Chapter 5

Conclusion

The underlying motivation behind much applied statistical and machine learning work is to guide us to make better decisions. In many cases, the actions that we take in response to the model will change the system from which the data was generated. It is critical that we are able to recognise the causal nature of such problems and appropriately model the decision making part of the process. If machine learning is to be as transformative in fields such as economics, medicine and social science as it has been for image recognition, voice processing and machine translation, we must develop methods to estimate the effect of, and optimally select, interventions that are as effective as those we have for pattern recognition. We need to bridge the gap between the theory driven models of economics and science and the black box prediction approach that has been so successful in machine learning. This will involve clarifying what information is (currently) best encoded by theory and what can be successfully inferred from data and developing methods that can incorporate the theory or structure required to allow models to generalise from one setting to another whilst retaining the flexibility to capture complex patterns in empirical data.

A better understanding of causality is also relevant for the discussions around transparency and ethics in machine learning, particularly with respect to the European Union's new General Data Protection Regulation, which requires that automated decision systems that significantly affect individuals provide provide *"meaningful information about the logic involved."* [**?**]. The recognition that there is a fundamental trade-off between accuracy and transparency, unless we can build perfect causal models, when the interests of individuals and society diverge has implications for they way we design and regulate systems that have the potential to have major impacts on people's lives. We must develop approaches to ensuring machine learning decisions are reasonable and ethical that allow effected individuals recourse to dispute or improve outcomes but do not undermine the ability of the system to function.

The observational and interventional viewpoints on learning to act contribute complementary components to a general approach. Observational causal inference provides, through the do-calculus, a formal means to map information from observational to interventional settings, as well as between different interventions. Bandit algorithms,

I have developed a framework that formally connects causal graphical models with bandit problems in a very natural way and demonstrated that this framework encode some key existing problems in the literature. I showed that knowledge of the causal structure (but not the functional relationships) between variables can induce a novel form of structure between alternate actions and that an algorithm that leverages this structure obtains better performance that one that does not.

This work represents an important first step towards a unified approach to causal inference and optimal decision making. There is much exciting work remaining to be done. Although the causal bandit framework can capture contextual information as well as post-action feedback, I have formally analysed and developed algorithms only for the latter. Additionally, to make the problem tractable, I made the (major) assumption that the interventional distribution over the parents of the outcome was known. This can be relaxed to assuming the interventional distribution over some Markov blanket with respect to the outcome is known. Information can then be shared between actions outside the blanket in the same way as in algorithm 3, whilst actions inside could be learned explicitly. Relaxing this assumption entirely is a much more challenging problem. However, as is demonstrated by the specific example of the parallel causal bandit problem, it is possible to develop algorithms that require only the casual structure of the graph and yield substantially lower regret.

Another interesting line of research is the question of off-policy evaluation for causal-bandit problems. As in the online case, knowledge of the causal structure between variables in the graph provides additional information about the reward for the actions that were not selected at each timestep. The problem differs from typical observational causal inference in several ways; the focus is on identifying an optimal policy, rather than unbiased estimation of all policies, the goal is to explore the value of interventions on a range of different variables, rather than the optimal setting of a single variable, and the data will, in general, be non-stationary in a rather special way due to the adaptive nature of policy that generated it. Existing work on off-policy evaluation focuses (at least implicitly) on estimating causal effects by adjusting for all variables that simultaneously effect both the action selection and outcome. If additional information is available about the causal relationships between variables involved, other approaches to identifying causal effects such as instrumental variables or the “front door method” [112] could be also be applied to off-policy evaluation.

Insights from the bandit literature can also be applied to more classical causal inference problems. In particular, estimators that are geared towards optimal action selection (rather than evaluation of all actions) and approaches to quantify the finite time properties, as opposed to asymptotic efficiency, of estimators, for example [102].

An important line of research, that is relevant to both the observational and interventional approaches to causal inference, is developing methodologies for model evaluation and selection that provide something equivalent to what cross-validation does for supervised learning.

My hope is that in the next years, combining the reinforcement learning approach to decision making, with causal graphical models and causal effect estimation techniques developed within statistics and economics, will revolutionise our ability to make good data driven decisions.