Brief Announcement: Towards an Abstract Model of User Retention Dynamics

Eli Ben-Sasson
Department of Computer Science, Technion, Haifa, Israel
eli@cs.technion.ac.il
https://orcid.org/0000-0002-0708-0483

Eden Saig
Department of Computer Science, Technion, Haifa, Israel
edens@cs.technion.ac.il
https://orcid.org/0000-0002-0810-2218

Abstract

A theoretical model is suggested for abstracting the interaction between an expert system and its users, with a focus on reputation and incentive compatibility. The model assumes users interact with the system while keeping in mind a single *"retention parameter"* that measures the strength of their belief in its predictive power, and the system's objective is to reinforce and maximize this parameter through "informative" and "correct" predictions.

We define a natural class of *retentive scoring rules* to model the way users update their retention parameter and thus evaluate the experts they interact with. Assuming agents in the model have an incentive to report their true belief, these rules are shown to be tightly connected to truth-eliciting "proper scoring rules" studied in Decision Theory.

The difference between users and experts is modeled by imposing different limits on their predictive abilities, characterized by a parameter called *memory span*. We prove the *monotonicity theorem* ("more knowledge is better"), which shows that experts with larger memory span retain better in expectation.

Finally, we focus on the intrinsic properties of phenomena that are amenable to collaborative discovery with a an expert system. Assuming user types (or "identities") are sampled from a distribution D, the *retention complexity* of D is the minimal initial retention value (or "strength of faith") that a user must have before approaching the expert, in order for the expert to retain that user throughout the collaborative discovery, during which the user "discovers" his true "identity". We then take a first step towards relating retention complexity to other established computational complexity measures by studying retention dynamics when D is a uniform distribution over a linear space.

2012 ACM Subject Classification Theory of computation → Models of computation

Keywords and phrases information elicitation, proper scoring rules, retention complexity

Digital Object Identifier 10.4230/LIPIcs.ICALP.2018.164

Related Version `https://eccc.weizmann.ac.il/report/2017/160/`

Funding This work was supported by the Israeli Science Foundation under grant number 1501/14.

Acknowledgements We thank Yuval Filmus for many helpful discussions.

45th International Colloquium on Automata, Languages, and Programming (ICALP 2018).
Editors: Ioannis Chatzigiannakis, Christos Kaklamanis, Dániel Marx, and Donald Sannella;
Article No. 164; pp. 164:1–164:4
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Motivation

Aspiring gurus face the problem of attracting new followers, and *retaining* existing ones, as they journey together to a better future. This is an old problem. Moses, for instance, raised it before The Lord before assuming leadership of the Israelite Exodus from Egypt, asking: *"What if they won't believe me or listen to me?" [Exodus 4:1].* Many gurus resolve the problem by *predicting unlikely events* as a demonstration of their powers; the Biblical Exodus story contains several such events (culminating with the crossing of the Red Sea), all of which were predicted correctly by Moses.

In today's information society, crowd-based automated gurus gather data from users on a voluntary basis in order to produce meaningful insights. The quality of insights greatly depends on the amount and quality of data provided by the users, but those users have limited attention, giving rise to the study of *attention economy* [2, 4], and the design of interactive systems taking limited attention span into account. By asking *"interesting questions"* and making *"meaningful predictions"*, an automated interactive system can retain users, but only if it "knows" how to ask "interesting" questions and provide "meaningful" feedback.

The phenomenon that motivated this research is that of *early child development*; the gurus are experts in this field and the followers are parents of newborn babies [1]. For the sake of concreteness we shall continue using this particular setting to describe our model but it may be conveniently replaced by the reader with physicians or psychologists playing the gurus as they interact with patients (followers) regarding a complex medical or mental problem, or with financial advisors as gurus and their follower clientele. In these and similar settings, gurus and followers discuss a complex phenomenon that evolves over time, which the followers wish to understand, and about which the guru claims to have an advantage of *"wisdom"* over them.

The main goal of this work is the development of a clean mathematical model which mimics the retention dynamics of interactive systems, and can be used to explain why "smarter" gurus tend to retain a larger following. A clean mathematical model often sheds light on the studied phenomena, and may facilitate the future design of more efficient and successful automated gurus.

2 The Collaborative Discovery Model

The *phenomenon* about which the guru and her followers interact is modeled by a *distribution* $\mathcal{T}$ over $\mathcal{X}^\Gamma$, where Γ is the set of properties manifested by the phenomenon and $\mathcal{X}$ is an arbitrary input space. In the context of childhood development, Γ is the set of developmental milestones (like *"first smile"*), and each follower is represented by a sample $u \in \mathcal{X}^\Gamma$ that describes the ages at which that child achieved each milestone.

The guru and follower interact over a number of *rounds*: At the start of each round of interaction, the guru picks an undisclosed property $\gamma_t \in \Gamma$, and makes a prediction by announcing a distribution P_{γ_t} over $\mathcal{X}$ that she claims is the true one for a latent attribute $\gamma_t \notin \Gamma_t$; the follower has a distribution Q_{γ_t} that he believes corresponds to γ_t. After announcing both distributions, the true value $u_{\gamma_t} \in \mathcal{X}$ is revealed.

After each round, the follower updates the strength of his *retention* by the guru. We assume this strength is given by a *retention parameter* r_t that starts with a fixed value r_0 and varies with time; once r_t turns negative, the follower will be said to have lost all faith in the guru and therefore terminate the interaction. The main objective of the guru is to maintain $r_t \geq 0$ for all $t \geq 0$.

3 Main Contributions

Retentive Scoring Rules

The retention parameter r_t mentioned above and the dynamics that form around it are the key ingredients that give the model its expressive strength. Defining the retention parameter update rules, and the surprising corollaries of the chosen definitions, are what dominates the first part of our study.

We assume the retention parameter changes in an additive manner after each round, according to a function $\mathcal{S}(\cdot,\cdot,\cdot)$ that is real-valued and takes three inputs: (i) the guru's predicted distribution P_{γ_t}; (ii) the follower's assessment of that distribution Q_{γ_t}; and (iii) the value u_{γ_t} that materialized, picked by Nature. Analysis is further facilitated by assuming that $\mathcal{S}$ belongs to a class of functions that *elicit* the true beliefs of both guru and follower regarding the distribution of the attribute γ_t. Truth eliciting rules are ones that incentivize rational players to supply the rule with what they believe to be the truth.

Truthfulness under rationality is a powerful assumption which often leads to non-trivial corollaries. We characterize the functions that can act as retentive scoring rules, showing that such rules can take a surprisingly simple form.

Memory Span and Monotonicity

To model the different predictive capacities of gurus and followers, we characterize the forecasting abilities of agents in the Collaborative Discovery model by a parameter called *memory span*.

A variety of psychological studies could be summarized by saying that the human short-term memory has a capacity of about "seven, plus-or-minus two" *chunks*, where each chunk can be roughly defined as a collection of elementary information relating to a single concept [5, 6], where the "information capacity" of a chunk depends on the knowledge of the person being tested.

We model the discrepancy between users and experts by imposing the different limits on their memory spans. In this context, we show that the definition of memory span is *monotone*, verifying that experts with larger memory span retain followers longer in expectation.

Retention Complexity and Linear Codes

A distribution $\mathcal{T}$ for which there exists a guru that, in expectation, manages to retain followers to eternity (or until $t = |\Gamma|$ for finite Γ) will be said to be r_0-*retainable* and the *retention complexity* of $\mathcal{T}$ will be the minimal r_0 such that $\mathcal{T}$ is r_0-retainable.

To initiate the study of the retention complexity of specific distribution, a class of simple-to-understand, but non-trivial distributions, is needed. Inspired by other initial works, like that of Valiant which studied machine learning in the "restricted, but nontrivial context" of boolean functions [7] and that of Goldreich, Goldwasser and Ron that studied property testing in the context of graph properties [3], we begin by studying retention complexity of uniform distributions over linear spaces.

Uniform distributions over linear spaces are such a family and our final object of discussion. While such distributions are far from ones appearing in the "real world", studying them in this context provide convenient tools and insights about the Collaborative Discovery model, and the intuition and techniques we develop here will be generalized as we move to more "applied" settings.

References

1 Ayelet Ben-Sasson, Eli Ben-Sasson, Kayla Jacobs, and Eden Saig. Baby CROINC: An online, crowd-based, expert-curated system for monitoring child development. In *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare*, PervasiveHealth '17, pages 110–119, New York, NY, USA, 2017. ACM. doi:10.1145/3154862.3154887.

2 Michael H. Goldhaber. The attention economy and the net. *First Monday*, 2(4), 1997. doi:10.5210/fm.v2i4.519.

3 Oded Goldreich, Shafi Goldwasser, and Dana Ron. Property testing and its connection to learning and approximation. *Journal of the ACM (JACM)*, 45(4):653–750, 1998.

4 Richard A Lanham. *The economics of attention: Style and substance in the age of information.* University of Chicago Press, 2006.

5 George A Miller. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, 63(2):81, 1956.

6 Endel Tulving and Fergus IM Craik. *The Oxford handbook of memory.* Oxford: Oxford University Press, 2000.

7 Leslie G Valiant. A theory of the learnable. *Communications of the ACM*, 27(11):1134–1142, 1984.