Report from Dagstuhl Seminar 12352

Information Flow and Its Applications

Edited by

Samson Abramsky[1], Jean Krivine[2], and Michael W. Mislove[3]

1 **University of Oxford, GB,** samson.abramsky@comlab.ox.ac.uk
2 **University of Paris Diderot, FR,** jean.krivine@pps.univ-paris-diderot.fr
3 **Tulane University, US,** mwm@math.tulane.edu

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 12352 "Information Flow and Its Applications". This seminar brought together mathematicians, computer scientists, physicists and researchers from related disciplines such as computational biology who are working on problems concerning information and information flow.

Seminar 26.–31. August, 2012 – www.dagstuhl.de/12352
1998 ACM Subject Classification H.1.1 Systems and Information Theory
Keywords and phrases Information flow, semantics of computation, quantum computing, systems biology, information theory, informatics
Digital Object Identifier 10.4230/DagRep.2.8.99

1 Executive Summary

Samson Abramsky
Jean Krivine
Michael W. Mislove

The seminar "Information Flow and Its Applications" that took place in Schloss Dagstuhl in August 2012, has been the latest in a series of meetings concerning information flow that began with the 2008 Clifford Lectures by Samson Abramsky at Tulane University, and continued with two further meetings on informatic phenomena at Tulane, as well as a previous Dagstuhl seminar on "The Semantics of Information"[1]. The seminar "Information Flow and Its Applications" brought together mathematicians, computer scientists, physicists and researchers from related disciplines such as computational biology who are working on problems concerning information and information flow.

The seminar gathered 21 participants in addition to the 3 organizers, in the studious but cosy atmosphere of Schloss Dagstuhl. Armed with slides and chalks, each speaker described in terms as simple as can be, the questions and problems they were trying to solve, which, as the title of the seminar suggests, had all in common the issue of the representation and analysis of information flows.

The hypothesis underlying the organization of the seminar was the following: information flows leave on substrates which transport and transform data along time and space. From the modeling, analysis or simulation of these substrates will emerge unifying techniques or

[1] http://www.dagstuhl.de/10232

concepts. It is understood that such substrate can be *artificial*, for instance in the case of an electrical circuit, or *natural*, as in the complex signaling pathways that govern cellular fate. Moreover, information may be treated by systems in a *designed* manner, for instance a computer that processes its inputs according to a determined program, or be the result of *evolution*, like the internet which is a perfect example of a system that carries and processes information in spite of the absence of a pre-existing specification.

Although traditionally information processing is studied by distinct communities, scattered along the *Artificial-Natural* and *Designed-Evolved* axes, it is noteworthy that this separation is, to some extent, a historical artifact in the sense that artificial systems may be the fruit of evolution (as the internet) while natural ones may be used in a purely specified manner (as in synthetic biology). It is therefore natural to expect that tools and techniques developed in one field may be also relevant to others.

Another unifying scheme of the seminar was the emphasis on the use of formal languages in the representation of information flows. Indeed once "a real world" computing system, such as the cell or a quantum circuit, is abstracted as a formal programming language, one may then start to apply techniques imported from theoretical computer science. In the study of evolved systems, these techniques may be used to *extract* a specification of what is being observed, while in the context of systems where a specification is *a priori* at disposal, one may use these techniques to *verify* that the way information is processed conforms to the expectation.

Over the 4 days of talks, which gave rise to feedback that went beyond the expectation of the organizers, the participants of the seminar "Information Flow and its Applications" have had the opportunity to listen to talks ranging from Systems Biology to Theoretical Physics, from Quantum Computing to the study of Ecological systems. As organizers, we believe that the original guess that Information Flow should be a topic of its own was largely a good one.

2 Table of Contents

3 Overview of Talks

3.1 Galois group of a symmetric measurement

Marcus Appleby (Perimeter Institute – Waterloo, CA)

The problem of proving (or disproving) the existence of symmetric informationally complete positive operator valued measures (SICs) has been the focus of much effort in the quantum information community during the last 12 years. In this talk we describe the Galois invariances of Weyl-Heisenberg covariant SICs (the class which has been most intensively studied). It is a striking fact that the published exact solutions (in dimensions 2–16, 19, 24, 35 and 48) are all expressible in terms of radicals, implying that the associated Galois groups must be solvable. Building on the work of Scott and Grassl (J. Math. Phys. 51 042203 (2010)) we investigate the Galois group in more detail. We show that there is an intriguing interplay between the Galois and Clifford group symmetries. We also show that there are a number of interesting regularities in the Galois group structure for the cases we have examined. We conclude with some speculations about the bearing this may have on the SIC existence problem.

3.2 Information and distributed computation

David Balduzzi (MPI für Intelligente Systeme – Tübingen, DE)

Main reference D. Balduzzi, "On the information-theoretic structure of distributed measurements," in Proc. of 7th Int'l Workshop on Developments of Computational Methods (DCM'11), EPTCS, vol. 88, pp. 28–42, 2012.
URL http://dx.doi.org/10.4204/EPTCS.88.3

Computations implemented in physical systems can be described at many different spatio-temporal granularities. For example, the work performed by the brain can be described at the level of atoms, molecules, neurons and potentially higher-order structures. I present an information-theoretic approach to coarse-graining distributed computations. The first step is to introduce effective information, which can be shown to incorporate Kolmogorov complexity, mutual information and VC-entropy (an important measure of complexity in statistical learning theory) as special cases. A second measure, excess information, provides a geometric characterization of indecomposable computations. The two measures are then applied to study coarse-grainings in Conway's Game of Life and Hopfield networks.

3.3 A Hypothesis Test For Bell's Inequality

Peter Bierhorst (Tulane University, US)

Experimental tests of Bell inequalities require statistical analysis. Usually, successive trials are taken to be independent and identically distributed. Thus the expectation quantities in Bell's inequality can be estimated by appealing to the Law of Large Numbers. Though

the i.i.d. assumption is a natural one, it need not be obeyed by a local hidden variable theory. Luckily, statistical methods can still distinguish Quantum Mechanics from local hidden variable theories over large numbers of trials.

3.4 Complex Information Systems

Robert J. Bonneau (AFOSR – Arlington)

The talk will provide an overview of complex information systems including quantifying, managing, and designing heterogeneous networked systems. Methods of measuring and assessing the performance of networked, software, and hardware integrated systems such as cloud architectures will be discussed including techniques of sparse approximation in systems measurements, and algebraic and topological statistical metrics for performance. Strategies of quantifying risk over different geometric and statistical classes of distributed systems will be examined as well as methods of tracking and coding dynamic information flows.

3.5 Structured Data Analysis

Gunnar Carlson (Stanford University, US)

We discuss methods for representing and "measuring" the shape of data sets. They are represented by simplicial complexes, and the shape is measured using homological signatures as extended to the world of point cloud data via the persistent homology methodology. Examples were given in both cases, and suggestions were made about how to represent even more complicated data types, involving dynamical systems and control systems.

3.6 Ensemble signaling in a MAP kinase cascade

Eric Deeds (University of Kansas, US)

Joint work of Deeds, Eric; Suderman, Ryan

A cell's ability to adapt to constantly changing environmental conditions is derived from its signaling networks. Despite there importance, there is currently no consensus regarding the nature of the protein complexes such networks employ. One prominent view involves signaling machines, while the inherent combinatorial complexity of such networks has led to the more recent proposal of pleiomorphic ensembles. In this work, we use rule-based modeling techniques to explore this question in the case of the yeast pheromone MAPK cascade. We constructed a model of this cascade based on current understanding of the interactions in the pathway. We found that, despite exhibiting considerable ensemble character, this model can replicate existing experimental data for the cascade. We also considered a model designed to

exhibit more machine-like character. This model could not replicate the behavioral changes observed in cascade when Ste5, the signaling scaffold, is overexpressed. These findings indicate that ensemble signaling can indeed produce "realistic" cellular behavior, and that machine and ensemble systems can exhibit distinctly different phenotypes.

3.7 Information Flow in Quantum Computing: Circuits, Entanglement, and MBQC

Ross Duncan (Université Libre de Bruxelles, BE)

The measurement based quantum quantum computer (MBQC) works by pushing quantum information through a network of entangled quantum states. The structure of these states, and the paths that the information takes within the network are easily studied using a high-level presentation of quantum theory based on symmetric monoidal categories and its graphical language. In this talk I'll show how to apply these techniques to derive a quantum circuit equivalent to a given MBQC program, thus verifying the correctness of the original program.

3.8 Formal model reduction

Jérôme Feret (ENS – Paris, FR)

Joint work of Camporesi, Ferdinanda; Danos, Vincent; Feret, Jérôme; Fontana, Walter; Harmer, Russell; Krivine, Jean

Modelers of molecular signaling networks must cope with the combinatorial explosion of protein states generated by post-translational modifications and complex formation. Rule-based models provide a powerful alternative to approaches that require an explicit enumeration of all possible molecular species of a system. Such models consist of formal rules stipulating the (partial) contexts for specific protein-protein interactions to occur. These contexts specify molecular patterns that are usually less detailed than molecular species. Yet, the execution of rule-based dynamics requires stochastic simulation, which can be very costly. It thus appears desirable to convert a rule-based model into a reduced system of differential equations by exploiting the lower resolution at which rules specify interactions.

In this talk, we present a formal framework for constructing coarse-grained systems. We track the flow of information between different regions of chemical species, so as to detect and abstract away some useless correlations between the state of sites of molecular species.

The result of our abstraction is a set of molecular patterns, called fragments, and a system which describes exactly the concentration (or population) evolution of these fragments. The method never requires the execution of the concrete rule-based model and the soundness of the approach is described and proved by abstract interpretation.

3.9 Hidden Bayesian networks

Tobias Fritz (Institute of Photonic Sciences – Castelldefels, ES)

Main reference T. Fritz, "Beyond Bell's theorem: correlation scenarios," New Journal of Physics, vol. 14, issue 10, Oct. 2012.
URL http://dx.doi.org/10.1088/1367-2630/14/10/103001

This will be an outline of a research program aiming at generalizing Bell's Theorem to arbitrary causal structures. It will probably be joint work with Rob Spekkens.

An arbitrary causal structure is given by a (finite) poset represented by a directed acyclic graph. The vertices in the graph represent spacetime events at which measurements are conducted. The edges are the wordlines in spacetime of physical systems being transmitted from one vertex to another; these messages are what allow the creation of correlations between measurements at different vertices. Such correlations are called classical if the correlations can be modeled in terms of classical messages, or, equivalently, in terms of a "hidden Bayesian network". (A hidden Bayesian network is like a hidden Markov model, just on an arbitrary causal structure.) The correlations are called quantum if they can be modeled by quantum systems being sent along the edges and quantum measurements being conducted on the vertices.

Standard Bell scenarios as well as the scenarios studied in arXiv:1206.5115 are subclasses of this formalism; interesting examples beyond these remain to be found.

3.10 Statistics, causality and Bell's theorem

Richard Gill (Leiden University, NL)

Main reference R.D. Gill, "Statistics, causality and Bell's theorem," submitted, arXiv:1207.5103 [stat.AP].
URL http://arxiv.org/abs/1207.5103v1

Bell's (1964) theorem is popularly supposed to establish the non-locality of quantum physics as a mathematical-physical theory. Building from this, observed violation of Bell's inequality in experiments such as that of Aspect and coworkers (1982) is popularly supposed to provide empirical proof of non-locality in the real world. My talk reviews recent work on Bell's theorem, linking it to issues in causality as understood by statisticians. The talk starts with a new proof of a strong (finite sample) version of Bell's theorem which relies only on elementary arithmetic and (counting) probability. This proof underscores the fact that Bell's theorem tells us that quantum theory is incompatible with the conjunction of three cherished and formerly uncontroversial physical principles, nicknamed here locality, realism, and freedom. I will argue that (accepting quantum theory) Bell's theorem should lead us to seriously consider relinquishing not locality, but realism, as a fundamental physical principle. In the talk I hope also to discuss statistical issues, in the interpretation of state-of-the-art Bell type experiments, related to post-selection in observational studies. Finally I state an open problem concerning the design of a quantum Randi challenge: a computer challenge to Bell-deniers. Can we prove useful probabilistic Bell inequalities for the situation that Alice and Bob's measurement stations receive all their measurement settings at once, and all signals from the source at once, and then (disconnected from the one another and the source) generate all their outputs at once. A "useful" Bell inequality in this context is one where the

probability of violation of the local realism expectation bound by some small amount delta decreases with N, the number of settings to be processed at a time.

3.11 Graphs, Rewriting and Pathway Reconstruction for Rule-Based Models

Jonathan Hayman (ENS – Paris, FR)

Joint work of Hayman, Jonathan; Danos, Vincent; Feret, Jérôme; Fontana, Walter; Harmer, Russell; Krivine, Jean; Thompson-Walsh, Chris; Winskel, Glynn

We introduce a novel way of constructing concise causal histories (pathways) to represent how specified structures are formed during simulation of systems represented by rule- based models. This is founded on a new, clean, graph-based semantics introduced in the first part of this paper for Kappa, a rule-based modeling language that has emerged as a natural description of protein-protein interactions in molecular biology. The semantics is capable of capturing the whole of Kappa, including subtle side-effects on deletion of structure, and its structured presentation provides the basis for the translation of techniques to other models. In particular, we give a notion of trajectory compression, which restricts a trace culminating in the production of a given structure to the actions necessary for the structure to occur. This is central to the reconstruction of biochemical pathways due to the failure of traditional techniques to provide adequately concise causal histories, and we expect it to be applicable in a range of other modeling situations.

3.12 Coherence in Hilbert's hotel

Peter Hines (University of York, GB)

Main reference A series of papers based on this talk is in progress. Please contact the author for drafts.

This talk describes the interaction of MacLane's categorical coherence with self- similarity $S \cong S \otimes S$ and untyped (i.e. single-object) monoidal structures.

A coherence result is presented, giving a decision procedure for commutativity of diagrams built up from typed and untyped monoidal tensors and structural isomorphisms, and canonical isomorphisms between the two settings.

Applications are discussed, including word problems in Thompson groups, and deciding equality of arithmetic expressions based on modular arithmetic and related operations.

3.13 Minimal glueings and unambiguous stoichiometry in Kappa

Ricardo Honorato-Zimmer (University of Edinburgh, GB)

Joint work of Honorato-Zimmer, Ricardo; Stucki, Sandro; Danos, Vincent

Rule-based modeling languages have been developed to represent biomolecular interactions in a concise way. They achieve this by using patterns that omit unnecessary details or context from the description of those interactions. Hence, combinatorial systems with a large or even infinite number of different species can then be studied and analyzed. However, in these languages it is usually not immediately clear when a rule creates or destroys an instance of a pattern or observable, that is, the stoichiometry of the rule for that observable is often ambiguous. In this work, we formally define the concept of minimal glueings in the category-theoretical framework developed by Harmer et al (2011) that allow us to check if a set of observables has an unambiguous stoichiometry with respect to a rule.

3.14 From Contextuality to Nonlocality

Shane Mansfield (University of Oxford, GB)

We outline a route to constructing n-partite Bell-type models from more general measurement configurations, e.g. Kochen-Specker configurations, using a sheaf theoretic formulation of measurement scenarios. The construction has the property that the resultant model is no-signaling and that it is nonlocal if and only if the original model is contextual. This could provide a new route to Bell tests for contextuality. It also raises an interesting an novel connection between the simplest possible contextual model, the triangle, and PR boxes, the maximally nonlocal (2,2,2) correlations.

3.15 Analyzing continuous channels

Michael W. Mislove (Tulane University, US)

Main reference Slides of the talk available at http://www.entcs.org/mislove/dagstuhl12.pdf
URL http://www.entcs.org/mislove/dagstuhl12.pdf

In this talk we describe how the Cantor Fan can be used as a basis for approximating continuous channels with discrete ones. The Cantor Fan is our term for the order-ideal compilation of the rooted full binary tree, with the path order, and is so named because the set of maximal elements is homeomorphic to the middle-third Cantor set, when the completion is regarded as an algebraic domain. We describe the effect of applying the Probability monad to this structure, and how it then gives a setting to describe how a channel with the Cantor set as input/output set can be approximated by a sequence of channels, C_n, in a hierarchy, where C_n has 2^n inputs/outputs for each n. We discuss how entropy and capacity of the

individual approximations can be viewed in this setting. Since this is very preliminary work, no definitive results are presented.

3.16 Differential Privacy: An Overview

Catuscia Palamidessi (Ecole Polytechnique – Palaiseau, FR)

Joint work of Palamidessi, Catuscia; Chatzikokolakis, Konstantinos; Alvim, Mario S.; Andrés, Miguel E.
Main reference M.S. Alvim, M.E. Andrès, K. Chatzikokolakis, and C. Palamidessi, "On the relation between Differential Privacy and Quantitative Information Flow," in Proc. of ICALP'11, LNCS 6756, Springer, pp. 60–76, 2011.
URL http://hal.inria.fr/inria-00627937/en

We discuss the general problem of protecting private information and we present differential privacy, a framework which has been recently – and quite successfully - introduced in the area of statistical databases. We discuss the trade-off between privacy and utility, and present some fundamental result in the area. Then, we generalize the notion of differential privacy so to make it applicable to domains other than databases. We start form the observation that the standard notion of differential privacy relies on the notion of Hamming distance on the set of databases, and we extend it to arbitrary metric spaces. We show various examples, and we revise some of the fundamental results of differential privacy in this extended setting. As a particular case study, we consider location-based applications, and the resulting notion of geo-indistinguishability.

3.17 Compact Closed Categories and Frobenius Algebras for Computing Natural Language Meaning

Mehrnoosh Sadrzadeh (University of Oxford, GB)

Joint work of Sadrzadeh, Mehrnoosh; Clark, Stephen; Coecke, Bob; Grefenstette, Edward; Kartsaklis, Dimitri; Pulman, Stephen
Main reference (1) B. Coecke, M. Sadrzadeh, S. Clark, "Mathematical Foundations for a Compositional Distributional Model of Meaning," Linguistic Analysis, volume dedicated to Lambek's Festschrift, 2010.
(2) E. Grefenstette, M. Sadrzadeh, "Experimental Support for a Categorical Compositional Distributional Model of Meaning," in Proc. of the Conf. on Empirical Methods in Natural Language Processing (EMNLP'11), pp. 1394–1404, ACL, 2011.
URL http://arxiv.org/abs/1003.4394
URL http://www.aclweb.org/anthology/D11-1129

Compact closed categories have found applications in modeling quantum information protocols by Abramsky-Coecke. They also provide semantics for Lambek's pregroup algebras, applied to formalizing the grammatical structure of natural language, and are implicit in a distributional model of word meaning based on vector spaces. In particular, in previous work, Coecke-Clark-Sadrzadeh used the product category of pregroups with vector spaces and provided a distributional model of meaning for sentences. We recast this theory in terms of strongly monoidal functors and advance it via Frobenius algebras over vector spaces. The former are used to formalize topological quantum field theories by Atiyah and Baez-Dolan, and the latter are used to model classical data in quantum protocols by Coecke-Pavlovic-Vicary. The Frobenius algebras enable us to work in a single space in which lives meanings of words,

phrases, and sentences of any structure. Hence we can compare meanings of different language constructs and enhance the applicability of the theory. We report on experimental results on a number of language tasks such as word sense disambiguation and term/definition extraction and show how our theoretical predictions are verified on real large scale date from British National Corpus.

3.18 Rigid geometric constraints for Kappa models

Sandro Stucki (University of Edinburgh, GB)

Joint work of Danos, Vincent; Honorato-Zimmer, Ricardo; Jaramillo-Riveri, Sebastian; Stucki, Sandro

Rule-based modeling languages such as Kappa and BNGL allow for a concise description of combinatorially complex biochemical processes. However, these languages do not provide means to directly express the three-dimensional geometry of chemical species. We propose an extension to the Kappa modeling language allowing the annotation of the structure of chemical species with three-dimensional geometric information. This naturally introduces rigidity constraints on the species and reduces the state space of the resulting model by excluding species that are not geometrically sound. We show that geometrically enhanced Kappa models can still be simulated efficiently, albeit at the cost of a greater number of null-events occurring during the simulation.

3.19 Computer and Information Science – Future Paradigm for Complex System Models?

Baltasar Trancon y Widemann (Universität Bayreuth, DE)

Joint work of Trancon y Widemann, Baltasar; Hauhs, Michael

We review the meta-theoretical situation of the scientific discipline of ecology, taken as a typical example of the family of 'complex system sciences'. We argue that its current relation to classical physics as paradigmatic supplier of concepts, and to computer science as operational supplier of tools is outdated. We propose to regard computer science on the same level as physics, and to explore the scientific potential of concepts imported from theoretical computer science. In particular, we predict a major impact of the richer CS concept of behavior, seen as primary target of empirics and formalization rather than a mechanistic consequence of state. We exemplify this claim by putting the concepts of safety and liveness into ecological context. We show how they can be used as both metaphors and mathematical entities, informing scientific discourse, methodology and modeling.

3.20 Coalgebraic Infinite Games without Discounting – Towards Reflexive Economics

Viktor Winschel (Universität Mannheim, DE)

After a short introduction into the economic field of optimal currency areas and the current problems of the European Monetary Union we sketch the many reflexive issues in economic theory that result from the fact that economic modeling takes place within the modeled system. As a first example towards reflexive structures in economics we present coalgebraic infinite games. The technical part introduces the predicate coinduction proof principle that allows to proof subgame perfect equilibria defined as predicates on the carrier of coalgebras. The economic important result shows that this proof principle does not depend on the usual discounting that essentially transforms infinite structures into finite ones where backwards induction can be applied.

3.21 Probabilistic event structures

Glynn Winskel (University of Cambridge, GB)

This talk presented a new definition of probabilistic event structures, extending existing definitions, and characterized as event structures together with a continuous valuation on their domain of configurations. Probabilistic event structures possess a probabilistic measure on their domain of configurations. This prepares the ground for a very general definition of a probabilistic strategies, which are shown to compose, with probabilistic copy-cat strategies as identities. The result of the play-off of a probabilistic strategy and counter- strategy in a game is a probabilistic event structure so that a measurable pay- off function from the configurations of a game is a random variable, for which the expectation (the expected pay-off) is obtained as the standard Legesgue integral.

Participants

- Samson Abramsky
University of Oxford, GB
- Marcus Appleby
Perimeter Institute – Waterloo, CA
- David Balduzzi
MPI für Intelligente Systeme – Tübingen, DE
- Peter Bierhorst
Tulane University, US
- Robert J. Bonneau
AFOSR – Arlington, US
- Gunnar Carlson
Stanford University, US
- Eric Deeds
University of Kansas, US
- Ross Duncan
Université Libre de Bruxelles, BE
- Jerome Feret
ENS – Paris, FR
- Tobias Fritz
Institute of Photonic Sciences – Castelldefels, ES
- Richard Gill
Leiden University, NL
- Jonathan Hayman
ENS – Paris, FR
- Peter Hines
University of York, GB
- Ricardo Honorato-Zimmer
University of Edinburgh, GB
- Jean Krivine
University Paris-Diderot, FR
- Shane Mansfield
University of Oxford, GB
- Michael W. Mislove
Tulane University, US
- Catuscia Palamidessi
Ecole Polytechnique – Palaiseau, FR
- Prakash Panangaden
McGill University – Montreal, CA
- Mehrnoosh Sadrzadeh
University of Oxford, GB
- Sandro Stucki
University of Edinburgh, GB
- Baltasar Trancón y Widemann
Universität Bayreuth, DE
- Viktor Winschel
Universität Mannheim, DE
- Glynn Winskel
University of Cambridge, GB