Communication, coordination and competition in causal problem solving

1. Motivation

1.1. Introduction

It has been argued that humans perceive and interpret the world through the lens of a causal model (e.g. Sloman, 2005; Steyvers et al., 2003). Such a causal model can help explain why observed events occurred (Coenen et al., 2015; Meder et al, 2014) and help predict what will happen next (Clark, 2014). Humans also have an innate willingness to ask "why" questions and seek reasons and explanations underlying the phenomena they encounter in the world (Bramley et al., 2015; Schmidhuber, 2010).

One important question is what is the relationship between this causal representation inside the head, and language as a means for asking causal questions and communicating causal insights? How children learn language and generate novel utterances is a key problem for cognitive science (e.g. Chater & Manning, 2006; Quine, 1960; Smith et al., 2017). Language helps each generation pass its discoveries on and helps people pass ideas between one another. Therefore, intuitively it must be a suitable medium for transferring causal representation from one mind to another.

In my PhD, I propose to explore how people communicate during group causal reasoning problems. Broadly, individuals must reason about one another's knowledge and attempt to share insights order to maximize the influence among other agents in the environment and ultimately task specific reward. My goal is to shed light on the understanding of how human share beliefs through communication during joint problem solving. I propose to analyse this problem using a combination of active learning (Settles, 2012), multi-agent reinforcement learning (Shoham, 2003) and game theory (Nash, 1950).

1.2. The multi-agent Reinforcement Learning framework

Reinforcement learning (RL) deals with agents acting in an environment with the goal of establishing a policy (i.e. state-behaviour mapping) that maximises their expected future rewards (Sutton & Barto, 2018). While a popular framework in theoretical neuroscience (from whence it originates) RL has not been applied extensively in the study of higher-level cognition. Recent advancements have explored intrinsic reward signals—like "curiosity" —that can drive learning in the absence of well-shaped extrinsic rewards (Pathak et al., 2017; Schmidhuber, 2010). Multi-agent RL generalises the formalism to settings with multiple agents that might share or differ in what they find rewarding. Each agent observes the environment, infers the state of the world based on the belief (i.e. agents state can differ at each time), and choose an action/policy maximizing expected cumulative reward. In multi-agent RL setting, it is possible to train agents to output statements in a shared or initially undetermined language, e.g. with the intrinsic goal of exerting causal influence on other agents in the environment. <u>Causal influence assessment can take place through comparison between observed outcomes and simulated counterfactual outcomes</u> (Lewis, 1979; Pearl, 2009; Jaques et al., 2018).[1] In this setting, agent language acquisition and understanding is grounded and embedded in collective group pursuit (e.g. Gauthier & Mordatch, 2016; Hermann et al., 2017; Mikolov et al., 2015), i.e. providing a potential explanation for how lexical concepts relate to the real world by acting into the world. This approach is called intrinsic social motivation via causal influence and has great potential as a framework within which to study communication and emergent coordination within natural human group settings.

2. Doctoral focus

My goal is to explore the interplay between causal problem solving, counterfactual reasoning and communication. First, I will explore communication in a novel multi-participant learning task in which each participant has only partial information and communication is limited to an unfamiliar system. In parallel, I plan to explore how communication systems emerge through interaction between artificial agents in the same problems, hoping this will provide insight into the algorithms and representations required for success. Finally, I plan to explore coordination and communication in environments involving both human and artificial agents where both must communicate in order to succeed. Both human and artificial agents must be able to infer goals, intentions and beliefs of each other for proper communication. Progress toward this ambitious goal will require a number of experiments and computational modeling.

2.1. proposed studies

Asking the right question is a crucial ability; i.e. humans can minimize the time for acquiring information by asking the right questions. For example, if someone wants to know about quantum gravity, she can either search the libraries for suitable references or directly ask experts for useful resources. But asking question comes with one drawback: you should be able to find a person with the right amount of knowledge (e.g. you should ask about quantum gravity from an expert or professors who is working in this area instead of a random person who is walking in the street). In this study set, I am going to explore how people assess the knowledge of others by observing their behaviour (Premack & Woodruff, 1978) and come up with the right question to acquire the maximum amount of information.

[1] A counterfactual is a conditional statement in which the conditional clause is false (Lewis, 1979). Counterfactuals are the inferences about what *would have happened* had the past differed in some way e.g. "I would have been on time if it hadn't been raining".

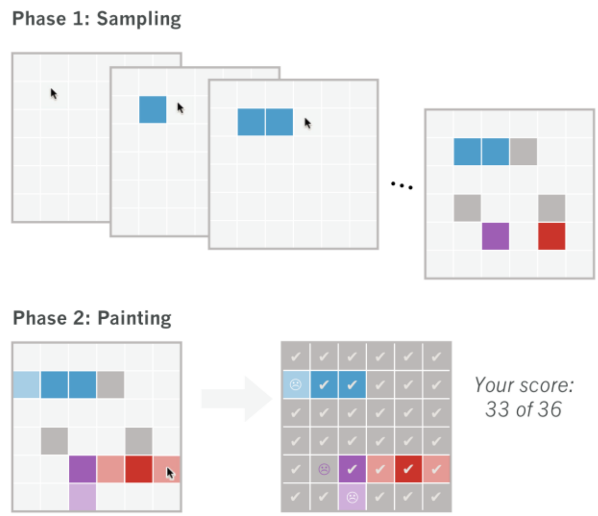

Figure 1: Battleship Game. In the sampling phase, participants can click on each tile and unveil the colour behind them. During painting phase, they should be able to determine the colours of the remaining tiles correctly in order to maximize their score (Rothe et al., 2018).

For my initial experiments, I plan to build on the battleship task explored in Rothe et al. (2018, see Figure 1). The goal of the task is to efficiently determine the location and size of three non-overlapping rectangular "ships" on a grid world by probing individual tiles. To explore multi-agent learning, I propose to have four participants learn simultaneously either cooperatively (Experiment 1) or competitively (Experiment 2-3). Each participant will be given different initial clues (e.g. perhaps blue ships are always oriented horizontally or blue and red ships are always touching each other). Thus, to a perform well as a group, participants must find ways to communicate their unique insight and infer that of others.

<u>Experiment 1. Cooperative communication</u>: Each participant will choose several tiles to reveal during an initial sampling phase as well as observing the tiles unveiled by others. Also, each person can ask a question in natural language from others. Participants are rewarded according to how accurately all four perform at locating the ships in the test phase. Thus, each participant should try to infer what others know about the game by observing their movement and based on that decide to ask the right question from the right person. The goal of this study is to find out how people make inference about others belief and knowledge and use this inference to come up with the most informative questions.

<u>Experiment 2. Adversarial communication:</u> In order to explore settings in which learners have mismatched goals, we propose to change participants reward structure e.g. such that_the number of ship tiles they colour in correctly plus the number of tiles others have coloured mistakenly. In this setting, each person can ask five questions during sampling and each person must answer honestly 60% of the time (i.e. must answer honestly 3 out of 5 question). The question here is again exploring how people behave in this setting and what is the optimal strategy. Moreover, compare this strategy with the multi-agents RL doing the same task.

Experiment 3. Repeated learning and the formation of trust and shared language: Trust becomes important if a scenario is faced repeatedly. In this experiment, group of participants repeat the task over ten trials and are rewarded only based on their own performance. In each iteration, each participant knows one of the hidden rules of the environment (e.g. iteration 1: there is one blue tile. iteration 2: red tiles touch the blue tile). Also, they can ask two questions during each trial. The score for determining who is the winner of the game is the score of the last painting phase. In this experiment, the participant can lye freely in order to win the game at the end. The aim of this experiment is to realize how people ask the right question in this repeated game and whether they trust each other or not? Another question to explore is what proportion of lies and truths emerge.

Experiment 4. Mixtures of human and artificial learners: An interesting issue for the near future is coordination between human and artificial agents. We propose to explore this in a line of studies building on those above but mixing real and artificial agents, asking under what conditions successful communication and coordination emerges, and what role human vs artificial agents play in successful teams.

2.2. proposed Timescale

1st year	2nd year	3rd year
Experiment 1. Comprehensive reviewing of multi-agent RL literature.	Experiments 2 & 3.	Experiment 4. Gathering all the finding and write up.

3. Summary

Linking embodied and grounded experience of the world and causal interpretation of the world to the language we are using to communicate is an important and understudied topic. During my PhD studies, I hope to clarify this link by proposing and testing models in which individuals and artificial agents communicate by inferring intentions, belief and goals while communicating and attempting to coordinate in joint problem solving and learning situations. This work promises to contribute to the development of more powerful and humanlike artificial intelligence, better equipped to integrate and contribute to society.

4. References

Bramley, N. R., Lagnado, D. A. & Speekenbrink, M. (2015). Conservative forgetful scholars: How people learn causal structure through interventions. Journal of Experimental Psychology: Learning, Memory & Cognition, 41(3), 708-731

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behavioral and brain sciences, 36(3), 181-204.

Coenen, A., Rehder, R., & Gureckis, T. M. (2015). Strategies to intervene on causal systems are adaptively selected. Cognitive Psychology, 79, 102-133.

Chater, N., & Manning, C. D. (2006). Probabilistic models of language processing and acquisition. Trends in cognitive sciences, 10(7), 335-344.

Gauthier, J., & Mordatch, I. (2016). A paradigm for situated and goal-driven language learning. arXiv preprint arXiv:1610.03585.

Hermann, K. M., Hill, F., Green, S., Wang, F., Faulkner, R., Soyer, H., ... & Wainwright, M. (2017). Grounded language learning in a simulated 3D world. arXiv preprint arXiv:1706.06551.

Jaques, N., Lazaridou, A., Hughes, E., Gulcehre, C., Ortega, P. A., Strouse, D. J., ... & de Freitas, N. (2018). Intrinsic Social Motivation via Causal Influence in Multi-Agent RL. arXiv preprint arXiv:1810.08647.

Lewis, D. (1979). Counterfactual dependence and time's arrow. Noûs, 455-476.

Meder, B., Mayrhofer, R., & Waldmann, M. R. (2014). Structure induction in diagnostic causal reasoning. Psychological Review, 121(3), 277.

Mikolov, T., Joulin, A., & Baroni, M. (2016). A roadmap towards machine intelligence. In International Conference on Intelligent Text Processing and Computational Linguistics, 29-61. Springer, Cham.

Nash, J. F. (1950). Equilibrium points in n-person games. Proceedings of the national academy of sciences, 36(1), 48-49.

Pathak, D., Agrawal, P., Efros, A. A., & Darrell, T. (2017). Curiosity-driven exploration by self-supervised prediction. In International Conference on Machine Learning (ICML) (Vol. 2017).

Pearl, J. (2009). Causality. Cambridge university press.

Pearl, J. & Mackenzie, D. (2018). The Book of Why: The New Science of Cause and Effect. New York: Basic Books (1st edition).

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? Behavioral and brain sciences, 1(4), 515-526.

Quine, W. V. O. (1960). Word & Object. MIT Press.

Rothe, A., Lake, B. M., & Gureckis, T. M. (2018). Do people ask good questions? Computational Brain & Behavior, 1(1), 69-89.

Schmidhuber, J. (2010). Formal theory of creativity, fun, and intrinsic motivation (1990–2010). IEEE Transactions on Autonomous Mental Development, 2(3), 230-247.

Settles, B. (2012). Active learning. Synthesis Lectures on Artificial Intelligence and Machine Learning, 6(1), 1-114.

Shoham, Y., Powers, R., & Grenager, T. (2003). Multi-agent reinforcement learning: a critical survey (pp. 1-13). Technical report, Stanford University.

Sloman, S. (2005). Causal models: How people think about the world and its alternatives. Oxford University Press.

Smith, K., Perfors, A., Fehér, O., Samara, A., Swoboda, K., & Wonnacott, E. (2017). Language learning, language use and the evolution of linguistic variation. Phil. Trans. R. Soc. B, 372(1711), 20160051.

Steyvers, M., Tenenbaum, J. B., Wagenmakers, E. J., & Blum, B. (2003). Inferring causal networks from observations and interventions. Cognitive science, 27(3), 453-489.

Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction. MIT press.