking up on the reflection on the pre-arrival questions. After a few minutes, the whole group stopped to share and discuss the mini-group conversations with all participants before we moved on to the next question and walking discussion (in different groups). This lead to creative discussions in changing groups that opened new perspectives and raised awareness of challenges with current mobility solutions. A challenge yet to solve for the next walk & talk session is to find a way to better document all the discussed aspects to have them available for subsequent sessions.

Concrete Discussions

The following paragraphs outline a few of the discussions and perspectives that we captured during the tour.

What is the experience or emotion or motivation of a journey? It is generally more than going from A to B. It is the experience to see beautiful sites, it is time for relaxation or for working and freedom of doing what the person wants to do. It is worth to think of motivations for a journey or commute when designing mobility solutions. Finding an added value to the journey itself is as much as important as fulfilling expectations of the journey. Commuters are expecting some compromises, but if it gets worse as planned, the motivation decreases and the subjective feeling of a negative mobility experience increases.

A discussion on boredom as a resource for creativity, that has become rare in the connected and fast world: Being bored comes from or means not being distracted by anything. By this a human being comes into a state of being oneself. It is rather the state of not being distracted that is the healthy and creative state that the boredom effect. Being not distracted means coming to a calmness and being open and receiving for inspirations.

A discussion on SAE Level 3 / 4 transfer phases, for example falling asleep and waking up to be able to take over responsibilities to drive a vehicle in traffic. Application idea: Using the system information (health data from seat or video sensors to relax the passenger individually perfectly to his/her individual needs and wake him or her the other way up: "Individual relaxing down and waking up" for well being. For liability reasons AV might require redundant confirmation to be ready to drive again– this is crucial from an ethical and legal point of view.

Discussion on seamless transportation experience: Important is a single point of planning and payment. An assistant might help to tell the traveller where to go and what vehicle to take, to receive information of what to expect during the journey, and while planning to inform the traveller and to meet these expectations. E.g.: would a positive customer experience require a transportation provider who gives proposals of alternative routes or vehicles when problems occur during the travelling?

4.3 Monday Movie Night

Andreas Riener (TH Ingolstadt, DE), and Bastian Pfleging (TU Bergakademie Freiberg, DE)

To stimulate discussions and find sources for inspiration related to mobility aspects, we asked participants before the seminar to collect links to videos that they found inspirational and worth sharing. After dinner on day 1 we met together for the movie night as another activity to get ideas for future mobility concepts based on concept studies, prototypes, research videos, etc. The playlist for the movie night was compiled from our participants' pre-seminar responses. After watching videos, the concepts were discussed and first ideas for the prototyping sessions of the coming days were brainstormed. The list below shows examples of the videos which participants found particularly inspiring.

- Robot Fail Compilation! Funniest on the internet, `https://youtu.be/iZD7Uuthb0E`
- Example clip from student seminar in Nottopia, `https://youtu.be/pukt02g-Wpc`

- The Future of Urban Mobility: an Arup animation, https://youtu.be/_HnLhmXSpUs
- Mobility 2030: Beyond transportation, https://youtu.be/4B7mZFU2sB4
- Cities Rise to the Challenge – Sustainable Mobility, https://youtu.be/8Fj2ARn1WMY
- Doppelmayr/Garaventa - Future Concept Urban (2021), https://youtu.be/bUKgkLSsAWI
- The next chapter in high-class mobility | The Audi urbansphere concept, https://youtu.be/9pQ9E4k8uck
- Grandma Freaks Out Self Driving Tesla! Funny Video, https://youtu.be/VvWGtgoTXAg
- Introducing the self-driving bicycle in the Netherlands, https://youtu.be/LSZPNwZex9s
- The making of... Introducing the self-driving bicycle in the Netherlands, https://youtu.be/6gOjRqlgk_Y
- The real T(h)OR: Evaluation of emergency take-over on a test track, https://youtu.be/LokjcYiF10I
- Notbremsassistent in PKW: Praxistest zeigt grosses Potenzial, aber auch Optimierungsbedarf, https://youtu.be/U4m0sL2RPgA
- ATHENA: supporting UX of conditionally automated driving with natural language reliability displays, https://youtu.be/RaFYpGRijns
- Mixed Reality Test Environment MiRE, https://youtu.be/QBfryCB6Dyk
- The F 015 Luxury in Motion Future City - Mercedes-Benz original, https://youtu.be/SlfpZmCCZ_U

4.4 What could go wrong? Autonomous Vehicle Edition

Wendy Ju (Cornell Tech – New York, US)

Joint work of Nikolas Martelaro, Ronald Schroeter, Wendy Ju

We introduced the game "What could go wrong? AV edition" to play in the evening at Dagstuhl. The goal of the game was to foster discussions about the potential negative aspects of autonomous cars in hopes of surfacing challenges that should be considered during the design process rather than after deployment. The goal is to consider the autonomous vehicle's benefits–improving safety, increasing mobility, reducing emissions–against potential drawbacks. By identifying potential harms and downsides, the community more broadly can design radical and innovative mobility solutions.

5 Panel discussions

5.1 Radical Innovation

Wendy Ju (Cornell Tech – New York, US)

Joint work of Wendy Ju, Virpi Roto, Gary Burnett, Helena Stromberg, Ronald Schroeter

Our first panel was focused on Radical Innovation, and the panelists were Virpi Roto, Gary Burnett, Helena Stromberg and Ronald Schroeter. The panel was structured around three priming questions:

Figure 7 Nikolas Martelaro introducing the *What Could Go Wrong?* game, *left*, the cards, *right*.

- What is your craziest idea of what might happen in the space of mobility and automotive interaction?
- What are ways that you go about exploring radical innovation?
- How should we act as a community to encourage the "unthinkable"

Figure 8 Radical Innovation Panelists Ronald Schroeter, Helena Stromberg, Garry Burnett and Virpi Roto, moderated by Wendy Ju.

To the question of the *craziest idea that they believed*, Roto stated that she believed in "slow traffic." She believed that AVs will be slow, and that though we talk about hyperloop and fast driving, more realistically AVs would be safest, best and most useful at slower speeds, and hopefully people would be willing to have delays in exchange for not having to have to drive. Burnett argued for color changing cars, and musical cars that could compose music based on how it is being driven– pleasing music when calm, bad driving leads to discord–"Carcophony." Stromberg said her craziest idea was what she *doesn't* want: connected everything and connected people with cars and tech (bionic etc.) Schroeter thought the craziest idea was mobility as a service–how can we actually realize it?–but also offered up the Metaverse as a crazy alternative.

How do you go about methodologically doing things that are innovative? Schroeter stated, that it was helpful to be in an environment that encourages risk-taking. He specifically mentioned the importance of having a multidisciplinary environment. He indicated that it was important to "go back to the roots and identifying what makes sense instead of going to the 'supermarket of existing methods and grabbing things from the shelves."' Stromberg emphasized the importance of going deep and applying the "5 whys" to understand the root cause of issues and then asking in which ways root issues could be resolved. The key mindset to do this: going back to the roots and knowing what we want to achieve. Burnett mentioned looking to science fiction for inspiration. He recognizes that sci-fi is not necessarily always right. Still, fiction can help us identify unconscious requirements as a springboard for radical thinking. Roto indicated that it is critical to start at the right question. What is the starting point for radical design? Most tech solutions start from a problem. She felt it was important to shift this perspective from tech problem to experiential problem.

Shadan Sadeghian Borojeni asked Burnett, "Industry looks at problem oriented approaches. Thinking of sci-fi solutions, nobody thought of efficiency or performance issues, but rather focus only on the hedonic and experiential aspects rather than the pragmatic ones. Is that indeed the right way to go?" Burnett responded, "There are no constraints in sci-fi. Things can "look cool" but the reality of how this actually would work (e.g. gestures in AR – error recovery, workflow, etc.) are not thought about." Stromberg stated that she personally uses sci-fi or design fictions to provoke ideas. Critical design/ design fiction are tools to determine and ideate where we could or should go rather than to provide as ground requirements.

Burnett reflected that What's not so radical now was probably radical 15-20 years ago, e.g. GPS/ satnav systems. That was science fiction back then. Lutz Mortiz asked for insights on behalf of those who work in industry how to bring radical sci-fi ideas to life and fruition? Wendy Ju pointed out that the movie Minority Report provided an example of an encapsulation of the research in that was then taking place at MIT Media lab. The fiction made the research a reality.

To the question, "How should we encourage the community to embrace this kind of radical thinking?" Schroeter recommended Stephen Johnson's book, *Where good ideas come from.* We are too strict as an academic research community in accepting novel and radical ideas. Thought-provoking concept or ideas even in position papers are often rejected. This can be addressed by cutting it some slack to papers presenting novel ideas even if they are deficient in other aspects. Burnett advocated being resilient and flexible to different scenarios. Bastian Pfleging suggested that we rethink conference venues, to make late breaking work more futuristic instead of being an outlet for rejected papers. Roto concluded the session by indicating that we can best encourage radical thinking by embracing multidisciplinary approaches.

5.2 Underrepresented users, accessibility, novel users

Andreas Riener (TH Ingolstadt, DE)

Joint work of Andreas Riener, Nik Martelaro, Regina Bernhaupt, Andreas Löcken, Mark Colley
Main reference Andreas Löcken, Philipp Wintersberger, Anna-Katharina Frison, Andreas Riener: "Investigating User Requirements for Communication Between Automated Vehicles and Vulnerable Road Users", in Proc. of the 2019 IEEE Intelligent Vehicles Symposium, IV 2019, Paris, France, June 9-12, 2019, pp. 879–884, IEEE, 2019.
URL https://doi.org/10.1109/IVS.2019.8814027

Designing mobility for the future requires in particular to look at underrepresented user groups, such as disabled people and populations with (any) accessibility problem. Lutz Morich has brought the problem in the open discussion to the point: "*Current practice is designing to the 5% woman and the 95% man. The spotlight is going where the money is. How do we design for small populations? (not yet). It is way cheaper to reduce the options radically (such as Tesla does) to reduce complexity in factories and maximize profit. For manufacturer it needs to be economically affordable to do the variance... An example that worked quit well: SUVs were designed for older people to be able to get in more easily and sit elevated; however, the design language is different – it doesn't look like a senior vehicle...?*".
In the course of the panel session the following **triggering questions** were discussed.

1. What new user groups can automated driving open up?
2. What would be prerequisites for a successful integration into future transport?
3. What accessibility problems are currently present in private/public transport?
4. What would be a radical innovation for blind people in future mobility?
5. How to ensure all people have access to future mobility?
6. Can remote operation help to compensate for physical or sensory disabilities?

In the following, only part of the question-answers are exemplary summarized. As for the first question "*What new user groups can automated driving open up?*", panelists responded:

- Andreas Loecken: Elderly people in rural areas. That is common. I'm not convinced. I like to think about people who have cognitive disabilities, because we don't often think about how to make the car easy to operate.
- Nikolas Martelaro: I'll echo that the people with cognitive difficulties, who usually have a caretaker, so an AV can open up independence. There are also people who have sensory disabilities: people who are blind, for example. One of the low vision people asked to have an autonomous motorcycle or convertible – funny because we often think of AVs as pods-because they would like to have the joy of driving, the physical sensation.
- Mark Colley: We often think about getting people from A to B. The people who are blind do not have a problem, they have support of the government because they can get a cab, they can talk to the cab driver. They plan their trips days in advance, know where they want to be, so they don't want to go someplace spontaneously, so they say they are skeptical and don't want it. But! We need to think about the whole journey, the cab driver comes to the door, they take them by the hand, they take the luggage. So there's a lot that the cab driver does that needs to be done. Automakers do not see that as a valid business case, it's far far away. Have not seen many scenarios of people with a variety of disabilities. People will need situation awareness. Apple has an app to help people walk around. There is the one paper on the last mile problem, also there is the first mile problem.

- Regina Bernhaupt: I think we should design the other way around. Yes we can design for the next user group and splitting them up, look at motor groups, look at adaptation potential. Design for inclusivity in the first place. For example, design for Regina, because she is too small for a typical car. By designing for all the people, by putting it on top of what is already doing, and additionally serving the groups, we will reach out further to the next one. For example, if you design a remote control for the people who are blind, that would make it a better design for everyone, and if it is haptic, it helps everyone. Sometimes these things are a hard sell in industry. So change the perspective, accessibility is for everyone.

More questions were discussed, also including the perspective of other seminar participants, leading to the final question to the panelists: "*Which accessibility problem would you like to solve first of accessibility?*" (one sentence)

- Andreas Loecken would like to give companies monetary incentives to steer the change process towards holistic, seemless transportation.
- Nik Martelaro wants more people from different backgrounds agency to be designers, to be on the design team.
- Mark Colley was intrigued by the idea of radically changing the design process itself, i.e., use tools to make it cheaper for companies to include this aspect.
- Regina Bernhaupt pointed out that for a person which is blind, maybe give them the butler, not an autonomous vehicle. Don't put technology first, keep it for later...

Figure 9 Panelists for the underrepresented users panel(from left to right): Regina Bernhaupt, Mark Colley, Nik Martelaro, Andreas Loecken.

References

1 Andreas Löcken and Philipp Wintersberger and Anna-Katharina Frison and Andreas Riener. *Investigating User Requirements for Communication Between Automated Vehicles and Vulnerable Road Users*. 2019 IEEE Intelligent Vehicles Symposium, IV 2019, Paris, France, pp. 879–884, June 9-12, 2019.

5.3 Teaching of Design Methods

Andreas Riener (TH Ingolstadt, DE)

We further wanted to use the seminar as an opportunity to communicate new design and usability evaluation methods. This activity was also organized as a community event – seminar participants could suggest methods that they find particularly innovative and profitable and that they would be willing to communicate to interested parties in a short session (15 min. to 90 min.) prior to arrival. Feedback was clustered and finally offered in 8 small groups on Monday afternoon and Tuesday morning (each seminar participant could participate in several activities). Here is an overview of idea generators and topics.

- Teach Design Methods, part I (Monday, 14.00-15.30 hrs)
  1. Virpi Roto: Experience goal-driven design (90 min.)
  2. David Sirkin: Collecting field data (from passers-by) (90 min.)
  3. Gary Burnett: VR to run design workshops (90 min.)
- Teach Design Methods, part II (Tuesday, 9.00-10.30 hrs)
  1. Nikolas Martelaro: Remote Wizard of Oz methods (for automotive interaction) (60 min.)
  2. Andreas Riener: User Enactments as a Design Method (60 min.)
  3. Regina Bernhaupt: Design-thinking workshop (60 min.)
  4. from 10.00AM (for everybody) David Goedicke: Driving Simulator Demo
  5. from 10.15AM Ronald Schroeter and Nik Martelero: What could go wrong? An interactive card game (two groups)

5.4 World Cafe: Discussions of major themes in the community

Wendy Ju (Cornell Tech – New York, US)

Joint work of All of the Dagstuhl 22222 participants

For the World Cafe event, we had people brainstorm topics they wanted to talk about with the entire group, and then grouped these. From the groupings, three main topics were selected.

Nik Martelaro, Shadan Sadeghian Borojeni and Regina Bernhaupt lead each of these themes. The test of the team broke into three groups, which rotated through each theme discussion. From this, we created a shared group sentiment on the area, interesting research questions and desired outcomes.

Nikolas Martelaro's theme focused on Personalized Design for Future Mobility. The advent of new forms of vehicles which could have modular blocks would make it possible to adapt for users who had different abilities, people who are on extremes of the human factor sizes, and families, who might have dogs or children. The modularity would make it possible to have reconfiguration for long trips or off-roading. Key questions are who is legally liable for manufacturing and safety validation for each vehicle?

Shadan led the discussion for new models and paradigms for interacting with automation. The research questions stemming from this area were:

Figure 10 David Sirkin, Alexander Meschtscherjakov, Ronald Schroeter, and Gary Burnett, brainstorming.

- RQ1: what are the needs for new models and paradigms? Potential answers were, new control schemes, new relations, expectation management/mental models, and responsibility management.
- RQ2: what are the implications of different models and paradigms for identity and for breakdowns in interaction.
- RQ3: How do we convey roles and responsibilities in changing levels of automation?
- RQ4: How acceptable are the defined roles for different users?
- RQ5: How to keep up with fast updates of car?
- RQ6: How do different interaction paradigms shape emerging mobility experiences?

Regina led a discussion of reinventing the automotive experience. The key stakeholders in this were the government, users and technology providers. The government is responsible for thinking about the economic impacts of mobility, the public health implications, and sustainability. The technology providers are thinking about the physical, emergent and virtual aspects of mobility technology. Users, on the other hand, are thinking about self-efficacy, user experience and their personal values. Interaction between technology and users is largely around automation and mobility services. Interaction between the government and technology companies have to do with policy and standards. And the government interacts with users through training and licensure. Key ideas in the arena have to do with improving self-efficacy for individual users, groups and community, having virtual butlers who can help individuals to customize their transportation, and visions of post-car transportation from the technology providers.

6 Prototyping Activities

6.1 Video Prototyping as a Way to Explore Radical Innovation

Wendy Ju (Cornell Tech – New York, US)

Video prototyping is a valuable tool that enables better exploration of scenarios, interactions, context, timing and experiences. The video outcomes can also support reflection and analysis, reinterpretation, and form the basis for collaboration.

On Wednesday morning in the workshop, I introduced the method in a presentation, providing several examples of video prototyping from my own research in automotive interaction and human-robot interaction design, and then presented some ideas of how video prototypes could be made.

We then broke into three groups and worked on brainstorming, quick and dirty prototyping, and putting together video prototypes. The activities culminate in three different videos from each of the three groups.

Figure 11 Wendy Ju presenting on video prototyping.

6.2 Video Prototype: Radical Innovations in Future Mobility and Virtual Assistants

Kai Holländer (LMU München, DE), Gary Burnett (University of Nottingham, GB), Wendy Ju (Cornell Tech – New York, US), Natasha Merat (University of Leeds, GB), Lutz Morich (frE3-innovations, DE), Virpi Hannele Roto (Aalto University, FI), Ronald Schroeter (Queensland University of Technology – Brisbane, AU), and David Sirkin (Stanford University, US)

Main reference Kai Holländer, Lutz Morich, Natasha Merat, Gary Burnett, Virpi Roto, Wendy Ju, David Sirkin: "Radical Innovations in Future Mobility and Virtual Assistants", in Proc. of the Adjunct 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI '22, p. 209 – 211, Association for Computing Machinery, 2022.
URL https://doi.org/10.1145/3544999.3551501

For the future of mobility it will be important to empower individuals in the selection of climate friendly and seamless interaction of different means of transport. In this work we present a video which shows an example use case of a virtual agent which helps a user to figure out the best transportation options for a specific individual journey. The virtual agent is in a dialogue with the user and can help to optimize selecting means of transportation which suit the purpose of the trip, optimize for the best possible occupancy rate of the suggested vehicles, the lowest energy consumption as well as costs, and account for special needs if a user has limitations of any sort.

This video was accepted to the videos track in AutoUI 2022, `https://dl.acm.org/doi/fullHtml/10.1145/3544999.3551501`.

6.3 Video Prototype: How to Manage Social Order in Shared Automated Vehicles

Nikolas Martelaro (Carnegie Mellon University – Pittsburgh, US), Jonas Andersson (RISE Research Institutes of Sweden – Göteborg, SE), Gary Burnett (University of Nottingham, GB), Debargha Dey (TU Eindhoven, NL), Andreas Löcken (TH Ingolstadt, DE), and Helena K. Strömberg (Chalmers University of Technology – Göteborg, SE)

Autonomous shared ride vehicles may be prone to similar social issues and non-ideal passenger behaviors as today's public transit. Such issues may include passengers littering, harassing others, and creating an environment that is generally unpleasant for riders. Transportation user experience designers should preemptively consider such scenarios early in their design work to help develop possible interfaces to manage social order and maintain good rider experience. Through a short video prototype, we present three possible non-ideal scenarios that may occur on shared autonomous shuttles and provide three potential solutions to begin a discussion around how to design for such non-ideal situations.

This video was accepted to the Videos track at AutoUI 2022, `https://dl.acm.org/doi/abs/10.1145/3544999.3550154`.

6.4 Video Prototype: A Critical Perspective on Radically Innovating Personal Mobility

Regina Bernhaupt (TU Eindhoven, NL), Mark Colley (Universität Ulm, DE), David Goedicke (Cornell Tech – New York, US), Alexander Meschtscherjakov (Universität Salzburg, AT), Bastian Pfleging (TU Bergakademie Freiberg, DE), Andreas Riener (TH Ingolstadt, DE), Shadan Sadeghian Borojeni (Universität Siegen, DE)

Automotive research today is mostly concerned with incremental improvement. With an automated vehicle in mind, increased safety, reduced fuel consumption, and the possibility of non-driving-related activities are anticipated. However, the challenges of future mobility require a critical rethinking of mobility in its entirety, including the availability of personalized and motorized mobility. With this video, we want to stimulate discussions on more radical innovation in mobility. In detail, we want the audience to imagine what challenges a world without individually driven cars would pose.

This video was accepted to the videos track at AutoUI 2022, `https://dl.acm.org/doi/abs/10.1145/3544999.3551689`.

7 Conclusion

7.1 Commitments and Future Plans

Wendy Ju (Cornell Tech, US)

Before the seminar closed, we had participants go around and make commitments for actions they planned to take following the workshop.

- Each of the video prototyping groups pledged to submit their videos to the AutoUI video track.
- There was also a commitment to incorporate the methods learned across the workshop in the curriculum at everyone's respective universities.
- There were discussions to create workshops or birds-of-a-feather sessions around accessibility in AVs, alternatives to vehicles, and micro-mobility at CHI.
- David Goedicke offered to share his virtual reality driving simulation platform with people who were interested in collaborating to use that platform to prototype alternative visions for transportation.
- Wendy Ju made plans to host Debargha Dey and Mark Colley at Cornell Tech for post-doctoral research and visiting researcher positions, respectively.
- Andreas Riener and Ronald Schroeter have agreed to submit a DAAD proposal to allow a PhD student to visit Prof. Riener's group to foster cooperation and scientific exchange. Furthermore, Schroeter and Riener have been discussing a long-term research visit and residency of Andreas Riener at CARRS-Q, in the frame of a sabbatical for several years now, which was delayed through COVID. Within seminar 22222, the planning of future exchanges that will deepen the engagement between the two institutions and staff was taken-up again.

7.2 Post-seminar Travel Experiences by Participants

Andreas Riener (TH Ingolstadt, DE)

Our seminar finished Friday, June 3rd at lunchtime. Most of the participants left Dagstuhl at around the same time. Using a WhatsApp group, we tracked everybody's travel experience back home. From the report below (see Figure 12) it can be very easily seen that individual car travel still seems to be the most convenient and relaxed way to reach home on-time. Traveling via plane to different places in Europe and even US lasts almost the same time then using a train within Germany... Plenty of opportunities to improve future mobility... ;-)

Acknowledgements. We would like to thank all participants of the Dagstuhl Seminar 22222 for their input on this spontaneous activity.

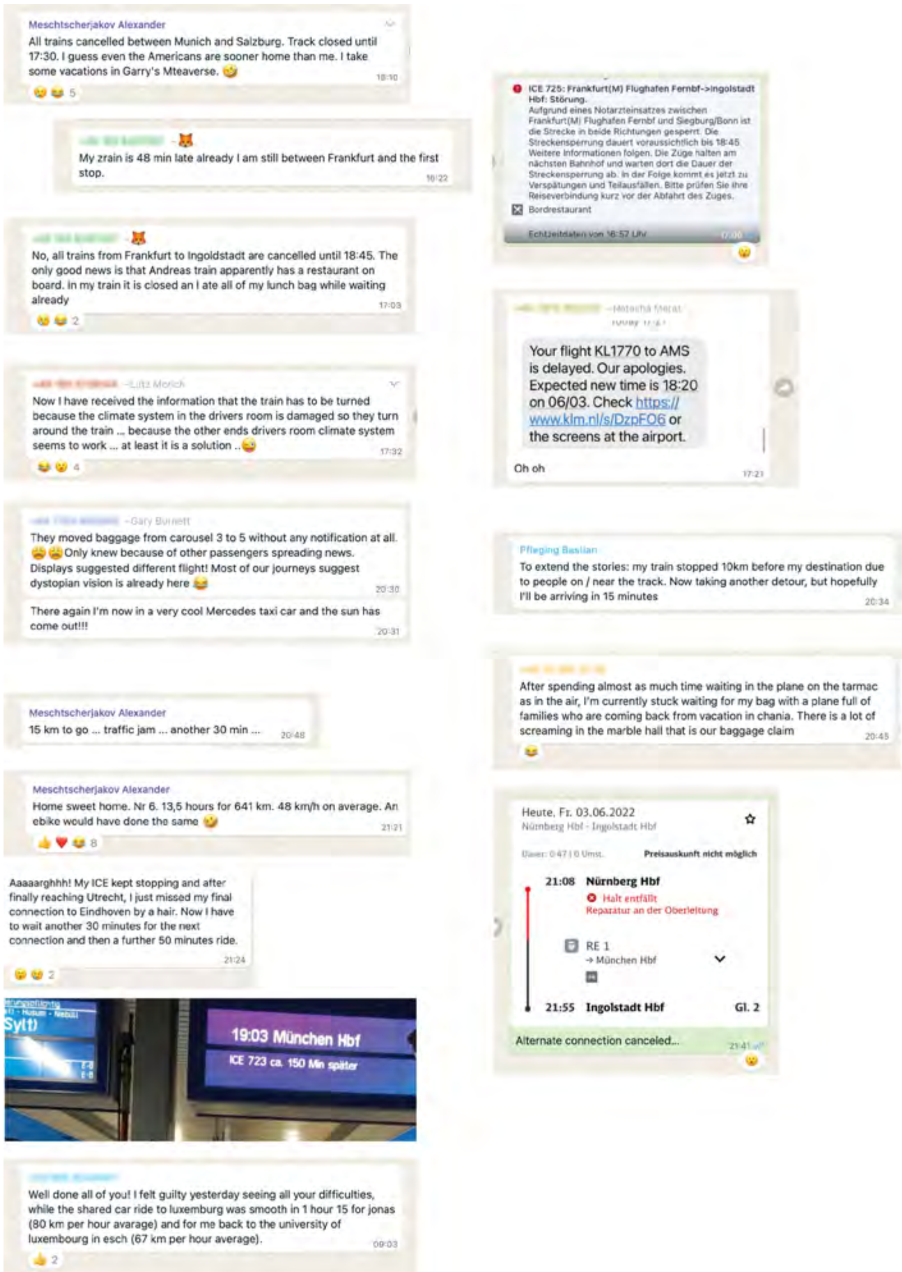

Figure 12 Travel experiences of participants on the way back from the Dagstuhl seminar.

Participants

- Jonas Andersson
RISE Research Institutes of Sweden – Göteborg, SE
- Regina Bernhaupt
TU Eindhoven, NL
- Gary Burnett
University of Nottingham, GB
- Mark Colley
Universität Ulm, DE
- Debargha Dey
TU Eindhoven, NL
- David Goedicke
Cornell Tech – New York, US
- Kai Holländer
LMU München, DE
- Wendy Ju
Cornell Tech – New York, US
- Andreas Löcken
TH Ingolstadt, DE
- Nikolas Martelaro
Carnegie Mellon University – Pittsburgh, US
- Nathasha Merat
University of Leeds, GB
- Alexander Meschtscherjakov
Universität Salzburg, AT
- Lutz Morich
frE3-innovations, DE
- Bastian Pfleging
TU Bergakademie Freiberg, DE
- Andreas Riener
TH Ingolstadt, DE
- Virpi Hannele Roto
Aalto University, FI
- Shadan Sadeghian Borojeni
Universität Siegen, DE
- Ronald Schroeter
Queensland University of Technology – Brisbane, AU
- David Sirkin
Stanford University, US
- Helena K. Strömberg
Chalmers University of Technology – Göteborg, SE