Project Report: ‘Toward Human-Machine Virtual Bargaining’

Alan Bundy (PI) Nick Chater Stephen Muggleton

1 Research Programme

1.1 Project Summary

This 9-month HLC Kick-Start project has aimed to connect recent discoveries in the study of human coordination with current work on logical theory inference and (in particular) automated theory repair.

The empirical work forming the backdrop of this project posits a highly efficient, reasoning-driven cognitive process (‘virtual bargaining’) for spontaneous creation and update of signalling conventions in low-bandwidth contexts [1][2]. No algorithm is specified in this literature, but modelling constraints can be extracted from both the experimental designs and observed behaviour of human participants. Collectively, these constraints (e.g. knowledge-based inference, vocabulary adaptation, cross-task transfer and limited sampling) argue in favour of logical reasoning, as opposed to sampling-based inference. This approach was demonstrated in our use of automated theory repair to update basic low-bandwidth signals in the *select/avoid* coordination game [3], where a Receiver guides a Sender to select or avoid, through spontaneous signals requiring inference over both player’s perspectives to interpret [1]. With this project, we extend the links between low-bandwidth signalling conventions and logical inference, to motivate an interdisciplinary research programme for replicating this behaviour.

Despite challenges to our team from the COVID-19 crisis, this work has produced: *(a)* proof of concept application of the ABC system [4] on deterministic signal creation and update in *select/avoid*, including mixed-knowledge contexts;*(b)* two follow-up grant proposals with HLC Network+ members.

1.2 Objectives

The high-level goal of our research programme is to test (and identify appropriate tests of) the claim that spontaneous low-bandwidth coordination behaviour can be understood as logical inference from premises, such as shared rules and facts, to a joint signal interpretation, with limited or no sampling.

However, the range of behaviours this covers is broad. Within the existing literature, coordination in deterministic vs. nondeterministic contexts, divisions of private vs. public knowledge, the degree of sampling, and the effect of prior conventions on future conventions [2] must all be accounted for. The three objectives we selected for this project represent a broad cross-section of these concerns.

I. Given the claim that efficient low-bandwidth signals could sustain flexible human rule-making in societies and workplaces [5] or on the road [6], our objective was an end-to-end proof of concept: a logical reasoner creating *and* updating simple signals, over novel rounds of the *select/avoid* game, extending our work on signal update for the deterministic *select/avoid* game to the creation of signals.

Target Deliverable: End-to-end model of deterministic signal inference in *select/avoid*.

II. Given the need to demonstrate broader applications of this methodology, our second objective was to extend the *select/avoid* signal update model to account for a changing balance of private and shared knowledge closer to real-world contexts, by more fully representing players’ common ground.

Target Deliverable: Model of deterministic signal update in mixed-knowledge *select/avoid*.

III. Given our aim to model broader behaviours, our final objective was to identify logical theory inference methods suitable for complex, non-deterministic contexts, and also identify the constraints such methods must capture to be sufficiently human-like, motivating applications for further funding.

Target Deliverable: EPSRC and ESRC follow-up grant proposals; follow-up journal article.

1.2.1 Challenges

The project start date of 1 April 2020 coincided with the start of the UK-wide lockdown in response to COVID-19, and the full span of the project was conducted via home working. During this 9-month period the RA faced significant disruption by a challenging home working environment, and recurring illness linked to the pandemic. Although work was able to progress on all the deliverables, changes were necessary to our projected timelines for the associated manuscript and proposal submissions.

2 Results

2.1 Deliverable I & II: Model Updates

Our general process of extending our models of the *select/avoid* game, for both the signal creation and common ground examples, was to consider two related questions. First, whether a richer input representation was required - a more complex theory representing the initial premises from which human players of *select/avoid* must produce and interpret a signal - or else a richer form of inference from a less rich input representation. Our input representation richness had been minimal, in light of implementation (keeping the input theory in Datalog) and cognitive workload (keeping working memory demands low) constraints. The question was whether or not to increase it from this baseline. A trade-off was thus established, between assuming that players *represent more information* to start with, than we did before, and assuming that players *reason more extensively* about that information.

The second question is whether our algorithm itself required modifications to implement theory creation and common ground inferences in *select/avoid*, relative to the input representation richness.

2.1.1 Deterministic end-to-end signalling

For **signal creation** in the *select/avoid* game, our best result was via inference not from a richer, but from a larger input theory than the signal update model. Leveraging the *belief revision* element of our ABC algorithm, a complete space of (incompatible) possible signal interpretations could be uniquely 'whittled down' to just the optimal mapping between signal and item content; producing the desired 'spontaneous' creation of select/avoid signalling from zero examples in strictly deterministic contexts. From thereon, our earlier, update model can take over the process of revising the established signal.

Logical theory inference, in the form of automated theory repair, can therefore create and flexibly maintain the 'spontaneous' *select/avoid* human signalling observed in strictly deterministic contexts. At the same time, our method for signal creation, in particular, is difficult to extend to nondeterministic contexts, where the full space of signal-meaning mappings has more than one (or no) optimal choice. Moreover, our approach assumes that signal creation occurs once, then is replaced by signal update: evidence of signals being remapped *de novo* would suggest a different signal inference mechanism.

2.1.2 Deterministic mixed-knowledge signal update

The goal of modelling **mixed-knowledge update** in *select/avoid* stems from the real-world structure of the various coordination phenomena (like ad hoc workplace collaboration and driving) that this abstract signalling game aims to capture. Unlike the *select/avoid* contexts we previously considered, where one agent has exclusive knowledge of a set of entities, while another has exclusive access to manipulate these entities, knowledge distribution in real-world tasks is 'messier', and ever-changing.

We thus aimed to model signal update in response to changes in the distribution of knowledge between players, distinct from the arrangement, content or features of the items they select or avoid. To allow this to happen, players must know what they each know, publicly and privately, independent of their own viewpoint: in cognitive psychology terms, they must represent their common ground [7]. It should be emphasised that reasoning-based approaches to *select/avoid* inherently require some representation of both player's viewpoints [1]. The distinction is between richer representations of players' knowledge, assigning each fact about the game to one or both players, and a more minimal representation of just their possible actions in respect to each other, as we have previously used [3].

For mixed-knowledge contexts, an approach we considered was relying on the ABC system's *conceptual change* element to enrich facts about item contents, such as '*Help(item)*' vs. '*Harm(item)*', adding arguments already used in action-based representations, such as '*Select(Receiver, item)*', to produce viewpoint-sensitive facts such as *'Help(Receiver, item)'*: Receiver thinks the item is helpful.

This approach may be valid in the long term, beyond the span of this Kick-Start project. Presently, we identified a challenge in automatically adding not just *some* argument, but *exactly* the arguments representing the viewpoints in action representations, to the necessary predicates. Overcoming this will require refinements to ABC's repair choice heuristics beyond what was achievable in this project.

Instead, we explored the inverse: enriching the input theory to represent players' knowledge via the above viewpoint arguments, so every fact about items in the game is assigned a viewpoint value for each player. We confirmed our existing signal update results, when player knowledge remained static, are compatible with this richer input; and that the target output for mixed-knowledge variations (e.g. Sender ignoring items that Receiver already knows about) is also achieved. So long as human players represent common ground inherently (vs. just selectively [7]) in *select/avoid*, our ABC model is enough to replicate 'spontaneous' signal update as both rules and perspectives change over time.

2.2 Deliverable III: Follow-up Grants and Articles

2.2.1 EPSRC Proposal: 'Virtual convention learning'

Our prior and current work on the potential for automated reasoning methods to model low-bandwidth conventions focuses on signalling for low-complexity, deterministic contexts. In such contexts, some optimal solution exists, and players can solve the puzzle of possible signals and meanings to find it. Where players solve this puzzle the same way, a 'spontaneous' convention emerges, without the players interacting, as observed empirically [1], and modelled by our methods (in [3], and the above).

However, deterministic contexts are most likely an exception rather than the rule for low-bandwidth conventions. And unlike signal update due to changing contexts, where automated theory repair fits most naturally as a methodology, humans' successful joint creation of signals in non-deterministic contexts can be intractable without some mix of sampling and induction (even in our source literature - where minor sampling can be observed as the complexity of *select/avoid* examples increases [1].)

Inductive logic programming overall, and inductive machine learning in particular, is a strong fit for logical theory inference that combines limited sampling and induction. Ongoing work at Imperial College London explores not just the learning, but the efficient *teaching* of conventions. One aspect in particular that our existing methodology does not capture is the efficient use of an interaction to inform, instruct or correct a collaborator in minimal steps, which this work is strongly focused on [8].

To further explore this domain, we co-authored and submitted an EPSRC responsive-mode grant proposal led by Imperial College London (Stephen Muggleton (Lead), Alan Bundy and Nick Chater PIs) to explore creating and adapting signalling conventions through a mix of reasoning and learning. If funded, this 3.5-year grant will employ two RAs, based in Imperial College London and Edinburgh.

2.2.2 ESRC Proposal: 'Spontaneous adaptation in low-bandwidth signalling'

A separate question from determinism and (formal) complexity is the effect of *cognitive factors*, such as precedent or working memory load, on how low-bandwidth signalling conventions are selected and especially how such conventions adapt over time. This links to open questions from the present project, such as the extent to which common ground is represented, and the choice to *replace* or *repair* a signalling convention; and to emerging work on sequences of spontaneous conventions [9].

Our second follow-up is a collaboration between Edinburgh and Warwick to explore this domain. Rather than focusing mainly on replicating human behaviour, the aim of this second collaboration is to understand what rules a sufficiently human-like replication must respect. Contrasting formal and cognitive factors in the selection of spontaneous conventions, we aim to more clearly situate low-bandwidth signalling in the larger domain of the cognitive psychology of reasoning and coordination.

In support of the primarily empirical nature of this follow-up, and in light of the ongoing restrictions on in-person laboratory experiments, we developed a fully remote version of the *select/avoid* game (originally developed for in-person testing) to use in the required extensive empirical data-gathering.

Due both to the challenges described in **§1.2.1** and generally longer development of the material for this second project, we aim to submit the relevant proposal as an ESRC responsive-mode grant in the second quarter of 2021. If funded, this 3-year grant will employ a single RA based in Edinburgh.

2.2.3 Journal Article: 'Automated theory Inference and spontaneous convention repair'

A manuscript covering the potential of automated theory inference for the modelling of convention repair in low-bandwidth settings is currently in preparation, drawing on the work from this Kick-Start.

3 Staffing

The project RA, Eugene Philalithis, is named as RA in both follow-up grant proposals and is attached to the School of Informatics in the University of Edinburgh as a visitor pending the councils' decisions.

References

[1] Misyak, J. B., Noguchi, T. & Chater, N. (2016). Instantaneous conventions: The emergence of flexible communicative signals. *Psychological Science*, *27*(12), 1550–1561.

[2] Misyak, J. & Chater, N. (2017). The spontaneous creation of systems of conventions. In *Proceedings of the 39th Annual Meeting of the Cognitive Science Society, London, UK, 16-29 July 2017.*

[3] Bundy, A., Philalithis, E. & Li, X. (2021). Modelling virtual bargaining using logical representation change. In S. S. Muggleton & N. Chater (Eds.), *Human-like machine intelligence.* OUP, Oxford.

[4] Li, X., Bundy, A. & Smaill, A. (2018). ABC repair system for datalog-like theories. In J. Bernardino, A. Salgado & J. Filipe (Eds.), *Proceedings of 10th International Conference on Knowledge Engineering and Ontology Development* (pp. 333–340). SCITEPRESS.

[5] Misyak, J. B., Melkonyan, T., Zeitoun, H. & Chater, N. (2014). Unwritten rules: Virtual bargaining underpins social interaction, culture, and society. *Trends in Cognitive Sciences*, *18*(10), 512–519.

[6] Chater, N., Misyak, J., Watson, D., Griffiths, N. & Mouzakitis, A. (2018). Negotiating the traffic: Can cognitive science help make autonomous vehicles a reality? *Trends in Cognitive Sciences*, *22*(2), 93 –95.

[7] Horton, W. S. & Keysar, B. When do speakers take into account common ground? *Cognition*, *59*(1), 91–117.

[8] Ai, L., Muggleton, S. H., Hocquette, C., Gromowski, M. & Schmid, U. (2020). Beneficial and harmful explanatory machine learning. *Machine Learning.* In Press, available http://arxiv.org/abs/2009.06410.

[9] Misyak, J. B. & Chater, N. (submitted). Instantaneous systems of communicative conventions.