# Incremental Semantic Judgements

Gijs Wijnholds, Matthew Purver, Mehrnoosh Sadrzadeh and Ruth Kempson
Queen Mary University of London

Dynamic Syntax (DS, Kempson et al., 2001) provides a framework for strictly incremental composition relating word sequences to semantic representations. One of its strengths is that the formalism is general enough to be agnostic as to the exact semantic representation framework. As a result, various versions of DS have used different approaches to semantic representation (e.g. the typed lambda calculus (Kempson et al., 2001), Type Theory with Records (Purver et al., 2010; Hough and Purver, 2012)). Recently, Sadrzadeh et al. (2018b) proposed the combination of *vector space* semantic models with DS, and Sadrzadeh et al. (2018a) showed how an instantiation of this approach with a suitable vector space could in theory be used to derive incremental predictions of quantitative semantic measures such as *plausibility*. This talk will present recent investigations which test this approach with real vector spaces derived from large corpora, and show that it has potential to incrementally give predictions of disambiguation and plausibility which improve on more commonly used models.

Sadrzadeh et al. (2018b) show how to map the standard semantic trees of DS to a vector-space version, in which the semantic decorations of tree nodes are correspond to general higher-order tensors, and the composition function is tensor contraction $\otimes$. Given a definition of suitable spaces (e.g. a word space W and sentence space S), together with a notion of requirement for a tensor of a particular order (e.g. $?W, ?W \otimes S$), the familiar incremental parsing dynamics of DS are unaffected:
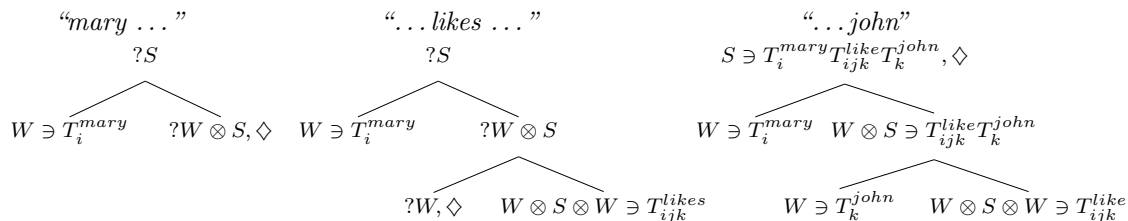
Among the advantages to providing a semantics of this kind are the abilities to move away from a view of semantic representations as fixed symbols, and to quantify judgements about word, phrase or sentence meaning in terms of distances between vectors or tensors in the appropriate space. By interpreting *requirements* as elements in the same space, we can even do this incrementally, and Sadrzadeh et al. (2018a) discuss some possible approaches to doing this, by regarding requirements variously as the union or sum of all possible continuations, or as a neutral (identity) element contributing no new information.

In this talk, we will describe recent work extending this model by implementing it computationally and evaluating it on real data. We take a corpus-based approach to word meaning, deriving vector spaces from large scale corpora, with noun meanings inferred from lexical co-occurrence via approaches now standard in computational linguistics (Mikolov et al., 2013; Baroni et al., 2014), and verbs either estimated from the vectors of their arguments (Grefenstette and Sadrzadeh, 2015) or learned directly to predict aspects such as plausibility (Polajnar et al., 2014). We show how to generate incremental judgements of plausibility and of disambiguation, and show that these can outperform standard additive baselines.

References

Baroni, M., Dinu, G., and Kruszewski, G. (2014). Don't count, predict! a systematic comparison of context-counting vs. context-predicting semantic vectors. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, pages 238–247, Baltimore, Maryland. Association for Computational Linguistics.

Grefenstette, E. and Sadrzadeh, M. (2015). Concrete models and empirical evaluations for the categorical compositional distributional model of meaning. *Computational Linguistics*, 41(1):71–118.

Hough, J. and Purver, M. (2012). Processing self-repairs in an incremental type-theoretic dialogue system. In *Proc. 16th SemDial Workshop*, pages 136–144, Paris, France.

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

Mikolov, T., Yih, W.-t., and Zweig, G. (2013). Linguistic regularities in continuous space word representations. In *Proceedings of NAACL-HLT*, pages 746–751.

Polajnar, T., Rimell, L., and Clark, S. (2014). Using sentence plausibility to learn the semantics of transitive verbs. *CoRR*, abs/1411.7942.

Purver, M., Gregoromichelaki, E., Meyer-Viol, W., and Cann, R. (2010). Splitting the 'I's and crossing the 'You's: Context, speech acts and grammar. In *Proc. 14th SemDial Workshop*, pages 43–50.

Sadrzadeh, M., Purver, M., Hough, J., and Kempson, R. (2018a). Exploring semantic incrementality with Dynamic Syntax and vector space semantics. In *Proceedings of the 22nd SemDial Workshop on the Semantics and Pragmatics of Dialogue (AixDial)*, pages 122–131, Aix-en-Provence.

Sadrzadeh, M., Purver, M., and Kempson, R. (2018b). A tensor-based vector space semantics for Dynamic Syntax. In *Proceedings of the 2nd Dynamic Syntax Conference*, Edinburgh, UK.