# Incrementality all the way up

Ellen Breitholtz, Christine Howes, Robin Cooper
University of Gothenburg, Sweden
ellen.breitholtz@ling.gu.se, christine.howes@gu.se, robin.cooper@ling.gu.se

September 5, 2017

A major challenge in formal analysis of discourse is that inferential connections between different parts of a text or dialogue are usually not logical in the technical sense. Instead they are based on knowledge (or beliefs) which are assumed to be shared between the agents involved in the discourse. Because of its interactive nature, allowing for feedback, clarification requests, etc. inference in dialogue may be even harder to describe formally. It is well known that language, in particular dialogue, is incremental at phonetic, syntactic and semantic levels. In this paper we will take a closer look at incremental interpretation of reasoning in dialogue. Our approach takes rhetorical reasoning as its point of departure and utilises some key concepts from classical rhetoric. We also suggest a way of analysing rhetorical arguments in dialogue using an information state update approach cast in Type Theory with Records, that will allow us to represent how dialogue participants draw inferences incrementally, that is tentatively make an inference, and then recompute it in the light of more specified information.

In (1) we see an example of how a dialogue participant interprets language incrementally. Following A's production of the noun phrase (NP) 'the doctor', i.e. before there is a complete proposition to interpret, B clarifies who is being referred to by A, thus indicating that for B, the referent of the NP is underspecified at this point in the exchange. Note that B could have produced a backchannel ('mm') here instead of the clarification request (CR), in which case we would not know whether or not the NP was underspecified for B – B may have correctly identified the specific doctor 'Chorlton' as the intended referent, or B might not have cared who the doctor was.

(1)  1. **A:** Er, the doctor

    2. **B:** Chorlton?

    3. **A:** Chorlton, mhm, he examined me . . .  [BNC: KPY 1006-1008]

Just as people are able to make low level interpretations incrementally from incomplete information, as in (1), we propose that higher level interpretation in dialogue is also incremental.

As illustrated in example (1), one of the key ways that we can investigate the incrementality inherent in dialogue is by examining examples of miscommunication. Miscommunication is often taken as something to be avoided in dialogue, as participants strive to align their mental representations with one another (Pickering and Garrod, 2004). However the local resolution of (potential) miscommunication at all levels is key in driving dialogue forward.

Examples of visible miscommunication also illustrate that interpreting what we have been given as early as possible means that incremental processing is non-monotonic. In example (1), B has offered a (tentative) fully specified interpretation in the CR ('Chorlton?'), but of course, this could be rejected by A (by saying e.g. "no, Fitzgerald"). In this case, B's initial interpretation is incompatible with A's subsequent turn, and will have to be abandoned and the new interpretation assimilated. Such mismatches can also occur in the reasoning undertaken by dialogue participants.

In example (2), we can see how one argument may be interpreted differently by two dialogue participants depending on the underpinning reasoning (topoi) they assume the argument to be based on. This is possible since arguments in dialogue are usually *enthymematic* i.e. drawing on tacit premises and principles of reasoning.

This kind of mismatch may go unnoticed in cases where consensus is reached. After all, if the interlocutors agree on the conclusion of the argument there is often no reason to argue about the rationale for agreeing. However, in example (2) it is made explicit that the speaker and the listener interpret the enthymeme in (2a) drawing on different topoi.

(2)   a.   **P:** Metal was actually the reason I started doing hip hop.

b.   **P:** ... Because I hated metal

c.   **J:** Oh, I thought you were going to say something completely different!

This snippet of dialogue is taken from a radio program where Swedish hip hop artist Petter is asked of his opinion of a song that has just been played. The example (2) demonstrates that incrementality in interaction concerns not only phonetic, syntactic and semantic aspects of language, but also pragmatic inferences. We propose a model to account for this using the notion of *underspecified* enthymemes (Breitholtz, 2014), illustrated using example (2).

Although great progress has been made over the past decade in the semantics and pragmatics of dialogue, some issues remain unsolved. One common problem for approaches to dialogue modelling is that they have an impoverished notion of context, since they are extensions of models developed to account for the truth value of sentences in terms of possible worlds rather than the meaning of utterances in interaction. There are, however, exceptions to this. By employing *information state* models of language (Ginzburg, 2012; Larsson and Traum, 2000) it is possible to account for how coordination of the dialogue gameboard (DGB) progresses with successive utterances. The DGB provides a structured characterisation of the information available to dialogue participants and offers a principled way in which asymmetries in shared knowledge can be represented. A common denominator for recent work along these lines (Ginzburg, 2012; Breitholtz, 2014; Cooper and Ginzburg, 2015) is that it uses Type Theory with Records (TTR) (Cooper and Ginzburg, 2015). TTR is based on the idea that humans perceive the world in terms of categories or *types*. In order to represent complex situations, TTR employs *record types*. A record type is a structure of pairs of labels and types where a label is like a variable associated with a type, e.g. individual (entity) and ptype (a type of situation constructed from a predicate and appropriate arguments).

The record types of TTR are used to model situation types and enable us to include virtually any feature from semantic context to phonology in our characterisation of a situation type, which makes TTR particularly well suited for modelling dialogue, where factors like context and intonation are often essential. Moreover, the notion of subtyping in TTR makes it possible to represent how a dialogue participant's take on the situation is incrementally built up during the course of a conversation. Record types offer a convenient way of doing this since we can easily add and remove fields, making a type more or less specified. Manifest fields in record types involve a restriction of the type in the field to a singleton type of one particular object of the type, thus enabling incremental partial specification of a record. TTR is also used by Purver et al. (2010); Eshghi et al. (2012) where it is combined with Dynamic Syntax (DS) (Kempson et al., 2001), an action-based grammar formalism which models the word-by-word incremental processing of linguistic input. In work by Breitholtz (2014); Breitholtz and Cooper (2011) dialogue semantics in TTR has been extended to include pragmatic reasoning, taking its point of departure in the Aristotelian concept of *enthymeme*. An enthymeme is an argument which relies on the listener supplying background knowledge or contextual information to correctly interpret the argument. If a dialogue participant presents the argument $P$ therefore $Q$, the interlocutor must supply a warrant that $P$ is a valid reason for $Q$ in order for the argument to be successful. The supressed principles of reasoning (such as $P$ implies $Q$ in the example above) warranting enthymemes are referred to as topoi (Aristotle, 2007; Ducrot, 1988). When we interact we expect topoi to be common ground, or to be accommodated (adopted by dialogue participants) during the course of the interaction.

The speaker typically assumes that the topos accessed by other conversational participants to interpret an argument is similar to that which the speaker himself has in mind. However, sometimes our individual takes on the conversation as well as the state of affairs in general, do not match. As previously

discussed agents involved in dialogue may accommodate different topoi which all satisfy the criteria for underpinning a particular enthymeme, while not being the ones assumed by the speaker. Although dialogue participants may have access to different topoi, and different topoi may be selected to plug into an enthymeme, it is not the case that the choice of topoi is entirely random or unpredictable. It is, however linked to situational context, domain and sociocultural context.

Before we have a look at the example of incremental reasoning in (2), let us consider (1) again, where the dialogue participants are incrementally specifying the interpretation of the noun phrase *the doctor*. In (3) we see the types of B's take on the concept at different points in the dialogue. First, the type B believes to be shared is that in (3a), where all that is specified is that the referent is a doctor. At the type of (3b) B has learnt, by means of a clarification question, that the doctor is Chorlton. At the type of (3c) B has also integrated the information that Chorlton was the one who examined A.

$$(3) \quad (a) \begin{bmatrix} x{:}Ind \\ c{:}doctor(x) \end{bmatrix} \quad (b) \begin{bmatrix} x{=}chorlton{:}Ind \\ c{:}doctor(x) \end{bmatrix} \quad (c) \begin{bmatrix} x{=}chorlton{:}Ind \\ c{:}doctor(x) \\ c_1{:}examined(x,B) \end{bmatrix}$$

A similar case regarding the incremental narrowing down of possible referents for proper names modelled in TTR can be found in Cooper (2017). We will now move on to see how we may model higher level incremental processing in a compatible framework.

$P$'s first utterance in (2a) – "metal was the reason I started doing hip hop" – constitutes an enthymeme – there is *something* about metal that made Petter start doing hip hop. We represent enthymematic inference and topoi on the DGBs of dialogue participants as dependent types, i.e. functions from situations to situation types. Thus the enthymeme conveyed in Petter's first utterance can be represented as a function from a situation where metal occurs to a type of situation of a type where $P$ starts "doing hip hop", as seen in (4). We refer to this enthymeme as $\mathcal{E}_{reason}$.

$$(4) \quad \mathcal{E} = \lambda r{:} \begin{bmatrix} T{=}music{:}Type \\ x{=}metal{:}T \\ z{=}Petter{:}Ind \\ c_1{:}relevant(T) \end{bmatrix} \cdot \begin{bmatrix} y{=}hiphop{:}r.T \\ c_2{:}do(r.z,y) \end{bmatrix}$$

$$(5) \quad \mathcal{T}_{similar} = \lambda r{:} \begin{bmatrix} T{:}Type \\ x{:}T \\ z{:}Ind \\ c_1{:}relevant(T) \\ c_2{:}like(z, x) \end{bmatrix} \cdot \begin{bmatrix} y{:}r.T \\ c_3{:}do(r.z, y) \end{bmatrix}$$

There might be several topoi accessible to $J$ which would underpin the enthymeme $\mathcal{E}$. Judging from $J$'s utterance she is surprised by $P$'s assertion that he hated metal. This example shows that $J$ begins to reason about $P$'s assertion even before the second utterance. We claim that $J$ accesses a topos to apply to the enthymematic argument as early as possible, which later must be abandoned in light of (2b). It seems reasonable to assume that $J$ expected metal being the reason for $P$ to do hiphop to be due to him liking metal. Thus, a possible topos could be one saying that if two activities are of the same type, and the speaker likes one of them, that may cause him to engage in the other activity. We see a formalisation of this topos, $\mathcal{T}_{similar}$ in (5).

Update rules for introducing rhetorical features onto the shared dialogue gameboard are presented in previous work (Breitholtz, 2014). Here we will not go into the formal details of these, but provide an informal account of the updates involved.

The basic principle for integrating a topos on the shared DGB is that the enthymeme first on the list of shared enthymemes[1] justifies the loading of a topos onto the shared gameboard if the enthymeme is a *specification* of the topos. This is defined as the antecedent part of the enthymeme being a subtype (less specific or identical to) the corresponding part of the topos. Secondly, the result of applying the enthymeme to a record $r$ must be a subtype of the result of applying the topos to the same record. In other words: if you hear an argument and you are aware of a more general principle of which this argument is a case, you will interpret the argument in accordance with that principle.

In the case of $\mathcal{E}_{reason}$ (4) the antecedent type is *not* a subtype of the antecedent type of $\mathcal{T}_{similar}$ (5) since it lacks the constraint $c_2$: like(z, x). However, since dialogue participants sometimes do accom-

---

[1]Following Ginzburg's terminology for questions under discussion Ginzburg (2012) p. 68-69, we refer to this as "max eud".

modate topoi based on underspecified enthymemes, our model should be able to integrate topoi based on less strict requirements. Thus, if a dialogue participant hears an argument for which he is not aware of any more general underpinning warrant, he might coerce an interpretation based on a topos that is a more specific version of the enthymeme in the discourse.

The type of $J$'s information state following the integration of $\mathcal{T}_{similar}$ (5) is shown in (6).

(6)
$$\left[ \text{shared: } \left[ \begin{array}{l} \text{eud=[ } \mathcal{E}_{metal\_reason} \text{ ]:list(}Enthymeme\text{)} \\ \text{topoi=[}\mathcal{T}_{similar}\text{]:list(}Topos\text{)} \end{array} \right] \right]$$

After $P$'s second utterance in (2b) – "Because I hated metal" – a new enthymeme, $\mathcal{E}'$ (7), is integrated at the top of the list of enthymemes under discussion.

(7)
$$\mathcal{E}' = \lambda r: \left[ \begin{array}{l} \text{T:}Type \\ \text{x=metal:T} \\ c_1\text{:relevant(T)} \\ \text{z=Petter:}Ind \\ c_{hate}\text{:hate(z, x)} \end{array} \right] \cdot \left[ \begin{array}{l} \text{y=hiphop:}r.\text{T} \\ c_2\text{:do(}r.\text{z,y)} \end{array} \right]$$

(8)
$$\mathcal{T}_{l\_t\_e} = \lambda r: \left[ \begin{array}{l} \text{T:}Type \\ \text{x:T} \\ \text{y:T} \\ \text{z:}Ind \\ c_1\text{:relevant(T)} \\ c_2\text{:hate(z, y)} \end{array} \right] \cdot \left[ e\text{:start\_doing(r.z, r.x)} \right]$$

In this case the topos available seems to be incompatible with the enthymeme: $\mathcal{E}'$ says that since Petter hated metal, he started doing hip hop, and the topos $\mathcal{T}_{similar}$ says that if someone likes something s/he might start doing something similar. The antecedents include concepts that we would want to model as mutually exclusive, namely *like* and *hate*. Following Cooper (prep) we say that if $T_1$ = hate(x,y) and $T_2$ = like(x,y), $T_1$ precludes $T_2$, that is, no situation can be of both types.

In conversation we normally try to interpret underspecified or implicit content by drawing on information already introduced on the dialogue gameboard. However, when $J$ realises the mismatch between the integrated topos $\mathcal{T}_{similar}$ (5) and the new enthymeme in (7), she searches her rhetorical resources (part of her long term memory) for a topos which could be drawn on to underpin $\mathcal{E}'$.

A topos that would work here would be one capturing the notion of "the lesser of two evils", $\mathcal{T}_{l\_t\_e}$ (8). The idea of this topos is that in a particular context or situation if you have two things to choose from, and one is bad, you pick the other one.

We have shown how a dialogue participant sets about interpreting an enthymematic utterance by trying to access a topos that may serve as underpinning for the enthymeme. We saw how dialogue participants can start reasoning before an argument is fully spelled out, in that there is a topos that warrants the enthymeme by being a generalised version of it. This indicates that the way we process rhetorical structure is analogous to the way we process sentential and non-sentential utterances as described in e.g. Eshghi et al. (2015) – by incrementally constraining the search space.

We can thus see how the pervasive incrementality at all levels of dialogue is also apparent in accessing the resources required for higher-level processes, i.e. reasoning.

We intend to continue this line of work, in particular in investigating to what degree underspecified enthymemes are actually used. Intuitively, in situations where dialogue participants know each other very well and/or the context allows it, they may infer topoi based on underspecified enthymemes, which turn out to be exactly the ones intended by the speaker. Furthermore, as with the possibility of repair at all levels of interaction, the possibility of asking follow up questions may make it efficient to reason based on underspecified enthymemes. We also plan to conduct a series of experiments to investigate the enthymemes people use in a dialogue, and to see what factors influence which argument people choose when more than one topos is available. This would also be a way of determining to what degree a particular topos is actually established in a particular dialogue.

# References

Aristotle (2007). *On Rhetoric, a theory of civic discourse*. Oxford University Press.

Breitholtz, E. (2014). *Enthymemes in Dialogue: A micro-rhetorical approach*. Ph. D. thesis, University of Gothenburg.

Breitholtz, E. and R. Cooper (2011). Enthymemes as rhetorical resources. In *SemDial 2011 - Proceedings of the 15th Workshop on the Semantics and Pragmatics of Dialogue (LosAngelogue)*, pp. 149–157.

Cooper, R. (2017). Proper names in interaction. In S. Kurahashi, Y. Ohta, S. Arai, K. Satoh, and D. Bekki (Eds.), *New Frontiers in Artificial Intelligence: JSAI-isAI 2016 Workshops, LENLS HAT-MASH, AI-Biz, JURISIN and SKL Kanagawa, Japan, November 14–16, 2016, Revised Selected Papers*, Number 10247 in Lecture Notes in Artificial Intelligence, pp. 5–18. Springer.

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`.

Cooper, R. and J. Ginzburg (2015). Type theory with records for natural language semantics. In S. Lappin and C. Fox (Eds.), *The Handbook of Contemporary Semantic Theory*, pp. 375–407. Wiley.

Ducrot, O. (1988). Topoï et formes topique. *Bulletin d'études de la linguistique française 22*, 1–14.

Eshghi, A., J. Hough, M. Purver, R. Kempson, and E. Gregoromichelaki (2012). Conversational interactions: Capturing dialogue dynamics. In *From Quantification to Conversation*, Volume 19 of *Tributes*, pp. 325–349. College Publications.

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*, London, UK.

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

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

Larsson, S. and D. Traum (2000). Information state and dialogue management in trindi dialogue move engine tool kit. *Natural Language Engineering 6*, 323–240.

Pickering, M. and S. Garrod (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences 27*, 169–226.

Purver, M., E. Gregoromichelaki, W. Meyer-Viol, and R. Cann (2010). Splitting the 'I's and crossing the 'You's: Context, speech acts and grammar. In *Proceedings of the 14th Workshop on the Semantics and Pragmatics of Dialogue*, Poznań, pp. 43–50.