From Distributional Semantics to Distributional Pragmatics?

Matthew Purver
Cognitive Science Research Group
Mehrnoosh Sadrzadeh
Theoretical Computer Science Research Group
School of Electronic Engineering and Computer Science
Queen Mary University of London
{m.purver,m.sadrzadeh}@qmul.ac.uk

To model meaning in dialogue, we must go beyond individual sentence meanings to look at their role in, and contribution to, a wider structure which emerges through the interaction. Formal models of dialogue tend to account for this dialogue structure in terms of either the relations between utterances (see e.g. Asher and Lascarides, 2003) or their function and effects on some model of context (see e.g. Ginzburg, 2012). These approaches have made great strides in understanding how to construct representations of meaning which depend appropriately on, and contribute appropriately to, the context. However, they are constrained in practice by many factors, including the need for broad-coverage semantic grammars to build sentence meanings, and the need to identify the illocutionary role or dialogue act of any given utterance, a task which depends not only on predicate-argument structure but lexical semantics, as well as on knowledge of the possible dialogue acts/roles to begin with.

Probabilistic dialogue models, with their ability to represent distributions over both utterance meanings and contexts, have shown more success in practical applications such as human-computer dialogue systems; for one thing, recent approaches circumvent the need for dialogue act recognition, instead using information present only in the utterance lexical (or acoustic) form, together with the context (Young et al., 2013). However, the light that they can shed on linguistic or semantic questions is currently limited, as is the degree to which they can exploit knowledge of fine-grained semantic structure.

Much recent work in computational semantics has explored the use of geometric representations: vector space models in which word meanings can be represented as vectors, with similarity in meaning relating to geometric closeness (see e.g. Clark, forthcoming). Models for the composition of sentential semantics from word vectors have been provided in various ways, including a view of predicates as tensors applying to their argument vectors (Coecke et al., 2010; Kartsaklis et al., 2012; Socher et al., 2012). Such a view can provide sentential representations which integrate lexical semantics and predicate-argument structureIn this talk, we will investigate ways in which this approach might be used to address some of the limitations of formal dialogue models outlined above. In particular, we will ask:

- How much can geometric models of semantics help us understand the dialogue act function of an utterance? We will discuss the insights from dialogue act tagging experiments such as those of Kalchbrenner and Blunsom (2013) and Milajevs et al. (2014).
- Many vector-space models of meaning are generated by exploiting the distributional hypothesis that words which appear in similar contexts have similar meanings (Harris, 1954). However, the notion of "context" used tends to be a narrow lexical or syntactic one, which can lead to surprising results such as an apparent similarity in meaning between terms like "happy" and "sad", or "hello" and "goodbye". Dialogue provides us with a different notion of context which includes interlocutors' responses – how can this be exploited to learn word meanings from distributions, and what difference does this make?
- Similarly, our intuitive concept of a dialogue act is governed to a large extent by the distribution of contexts in which it can be used, and in which it results. Can distributional methods allow us to investigate – and measure – the similarity in pragmatic meaning between dialogue act types, and perhaps learn dialogue act taxonomies?

- Tensor-based compositional models (e.g. Kartsaklis et al., 2012) provide us with a way to articulate predicate-argument structure in vector spaces, e.g. viewing verbs as distributional relations over their arguments. Can we apply this to model the context update functions of utterances in dialogue, by associating utterances with distributions over the contexts that they relate?

- As well as utterances, words and phrases have their own context update effects, associated with different distributions of contexts and responses, and this can give us insights into their meaning (see e.g. Purver and Ginzburg, 2004). Can the integration of lexical and sentential meaning embodied in geometric semantic models help us approach these insights empirically?

- Finally, dialogues consist of utterances often categorised as grammatically ill-formed: they contain fragmentary or unfinished contributions uttered for purposes such as pause, repair, and clarification. However, these acts (and as a result, a dialogue as a whole) have grammatical structures of their own, and efforts have been made to formalise these in grammars (see e.g. Kempson et al., 2001) and implement them in incremental parsers (e.g. Purver et al., 2011). Can one develop a functorial passage in the style of Coecke et al. (2010) to homomorphically transfer the grammatical structures of dialogue acts to linear maps in vector spaces and hence define a theoretical notion of compositional distributionality of meaning for dialogue acts?

References

Asher, N. and A. Lascarides (2003). *Logics of Conversation*. Cambridge University Press.

Clark, S. (forthcoming). Vector space models of lexical meaning. In S. Lappin and C. Fox (Eds.), *Handbook of Contemporary Semantics* (2nd ed.). Wiley-Blackwell.

Coecke, B., M. Sadrzadeh, and S. Clark (2010). Mathematical foundations for a compositional distributional model of meaning. *Linguistic Analysis 36*(1–4), 345–384.

Ginzburg, J. (2012). *The Interactive Stance: Meaning for Conversation*. Oxford University Press.

Harris, Z. S. (1954). Distributional structure. *Word 10*(23), 146–162.

Kalchbrenner, N. and P. Blunsom (2013). Recurrent convolutional neural networks for discourse compositionality. In *Proceedings of the Workshop on Continuous Vector Space Models and their Compositionality*, Sofia, Bulgaria, pp. 119–126.

Kartsaklis, D., M. Sadrzadeh, and S. Pulman (2012). A unified sentence space for categorical distributional-compositional semantics: Theory and experiments. In *Proceedings of COLING 2012: Posters*, Mumbai, India, pp. 549–558.

Kempson, R., W. Meyer-Viol, and D. Gabbay (2001). *Dynamic Syntax: The Flow of Language Understanding*. Blackwell.

Milajevs, D., D. Kartsaklis, M. Sadrzadeh, and M. Purver (2014). Evaluating neural word representations in tensor-based compositional settings. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar, pp. 708–719.

Purver, M., A. Eshghi, and J. Hough (2011). Incremental semantic construction in a dialogue system. In *Proc. 9th International Conference on Computational Semantics (IWCS)*, Oxford, UK, pp. 365–369.

Purver, M. and J. Ginzburg (2004). Clarifying noun phrase semantics. *J. Semantics 21*(3), 283–339.

Socher, R., B. Huval, C. D. Manning, and A. Y. Ng (2012). Semantic compositionality through recursive matrix-vector spaces. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pp. 1201–1211.

Young, S., M. Gašić, B. Thomson, and J. D. Williams (2013). POMDP-based statistical spoken dialog systems: A review. *Proceedings of the IEEE 101*(5), 1160–1179.