Balancing utility and cognitive cost in social representation

Max Taylor-Davies
School of Informatics
University of Edinburgh
s2227283@ed.ac.uk

Christopher G. Lucas
School of Informatics
University of Edinburgh

Abstract

To successfully navigate its environment, an agent must construct and maintain representations of the other agents that it encounters. Such representations are useful for many tasks, but they are not without cost. As a result, agents must make decisions regarding how much information they choose to store about the agents in their environment. Using selective social learning as an example task, we motivate the problem of finding agent representations that optimally trade off between downstream utility and information cost, and illustrate two example approaches to resource-constrained social representation.

Representing agents under cost constraints

In order to produce adaptive behaviour, an agent must acquire and maintain an internal representation of its environment [Craik, 1943, Tolman, 1948, Wilson et al., 2014]. Furthermore, unless condemned to an entirely solitary existence, we can expect that many environments encountered by a hypothetical agent will contain *other agents*. Much as our agent should represent the rest of the environment, we expect that it ought also to be able to represent these other agents. Humans do this—in fact, it seems we are intrinsically motivated to form mental representations of the other people we encounter [Dennett, 1987, Malle, 2008, Baker et al., 2017]. We use these representations for a variety of different purposes: understanding the strengths and weaknesses of a colleague to effectively collaborate with them; determining whether a stranger should be treated as friend or foe; or predicting the plays of a chess opponent in order to beat them. In an ideal world, we would maintain representations that encode all possible available information about every agent in our environment. But real agents, whether biological or artificial, will inevitably have to contend with limits on their cognitive resources. Inspired by a growing body of work in computational cognitive science that models human cognition through the lens of resource rationality [Lieder and Griffiths, 2019, Bhui et al., 2021], we consider the problem of developing social representations that balance downstream utility against cognitive cost.

Social decision tasks

To consider the idea of using social representations for decision-making, we develop the general construct of a *social decision task*—which describes any decision task where the optimal strategy depends on some representation of some set of other agents. That is, adapting the notation of Markov Decision Processes (MDPs) [Sutton and Barto, 2018, Puterman, 1994], the optimal policy (in terms of maximising expected return) is conditional on the value of some social representation χ: $\pi^*(s, a) = \pi^*(s, a; \chi)$. To consider the utility of a particular social representation χ, first let $\Pi(\chi)$ be the class of possible policies whose output depend on the value of χ. We can then consider the

Information-Theoretic Principles in Cognitive Systems Workshop at the 37th Conference on Neural Information Processing Systems (NeurIPS 2023)

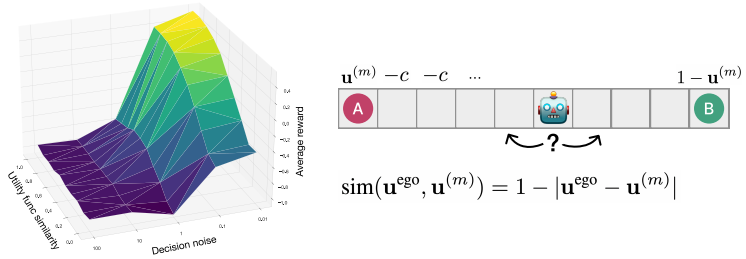

Figure 1: We generate several agents with randomly sampled scalar utility functions $\mathbf{u}^{(m)} \in [0, 1]$, which control each agent's relative preference for the two states labelled A and B. From each agent we sample a number of trajectories using different values of the decision noise parameter $\beta^{(m)}$. We then plot social learning reward as a function of the *similarity* between $\mathbf{u}^{(m)}$ and $\mathbf{u}^{\text{ego}}$, and $\beta^{(m)}$, averaged over both repeated trials and different values of $\mathbf{u}^{\text{ego}}$.

utility of χ as the expected return induced by the *best* member of $\Pi(\chi)$:

$$U(\chi) = \max_{\pi \in \Pi(\chi)} \mathbb{E}[R \,|a \sim \pi(s, a; \chi)] \tag{1}$$

Where R is the total (possibly discounted) return achieved from executing actions a in the environment. Equation 1 tells us, given access to social representation χ, how well we expect to perform on a particular task assuming that we're able to make the best possible use of the information in χ? Given some function $C : \chi \to \mathbb{R}$ that measures cognitive cost, we then define the cost-adjusted utility of χ as

$$U'(\chi) = (1 - \lambda)U(\chi) - \lambda C(\chi) \tag{2}$$

where $\lambda \in [0, 1]$ is a parameter that trades off between utility and cost. A representation χ^* is then considered optimal if it satisfies $\chi^* = \arg\max_\chi U'(\chi)$. This optimality criterion is similar to the objectives used within work on capacity-limited Bayesian decision-making and RL, such as Arumugam et al. [2023]. The key difference (beyond our explicit focus on social representations) is that we are interested not so much in the cognitive cost of converting representations into behaviour, but in the cost the representations themselves. In general, we expect this be a combination of the cost involved in *acquiring* a representation (i.e. inferring it from observation), and the cost involved in *storing* it—for now we simply group both under entropy (for more details see Appendix A), i.e. $C(\chi) = H[\chi]$. This makes intuitive sense as an objective—essentially saying that we want to maximise our expected task performance while minimising the amount of information we store about other agents. This is related to the objective used in Abel et al. [2019], which considers a similar tradeoff for the problem of state abstraction in apprenticeship learning.

Selective social learning as an example downstream task

As a simple example of a social decision task, we consider *selective social learning*, a common component of human and animal social behaviour which involves the problem of identifying the 'best' agent to learn from within a given environment [Rendell et al., 2011, Heyes, 2016]. In our framing of this task, the learning agent α^{ego} is situated within some environment populated by a number of other agents, with the ability to observe their behaviour (in the form of state-action transitions). At each trial, α^{ego} selects an agent α^{target}. All agents then execute some sequence of behaviour within the environment, visible to α^{ego}. α^{ego} executes the same behaviour observed from the chosen α^{target}, and receives some reward corresponding to their own internal utility function $\mathbf{u}^{\text{ego}}$. A general policy for selecting social learning targets on the basis of representations is given by

$$\Pr(\alpha^{\text{target}} = \alpha^{(m)}) \propto \exp\left(\frac{w(\alpha^{(m)}; \chi)}{\beta^{\text{ego}}}\right) \tag{3}$$

where w is some function that weights agents based on the value of χ and β^{ego} gives the *decision noise* of the social learning agent.

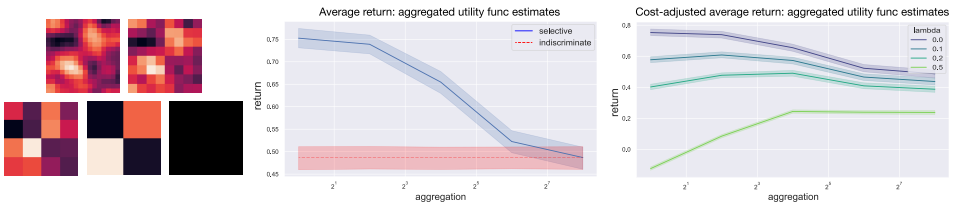

Figure 2: Left: illustration of state-aggregated estimates of an example 2D utility function (lighter grid squares indicate higher values). Middle: average return from selective social learning using state-aggregated value function estimates as agent representation (with indiscriminate selection as a baseline). Right: cost-adjusted return for aggregated representations at different values of λ. Since we are interested only in the tradeoff between return and cost (entropy), rather than each quantity's absolute values, both are individually normalised to lie in [0,1].

We can use Equation 3 to express different selection strategies by changing the weighting function w. Of course, the optimal choice of w will depend on the type of decision-making employed by the other agents. We assume that each observed agent attempts to maximise its own return in a noisily rational manner, according to a Boltzmann-style policy with decision noise $\beta^{(m)}$. Under this model, the social learner's expected reward should be optimised by selecting an agent such that we maximise some measure of the *similarity* between $\mathbf{u}^{\text{target}}$ and $\mathbf{u}^{\text{ego}}$ while minimising $\beta^{(m)}$. We can demonstrate this through simulation using a simplified one-dimensional gridworld environment—Fig. 1 shows that the social learner's return indeed increases with increasing utility function similarity and decreasing $\beta^{(m)}$. For the remainder of this paper, we will make the assumption that all observed agents are equivalently rational, and thus set aside the decision noise parameter. We therefore consider optimal social representations of the form $\chi^{(m)} = \hat{\mathbf{u}}^{(m)}$ where $\hat{\mathbf{u}}^{(m)}$ is some approximation of $\mathbf{u}^{(m)}$.

To study strategies for selective social learning, we use a contextual bandit setting. At each trial, each observed agent m is placed at a random tile within a 2D grid, chooses an destination tile to travel to, and then receives a reward $r = \mathbf{u}^{(m)}_{\text{dest}} - c\big(|x_{\text{dest}} - x_{\text{start}}| + |y_{\text{dest}} - y_{\text{start}}|\big)$ where $c \geq 0$ is a constant scalar that determines the cost of taking one step on the grid. Following the design of Wu et al. [2018] and Witt et al. [2023], we generate spatially correlated 2D utility functions (i.e. nearby tiles generally yield similar rewards) by sampling from a Gaussian process prior with a radial basis function kernel. The social learner selects targets according to Eq. 3, with weights given by

$$w(\alpha^{(m)}; \chi) = w(\alpha^{(m)}; \hat{\mathbf{u}}^{(m)}) = \text{sim}(\mathbf{u}^{\text{ego}}, \hat{\mathbf{u}}^{(m)}) = \frac{\mathbf{u}^{\text{ego}} \cdot \hat{\mathbf{u}}^{(m)}}{|\hat{\mathbf{u}}^{(m)}||\hat{\mathbf{u}}^{(m)}|} \tag{4}$$

Experiment 1: state aggregation

If α^{ego} has to choose between only a small number of agents, or faces only a small state space, then it may be feasible to represent utility functions exactly (i.e. use $\hat{\mathbf{u}}^{(m)} = \mathbf{u}^{(m)}$), even for $\lambda > 0$. But if the agent population or state space is large, then the optimal representation in terms of cost-adjusted utility (Eq. 2) will likely be an approximation that discards some information for the sake of lower entropy (see Appendix A). A simple way to obtain such an approximation is through state aggregation—i.e. group all states within a given-sized 'patch' of the grid under a single value [Sutton and Barto, 1998, Abel et al., 2019]. Fig. 2 shows how the cost-balanced utility of aggregated value function representations under various values of the tradeoff parameter λ. As expected, we observe that increasing the amount of aggregation leads to a decrease in both average return and representation cost, with the relative advantage or disadvantage determined by the tradeoff parameter.

Experiment 2: representing social groups

So we have seen how naive compression of utility function information can help a social learner navigate this tradeoff. But rather than merely taking individual representations and making them coarser, a more *humanlike* approach might involve using information about the agent population *as a whole*, and how individual agents within the population relate to one another. For instance, we often represent people that we're unfamiliar with in terms of their membership of certain social *groups* or

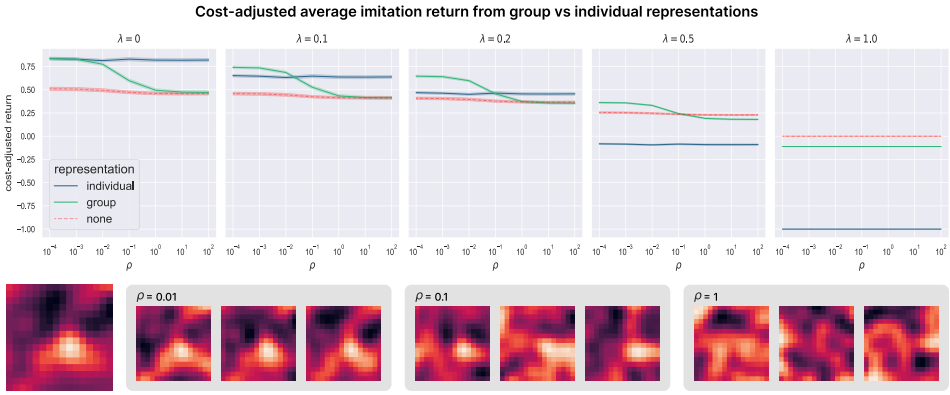

Figure 3: Top: cost-adjusted return for group-only and individual representations (with indiscriminate baseline) for different values of the group variance ratio ρ and the tradeoff parameter λ. As before, both return and cost (entropy) are normalised to lie in [0,1]. Bottom: illustration of example utility functions sampled for a 3-member group for three different values of the variance ratio ρ.

categories, using our knowledge of those groups to make inferences about the unknown properties of their members [Rhodes and Baron, 2019, Liberman et al., 2017]. To see how this can produce cheap agent representations, imagine that we have a known fixed number of groups K, and a population of M agents that are partitioned such that each agent m belongs to only one group (given by $\mathbf{z}_m$). Continuing our example of selective social learning, we will say that each group k is associated with a corresponding distribution over utility functions, $p(\mathbf{u}^{(m)}|\mathbf{z}_m = k)$, with mean value $\mu^{(k)}$. As an extreme, if we treat all members of group k as having $\mathbf{u} = \mu^{(k)}$ (i.e. representing zero intragroup variance), then we need only represent each agent with a single scalar value $\mathbf{z}_m$. Including the group means, this would yield an overall representation cost

$$C(\chi_{\text{grp}}) = KH[\mu^{(k)}] + MH[\mathbf{z}_m] \tag{5}$$

Intuitively, the extent to which this representation is useful will be primarily determined by the ratio between inter- and intragroup variation. When groups are very different, and agents within each group are very similar, then knowing an agent's group assignment tells you a lot about their individual properties; when groups are similar or agents vary a lot within each group, then group assignments are less informative. We can test this using the same task setup as in the previous experiment. We first sample a number of group mean value functions $\mu^{(k)}$ from the original GP prior; for each group we then sample a number of agents from a multivariate Gaussian distribution with mean $\mu^{(k)}$ and covariance $\Sigma = \rho\Sigma_{\text{GP}}$ where $\rho > 0$ is a scalar and Σ_{GP} is the covariance of the GP prior. Fig. 3 compares the cost-adjusted utility of the groups-only and individuals-only representation strategies for different values of ρ and λ. We can see that, below a certain λ threshold, individual representations are always as good or better than group representations; above a certain λ threshold, the reverse is true. Otherwise, as predicted, the optimal representation depends on the group variance ratio, with group-only representations becoming less useful as ρ increases.

Discussion

In this paper, we have briefly laid out and motivated the problem of cost-constrained social representation; that is, representing other agents in such a way that optimally trades off between downstream utility and information cost. Using selective social learning as an example task, we explored two potential approaches to this problem: compressing individual representations, and representing agents in terms of group identity. While we hope the work we present here is of some value, it is of course very much a preliminary step—in future work we hope to consider how agents can *learn* good social representations by optimising directly for this tradeoff, as well as broadening our analysis to encompass additional example tasks and agent features. As hinted at earlier, we also hope to explore a more detailed conception of representation cost that considers both acquisition and storage of information.

References

D. Abel, D. Arumugam, K. Asadi, Y. Jinnai, M. L. Littman, and L. L. S. Wong. State abstraction as compression in apprenticeship learning. In *AAAI Conference on Artificial Intelligence*, 2019. URL https://api.semanticscholar.org/CorpusID:53689696.

D. Arumugam, M. K. Ho, N. D. Goodman, and B. V. Roy. Bayesian reinforcement learning with limited cognitive load. *ArXiv*, abs/2305.03263, 2023. URL https://api.semanticscholar.org/CorpusID:258546889.

C. L. Baker, J. Jara-Ettinger, R. Saxe, and J. B. Tenenbaum. Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, 1(4):1–10, 2017.

R. Bhui, L. Lai, and S. J. Gershman. Resource-rational decision making. *Current Opinion in Behavioral Sciences*, 41:15–21, 2021.

T. M. Cover and J. A. Thomas. *Elements of information theory / Thomas M. Cover, Joy A. Thomas.* Wiley-Interscience, Hoboken, N.J, second edition. edition, 2006. ISBN 0471241954.

K. Craik. Hypothesis on the nature of thought, 1943.

D. C. Dennett. The intentional stance. 1987. URL https://api.semanticscholar.org/CorpusID:143977161.

C. Heyes. Who knows? metacognitive social learning strategies. *Trends in Cognitive Sciences*, 20: 204–213, 2016.

Z. Liberman, A. L. Woodward, and K. D. Kinzler. The origins of social categorization. *Trends in Cognitive Sciences*, 21:556–568, 2017.

F. Lieder and T. L. Griffiths. Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, 43, 2019.

B. F. Malle. Chapter 12 - the fundamental tools, and possibly universals, of human social cognition. In R. M. Sorrentino and S. Yamaguchi, editors, *Handbook of Motivation and Cognition Across Cultures*, pages 267–296. Academic Press, San Diego, 2008. ISBN 978-0-12-373694-9. doi:https://doi.org/10.1016/B978-0-12-373694-9.00012-X. URL https://www.sciencedirect.com/science/article/pii/B978012373694900012X.

M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc., 1st edition, 1994. ISBN 0471619779.

C. E. Rasmussen and C. K. I. Williams. *Gaussian processes for machine learning / Carl Edward Rasmussen, Christopher K.I. Williams.* Adaptive computation and machine learning. MIT Press, Cambridge, Mass. ;, 2006. ISBN 026218253X.

L. Rendell, L. Fogarty, W. J. E. Hoppitt, T. J. H. Morgan, M. M. Webster, and K. N. Laland. Cognitive culture: theoretical and empirical insights into social learning strategies. *Trends in Cognitive Sciences*, 15:68–76, 2011.

M. Rhodes and A. S. Baron. The development of social categorization. *Annual review of developmental psychology*, 1:359–386, 2019.

R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. Adaptive Computation and Machine Learning series. MIT Press, Cambridge, 1998.

R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. The MIT Press, second edition, 2018. URL http://incompleteideas.net/book/the-book-2nd.html.

E. C. Tolman. Cognitive maps in rats and men. *Psychological review*, 55 4:189–208, 1948. URL https://api.semanticscholar.org/CorpusID:42496633.

R. C. Wilson, Y. K. Takahashi, G. Schoenbaum, and Y. Niv. Orbitofrontal cortex as a cognitive map of task space. *Neuron*, 81:267–279, 2014. URL https://api.semanticscholar.org/CorpusID:2713112.

A. Witt, W. Toyokawa, W. Gaissmaier, K. N. Lala, and C. M. Wu. Social learning with a grain of salt. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 45, 2023.

C. M. Wu, E. Schulz, M. Speekenbrink, J. D. Nelson, and B. Meder. Generalization guides human exploration in vast decision spaces. *Nature Human Behaviour*, 2:915–924, 2018. URL `https://api.semanticscholar.org/CorpusID:53278852`.

A Representation entropy

Here we provide some additional details on our use of representation entropy as a measure of cognitive cost. For a comprehensive introduction to entropy (and other related information-theoretic quantities) we direct the reader to Cover and Thomas [2006]. For a discrete random variable $X \in \mathcal{X}$ with probability mass function $p(x) := \Pr(X = x)$, the entropy of X is given by

$$H_b(X) = -\sum_{x \in \mathcal{X}} p(x) \log_b(p(x)) \tag{6}$$

When $b = 2$ (as it typically is), the entropy has units of bits. For a continuous random variable $Y \in \mathcal{Y}$ with probability density function $p(y)$, the differential entropy of Y is given by

$$H_b(Y) = -\int_{y \in \mathcal{Y}} p(y) \log_b(p(y)) \tag{7}$$

If Y follows a multivariate Gaussian distribution with covariance Σ, then $H(Y)$ is computed as

$$H_2(Y) = \frac{1}{2} \log_2 |\Sigma| + \frac{n}{2}\big(\log_2(2\pi e)\big) \tag{8}$$

where n is the dimensionality of Y [Rasmussen and Williams, 2006]. This allows us to compute the entropy of our agent value functions $\mathbf{v}$, since each value function is in effect a continuous random variable distributed according to a multivariate Gaussian with known Σ. The total entropy of a representation χ_{ind} that consists of the value functions of M agents over n states is thus given by

$$H_2[\chi_{\text{ind}}] = \frac{M}{2}\big[\log_2 |\Sigma| + n\big(\log_2(2\pi e)\big)\big] \tag{9}$$

To compute the entropy of a state-aggregated value function estimate $\hat{\mathbf{v}}$, we can treat $\hat{\mathbf{v}}$ as a new continuous RV with a lower-dimensional multivariate Gaussian distribution whose covariance Σ_{agg} is determined entirely by Σ and the level of state aggregation. Determining Σ_{agg} is then sufficient to compute $H(\hat{\mathbf{v}})$. For the group-only representation cost, we need simply to compute the entropy of the scalar discrete random variable $\mathbf{z}$ (the agents' group assignments). With a uniform distribution over K groups, this is simply given (following Eq. 6) by

$$H_2(\mathbf{z}) = -\sum_{k=1}^{K} \frac{1}{K} \log_2\big(\frac{1}{K}\big) = \log_2(K) \tag{10}$$

The total entropy of a groups-only representation χ_{grp} is given by combining this with the total entropy of the group mean value functions:

$$H_2[\chi_{\text{grp}}] = \frac{K}{2}\big[\log_2 |\Sigma| + n\big(\log_2(2\pi e)\big)\big] + M \log_2(K) \tag{11}$$