Incentivizing and Coordinating Exploration

(Tutorial proposal for *ALT 2019*)

Robert D. Kleinberg* Aleksandrs Slivkins†

October 2018

While exploration-exploitation tradeoffs are well-studied, in many scenarios exploration is performed by self-interested individuals (*agents*) who make their own decisions. For example, a decision to dine in a given restaurant may reveal some observations about this restaurant that others may benefit from (either directly, via photo, review, tweet, etc., or indirectly through rankings and recommendations). The exploring agent usually bears the full cost of exploration, either as a direct monetary cost or as an opportunity cost, but the full benefit is shared across many. If a social planner were to guide the agents, she'd balance exploration and exploitation for the sake of the common good. A platform such as Yelp or Amazon may wish to coordinate the agents in a similar way, as increasing the overall customer satisfaction would align with the platform's business objectives.

However, when the decisions are made by agents, we face under-exploration and selection bias. Agents' incentives are typically skewed in favor of exploitation, as many people prefer to benefit from exploration done by others, so the society as a whole may suffer from insufficient exploration. In particular, the best alternative may remain unexplored if it appears suboptimal initially. Further, agent's properties (observable or not) may affect both her decisions and the outcomes of these decisions. This leads to selection bias, which may render the data invalid for statistical analysis.

These issues are broadly relevant to the numerous recommendation platforms.[1] They also arise in economics, when auction/market participants face uncertainty that can be resolved by exploration: *e.g.,* in "matching markets" for college admissions, medical residency admissions and academic jobs, and in large-scale deals such as start-up acquisitions and real-estate purchases.

Thus, there is a need for algorithms that incentivize and coordinate exploration while respecting agents' incentives. A recent surge of research [10, 3, 5, 6, 11, 12, 1, 9, 2, 13, 4, 7, 8], coming from computer science, economics and operations research, has achieved remarkable progress on the theoretical side, making substantial modeling assumptions for the sake of elegance and tractability. However, bringing these approaches closer to practice presents new and interesting challenges, and we believe machine learning community is the right one to address them.

The proposed tutorial introduces this problem space to the machine learning audience. Throughout, an algorithm interacts with self-interested agents whose actions may reveal information not

*Computer Science Department, Cornell University, Ithaca NY, USA. Email: rdk@cs.cornell.edu.

†Microsoft Research, New York NY, USA. Email: slivkins@microsoft.com.

[1] *E.g.,* for movies (*Netflix*), restaurants (*Yelp*), vacations (*TripAdvisor*), products (*Amazon*), driving (*Waze*), etc.

previously known to themselves, the algorithm, or other agents. The information-revealing actions are directly controlled by the agents, whereas the principal can only influence them via signals (*e.g.,* recommendations) and/or monetary transfers. Principal and/or the agents can *learn*: aggregate and subsequently use the new information revealed by agents' actions. Absent incentives, these models reduce to various multi-armed bandit problems. Our focus is on a three-way tradeoff between exploration, exploitation, and agents' incentives.

The tutorial is structured as follows. One half covers the work involving time-discounted rewards and monetary transfers [5, 6, 9, 4], focusing on the material in [5, 9] and drawing on a deep generalization of the classic Gittins algorithm for multi-armed bandits. The other half is on scenarios with no time-discounting and no monetary transfers, covering the progression of papers [10, 11, 12, 1, 7, 8]. While this work uses Bayesian priors to model beliefs and incentives, their algorithms build on the rich literature on regret minimization. Both halves carefully discuss motivating examples and connections to ML, and emphasize opportunities for further progress.

Previously taught tutorials. Earlier version of this tutorial was presented at ACM EC 2017 for the algorithmic economics audience. Slides can be found at `https://www.microsoft.com/en-us/research/people/slivkins/#!teaching`. Alex Slivkins has co-taught a tutorial at ACM EC 2015 on dynamic pricing [14].

Tutor biographies

Bobby Kleinberg is an Associate Professor of Computer Science at Cornell University. He was also a researcher at Microsoft Research New England from 2014 to 2016. His research in general pertains to the design and analysis of algorithms, and their applications to economics, machine learning, networking, and other areas. Prior to receiving his doctorate from MIT in 2005, Kleinberg spent three years at Akamai Technologies, where he assisted in designing the world's largest Internet Content Delivery Network. He is the recipient of a Microsoft Research New Faculty Fellowship, an Alfred P. Sloan Foundation Fellowship, and an NSF CAREER Award. His research has received the best paper awards at ACM EC 2010 and 2014.

Alex Slivkins is a Senior Researcher at Microsoft Research New York. Previously he was a researcher at MSR Silicon Valley in 2007-2013, after receiving his Ph.D. from Cornell in 2006 and a brief postdoc at Brown. His research interests are in algorithms and theoretical computer science, spanning machine learning theory, algorithmic economics, and networks. Alex is particularly interested in exploration-exploitation tradeoff and online machine learning, and their manifestations in mechanism design and human computation. His work has been recognized with the best paper award at ACM EC 2010, the best paper nomination at WWW 2015, and the best student paper award at ACM PODC 2005.

References

[1] G. Bahar, R. Smorodinsky, and M. Tennenholtz. Economic recommendation systems. In *16th ACM Conf. on Electronic Commerce (EC)*, 2016.

[2] K. Bimpikis, Y. Papanastasiou, and N. Savva. Crowdsourcing exploration. *Management Science*, 64:1477–1973, 2018.

[3] Y.-K. Che and J. Hörner. Optimal design for social learning. *Quarterly Journal of Economics*, 2018. Forthcoming. First published draft: 2013.

[4] B. Chen, P. I. Frazier, and D. Kempe. Incentivizing exploration by heterogeneous users. In *Conf. on Learning Theory (COLT)*, pages 798–818, 2018.

[5] P. Frazier, D. Kempe, J. M. Kleinberg, and R. Kleinberg. Incentivizing exploration. In *ACM Conf. on Economics and Computation (ACM EC)*, pages 5–22, 2014.

[6] L. Han, D. Kempe, and R. Qiang. Incentivizing exploration with heterogeneous value of money. In *11th Intl. Conf. on Web and Internet Economics (WINE)*, pages 370–383, 2015.

[7] N. Immorlica, J. Mao, A. Slivkins, and S. Wu. Bayesian exploration with heterogenous agents. Working paper, 2018.

[8] N. Immorlica, J. Mao, A. Slivkins, and S. Wu. Incentivizing exploration with unbiased history. Working paper, 2018.

[9] R. D. Kleinberg, B. Waggoner, and E. G. Weyl. Descending price optimally coordinates search. Working paper, 2016. Preliminary version in *ACM EC 2016*.

[10] I. Kremer, Y. Mansour, and M. Perry. Implementing the wisdom of the crowd. *J. of Political Economy*, 122:988–1012, 2014. Preliminary version in *ACM EC 2014*.

[11] Y. Mansour, A. Slivkins, and V. Syrgkanis. Bayesian incentive-compatible bandit exploration. Working paper, 2018. Available at `https://arxiv.org/abs/1502.04147`. Preliminary version in *ACM EC 2015*. To be published in *Operations Research*.

[12] Y. Mansour, A. Slivkins, V. Syrgkanis, and S. Wu. Bayesian exploration: Incentivizing exploration in bayesian games. Working paper, 2018. Available at `https://arxiv.org/abs/1602.07570`. Preliminary version in *ACM EC 2016*.

[13] S. Schmit and C. Riquelme. Human interaction with recommendation systems. In *Intl. Conf. on Artificial Intelligence and Statistics (AISTATS)*, pages 862–870, 2018.

[14] A. Slivkins and A. Zeevi. Dynamic pricing with model uncertainty, 2015. Tutorial at *ACM EC 2015*. Slides can be found at `http://slivkins.com/work/EC15-tutorial-merged.pdf`.