Distributed Hypothesis Generation & Evaluation

Jordan Robinson, supervised by Katie Atkinson, Simon Maskell, Chris Reed, and Todd Robinson (Dstl)

EPSRC Centre for Doctoral Training in Distributed Algorithms, University of Liverpool, Liverpool, UK

Background & Aims

Intelligence analysis is currently conducted by distributed teams of expert human agents who use their domain knowledge, combined with a variety of structured analytical techniques, to generate and evaluate sets of conflicting hypotheses to inform potential high-stake decision making. Analysis can be tedious but it requires the full attention of human agents as the context can be such that what would otherwise be a minor detail has a significant impact on the likelihood of a hypothesis, and so on the downstream decision making. This project aims to enhance the speed and scale of intelligence analyses through the development of decision-support tools which combine explainable artificial intelligence algorithms with human expertise, in the form of human-machine teams, to aid intelligence analysts in evaluating complex and competing hypotheses. The tools created will combine techniques found within the Natural Language Processing (NLP) and Computational Argumentation (CA) literature and should enable analysts to focus their attention where it's needed most by assisting them throughout the analytical pipeline.

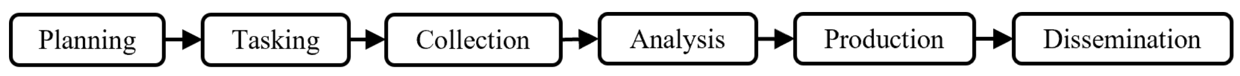


Fig. 1: A typical intelligence analysis pipeline.

The stages within a typical intelligence cycle are *planning*, *tasking*, *collection*, *analysis*, *production*, and *dissemination* of a finished intelligence product to decision makers, stakeholders and policymakers (Fig. 1) [1]. All information is funnelled through this process which aims to produce succinct and accurate intelligence. After *planning*, an All-Source Analyst (ASA) raises a Request For Information (RFI), in response to which a Single-Source Analyst (SSA) is *tasked* to *collect*, *analyse* and *produce* a summary report which is *disseminated* back to the all-source analyst. SSAs gather their information using the five disciplines of intelligence collection, which are defined in [2] as: Human Intelligence, Signals Intelligence, Geo-spatial Intelligence, Measurement and Signature Intelligence, and Open-Source Intelligence. Once the RFI has been satisfied, the ASA's job is to produce a finished intelligence product using all the provided data and a series of structured analytical techniques [3].

Ongoing Work

To date, this research has modelled the argumentative mechanics between an ASA and a set of SSAs, using a multi-agent system. The approach captured this using Dung-style argument frameworks [4], labelling-based semantics [5], and an uncertainty quantification about the inclusion and acceptance status of arguments within frameworks. The algorithm is currently being extended to a real-world scenario through the incorporation of an argumentation mining pipeline (Fig. 2) and will evaluate the hypothesis: *"The Malaysian Government will conduct another search for MH370 at some point within the next 6 months on the discovery of new, credible information as to the aircraft's whereabouts."*

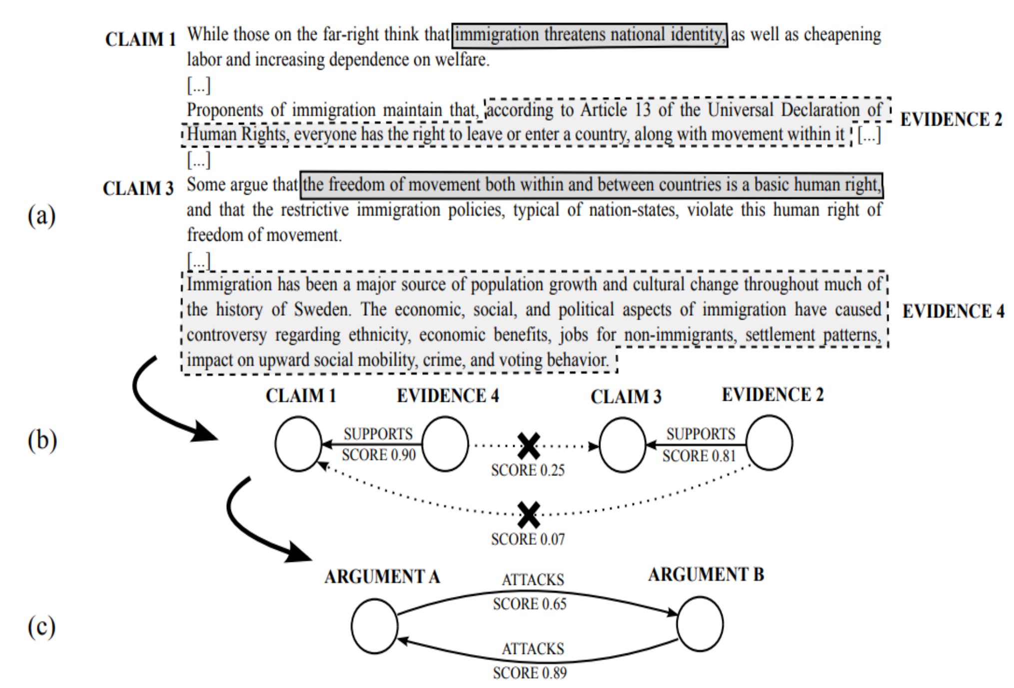
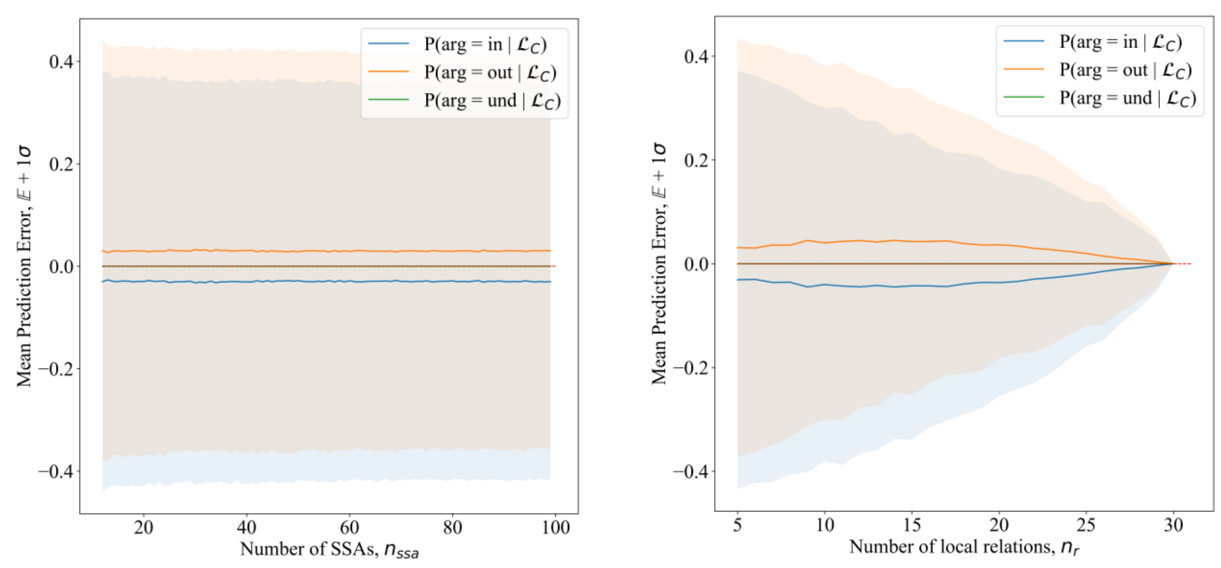


Fig. 2: An argumentation mining pipeline for an example text corpus (from [6]).

Preliminary Results



(a) (b)

Fig. 3: Measures the change in the ASA's mean prediction error when increasing the number of: (a) SSAs n_{ssa} from 10 to 100, where $n_r = 5$; and (b) local relations n_r from 5 to 30, where $n_{ssa} = 10$.

Some of the preliminary results are presented for one global Dung abstract argumentation framework which had 25 arguments and 30 total relations with no symmetric attacks and one complete labelling only, for 100 test runs. This experiment aims to assess whether increasing the number of SSAs n_{ssa} (Fig. 3a) or the amount of local knowledge an SSA possesses (i.e., increasing the number of local relations n_r sampled) (Fig. 3b), increases the ASA's prediction accuracy. Increasing the number of local relations available to each SSA dramatically reduced the ASA's prediction error, when compared to increasing the number of SSAs, which seems to affirm an old saying, *quality over quantity*.

Future Work

- **Argumentation Mining Pipeline:** Extend the current implementation to evaluate a hypothesis for a real scenario.
- **Diagnostic Argument Identification:** Develop an algorithm which can discover the most important arguments for intelligence analysts to analyse in situations which are time critical.
- **Hypothesis Generation:** Combine abstractive and extractive summarisation algorithms, found within the NLP literature, with argumentation to generate the set of Mutually Exclusive and Collectively Exhaustive (MECE) hypotheses using obtained data.
- **Application:** Development of a multi-agent system which can generate and evaluate MECE hypotheses, and automatically identify diagnostic arguments, providing a decision-support capability for a distributed set of intelligence analysts.

Conclusion

This project is developing generation-after-next decision-support tools, using techniques found within the CA and NLP literature, to alleviate cognitive load and bias by aiding intelligence analysts throughout a typical analytical pipeline.

References

[1] U.S., I. O. S. S. S. (2000). Intelligence threat handbook.
[2] Lo Clwenthal, M.M. andark, R. (2015). The Five Disciplines of Intelligence Collection. SAGE Publications.
[3] Heuer Jr., R. and Pherson, R. (2014). Structured Analytic Techniques for Intelligence Analysis. CQ Press.
[4] Dung, P. M. (1995). On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artificial Intelligence, 77(2):321–357.
[5] Wu, Y. and Caminada, M. (2010). A labelling based justification status of arguments. Studies in Logic, 3:12–29
[6] Lippi, M. and Torroni, P., 2016. Argumentation Mining. *ACM Transactions on Internet Technology*, 16(2), pp.1-25.

Engineering and Physical Sciences Research Council

UNIVERSITY OF LIVERPOOL

The Science Inside