Interactive Meaning Construction
A Workshop at IWCS 2015
Proceedings

Edited by:
Ruth Kempson, Robin Cooper, Matthew Purver

Queen Mary University of London
14th April 2015

IMC (Interactive Meaning Construction) was a workshop held in conjunction with the 11th International Conference on Computational Semantics (IWCS 2015), at Queen Mary University of London, UK on 14th April 2015.

Despite being the mainstay of our language experience, the data of conversational dialogue pose very considerable challenges for semantic modelling. They violate expectations provided by standard frameworks, with apparently incomplete and/or highly context-dependent fragments widespread. Conventional grammar frameworks are poorly set up to reflect these dynamics, but the goal of defining models able to reflect them is an active research area. However, dialogue phenomena and data provide us with evidence about intended and understood meaning which can help define more suitable approaches, either through inspection or through computational methods (especially given the recent progress in distributional and inferential methods for deriving semantic representations and parsers from context). This workshop was conceived to bring together researchers addressing these issues and assess the significance of this ongoing work both for approaches to semantics and for computational modelling.

Contents

Computing Discourse Structures for Dialogues from the Stac Corpus

Nicholas Asher
IRIT, Toulouse
nicholas.asher@irit.fr

My talk will give an interview of the work colleagues and I have done on the STAC project with a particular focus on the computing full discourse structures for negotiation dialogues. The dialogues come from chat interactions between player of an online version of the board game Settlers of Catan. The dialogues are multiparty chat conversations and exhibit more complex structures and offer more challenging annotation tasks than monologue for several reasons that I will explore in the talk. I will then discuss some discourse parsing efforts and how they fit into the broader context of the project.

Language as a set of mechanisms for interaction

Stergios Chatzikyriakidis
LIRMM, University of Montpellier 2
stergios.chatzikyriakidis@lirmm.fr

Eleni Gregoromichelaki
King's College London
elenigregor@gmail.com

Ruth Kempson
King's College London
ruth.kempson@kcl.ac.uk

We will be using phenomena of conversational interaction to highlight the need for a reconsideration of the standard conception of the syntax/semantics interface. We shall argue that the modelling of such phenomena requires that natural languages need to be seen procedurally as mechanisms enabling participant-coordination during interaction (Dynamic Syntax, DS, (Kempson et al., 2001; Cann et al., 2005; Gregoromichelaki et al., 2011)).

The argument will proceed in five steps:

(1) the observation that ALL syntactic-semantic dependencies can be distributed across more than one participant, a phenomenon problematic for conventional grammar frameworks;

(2) a sketch of DS, illustrated by a derivation of English *wh*-questions, demonstrating successful modelling of local and non-local discontinuity effects across more than one participant, in virtue of the assumption of "syntax" as mirrored actions for parsing/linearising strings and meaning representations, performed by parser and producer in parallel;

(3) a prediction of parallelism between mechanisms underpinning discontinuity effects and anaphora, both modelled as licensing resolution for some initial underspecification in three ways: (a) from some linguistically-provided antecedent source, (b) indexically, (c) by subsequent structurally-provided resolution;

(4) a prediction that in all cases this range of update effects can be distributed across participants;

(5) arguing that the confirmed generalisation of (3)-(4) is available only to models of syntax defined non-representationally as driving the time-linear participant coordination.

If time, we shall round out the theoretical claim with an extension of the same form of analysis to morpho-syntax, specifically *clitic-clustering* (one of the most notoriously opaque morphosyntactic phenomena (Rezac, 2010; Chatzikyriakidis and Kempson, 2011, among many others). We argue that clitic clusters are diachronic reflexes of local discontinuity effects subject to the same constraints on processing as applicable to discontinuity effects (only one unfixed node of a type at a time), despite their different status vis-a-vis lexical storage. Hence the overall claim that language as a general system should be seen as a set of mechanisms for interaction.

References

Cann, R., R. Kempson, and L. Marten (2005). *The Dynamics of Language*. Oxford: Elsevier.

Chatzikyriakidis, S. and R. Kempson (2011). Standard modern and pontic greek person restrictions: A feature-free dynamic account. *Journal of Greek Lingusitics*, 127–166.

Gregoromichelaki, E., R. Kempson, M. Purver, G. J. Mills, R. Cann, W. Meyer-Viol, and P. G. T. Healey (2011). Incrementality and intention-recognition in utterance processing. *Dialogue and Discourse 2*(1), 199–233.

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

Rezac, M. (2010). *Phi-features and the modular architecture of language*, Volume 81. Springer Science & Business Media.

Proper Names in Interaction

Robin Cooper
University of Gothenburg
cooper@ling.gu.se

The traditional treatment of proper names in formal semantics (following Montague's classic work in PTQ, Montague, 1973) tells us little about the communicative processes associated with utterances of proper names. In Cooper (2013) we pointed out that this kind of analysis does not give us any way of placing the requirement on the interlocutor's gameboard that there already be a person named Sam available in order to integrate the new information onto the gameboard. As Ginzburg (2012) points out, the successful use of a proper name to refer to an individual a requires that the name be publically known as a name for a. We will follow the analysis of Cooper (2013) in parametrizing the content. A *parametric content* is a function which maps a context to a content. As such it relates to Montague's technical notion of *meaning* in his paper 'Universal Grammar' (Montague, 1970, 1974) where he regarded meaning as a function from possible worlds and contexts of use to denotations. This also corresponds to the notion of *character* in Kaplan (1978).

Our basic proposal for the compositional semantics in TTR (Cooper, 2012, prep) is to associate utterances of a proper name like *Sam* with the paramtric content in (1).

(1) $$\lambda r{:}\begin{bmatrix}\text{x}{:}\textit{Ind}\\ \text{e}{:}\text{named(x, “Sam”)}\end{bmatrix}\ .\\ \lambda P{:}\textit{Ppty}\ .\ P(r)$$

That is, a function from contexts where there is an individual named Sam to a function from properties to the result of applying the property to that context. A property such as (2) will predicate 'run' of the individual in the context.

(2) $$\lambda r{:}\begin{bmatrix}\text{x}{:}\textit{Ind}\end{bmatrix}\ .\ \begin{bmatrix}\text{e}{:}\text{run}(r.\text{x})\end{bmatrix}$$

This is closely related to treatments of proper names that were proposed earlier in situation semantics (Gawron and Peters, 1990; Cooper, 1991; Barwise and Cooper, 1993). A more recent close relation is Maier's (2009) proposal for the treatment of proper names in terms of layered discourse representation theory (LDRT).

The domain of this function plays something of the role of a presupposition. However, it is not a presupposition in the sense of requiring there to be an individual named Sam, but rather that there should be a match on the dialogue participant's gameboard for an individual named Sam. If a match is not found on the shared commitments (FACTS) of the gameboard then the agent looks for a match in long term memory and if this too fails then the agent adds a new (unanchored) match to the shared commitments on the gameboard. We show how to make this precise using the tools of TTR and argue that it yields a formal approach to a notion of salience.

References

Barwise, J. and R. Cooper (1993). Extended Kamp Notation: a Graphical Notation for Situation Theory. In P. Aczel, D. Israel, Y. Katagiri, and S. Peters (Eds.), *Situation Theory and its Applications*, Volume 3. Stanford: CSLI.

Cooper, R. (1991). Three lectures on situation theoretic grammar. In M. Filgueiras, L. Damas, N. Moreira, and A. P. Tomás (Eds.), *Natural Language Processing, EAIA 90, Proceedings*, Number 476 in Lecture Notes in Artificial Intelligence, pp. 101–140. Berlin: Springer Verlag.

Cooper, R. (2012). Type theory and semantics in flux. In R. Kempson, N. Asher, and T. Fernando (Eds.), *Handbook of the Philosophy of Science*, Volume 14: Philosophy of Linguistics, pp. 271–323. Elsevier BV. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

Cooper, R. (2013). Update conditions and intensionality in a type-theoretic approach to dialogue semantics. In R. Fernández and A. Isard (Eds.), *Proceedings of the 17th Workshop on the Semantics and Pragmatics of Dialogue*, pp. 15–24. University of Amsterdam.

Cooper, R. (in prep). Type theory and language: from perception to linguistic communication. Draft of book chapters available from `https://sites.google.com/site/typetheorywithrecords/drafts`.

Gawron, J. M. and S. Peters (1990). *Anaphora and Quantfication in Situation Semantics*. CSLI Publications.

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

Kaplan, D. (1978). On the Logic of Demonstratives. *Journal of Philosophical Logic 8*, 81–98.

Maier, E. (2009). Proper names and indexicals trigger rigid presuppositions. *Journal of Semantics 26*, 253–315.

Montague, R. (1970). Universal Grammar. *Theoria 36*, 373–398.

Montague, R. (1973). The Proper Treatment of Quantification in Ordinary English. In J. Hintikka, J. Moravcsik, and P. Suppes (Eds.), *Approaches to Natural Language: Proceedings of the 1970 Stanford Workshop on Grammar and Semantics*, pp. 247–270. Dordrecht: D. Reidel Publishing Company.

Montague, R. (1974). *Formal Philosophy: Selected Papers of Richard Montague*. New Haven: Yale University Press. ed. and with an introduction by Richmond H. Thomason.

Inferring Meaning From Disfluencies in an Incremental Dialogue Framework

Julian Hough and David Schlangen
Bielefeld University
{julian.hough,david.schlangen}@uni-bielefeld.de

It has been established from psycholinguistics that disfluencies such as self-repairs, filled pauses and hesitations have meaning in dialogue which requires computing, rather than being filtered out as noise (Brennan, 2000; Brennan and Schober, 2001; Arnold et al., 2007). Computing this meaning for a dialogue system or formal dialogue model is challenging: it requires breaking free of traditional notions of grammar, and of the traditionally conceived competence-performance distinction (Chomsky, 1965). It requires modelling semantic update for dialogue participants' states on at least as fine-grained a level as word-by-word and taking linguistic *actions* to be first class citizens of the context, in addition to *time*.

Here we present a formal model to do this using elements of the Incremental Unit (IU) framework (Schlangen and Skantze, 2011) and Dynamic Syntax with Type Theory with Records (DS-TTR) (Purver et al., 2011), an inherently incremental grammar formalism. We then discuss an approach to modelling real-time probabilistic inference from disfluency in accordance with Brennan and Schober (2001)'s results in a simple referring expression game, namely by integrating this model into probabilistic TTR with a simple notion of relevance for questions under discussion, as presented in Hough and Purver (2014).

We go on to discuss the consequences of such an approach for more general dialogue models, and how one attempt to integrate meaning of disfluency into the dialogue state of the KoS framework in Ginzburg et al. (2014). Finally we report the current progress of the DUEL ('Disfluencies, Exclamations and Laughter in dialogue') project (Ginzburg et al., 2014) and our plans for implementation into a working dialogue system.

References

Arnold, J. E., C. L. H. Kam, and M. K. Tanenhaus (2007). If you say 'thee uh' you are describing something hard: The on-line attribution of disfluency during reference comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition 33*(5), 914.

Brennan, S. and M. Schober (2001). How listeners compensate for disfluencies in spontaneous speech* 1. *Journal of Memory and Language 44*(2), 274–296.

Brennan, S. E. (2000). Processes that shape conversation and their implications for computational linguistics. In *Proceedings of the 38th annual meeting of the ACL*, pp. 1–11. ACL.

Chomsky, N. (1965). *Aspects of the Theory of Syntax.* Cambridge, MA: MIT.

Ginzburg, J., R. Fernández, and D. Schlangen (2014, June). Disfluencies as intra-utterance dialogue moves. *Semantics and Pragmatics 7*(9), 1–64.

Ginzburg, J., Y. Tian, P. Amsili, C. Beyssade, B. Hemforth, Y. Mathieu, C. Saillard, J. Hough, S. Kousidis, and D. Schlangen (2014). The Disfluency, Exclamation and Laughter in Dialogue (DUEL) Project. In *Proceedings of the 18th SemDial Workshop (DialWatt)*, Herriot Watt University, Edinburgh, pp. 176–178.

Hough, J. and M. Purver (2014). Probabilistic type theory for incremental dialogue processing. In *Proceedings of the EACL 2014 TTNLS Workshop*, Gothenburg, Sweden, pp. 80–88. ACL.

Purver, M., A. Eshghi, and J. Hough (2011). Incremental semantic construction in a dialogue system. In J. Bos and S. Pulman (Eds.), *Proceedings of the 9th IWCS*, Oxford, UK, pp. 365–369.

Schlangen, D. and G. Skantze (2011). A General, Abstract Model of Incremental Dialogue Processing. *Dialoge & Discourse 2*(1), 83–111.

What we ought to know: Making and breaking common ground

Christine Howes
Centre for Language Technology
University of Gothenburg
christine.howes@gu.se

Patrick G. T. Healey
Cognitive Science Research Group
Queen Mary University of London
p.healey@qmul.ac.uk

Turn construction in dialogue is a fundamentally incremental and interactive process (Goodwin, 1979), and the coordination of common ground is crucial to understanding. However, although the establishment of common ground is known to be influenced by a number of factors in dialogue, such as the context in which information was mentioned and partner commitment and engagement (Brown-Schmidt, 2012), many accounts assume that interaction plays only a peripheral role (Keysar, 2007). Additionally, contributions to dialogue are often fragmentary or incomplete (Fernández and Ginzburg, 2002) and these incomplete contributions may also be grounded (Eshghi et al., 2015), clarified or subsequently completed (Gregoromichelaki et al., 2011). Despite these observations, there has been little work that experimentally tests the influence of common ground on the interactive building up of meanings in dialogue at the sub-sentential level, or to what extent we take account of shared context when we are constructing a turn.

Using the DiET chat tool (Healey et al., 2003), we report a series of experiments that alter turns in an ongoing dialogue to see how this collaborative process of building common ground with an interlocutor affects people's interpretations of and responses to clarification questions and incomplete utterances. In two experiments we systematically intervene in a real time text chat, by targeting a noun phrase that has already been talked about in the current dialogue (i.e. it is *given* information) or one that has not been mentioned previously (i.e. it is *new*). In the first experiment, we introduce spoof clarification requests, querying a given or a new noun phrase, and appearing to come from either the other person in the conversation or an external source. The second experiment truncates genuine turns in between a determiner and noun in a given or new noun phrase (following Howes et al., 2012).

Results show that whether something has been previously parsed or produced affects responses. However, there are additional effects that can only be accounted for by taking into account the joint action of building common ground. A formal model of dialogue needs to take into account not just what is said and how, but also who is actively involved in the process of doing so.

References

Brown-Schmidt, S. (2012). Beyond common and privileged: Gradient representations of common ground in real-time language use. *Language and Cognitive Processes 27*(1), 62–89.

Eshghi, A., C. Howes, E. Gregoromichelaki, J. Hough, and M. Purver (2015). Feedback in conversation as incremental semantic update. In *Proceedings of the 11th International Conference on Computational Semantics (IWCS 2015)*, London, UK. Association for Computational Linguisitics.

Fernández, R. and J. Ginzburg (2002). Non-sentential utterances: A corpus-based study. *Traitement Automatique des Langues 43*(2).

Goodwin, C. (1979). The interactive construction of a sentence in natural conversation. In G. Psathas (Ed.), *Everyday Language: Studies in Ethnomethodology*, pp. 97–121. New York: Irvington Publishers.

Gregoromichelaki, E., R. Kempson, M. Purver, G. J. Mills, R. Cann, W. Meyer-Viol, and P. G. T. Healey (2011). Incrementality and intention-recognition in utterance processing. *Dialogue and Discourse 2*(1), 199–233.

Healey, P. G. T., M. Purver, J. King, J. Ginzburg, and G. Mills (2003, August). Experimenting with clarification in dialogue. In *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*, Boston, Massachusetts.

Howes, C., P. G. T. Healey, M. Purver, and A. Eshghi (2012, August). Finishing each other's ... responding to incomplete contributions in dialogue. In *Proceedings of the 34th Annual Meeting of the Cognitive Science Society (CogSci 2012)*, Sapporo, Japan, pp. 479–484.

Keysar, B. (2007). Communication and miscommunication: The role of egocentric processes. *Intercultural Pragmatics 4*(1), 71–84.

A Model for Attention-Driven Judgements in Type Theory with Records

John D. Kelleher
Dublin Institute of Technology
School of Computing, Dublin, Ireland
john.d.kelleher@dit.ie

Simon Dobnik
University of Gothenburg
Dept. of Philosophy, Linguistics and Theory of Science
Gothenburg, Sweden
simon.dobnik@gu.se

Recently, Type Theory with Records (TTR, (Cooper, 2012; Cooper et al., 2014)) has been proposed as a formal representational framework and a semantic model for embodied agents participating in situated dialogues (Dobnik et al., 2014). Although TTR has many potential advantages as a semantic model for embodied agents, one problem it faces is the combinatorial explosion of types that is implicit in the framework and is due the fact that new types can be created or learned by an agent dynamically. Types are *intensional* which means that a given situation in the world may be assigned more than one record type. A sensory reading of a particular situation in the world involving spatial arrangement of objects may be assigned several record types of spatial relations simultaneously, for example *Left*, *Near*, *At*, *Behind*, etc. TTR also incorporates the notion of *sub-typing* which allows comparison of types. A situation judged as being of a particular record type may also be judged of potentially infinite number of its sub-types: a situation of type *Table-Left-Chair* is also of type *Table* and *Left*, etc.

The rich type system of TTR gives us a lot of flexibility in modelling natural language semantics. However, unfortunately, the flexibility with which types are assigned to records of situations and which is also required by modelling natural language and human cognition comes with a computational cost. Since each type assignment involves a binary judgement (something is of a type T or not) for each record of situation an agent having an inventory of n types can make n assignments with 2^n possible outcomes, hence for $n = 3$, $2^3 = 8$: $\{\}$, $\{T_1\}$, $\{T_2\}$, $\{T_3\}$, $\{T_1, T_2\}$, $\{T_1, T_3\}$, $\{T_2, T_3\}$ and $\{T_1, T_2, T_3\}$. Such combinatorial explosions of possible outcomes of type assignments or judgements present a great difficulty for an agent that is trying to learn what types to assign to a situation from the linguistic behaviour of another agent.

In this presentation we argue that agents need (i) a judgement control mechanism and (ii) a method for organising their type inventory. For (i) we propose the Load Theory of selective attention and cognitive control (Lavie et al., 2004) to be a suitable candidate. This model of attention distinguishes between two mechanisms of selective attention: *perceptual selection* and *cognitive control*. Perceptual selection is a mechanism that excludes the perception of task irrelevant distractors under situations of high perceptual load; however, in situations of low perceptual load any spare capacity will spill over to the perception of distractor objects. The cognitive control mechanism is an active process that reduces the interference from perceived distractors on task response. It does so by actively maintaining the processing prioritisation of task relevant stimuli within the set of perceived stimuli. It follows, that agents make judgements of several different kinds which we call (i) pre-attentive, (ii) task induced, and (iii) context induced judgements. Pre-attentive judgements (the segmentation of a visual scene into entities and background) are controlled by the perceptual selection mechanism of Load Theory. Task induced and context induced judgements require conscious attention. As such, they are controlled by the cognitive control mechanisms of Load Theory. These judgements are applied to types that are in working memory and result in new types being introduced to working memory. Task induced and context induced judgements are primed by the types associated (via memory) with the current activities that the agent is currently engaged in (making a cup of tea) and their physical location (the plate beside the kettle is very hot).

For the requirement (ii) above we propose that agents organise their type inventory into subsets or bundles of types that are represented as cognitive states. These can be thought of as sensitivities towards certain objects, events, and situations where the mapping between states and entities has been learned from experience. More than one cognitive state may be active at any moment. We propose that a (probabilistic) POMDP framework (Partially Observable Markov Decision Processes, (Kaelbling et al., 1998)) provides a useful mathematical model for implementations of a control structure for judgements in an embodied agent/robot using TTR that has learned or been given by its designer a number of cognitive states. We map the problem of controlling judgements within TTR to a POMDP control problem as follows: (i) the cognitive states of the agent are mapped to the states in the belief state of the POMDP, (ii) the priming of an agent to observe certain types is mapped to the action specified for the current belief state by the policy driven by the attention mechanisms, (iii) the types an agent actually perceives and processes are mapped to the observations the agent receives from its sensors, (iv) the benefits to the agent of being primed to comprehend the world are mapped to the reward function.

Overall, we hope that the account presented here provides a move towards linking the formal semantic representation of TTR with cognitive attentional mechanisms.

References

Cooper, R. (2012). Type theory and semantics in flux. In R. Kempson, N. Asher, and T. Fernando (Eds.), *Handbook of the Philosophy of Science*, Volume 14 of *General editors: Dov M Gabbay, Paul Thagard and John Woods*. Elsevier BV.

Cooper, R., S. Dobnik, S. Lappin, and S. Larsson (2014, 27 April). A probabilistic rich type theory for semantic interpretation. In *Proceedings of the EACL 2014 Workshop on Type Theory and Natural Language Semantics (TTNLS)*, Gothenburg, Sweden, pp. 72–79. Association for Computational Linguistics.

Dobnik, S., R. Cooper, and S. Larsson (2014, August 31st). Representing different kinds of reference with type theory with records. In *Proceedings of RefNet Workshop on Psychological and Computational Models of Reference Comprehension and Production*. Edinburgh University Informatics Forum.

Kaelbling, L. P., M. L. Littman, and A. R. Cassandra (1998). Planning and acting in partially observable stochastic domains. *Artificial intelligence 101*(1), 99–134.

Lavie, N., A. Hirst, J. W. de Fockert, and E. Viding (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General 133*(3), 339–354.

Situated Construction and Coordination of Perceptual Meaning

Staffan Larsson
University of Gothenburg
sl@ling.gu.se

Part of learning a language, it seems, is learning to identify the individuals and situations that are in the extension of the phrases and sentences of the language. For many concrete expressions, this identification relies crucially on the ability to perceive the world, and to use perceptual information to classify individuals and situations as falling under a given linguistic description or not. This view was first put forward by Harnad (1990) as a way of addressing the "symbol grounding problem" in artificial intelligence.

This talk will take as its starting point the formalisation of perceptual meaning in Larsson (2013). There, agents' judgements about situations as being of different situation types are regarded as being the result of applying statistical classifiers to low-level perceptual input. Crucially, these classifiers are trained in interaction with other agents in a process of semantic coordination, often taking place in spoken dialogue. To integrate classification of perceptual data with formal semantics, we are using TTR (Type Theory with Records), a framework developed with a view to giving an abstract formal account of natural language interpretation Cooper (2012), as our formalism and foundational semantic theory. TTR starts from the idea that information and meaning is founded on our ability to perceive and classify the world, i.e., to perceive objects and situations as being of types. TTR allows perceptual classifier functions to be formalised and used in representing meanings of linguistic expressions together with the high-level conceptual aspects of meaning traditionally studied in formal semantics.

We will focus on how processes of contextual interpretation, construction, and coordination on perceptual meanings can be described in TTR. Very roughly, this involves the following steps:

1. Compositional construction of the "literal" meaning of the utterance in question
2. Situated interpretation of utterance given the context
3. Evaluation of the acceptability of the utterance in context
4. Updating the (agent's take on) the context
5. Updating the agent's take on the meanings of the expressions used in the utterance

We will discuss the challenges involved in the effort to give a comprehensive account of these processes. We will also discuss how interaction can come in at various stages of this process to resolve problems, and how interaction can be included in the formal account of utterance interpretation, construction, and coordination. Finally, we will discuss open problems and plans for future work.

References

Cooper, R. (2012). Type theory and semantics in flux. In R. Kempson, N. Asher, and T. Fernando (Eds.), *Handbook of the Philosophy of Science*, Volume 14: Philosophy of Linguistics. Elsevier BV. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena 42*(1990), 335–346.

Larsson, S. (2013). Formal semantics for perceptual classification. *Journal of Logic and Computation*.

Deep Reinforcement Learning for constructing meaning by 'babbling'

Oliver Lemon
Interaction Lab
Heriot-Watt University
o.lemon@hw.ac.uk

Arash Eshghi
Interaction Lab
Heriot-Watt University
a.eshghi@hw.ac.uk

Even within a simple dialogue domain, there's a lot of variation in language use that does not ultimately affect the overall communicative goal. For example, in the travel domain, the following dialogues all lead to a context in which A is committed to finding a ticket for B from London to Paris: (a) *A: Where would you like to go? B: Paris, from London;* (b) *B: I would like to go to Paris; A: Sure, where from? B: London*; (c) *B: I need to get to Paris from London; A: OK.* These dialogues can be said to be *pragmatically synonymous* modulo the travel domain, since they all achieve the same communicative goal – but they consist of different sequences of "dialogue acts" (see below). Humans are able to learn how to use such different word and interaction sequences to achieve their goals in this domain, and in many others, and they are able to transfer this learning to use the same actions to achieve related goals in other domains.

However, current dialogue technology doesn't learn about pragmatic synonymy. Rather, the pragmatic equivalence of word sequences is enforced by the use of domain-specific representations, rendering today's systems and tools difficult to re-use or extend for new applications.

Regarding learning methods, statistical approaches to Spoken Language Understanding (SLU), Dialogue Management (DM), and Natural Language Generation (NLG) have received considerable attention in the last decade or so. An important area of progress has been in the use of Reinforcement Learning methods with Markov Decision Process models for automatic, data-driven optimisation of DM and NLG, e.g. Rieser and Lemon (2011); Gasic et al. (2010); Lemon and Pietquin (2012). There are three main potential advantages claimed for such approaches: (1) the ability to robustly handle noise and uncertainty; (2) automatic optimisation of action-selection, i.e. minimising human involvement in SDS design and development; and (3) easier portability to new domains, with the possibility of reusable statistical methods for constructing / training system components from data. However, the last two of these advantages have yet to come to full fruition: a major challenge has been that current statistical learning approaches demand substantial amounts of dialogue data annotated with domain-specific semantic and pragmatic information, in the form of handcrafted and domain-specific "dialogue acts" (DAs), e.g. Young et al. (2009); Henderson et al. (2008); Gasic et al. (2010); Rieser and Lemon (2011).

1 Domain-Specificity of Dialogue Acts

Semantic processing in real-world SDS remains invariably domain-specific, with pragmatic synonymy being enforced by the use of hand-crafted representations using a variety of 'Dialogue Act' (DA) schemes, rather than being learned through interaction. DAs have been used because they provide a level of representation which *abstracts over* the variability in human dialogue behaviour (e.g. use of different words, phrases etc.) while providing a representation of the specific information that is needed to complete tasks in a domain. They have thus performed an important role in reducing the size of the state spaces and action sets to be considered in dialogue processing, and in learning DM and NLG policies. DAs have also been important in defining interfaces between SLU, DM, and NLG components.

Even with modern trainable approaches using machine learning methods, DAs are still used, and large amounts of (expensive) domain-specific annotated data are required to train SLU, DM, and NLG components. For example, in a tourism domain, an utterance such as "I want a fancy Thai restaurant" is mapped to a hand-crafted DA representation such as $INFORM(foodType = Thai, price = expensive)$, e.g.Young et al. (2009); Rieser and Lemon (2011); Gasic et al. (2010). Similar issues arise for NLG.

This use of hand-crafted, domain-specific DAs has three main disadvantages: (A) transfer of a SDS to a new application requires new DAs to be defined by domain experts; (B) data for learning in the new domain needs to be annotated with the new DAs; (C) the DA representation may either under- or over-estimate the features required for learning good DM and/or NLG policies for the domain.

2 Deep Learning for successful and grammatical "Babble"

We propose to overcome these problems by bypassing the need for hand-crafted DA representations, instead allowing systems to learn / explore how to achieve tasks by 'babbling' at the level of generating grammatical and task-effective word sequences, rather than by planning DAs (Eshghi and Lemon, 2014). This idea is related to a compositional version of Ramsey's "Success Semantics" (Blackburn, 2010).

In the BABBLE project[1], we will explore deep reinforcement learning methods (Mnih et al., 2015) which can avoid the use of DAs in dialogue processing. Deep learning allows efficient representations to be derived automatically, for example starting with basic semantic units which are delivered by potentially wide-coverage / domain-general semantic parsers (such as Dynamic Syntax or CCG). Effectively, these units will form the "atoms" of meaning, which are then combined on-the-fly to create structures which achieve goals in a domain (e.g. booking a flight or finding a restaurant). Given data for successful dialogues in a domain, learning how to combine these meaning-atoms successfully will be done using deep reinforcement learning combined with with an incremental semantic grammar (i.e. word-by-word "babbling" of grammatical strings from the language). This method would not require DAs, instead allowing an appropriate level of meaning representation for a domain to emerge through deep reinforcement learning, therefore avoiding the problems (A, B, C) described above.

Acknowledgements

This project is funded by the EPSRC, under grant number EP/M01553X/1.

References

Blackburn, S. (2010). Success semantics. In *Practical Tortoise Raising*. Oxford University Press.

Eshghi, A. and O. Lemon (2014). How domain-general can we be? Learning incremental Dialogue Systems without Dialogue Acts. In *Proceedings of Semdial*.

Gasic, M., F. Jurcicek, S. Keizer, F. Mairesse, B. Thomson, and S. Young (2010). Gaussian Processes for Fast Policy Optimisation of POMDP-based Dialogue Managers. In *Proc. SIGDIAL*.

Henderson, J., O. Lemon, and K. Georgila (2008). Hybrid Reinforcement / Supervised Learning of Dialogue Policies from Fixed Datasets. *Computational Linguistics 34*(4), 487 – 513.

Lemon, O. and O. Pietquin (Eds.) (2012). *Data-driven Methods for Adaptive Spoken Dialogue Systems: Computational Learning for Conversational Interfaces*. Springer.

Mnih, V., K. Kavukcuoglu, D. Silver, A. Rusu, J. Veness, M. Bellemare, A. Graves, M. Riedmiller, A. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, and D. Hassabis (2015). Human-level control through deep reinforcement learning. *Nature 518*, 529–533.

Rieser, V. and O. Lemon (2011). Learning and evaluation of dialogue strategies for new applications: Empirical methods for optimization from small data sets. *Computational Linguistics 37*(1), 153–196.

Young, S., M. Gasic, S. Keizer, F. Mairesse, J. Schatzmann, B. Thomson, and K. Yu (2009). The Hidden Information State Model: a practical framework for POMDP-based spoken dialogue management. *Computer Speech and Language 24*(2), 150–174.

[1] http://sites.google.com/site/hwinteractionlab/babble

Two accounts of ambiguity in a dialogical theory of meaning

Paul Piwek
Centre for Research in Computing
The Open University, UK
paul.piwek@open.ac.uk

Abstract

Recent work on dialogue has emphasised the procedural and, more specifically, incremental nature of meaning construction (e.g., Gregoromichelaki et al., 2011): semantic representations are built up on a word-by-word basis, with each word viewed as an action operating on the representation. In a complementary effort, inspired by Robert Brandom's inferentialist philosophy (Brandom, 1994), an account of the semantic import of such representations is proposed. In contrast with traditional declarative accounts of meanings as inhabitants of a set-theoretically given universe (e.g., truth conditions), the current semantics is dialogical and algorithmic. It combines the proposals put forward in Piwek (2011) and Piwek (2014).[1]

In the second part of the talk, we examine how the proposed machinery can accommodate empirical findings on the use of lexically ambiguous expressions. We focus on the extremely strong tendency for meanings to be constant: within a single discourse, different occurrences of the same expression tend to have the same meaning (see Gale et al., 1992). We examine potential explanations for these findings, whilst taking into account rare, but real, exceptions as in 'The pitcher was drinking wine from a pitcher'. In this example, the first occurrence of the word 'pitcher' concerns a baseball player and the second a jug (see van Deemter, 1996).

We distinguish between two possible strategies:

1. a fairly common strategy, where ambiguity is encoded by the semantic representations, using some form of indexing and underspecification; and

2. a strategy (which, as far as we know, is novel) in which ambiguity is modelled by how language users reason with the expressions: disambiguation is no longer thought of as a refinement of the semantic representations (e.g., by fixing the value of a variable or addition of a restriction). Rather, disambiguation is modelled in terms of how the inferential machinery that allows language users to reason with semantic representations is configured. Roughly speaking, choosing an interpretation for an expression corresponds with activating the inference rules that go with that interpretation (whilst blocking rules that correspond with alternative interpretations).

We will see that the second approach suggests a powerful explanation for the constancy of meanings. It is, however, less well-equipped when it comes to dealing with exceptions to constancy. This leads us into a discussion of the merits and downsides of different ways in which theories account for exceptions. In particular, we will contrast two different approaches, which roughly correspond with the aforementioned two strategies:

- different representations, but uniform processing mechanisms *versus*

- uniform representations, but multiple processing mechanisms (i.e., different processing mechanisms for frequent cases versus exceptions).

[1] See also Kibble (2006) and Millson (2014) for partial formalisations of Brandom's work.

References

Brandom, R. (1994). *Making It Explicit: reasoning, representing, and discursive commitment*. Cambridge, Mass.: Harvard University Press.

Gale, W., K. Church, and D. Yarowsky (1992). One sense per discourse. In *Proceedings of the 4th DARPA Speech and Natural Language Workshop*, pp. 233–237.

Gregoromichelaki, E., R. Kempson, M. Purver, G. Mills, R. Cann, W. Meyer-Viol, and P. Healey (2011). Incrementality and Intention-recognition in Utterance Processing. *Discourse & Dialogue 2*(1), 199–233.

Kibble, R. (2006). Reasoning about propositional commitments in dialogue. *Research on Language and Computation 4*(2-3), 179–202.

Millson, J. (2014). Queries and assertions in minimally discursive practices. In *Proceedings of 50th Annual Convention of the AISB (Symposium on Questions, Dialogue and Discourse)*, London, UK.

Piwek, P. (2011). Dialogue structure and logical expressivism. *Synthese 183*(1), 33–58.

Piwek, P. (2014). Towards a computational account of inferentialist meaning. In *Proceedings of 50th Annual Convention of the AISB (Symposium on Questions, Dialogue and Discourse)*, London, UK.

van Deemter, K. (1996). Towards a logic of ambiguous expressions. In K. van Deemter and S. Peters (Eds.), *Semantic ambiguity and underspecification*, Stanford, CA, pp. 203–237. CSLI Publications.

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.

Dialogue as Computational Metasemantics

Matthew Stone
Rutgers University
matthew.stone@rutgers.edu

Idealized notions of computation have long served as intuition pumps for exploring our concepts of meaning (Turing, 1950; Searle, 1980; Shieber, 2007). But progress in language technology and cognitive science has now made it possible to implement theoretically-informed systems that people can actually talk to. We can now use these systems—and people's responses to them—to sharpen our understanding of the sources of meaning. I will describe some new ways such efforts can help to substantiate the fundamental importance of dialogue to a science of meaning.

Dialogue is the natural setting for language use, so it's not surprising that the study of dialogue can yield diverse insights into linguistic structure and function. Dialogue gives important evidence about the rules of meaning (Ginzburg and Cooper, 2004), and reveals profound constraints on the relationship of meaning and grammar (Gregoromichelaki et al., 2011). But philosophers offer diverse reasons to think interaction is essential to human meaning (Wittgenstein, 1953; Kripke, 1972; Ludlow, 2014); the same considerations apply to meaning in interactive computer systems (DeVault et al., 2006). In fact, I argue here, ongoing computational research can finally let us explore the interactions that make meaning possible—in humans and machines—and get clearer on the knowledge, expectations, choices and relationships that must be involved.

My starting point is work that I have been doing with my collaborators Bert Baumgaertner, Raquel Fernandez, Brian McMahan, Timothy Meo and Joshua Gang on broad coverage models of the meanings of English color terms (Baumgaertner et al., 2012; Meo et al., 2014; McMahan and Stone, 2015). Our data, results, models and visualization software are available online at `http://mcmahan.io/lux/`. These models make it possible to ask find-grained computational questions about how people make meaningful choices, update the context and pursue meaning in dialogue. We are particularly excited about these questions:

- Can we find better ways to distinguish between semantics and pragmatics? We built a machine learning model with the surprising but apparently accurate prediction that there should be NO Gricean implicatures in our conversational setting, because semantics is already a NASH EQUILIBRIUM—following Lewis (1969) and Cumming (2013). So the information the hearer recovers from an utterance is exactly the information the speaker intends to express. Computational models show how problematic it can be to characterize speakers' communicative goals and collaborative rationality: different approaches lead to strikingly different predictions about the scope and limits of meaning. But they also give the tools to test which assumptions best fit ordinary conversation—with sometimes counterintuitive results.

- How do agents develop a consistent and coherent interpretation of one another across interactions? Researchers like Lewis (1979) and Barker (2002) posit that vague language is negotiated across interactions, but hypothesize that people seem often to aim to agree on meaningful standards of interpretation that they can apply to all cases. Our models let us simulate, quantify and assess these effects—to show how far people are willing to work to develop a shared understanding of the distinctions that matter to them. Do people always build on what's come before? Do they aim to resolve future cases in natural and meaningful ways?

- What must agents do to pursue meaning in problematic cases? Famously, in *The Princess Bride*, after Vizzini has repeatedly expressed his incredulity at the feats of the Dread Pirate Roberts with the proclamation “Inconceivable!”, Inigo Montoya eventually responds, “You keep using that word. I do not think it means what you think it means.” Failure is as much a part of meaning as success. Among the many speakers we’ve seen who use *vermilion* to evoke a striking orangish-red, there are a few who use the word to describe emerald greens. Have those people mistakenly extrapolated from false friends like *verdigris* and *viridian*? Or are they party to another convention, which we should in time learn to defer to? How should a computer respond?

Across all these cases, we find that computational modeling, and the related methodologies of Bayesian cognitive science, make it possible to ground a view of meaning as an abstract goal of communication that’s only indirectly related to the antecedent knowledge and capabilities of interlocutors—not something we TAKE TO dialogue, but something we MAKE THROUGH dialogue.

References

Barker, C. (2002). The dynamics of vagueness. *Linguistics and Philosophy 25*(1), 1–36.

Baumgaertner, B., R. Fernández, and M. Stone (2012). Towards a flexible semantics: colour terms in collaborative reference tasks. In *Proceedings of the First Joint Conference on Lexical and Computational Semantics-Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation*, pp. 80–84. Association for Computational Linguistics.

Cumming, S. (2013). Coordination and content. *Philosophers’ Imprint 13*(4), 1–16.

DeVault, D., I. Oved, and M. Stone (2006). Societal grounding is essential to meaningful language use. In Y. Gil and R. Mooney (Eds.), *Proceedings of the Twenty-First National Conference on Artificial Intelligence (AAAI)*, pp. 747–754.

Ginzburg, J. and R. Cooper (2004). Clarification, ellipsis and the nature of contextual updates in dialogue. *Linguistics and Philosophy 27*(3), 297–366.

Gregoromichelaki, E., R. Kempson, M. Purver, G. J. Mills, R. Cann, W. Meyer-Viol, and P. G. T. Healey (2011). Incrementality and intention-recognition in utterance processing. *Dialogue and Discourse 2*(1), 199–233.

Kripke, S. (1972). Naming and necessity. In G. Harman and D. Davidson (Eds.), *Semantics of Natural Language*, pp. 253–355. Dordrecht: Reidel.

Lewis, D. (1979). Scorekeeping in a language game. *Journal of Philosophical Logic 8*, 339–359.

Lewis, D. K. (1969). *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press.

Ludlow, P. (2014). *Living Words: Meaning Underdetermination and the Dynamic Lexicon*. Oxford: Oxford University Press.

McMahan, B. and M. Stone (2015). A bayesian model of grounded color semantics. *Transactions of the Association for Computational Linguistics 3*, 103–115.

Meo, T., B. McMahan, and M. Stone (2014). Generating and resolving vague color references. In *SEMDIAL 2014: THE 18th Workshop on the Semantics and Pragmatics of Dialogue*, pp. 107–115.

Searle, J. R. (1980). Minds, brains and programs. *Behavioral and Brain Sciences 3*(3), 417–457.

Shieber, S. M. (2007). The Turing test as interactive proof. *Noûs 41*(4), 686–713.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind 49*, 433–460.

Wittgenstein, L. (1953). *Philosophical Investigations*. New York: The Macmillan Company.