Information parsimony in collaborative interaction

Dari Trendafilov and Daniel Polani

Adaptive Systems Research Group
University of Hertfordshire, Hatfield, UK
{d.trendafilov,d.polani}@herts.ac.uk

Introduction

We propose an information-theoretic approach for deriving predictive collaborative strategies in a dyad. Social interaction and coordination in robotics have been initially studied by Walter (1950) in natural and artificial agents and more recently by Dautenhahn (1995); Ikegami and Iizuka (2007); Paolo et al. (2008); Goldstone and Janssen (2005). Our framework builds on the concept of relevant information (Polani et al. (2001)), while not assuming any intrinsic dynamics nor a particular metabolism, instead of imposing certain information processing constraints. Information theory provides a universal language to quantify conditions and invariants for a large class of models in a generic and principled way and allows to compare quantities of models that are otherwise not directly comparable.

To study agent coordination from an information-theoretic perspective towards a predictive and quantitative theory of agent interactions, we consider embodied agents with independent controllers and a redundant set of actions for controlling an object in a grid-world, motivated by earlier studies (see Wahn et al. (2016)). Using information-theoretic tools we derive theoretical predictions for the agents' behavior, constrained by the level of cooperation of its partner.

When agents interact socially, sometimes they act in a complete agreement towards a shared goal, and other times their goals may diverge or be completely incompatible with each other. Limitations in the perception–action loop or in the decision-making capabilities (e.g. autistic behavior) could make the agents behave irrationally during joint action. We investigate how a rational agent would adapt its behavior to its irrational (or adverse) partner under information-processing constraints, using the Relevant Information method (Polani et al. (2006)), which provides a measure quantifying the minimal amount of information an agent needs to process in order to achieve a certain level of utility as specified by a reward function.

Experiment

In our experimental scenario two agents with independent controllers jointly perform a task using redundant control. Assuming a prediction for the policy of one agent we compute the optimal policy of the other using the Relevant Information method and explore how one's behavior influences the other's. The goal is to move an object from one corner of a 2-D grid to the opposite corner along the diagonal by applying four actions (up, down, left, right) (see Figure 1). The state of the environment is denoted with the random variable S and the action with A. The state transition model $p(s_{t+1}|a_t, s_t)$, representing the movement of the object, is deterministic. The reward is -1 everywhere but the goal where it is 0. Thus, a policy maximising the expected utility is one that takes the shortest path to the goal. The shortest path in this case is not unique, however they differ in informational cost.

We investigate how one agent's behavior adapts to various strategies of the other ranging from fully cooperative to completely antagonistic (a similar treatment can be found in Ortega and Braun (2013)). Assuming a static prediction for the strategy of agent B, agent A optimizes its own policy under informational constraints.

The relevant information represents the minimal level of information required for achieving a certain level of performance. The problem of informational parsimony can be formulated as search for a value-optimal strategy $\pi^*(a|s)$, which at the same time minimises the required relevant information $I(S;A)$, i.e. is also information-optimal. This double optimisation is transformed into an unconstrained

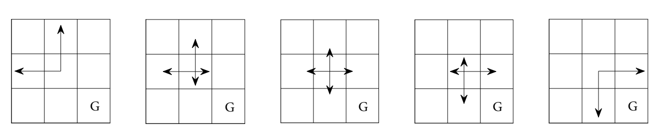

Figure 1: Various strategies ranging from completely antagonistic to fully cooperative (from left to right). G denotes the goal state and the arrow length represents the action probability (0.5 at the extremes and 0.25 each in the middle).

minimisation problem via Lagrange multiplier

$$\min_{\pi(a|s)} (I(S;A) - \beta E[U(S,A)]),$$

which corresponds to trading in utility $U(S,A)$ for a reduction of relevant information.

The parsimony pressure tries to minimise this quantity while the efficiency pressure drives towards higher performance. Therefore for $\beta \to \infty$ the resulting optimal policy maximizes the expected utility and at the same time minimizes the relevant information $I(S;A)$. This optimization provides the policy of agent A while taking into account the specific behavior of agent B.

Discussion

We computed the optimal policies of agent A for a range of antagonistic and cooperative behaviors of agent B (see Figure 1) and various levels of β.

The simulated trajectories of agent A's optimal policies for five collaboration rates and two β levels (a high driving towards higher performance and a low towards information parsimony) are shown in Figure 2, which reveal the influence increasing cooperation has on the optimal path. For antagonistic behaviors of agent B agent A feels more confident following the wall, which protects the object from moving back as it invalidates certain actions. This period of initial uncertainty gets shorter as cooperation increases and agent A can increasingly rely on its partner and gradually moves away from the wall onto the diagonal, which corresponds to the optimal informationally parsimonious policy. The results of this study suggest that when facing an antagonistic partner the cooperative agent exploits the particular embodiment of its environment in order to maximize its control abilities. This demonstrates that the agent increasingly transfers 'trust' from the indifferent environment to a trustworthy agent.

The results also revealed, as anticipated, that when the agents cooperate towards a common goal they achieve higher utility at lower informational cost and vice versa, when their goals contradict the utility level drops and the informational cost increases. This scenario could apply to various situations, for example when the controllers are stochastic, the precise knowledge of the other agent's model is unavailable or compromised, or the other agent is adverse or autistic.

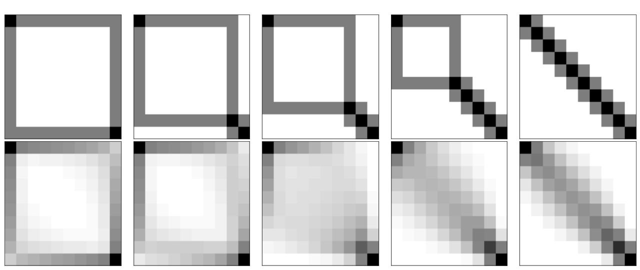

Figure 2: Simulated trajectories of agent A's optimal policies for five collaboration rates (increasing from left to right) and for high (top) and low (bottom) β. Darker color denotes higher recurrence. For antagonistic behaviors of agent B agent A initially follows the wall before switching onto the diagonal. The goal is in the bottom-right cell.

In this particular example the proposed information-theoretic approach highlights two types of optimal trajectories, one following the walls and another one traversing the diagonal, and trades off these two paths on the base of the cooperation level.

This paper presents an application of an information-theoretic framework to a reward driven decision process in the perception–action cycle of a dyad. It demonstrates how the Relevant Information method could provide informationally parsimonious optimal policies for collaborative behaviors ranging from fully cooperative to fully antagonistic. Furthermore, it presents an information-theoretic characterisation of the trade-off between the level of collaboration and the informational cost, elucidating the key role of the embodiment.

Acknowledgements

The authors would like to acknowledge support by H2020-641321 socSMCs FET Proactive project.

References

Dautenhahn, K. (1995). Getting to know each other - artificial social intelligence for autonomous robots. *Robotics and Autonomous Systems*, (16):333–356.

Goldstone, R. and Janssen, M. (2005). Computational models of collective behavior. *Trends in Cognitive Sciences*, 9(9):424–430.

Ikegami, T. and Iizuka, H. (2007). Turn-taking interaction as a cooperative and co-creative process. *Infant Behavior and Development*, (30):278–288.

Ortega, P. and Braun, D. (2013). Thermodynamics as a theory of decision-making with information-processing costs. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 469(2153).

Paolo, E. D., Rohde, M., and Iizuka, H. (2008). Sensitivity to social contingency or stability of interaction? modelling the dynamics of perceptual crossing. *New Ideas in Psychology*, 26(2):278–294.

Polani, D., Martinetz, T., and Kim, J. T. (2001). An information-theoretic approach for the quantification of relevance. *In ECAL '01: Proceedings of the 6th European Conference on Advances in Artificial Life*, pages 704–713.

Polani, D., Nehaniv, C., Martinetz, T., and Kim, J. T. (2006). Relevant information in optimized persistence vs. progeny strategies. In *Artificial Life X: Proceedings of the Tenth International Conference on the Simulation and Synthesis of Living Systems*, pages 337–343. MIT Press.

Wahn, B., Schmitz, L., König, P., and Knoblich, G. (2016). Benefiting from being alike:interindividual skill differences predict collective benefit in joint object control. In *Proceedings of the 38th Annual Conference of the Cognitive Science Society*, pages 2747–2752.

Walter, W. (1950). An imitation of life. *Scientific American*, May:42–45.