Cognitive Ranking of Information Patches by Equiprobable Pure Exploration and Uncertainty-directed Search

Thomas WIEBRINGHAUS

Institute for Empiricism and Statistics (ifes)
FOM University of Applied Sciences for Economics and Management
Herkulesstraße 32, D-45127 Essen

Abstract: To optimize the increasing cooperative work between humans and machines, it is essential to examine the cognitive user behavior to adapt systems productively, e.g. for information retrieval or collaborative robotics. Learning from an environment requires a proper balancing and alternation of seeking modes. As too much *exploration* consumes resources and decreases reward, too much *exploitation* decelerates or prevents generalization. Empirical evidence is presented that a ranking of information complexity can already converge significantly after two rounds across only five participants in an information environment. Results suggest that a uniform exploration was followed by an uncertainty-directed exploitation activity to maximize cumulative information by maintaining individual aspiration level (unconsciously) relatively constant.

Keywords: explore/ exploit, uncertainty, aspiration level

1. Introduction

Based on estimations, artificial agents would need about 25,000 samples to learn with *certainty* in a two-alternative event (Vul et al. 2014; Stigler 1986). Human cognition, equipped with a diverse set of heuristics, only needs very few samples or a few episodes of learning by interaction to make a near-optimal decision (Stigler 1986; Goodman et al. 2008; Mozer et al. 2008; Vul & Pashler 2008; Vul et al. 2014).

Deriving an algorithm from a cognitive strategy is fruitful not only for cognitive sciences, but also for machine learning and its applications (Neftci & Averbeck 2002). As e.g. Reinforcement Learning enters real world applications like healthcare, the safety and effectiveness of a new substance in a medical trial has to be evaluated in a short time with very few patients. Algorithms therefore need to generalize quickly (e.g. few-shot learning (Kadam & Vinay 2018)).

Learning from an uncertain environment requires *exploration* of alternatives to gather knowledge and *exploitation* of already known items to deepen knowledge or to gain reward (Cohen et al. 2007). Human beings deal with this *explore-exploit dilemma* with a directed exploration targeted towards options with highest uncertainty and a random undirected exploration (Wilson et al. 2014). The balancing (Wiebringhaus 2020) and proper alternation of the seeking modes determine algorithm performance and fast convergence.

In addition to individual working memory capacity (Wiebringhaus 2018a & 2018b) that affects the individual sampling size (Vul et al. 2014), users inspect information patches with diverse aspiration levels, different thoroughness or reward expectations. However, a heterogenous combination of people might explore the hypothesis space rational (wisdom-of-the-crowd) (Mozer et al. 2008; Vul & Pashler 2008).

It is currently unknown how human agents evaluate the depleting resource '*information'* within only a few steps of interaction. The study aims to examine user strategies for a ranking of depleting information patches observed in Wiebringhaus (2019). As human cognition is sophisticated, very small numbers of participants and episodes might be sufficient to evaluate the unknown state space.

2. Methods

The task was to inspect and estimate the optimal patch leaving time for four consecutive video-tutorials for learning basic steps of an unknown software. Each tutorial was looped as a GIF (four loops = one round). Users accomplished two rounds without skipping back, with the rules to skip to next when satisficed (Simon 1955) (e.g. when repetition predominates, ~50% understood) in round one, and to skip under a maximizing policy when feeling that a tutorial content is understood to ~90% in round two. Afterwards participants were asked to estimate durations for the four tutorials that might be optimal regarding patch complexity. Details of the experiment are described in Wiebringhaus (2019).

Undirected explorative search is viewed here under a maximum entropy (Jaynes 1957) assumption *1/k* with *k* as number of patches. Users have no information about the density of certain patches and thus might allocate similar time. User behavior converges as patch times significantly deviate from a uniform equiprobable distribution across participants. Convergence is observed here as an increase in the number of significant deviating patches per round. Shapiro-Wilk normality test and one-sample t-test were applied to analyze deviation from uniform distribution *1/k*.

3. Results

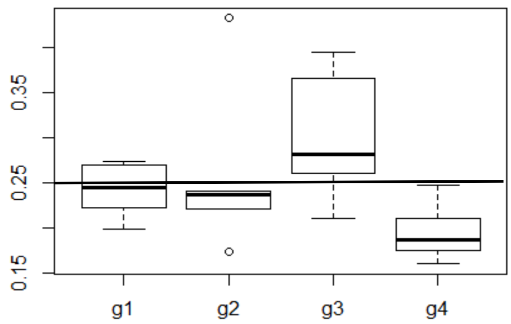
The aggregated mean intra-patch time of participants is shown in table 1. The allocated time mirrors patch density with the ranking of *3>2>1>4*. The ranking is already established in round 1 and stable through round 2, and persists also in the estimation of the optimum (est) in the same order.

Table 1. *Relative times: Relative satisficed (round 1), relative maximized (round 2), relative estimate for an optimum (est); *significance level 0.05; ** significance level 0.01;*

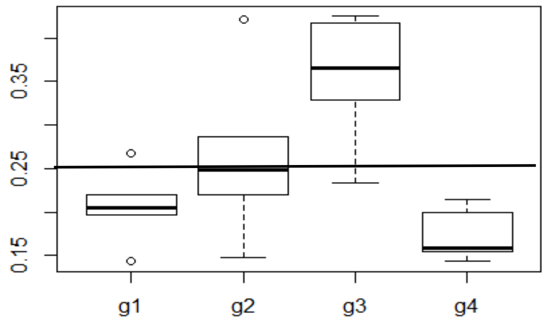
n=5	round 1	round 2	est	mean	sum	max-sat
1	0,254	**0,200***	**0,21***	0,227	0,454	-0,054
2	0,266	0,254	0,25	0,26	0,52	-0,012
3	0,283	**0,350***	**0,35***	0,317	0,633	0,067
4	**0,197****	**0,196****	**0,19****	0,197	0,393	0,001
#sig	1	3	3			

The three first patch visits are close to *1/k* and not significantly different from 0.25. The first deviation begins with the visit of patch number 4.

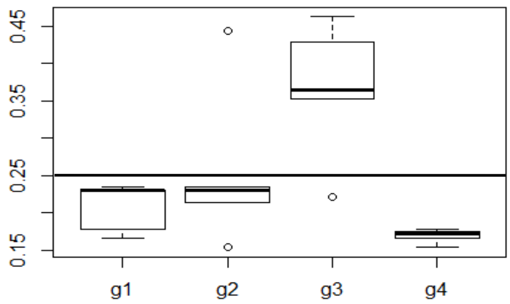
Uniform exploration was followed by an uncertainty-directed fine-tuning activity to maximize cumulative information by maintaining individual cumulative time costs relatively constant. Depleting resources quickly converge significantly across five participants in less than two rounds (table 1 and figure 1).



1a



1b



1c

Figure 1a-c. *Boxplots. a) round one satisficed ~50%; b) round two maximized ~90%; c) estimated optimum. Exploration ratio is shown on y-axis. Solid line represents 1/k at 0.25*

4. Discussion

Empirical evidence is shown that user judgment about patch density (information complexity/ cognitive workload) already converges significantly after two rounds across five participants. Results suggest that users begin with an allocation of intra-individually approximately equal time for a pure- exploration of each patch and switch to uncertainty-driven exploitation.

The exploration phase observed here is reminiscent to the fixed-time strategy found in animal optimal foraging theory in ecology (Stephens & Krebs 1986). The difference here is that invested time is not fixed absolutely but rather relative at the beginning and adapts immediately to new information.

Training und performance are not separated processes here, as typical in machine learning algorithms, e.g. in supervised learning (Bishop 1995). The process here is more similar to *active learning* (Osugi et al. 2005; Settles 2009; Bouneffouf et al. 2014) and confirms also the usefulness of the broad employment of the equiprobable random policy in Reinforcement Learning (Sutton & Barto 1998).

The results might be utilized as a usability test for optimizing video-tutorial cognitive workload. As the wisdom-of-(small)-crowds-effect is observed as valuable here, and designing reward functions in Few-Shot Learning (FSL) and Small Sample Learning (SSL) in Reinforcement Learning is difficult for real world applications, further studies might focus on mimicking several individual fixed-time strategies to estimate the optimal policy with a multi- agent system.

5. References

Bishop, Christopher M. *Neural networks for pattern recognition*. Oxford university press, 1995.

Bouneffouf, Djallel, et al. "Contextual bandit for active learning: Active thompson sampling." International Conference on Neural Information Processing. Springer, Cham, 2014.

Cohen, Jonathan D., Samuel M. McClure, and Angela J. Yu. "Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration." *Philosophical Transactions of the Royal Society B: Biological Sciences* 362.1481 (2007): 933-942.

Goodman, Noah D., et al. "A rational analysis of rule-based concept learning." *Cognitive science* 32.1 (2008): 108-154.

Jaynes, Edwin T. "Information theory and statistical mechanics." *Physical review* 106.4 (1957): 620.

Kadam, Suvarna, and Vinay Vaidya. "Review and Analysis of Zero, One and Few Shot Learning Approaches." *International Conference on Intelligent Systems Design and Applications*. Springer, Cham, 2018.

Mozer, Michael C., Harold Pashler, and Hadjar Homaei. "Optimal predictions in everyday cognition: The wisdom of individuals or crowds?." *Cognitive Science* 32.7 (2008): 1133-1147.

Neftci, Emre O., and Bruno B. Averbeck. "Reinforcement learning in artificial and biological systems." *Nature Machine Intelligence* 1.3 (2019): 133-143.

Osugi, Thomas, Deng Kim, and Stephen Scott. "Balancing exploration and exploitation: A new algorithm for active machine learning." Fifth IEEE International Conference on Data Mining (ICDM'05). IEEE, 2005.

Pirolli, Peter, and Stuart Card. "Information foraging." *Psychological review* 106.4 (1999): 643.

Simon, Herbert A. "A behavioral model of rational choice." *The quarterly journal of economics* 69.1 (1955): 99-118.

Settles, Burr. *Active learning literature survey*. University of Wisconsin-Madison Department of Computer Sciences, 2009.

Stephens, David W., and John R. Krebs. *Foraging theory*. Princeton University Press, 1986.

Sutton, Richard S., and Andrew G. Barto. *Introduction to reinforcement learning*. Vol. 2. No. 4. Cambridge: MIT press, 1998.

Stigler, Stephen M. *The history of statistics: The measurement of uncertainty before 1900*. Harvard University Press, 1986.

Vul, Edward, and Harold Pashler. "Measuring the crowd within: Probabilistic representations within individuals." *Psychological Science* 19.7 (2008): 645-647.

Vul, Edward, et al. "One and done? Optimal decisions from very few samples." *Cognitive science* 38.4 (2014): 599-637.

Wiebringhaus, Thomas “Stability of Verbal Working Memory Capacity Limitations under Standard Language Demands” 11th Forum of Neuroscience FENS Berlin, Germany, 2018a

Wiebringhaus, Thomas “Exploring Heterogeneity of Individual Cognitive Workload and Capacity Limitations in a Consecutive Ascending Task” Gesellschaft für Arbeitswissenschaften (GfA) Conference Proceedings B.4.6 Frankfurt, Germany, 2018b

Wiebringhaus, Thomas “Studying Individual Satisficing-Optimizing Decisions in an Uncertain Explore/ Exploit Information Environment” Gesellschaft für Arbeitswissenschaften (GfA) Conference Proceedings B.7.4 Dresden, Germany, 2019

Wiebringhaus, Thomas “Human Strategic Balancing or Random Sampling in Exploration and Exploitation Decisions under Continuous Uncertainty in Web Search Interaction” Gesellschaft für Arbeitswissenschaften GfA Conference Proceedings B.17.7 Berlin, Germany, 2020

Wilson, Robert C., et al. "Humans use directed and random exploration to solve the explore–exploit dilemma." *Journal of Experimental Psychology: General* 143.6 (2014): 2074.

Please cite as: Wiebringhaus, Thomas “Cognitive Ranking of Information Patches by Equiprobable Pure Exploration and Uncertainty-directed Search” Gesellschaft für Arbeitswissenschaften (GfA) Conference Proceedings B.17.6 Berlin, Germany, 2020

Gesellschaft für
Arbeitswissenschaft e.V.

Digitale Arbeit, digitaler Wandel, digitaler Mensch?

66. Kongress der
Gesellschaft für Arbeitswissenschaft

TU Berlin
Fachgebiet Mensch-Maschine-Systeme

HU Berlin
Professur Ingenieurpsychologie

16. – 18. März 2020, Berlin

GfA-Press

Bericht zum 66. Arbeitswissenschaftlichen Kongress vom 16. – 18. März 2020

TU Berlin, Fachgebiet Mensch-Maschine-Systeme
HU Berlin, Professur Ingenieurpsychologie

Herausgegeben von der Gesellschaft für Arbeitswissenschaft e.V.
Dortmund: GfA-Press, 2020
ISBN 978-3-936804-27-0

NE: Gesellschaft für Arbeitswissenschaft: Jahresdokumentation

Als Manuskript zusammengestellt. Diese Jahresdokumentation ist nur in der Geschäftsstelle erhältlich.

Schriftleitung: Matthias Jäger
im Auftrag der Gesellschaft für Arbeitswissenschaft e.V.

Screen design und Umsetzung
© 2020 fröse multimedia, Frank Fröse
office@internetkundenservice.de · www.internetkundenservice.de