Two accounts for recording lexical disambiguations*

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

April 16, 2015

1 Introduction

Given a discourse with one or more lexically ambiguous expressions, depending on how these expressions are disambiguated, certain inferences will go through (the ones associated with the disambiguation) and others won't. After a high-level description of an inference system for drawing such inferences, I describe two possible ways to record the result of disambiguation in discourse: a representational and a non-representational one. I will argue that the non-representational, but still computational, approach has certain desirable properties.

2 An inference system

In this section, I sketch a simple inference system. Nothing much will depend on the precise shape of this system – it is introduced primarily so I can subsequently illustrate the key ideas behind the proposed approach. The inference system consists of a set of rules of the following form:

(1)

$$\frac{S_1\ S_2\ \ldots\ S_n}{S_{n+1}}\ n \geq 1$$

Here, each S_x represents a sentence or sentence pattern. More precisely, e.g. S_x is a non-empty sequence of items, where each item is either a variable ($X, Y, Z, \ldots$) or a sequence of one or more words.

We refer to the set of these rules with the letter **B** (for Background). These rules represent material inferences as in:

(2)

$$\frac{X \text{ is bald.}}{X \text{ has no hair.}}$$

The rules also spell out the inferences that result from different disambiguations $e_1, ..., e_n$ of an expression e. For instance, here are inference rules for the two senses $drug_1$ and $drug_2$ of $drug$.

*This note is based on a talk I gave at the IWCS workshop on Interactive Meaning Construction at Queen Mary University of London on April 14, 2015. The note covers the second half of the talk. In the first half of the talk, I described a dialogical theory of meaning. Apart from the assumption that meaning is underwritten by inferences, the proposal here is however independent of that theory which, in any event, is also described in Piwek (2014). I am indebted to Nicholas Asher, Robin Cooper, Ruth Kempson and Shalom Lappin for helpful feedback and stimulating comments at the workshop.

(3)

$$\frac{X \text{ was killed by a drug}_1.}{X \text{ was killed by a medicine.}}$$

(4)

$$\frac{X \text{ was killed by a drug}_2.}{X \text{ was killed by an illicit substance.}}$$

Before proceeding, let us explicitly define the two key notions that the remainder of this note depends on:

Definition 1. *(Background) The background* **B** *is a collection of rules for inferring a sentence from one or more other sentences.*[1]

Definition 2. *(Background expressions) The set of background expressions expr(***B***) is equal to* $\{i : i$ is an item that occurs in **B** and i is not a variable$\}$

3 Recording the results of word sense disambiguation

There are two aspects to lexical discourse disambiguation. Firstly, there is the Word Sense Disambiguation (WSD) problem of determining, given an expression e that is ambiguous between the members of $\{e_1 \ldots e_n\}$ what the intended or most plausible sense(s) is/are (given its occurrence at some position in a discourse d). Secondly, given a partial or complete disambiguation, there is the question of how an interpreter of a discourse keeps a record of these disambiguations, such that they can draw the inferences on the basis of the discourse that are consistent with the disambiguations.

Our focus is primarily on the second problem. In particular, we will have nothing to say about the WSD problem that occurs when an interpreter encounters an expression for the first time in a discourse. We will, however, see that the way a disambiguation is stored by the interpreter can interact with how subsequent uses of the expression in the same discourse are interpreted.

In what follows, we assume a discourse $d = \ldots e \ldots e \ldots$ with one or more occurrences of the lexically ambiguous expression e. Suppose e is ambiguous between $e_1, \ldots e_n$. The background **B**, as defined earlier on, specifies appropriate inferences for each of these disambiguations. For this to be possible, each of $e_1, \ldots e_n$ has to occur in **B**. So, $\{e_1, \ldots e_n\} \subseteq expr(\mathbf{B})$.

3.1 Disambiguation with representation

Let us first consider how to explicitly record disambiguations in the discourse representation that an interpreter constructs as a result of processing the discourse. A disambiguated representation of d is obtained by replacing each of the occurrences of e in d with an e_x (where $1 \leq x \leq n$). We may also introduce underspecified representations, e.g. $e_1|e_5$ to express that both e_1 and e_5 are plausible senses of e according to the context so far (but we don't yet know which of the two is the final interpretation). In this case, occurrences of e would be replaced with these underspecified representations.

3.2 Disambiguation without representation

Let us now consider an alternative, to our knowledge novel, non-representational approach. In this approach, the interpreter stores the discourse verbatim as a sequence of sentences. Lexically ambiguous expressions are left unaltered. Of course, it may be that we do need to do some processing of the sentences, arriving at computationally more convenient representation (e.g. a Discourse Representation Structure). What matters here is, however, that any occurrences of lexically ambiguous expression can

[1]Note that rules which involve sentence patterns also yield inferences with sentences as their premises and conclusions. These sentences will be instances of the sentence patterns in the rule.

be used in the discourse representation without any alteration (i.e. explicit representation of different disambiguations).

So, how can we ensure that the inferences that are drawn from d are consistent with the interpreter's (partial) disambiguation? Since this information is not recoverable from (the representation of) d itself, we propose to record it through via **B**, i.e. the rules of the inference system. This means that we have to take a somewhat unconventional view of **B**: it is no longer a static resource – instead as a result of discourse interpretation it is updated and this affects which inferences the interpreter can draw on the basis of the discourse. In the remainder of this section, I describe the updates in question.

3.2.1 Unmarked expressions

Let's assume that the discourse $d = \ldots e \ldots e \ldots$ is interpreted incrementally from left to right and that none of the e is marked (e.g., by means of pitch accent). As the interpreter receives more information, they adjust their interpretation of e. Let's assume that at the outset, time t_0, e has n different possible senses: $e_1, \ldots, e_n$. We write $\sigma(t_0, e) = \{e_1, \ldots, e_n\}$. Generally, we refer to the senses that are still available at time t_k with $\sigma(t_k, e)$, where $\sigma(t_k, e) \subseteq \sigma(t_0, e)$. In words, as the interpreter receives more information, the number of available senses may be reduced.[2]

In terms of **B**, at t_0 the set $e_1, \ldots, e_n \subseteq expr(\mathbf{B})$. If contextual information rules out one of these sense e_x, this is modelled by replacing e_x everywhere it occurs in **B** with $inh(e_x)$. Here, 'inh' stands for inhibited.

We remain agnostic about the specific method for determining which e_x to rule out based on the context. A suitable machine learning algorithm may be used for this purpose, possibly in combination with symbolic methods (e.g. by considering whether the contextual information gives rise to inferences that are inconsistent with a particular sense).

Definition 3. *(Non-representational disambiguation storage) If, as a result of processing discourse d, sense e_x of e is ruled out, this is stored by updating* **B** *with the following substitution:* $\mathbf{B}[inh(e_x)/e_x]$.

We stipulate that the expression $inh(e_x)$ is opaque. This means that as a result of the update of **B**, e_x is no longer a member of $expr(\mathbf{B})$, whereas $inh(e_x)$ now is a member of $expr(\mathbf{B})$.

Full disambiguation of e is achieved, when all $e_x \in expr(\mathbf{B})$ have the same index x. In that case, we can suppress the index altogether:

Definition 4. *(Full non-representational disambiguation storage)* $\forall\ e_x\ e_y \in expr(\mathbf{B}) : \ (x = y \Rightarrow \mathbf{B}[e/e_x])$.

Now suppose that the discourse contains the sentence $a\ e\ b$ and $\mathbf{B}$ contains the inference rule:

(5)
$$\frac{X\ e_1\ Y}{c}$$

This rule tells us that we can infer c if our input matches with $X\ e_1\ Y$. Now, at this point, we can *not* draw any inferences from our discourse using this rule (since e in the discourse doesn't match with e_1). But now suppose that e has been fully disambiguated. That is, for all e_x in $\mathbf{B}$: $e_x = e_1$. In that case, Definition 4 tells us that we can replace e_1 with e in **B**, so we get:

(6)
$$\frac{X\ e\ Y}{c}$$

Now, based on discourse d, we can infer c, since $X\ e\ Y$ matches with $a\ e\ b$ in the discourse.

[2]We assume that this is the normal way a discourse is processed. An interpreter may, however, on occassion discover that they've committed to the wrong disambiguation and need to go back and revise previous disambiguation. Here, we won't consider such non-monotonic adjustments.

3.2.2 Marked expressions

We assume that an expression which is marked (e.g. by pitch accent), receives a 'marked' disambiguation (which will deviate from the interpretations of unmarked tokens of the expression). Consider this example from van Deemter (1996), where one occurrence of pitcher concerns a baseball player and the other a jug.

(7) A pitcher was drinking wine from a pitcher.

We assume that the second occurrence is marked, and thus represented as:

(8) A pitcher was drinking wine from a marked(pitcher).

Generally, we will represent a marked expression e with $marked(e)$.

Definition 5. *(Non-representational disambiguation storage for marked expressions) If, as a result of processing discourse d, sense e_x of $marked(e)$ is ruled out, this is stored by updating* **B** *with the following substitution:* $\mathbf{B}[inh(inh(e_x))/inh(e_x)]$.

Note that only the senses of e that have already been inhibited are candidate senses for $marked(e)$.[3] Again, we assume that X in expression $inh(X)$ is opaque: we can't match on subsexpressions of X.

Definition 6. *(Full non-representational disambiguation storage)* $\forall\, inh(e_x)\, inh(e_y) \in expr(\mathbf{B}) : (x = y \Rightarrow \mathbf{B}[marked(e)/inh(e_x)])$.

3.2.3 An example

Let us look at an example. Consider a discourse $d = e \ldots e \ldots marked(e) \ldots e$. Let's assume that $\sigma(t_0, e) = \{e_1, e_2, e_3\}$. That is, e has three senses: e_1, e_2 and e_3. We also need some way to refer to the different points in time as processing of the discourse proceeds: t_0 represents the point in time just before the discourse is processed. t_1, t_2, t_3 and t_4 represent the point in time just after the first, second, third and fourth occurence of e (actually $marked(e)$ in the case of t_3) has been processed.

We assume a number of disambiguations which have been provided by word sense disambiguation oracle. What we're focussing on here is how this disambiguation then gets recorded by an update to **B**. The update can summarised in terms of expr(**B**). In what follows we represent expr(**B**) only for the senses we're interested in, i.e. e_1, e_2 and e_3.

TIME	DISAMBIGUATION	$expr(\mathbf{B})$	full disambiguation $expr(\mathbf{B})$
t_0	n/a	$\{e_1, e_2, e_3\}$	
t_1	e_3 ruled out	$\{e_1, e_2, inh(e_3)\}$	
t_2	e_2 ruled out	$\{e_1, inh(e_2), inh(e_3)\}$	$\{e, inh(e_2), inh(e_3)\}$
t_3	e_2 ruled out	$\{e_1, inh(inh(e_2)), inh(e_3)\}$	$\{e, inh(inh(e_2)), marked(e)\}$
t_4	n/a	$\{e_1, inh(inh(e_2)), inh(e_3)\}$	$\{e, inh(inh(e_2)), marked(e)\}$

Initially, at t_0 all three senses are available. We can't draw any inferences about e, since **B** has no occurrences of e.[4] At t_1, e_3 is ruled out. Note that at this point we still *cannot* reason about e directly in **B**. One can think of **B** as underspecifying the interpretation of e (though it is more specific than it was at t_0). Finally, at t_2 we arrive at a single interpretation for e. e_1 is the only sense of e that hasn't been inhibited. This licenses us to replace e_1 with e. Now, we can reason about e in **B**. At t_3, we've reached marked(e). Its possible senses are $inh(e_2)$ and $inh(e_3)$. $inh(e_2)$ is ruled out by the context (by the oracle). This means there is only one sense left, $inh(e_3)$ and it is replaced with $marked(e)$ in **B**. So we can now reason about marked(e) with **B**. Finally, the last occurrence of e (at t_4) is interpreted relative to **B**. As a result of previous updates, **B** has occurrences of e in some of its inference rules (since $e \in expr(\mathbf{B})$), which means that these rules may get triggered.

[3] For an application of a similar idea to the interpretation of anaphora with pitch accent, see for example Piwek (1997).

[4] Though an interpreter could of course hypothetically replace any of e_x with e to find out what the consequences of that particular disambiguation would be (in terms of the inferences that become available).

3.2.4 Desirable properties

The non-representational storage of disambiguations as proposed here has a couple of desirable properties.

Firstly, according to Gale et al. '. . . senses tend to appear in clumps. In particular, it appeared to be extremely unusual to find two or more senses of a polysemous word in the same discourse.' (Gale et al., 1992, p. 235). This one sense per discourse finding falls out naturally. In our incremental account of discourse processing, once an expression e has been fully disambiguated, subsequent uses of the expression will automatically yield the inferences associated with e (since e has been activated in $\mathbf{B}$ and other senses have been inhibited).

Secondly, the account provides a natural way to deal with underspecification. As long as there are several active senses of e, the inference rules in $\mathbf{B}$ represent the consequences of adopting any one of them.

Finally, global effects of disambiguation fall out naturally. Once an expression e has been disambiguated any inferences that are licensed by previous uses of e (*within the same discourse*) will automatically be available. The interpreter doesn't need to go back to the representations they constructed for these occurrences of the lexical item e (since they're all represented by e, rather than say $e_1|e_2|e_3$ at the beginning of the discourse, $e_1|e_2$ after e_3 has been ruled out, and e_1 after e_2 has been ruled out).

4 Concluding remarks

In our non-representational proposal for recording lexical disambiguations, rather than map the input discourse to a representation that matches with the appropriate inferential patterns, the patterns or rules themselves are adjusted (during discourse interpretation) such that some inferences are inhibited, whereas others (that are consistent with the disambiguation) are enabled. In other words, as a result of word sense disambiguation, inferences that go with the disambiguation are 'activated'.

Speaking more generally, the current proposal adheres to a picture of cognitive processing where new information gives rise to adjustments of the interpretation system itself (rather than some representation of the content that has been conveyed) and this makes certain inferences available that weren't available previously. A view of discourse interpretation emerges where interpretation manifest itself in a continual change of the frame of mind of the interpreter. Interestingly, this change of their frame of mind has global effects: it not only affect how the interpreter thinks about what they are currently interpreting but also the discourse that they encountered previously (e.g., once an interpreter has found out that the speaker was talking about a medicine when they used the expression 'drug', previous uses of the word 'drug' in the same discourse will also be viewed in that way, and results in certain inferences becoming available based on those previous uses of word).

References

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.

Piwek, P. (1997). Accent interpretation, anaphora resolution and implicature derivation. In *The proceedings of the 11th Amsterdam Colloquium*, pp. 55–60. University of Amsterdam, ILLC/Department of Philosophy.

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.