Making sense of meanings in dialogue

August 2, 2016

Flex Fund Project Summary:
Project Id:
RA: Joseph Corneli
Email: j.corneli@gold.ac.uk
Host site: Goldsmiths
Start date: 1/10/16
Project duration: 6 months, full time

1 Proposed programme of work

This proposal seeks support for six months focused on grant development for a project with applications in reasoning about textual data. The methods have implications for mathematical reasoning within a social framework. This work will consolidate results from the EU FP7-ICT COINVENT project (611553). This short project will be based at Goldsmiths, University of London, with Professor Simon Colton as PI and Dr Joseph Corneli as named researcher.

The primary deliverable in the project is a new multisite grant proposal. Discussions about the content of proposal are already underway with Simon Colton and Teresa Llano (Goldsmiths/Falmouth), Alison Pease (Dundee) and Daniel Winterstein (Winterwell Associates). One of the central themes to be developed in the grant is the role of *argumentation* in making sense of a given obscure text. Pease et al. (unpublished) have been developing a Lakatos-inspired theory of argumentation in mathematics that will offer a relevant foundation. A more restricted prototype of our argumentative theory was previously applied to evaluate simple concept blends [3]. The proposal builds on earlier research by the RA and close collaborators into dialogues and other forms of social interaction in mathematics [10, 6, 9] and aims to combine theories of discourse with technical knowledge from mathematics domains. The expected long-term impact of this small-scale project is that it will foster the development of computational agents that simulate or participate in MiniPolymath-style discussions. Grant development work will be supported by a small-scale technical pilot study that will help shape the direction of the grant. The distribution of effort will be 3 days a week devoted to grant-writing, 2 days a week to technical work, over a total of 24 work weeks (6 months). A work plan follows.

Tasks:

1. **Develop a new EPSRC responsive mode bid.** This constitutes the critical core of the project. See the outline of work packages below; this sketch will be refined with project partners. [**85 days**]

2. **Carry out a pilot study.** Stack Exchange provides a corpus with rich content from various technical and non-technical domains. One task for which success is easy to quantify is matching existing Stack Exchange questions with their answers (henceforth, SXMATCH). SXMATCH can be understood as a much simplified version of the full question-answering problem, but one that is nevertheless grounded in real-world data (which makes it different from a previous important strand of research on dialogue [13]). This pilot study will compare simple computational approaches to SXMATCH, using industry-standard distributional semantics systems GloVe [11] and word2vec [8] and a next generation competitor [7]. The aim is to provide a baseline success rate for SXMATCH, as a proof of concept motivating further work on this task using more intricate models of technical discourse in the 3 year grant (see **WP3**, below). [**35 days**]

Tentative sketch of work packages for the 3 year grant (2017-2020):

[**WP1**] **Obfuscating fictional and mathematical ideas through rhetoric.** Most interesting questions include obscure components. This work package will use rhetorical techniques to generate a set of aesthetically or pedagogically interesting test cases to be disambiguated in **WP2**.

[**WP2**] **Argumentation for resolving obscure ideas.** Reasoning about typical and non-typical meanings can be embedded in an argumentative framework, in which defeasible support for different understandings (identified, e.g., via corpus methods) can be discussed.

[**WP3**] **Models of technical discourse** This work package will make use of *compositional semantics* [4, 2] and the *distributional compositional semantics* [1] approach, along with social network models, and domain-specific knowledge encodings for mathematics to improve the success rate for SXMATCH on two mathematics-specific sites in the StackExchange network, MathOverflow, and math.stackexchange.net.

[**WP4**] **Agent-based modelling.** Build agents that can generate and participate in argumentative dialogues by asking and answering questions. The usefulness of the generated questions and answers will be validated in a study with Stack Exchange users.

[**WP5**] **Industrial applications.** Develop applications of the above, drawing also on recent advances in "Long Short-Term Memory" (LSTM) neural nets [12], within, e.g., automatic email answering systems. The flexible syntax and loose semantics of poetry make it an attractive focus for test cases.

Risk management: The team has been involved in earlier collaborative research exploring the various facets of the proposal [5, 14, 3, 9]. The pilot study is intended to ensure that the specific thrust of the proposal current is validated. Depending on the results from that study and developments at other project sites, the details of the 3 year EPSRC bid would be able to shift: e.g., we might see that distributional semantics is not useful for working with technical content, and emphasise a different approach to knowledge representation in **WP3**. The named researcher will give a DReaM talk at the mid-point of the project to share preliminary results, and will gather feedback from the wider community.

2 Expected deliverables

- A conference paper that addresses the question-answer matching problem SXMATCH within a pilot study that compares different distributional semantics models and outlines future work.

- A multi-site EPSRC responsive mode bid that will develop this plan of work in detail and describe potential applications within dialogue systems.

- Two DReaM talks summarising the state of work, at the mid-point and at the end of the project.

3 Actual deliverables

4 Summary

References

[1] Marco Baroni, Raffaela Bernardi, and Roberto Zamparelli. Frege in space: A program of compositional distributional semantics. *LiLT (Linguistic Issues in Language Technology)*, 9, 2014.

[2] Jonathan Berant, Andrew Chou, Roy Frostig, and Percy Liang. Semantic Parsing on Freebase from Question-Answer Pairs. In *EMNLP*, volume 2, page 6, 2013.

[3] Roberto Confalonieri, Joseph Corneli, Alison Pease, Enric Plaza, and Marco Schorlemmer. Using argumentation to evaluate concept blends in combinatorial creativity. In S. Colton, H. Toivonen, M. Cook, and D. Ventura, editors, *Proceedings of the Sixth International Conference on Computational Creativity, ICCC 2015*. 2015.

[4] Percy Liang, Michael I Jordan, and Dan Klein. Learning dependency-based compositional semantics. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1*, pages 590–599. Association for Computational Linguistics, 2011.

[5] Maria Teresa Llano Rodriguez, Christian Guckelsberger, Rose Hepworth, Jeremy Gow, Joseph Corneli, and Simon Colton. What If A Fish Got Drunk? Exploring the Plausibility of Machine-Generated Fictions. In Amilcar Cardoso, François Pachet, Vincent Corruble, and Fiammetta Ghedini, editors, *Proceedings of the Seventh International Conference on Computational Creativity, ICCC 2016*, 2016.

[6] Ursula Martin and Alison Pease. What does mathoverflow tell us about the production of mathematics? In *SOHUMAN, 2nd International Workshop on Social Media for Crowdsourcing and Human Computation, at ACM Web Science 2013, May 1, 2013, Paris.* 2013.

[7] Stephen McGregor, Kat Agres, Matthew Purver, and Geraint Wiggins. From distributional semantics to conceptual spaces: A novel computational method for concept creation. *Journal of Artificial General Intelligence*, 6(1):55–86, 2015.

[8] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed representations of words and phrases and their compositionality. In *Adv. in Neural Information Processing Systems*, pages 3111–3119, 2013.

[9] Dave Murray-Rust, Joseph Corneli, Alison Pease, Ursula Martin, and Mark Snaith. Synchronised multi-perspective analysis of online mathematical argument. In Sally Jackson, Dima Mohammed, Lilian Bermejo-Luque, and Steve Oswald, editors, *1st European Conference on Argumentation: Argumentation and Reasoned Action, 9-12 June 2015, Lisbon, Portugal [Extended Abstracts].* 2015.

[10] Alison Pease and Ursula Martin. Seventy four minutes of mathematics: An analysis of the third mini-polymath project. In *Proc. of AISB/IACAP 2012, Symposium on Mathematical Practice and Cognition II*, pages 52–105, 2012.

[11] Jeffrey Pennington, Richard Socher, and Christopher D. Manning. GloVe: Global Vectors for Word Representation. In *Empirical Methods in Natural Language Processing (EMNLP)*, pages 1532–1543, 2014.

[12] Kai Sheng Tai, Richard Socher, and Christopher D Manning. Improved semantic representations from tree-structured long short-term memory networks. *arXiv preprint arXiv:1503.00075*, 2015.

[13] Jason Weston, Antoine Bordes, Sumit Chopra, Alexander M Rush, Bart van Merriënboer, Armand Joulin, and Tomas Mikolov. Towards AI-Complete Question Answering: A Set of Prerequisite Toy Tasks. *arXiv preprint arXiv:1502.05698*, 2015.

[14] Daniel Winterstein and Joseph Corneli. X575: writing rengas with web services, 2016. http://arxiv.org/abs/1606.07955.