Chapter 16

Dynamic Semantics and Ambiguity

CHRISTOF MONZ

ABSTRACT. Ambiguity is one of the most difficult problems in natural language processing. We present a dynamic semantics which gives a formal account of the behavior of ambiguous expressions occurring in a sequence of sentences. Ambiguous sentences are represented in an underspecified way and the update potential of underspecified representations is defined. Updating information states with underspecified semantic representations is interleaved with partial disambiguation, helping to avoid the problem of state explosion. In general, it is considered how ambiguity survives throughout a discourse and how contextual information can restrict ambiguity.

1 Introduction

Natural language expressions can be highly ambiguous, and this ambiguity may have various faces. Well-known phenomena include lexical and syntactic ambiguities. In this paper we focus on constructing a discourse semantic for quantificational ambiguities and ambiguities evolving in pronoun resolution, as exemplified in (16.1.a) and (16.1.b), respectively.

(16.1) (a) Every man loves a woman.
        (b) A man saw a boy in the park. He whistled.

The different readings of (16.1.a) correspond to the two logical representations in (16.2) and the pronoun in (16.1.b) can be resolved in two ways as illustrated in (16.3)

(16.2) (a) $\forall x (man(x) \rightarrow \exists y (woman(y) \land love(x, y)))$.
        (b) $\exists y (woman(y) \land \forall x (man(x) \rightarrow love(x, y)))$

(16.3) (a) A man, saw a boy in the park. He, whistled.
(b) A man saw a boy in the park. He whistled.

We refer the reader to [vDP96] for extensive discussions of these and other examples of quantificational ambiguity. Often ambiguous sentences have a preferred reading, cf. [KM93, JK96] for more details, but adding preference selection would result in a nonmonotonic framework. For instance, assume that a wide scope reading for the universal quantifier is preferred in (16.1), then (16.2.a) would be an appropriate semantic representation, but if (16.1) is followed by

(16.4) But she is already married.

(16.2.b) would be more appropriate. If we want to process a discourse in a monotone fashion, we must initially allow for both possibilities.

The problem of ambiguity and underspecification has recently enjoyed a considerable increase in attention from computational linguists, computer scientists, and logicians (see, for instance, [vDP96]). However, the focus has mostly been on giving a semantic representation for single sentences. Although some of the theories are couched in a dynamic setting (cf. for example [Bos95, Rey93]), most of them do not give a formal account of the behavior of ambiguous expressions in a discourse, but see [Fer97, Mon98]. Roughly, two things are lacking: First, there is no detailed account of pronoun resolution. In [Bos95] coindexation is assumed and in URT ([Rey93, Rey95]) only one way of binding is considered, but not different ways of binding in parallel. Second, there is no notion of an ambiguous context and ambiguous updating. In other words, the update potential of ambiguous sentences is not given.

The aim of this paper is to give a dynamic semantics for a language that includes ambiguous expressions, where ambiguity is represented in a non-redundant, underspecified way, as it is done in [Rey93, Bos95].

The rest of this paper is structured as follows. Section 2 discusses representational issues of ambiguity and the notion of underspecification is introduced. In Section 3 a dynamic semantics for a language including underspecified representations is given and some of its properties are discussed. Finally, conclusions and prospects for further work are provided in Section 4.

2 Representing Ambiguity

It is possible to represent quantificationally ambiguous formulas in a compact way as upper semi-lattices, cf. [Rey93]:

213
1 Definition (Underspecified Representation) An underspecified representation (UR) is a quadruple $\langle LHF, L, H, C \rangle$, where $LHF$ is the set of labeled $h$-formulas, and $C$ is a partial order on the labels ($L$) and holes ($H$) occurring in $LHF$. A hole indicates that the concrete form of this subformula is not yet known. For instance, we do not know what the scope of the universal quantifier is, so the succedent of the implication is a hole. $h$-formulas are the formulas possibly containing holes. In (16.5) $LHF = \{l_0 : h_0, l_1 : \forall x (man(x) \rightarrow h_1), l_2 : \exists y (woman(y) \land h_2), l_3 : love(x, y)\}$ and $C = \{l_1 \leq h_0, l_2 \leq h_0, l_3 \leq h_1, l_3 \leq h_2\}$.

There are two possible sets of instantiations, $\iota_1$ and $\iota_2$, of the holes $h_0, h_1, h_2, h_3$ in (2) which obey the constraints:

- $\iota_1 = \{h_0 := l_1, h_1 := l_2, h_2 := l_3\}$
- $\iota_2 = \{h_0 := l_2, h_2 := l_1, h_1 := l_3\}$

For each instantiation of the holes there is a corresponding substitution $\sigma(\iota)$ which is like $\iota$ but $h := \varphi \in \sigma(\iota)$ if there is a $l$, such that $l : \varphi \in LHF$ and $h := l \in \iota$.

An obvious question at this point is, how does one associate a UR with a given natural language expression? We will not address this issue here, but we will assume that there exists some mechanism for arriving at UR's, see for example [Kön94].

The next step is to define an extension of the language of first-order logic, $L$, in which both standard (unambiguous) expressions occur side by side with the above underspecified representations. The resulting language of the language of underspecified logic, or $L''$ for short, is the language in which we will use for constructing semantic representations.

2 Definition (Underspecified Logic) A formula $\varphi$ is a formula of our underspecified logic $L''$, or a $u$-formula, that is, a formula possibly containing underspecified representations, if it is built up from underspecified representations and the usual atomic formulas from standard first-order logic using the familiar boolean connectives and quantifiers.

3 Example As an example of a more complex $u$-formula consider the semantic representation of *if every boy didn't sleep and John is a boy, then John didn't sleep.*
4 Definition (Total Disambiguations) Then, by $\delta(\varphi)$ we denote the set of total disambiguations of the u-formula $\varphi$, where for all $d \in \delta(\varphi)$, $d \in L$. For complex u-formulas $\delta$ is defined recursively:

1. $\delta(\langle LHF, L, H, C \rangle) = \text{the set of } LHF \sigma(\iota) \text{ such that}$
   
   (i) $\iota$ is an instantiation and $\sigma(\iota)$ is the corresponding substitution
   
   (ii) $H_{\iota} = L$
   
   (iii) for all $l, l' \in L$, $l \leq l' \in \text{closure}(C_{\iota})$ or $l' \leq l \in \text{closure}(C_{\iota})$

2. $\delta(\neg \varphi) = \{ \neg d \mid d \in \delta(\varphi) \}$

3. $\delta(\varphi \circ \psi) = \{ d \circ d' \mid d \in \delta(\varphi), d' \in \delta(\psi) \}$, where $\circ \in \{ \land, \lor, \rightarrow \}$

4. $\delta(\exists x \varphi) = \{ \exists x d \mid d \in \delta(\varphi) \}$, where $\exists \in \{ \lor, \exists \}$.

To see whether an underspecified representation holds in a model or not we check whether this holds for all its total disambiguations.

5 Definition (Validity and Falsity) Given an arbitrary model $M$ and a u-formula $\varphi$

\[ M \models \varphi \text{ iff } M \models d \text{ for all } d \in \delta(\varphi) \]

\[ M \not\models \neg \varphi \text{ iff } M \not\models d \text{ for all } d \in \delta(\varphi) \]

3 A Dynamic Semantics for Ambiguity

To combine underspecified representation and update semantics we take Dekker's EDPL [Dek93] as a basis and adapt it in a way such that it fits our purposes. We chose EDPL because it is both, eliminative and distributive and the notion of an information state is explicit.

The adaptation proceeds mainly along two lines. First, contrary to approaches in the tradition of dynamic predicate logic (DPL), see [GS91], we do not take resolution of pronouns as granted; i.e., coindexation of pronouns and their antecedents is not presupposed. We dispense with this — admittedly quite handy — restriction on our input, because it is unrealistic and anaphora resolution is an important source of ambiguity in natural language texts,
which we would like to account for. If we do not assume coindexation, it is necessary to deal with pronouns in a different way. We do so by binding pronominal arguments by three new operators he, she and it, similar to [Bea93], that roughly say that its arguments have to be identified with an accessible antecedent in a given context. Second, whereas in EDPL contexts are identified with states, we have to think about the notion of an ambiguous context. In this paper we define an ambiguous context, σ, simply by a set of states, where each state in a set of states represents an unambiguous reading of the context. Of course the notion of an update has to be redefined appropriately. A context is said to be unambiguous if it contains exactly one state.

3.1 EDPL

Before adapting EDPL, we briefly repeats its semantic definitions.

6 Definition (Semantics of EDPL)

\[
\begin{align*}
s[[R(x_1 \ldots x_n)]] & = \{i \in s \mid (i(x_1), \ldots, i(x_n)) \in F(R)\} \text{ if } x_1 \ldots x_n \in D(s) \\
s[x = y] & = \{i \in s \mid i(x) = i(y)\} \text{ if } x, y \in D(s) \\
s[\neg \varphi] & = s - s[[\varphi]] \\
s[\exists x \varphi] & = s[x][[\varphi]] \text{ if } x \notin D(s) \\
s[[\varphi \land \psi]] & = s[[\varphi]][[\psi]]
\end{align*}
\]

where \( s \rightarrow s' = s\{i \upharpoonright D(i) \mid i \in s'\} \text{ and } s[x] = \{j \mid \exists i \in s : i \leq \{x\} j\} \), such that \( i \leq x j \) iff

\( D(j) = D(i) \cup X \) and \( i = j \upharpoonright D(i) \)

The other connectives, \( \forall, \lor, \rightarrow \), can be defined in terms of the definitions given above.

3.2 Pronoun Resolution

The first difference between EDPL and our dynamic predicate logic with ambiguity (DPLA), concerns the treatment of pronouns. In EDPL it is assumed that pronouns are resolved and therefore pronouns and there antecedents use the same variable. In our approach pronouns are bound by a pronominal operator P, \( P \in \{\text{he, she, it}\} \), so \( \text{she } x \varphi \) means that the argument \( x \) in \( \varphi \) is pronominal and its antecedent must be female.\footnote{This is of course a simplified picture of the whole issue of grammatical and natural gender and the way it restricts pronoun resolution.} Because pronom-
inal resolution might be ambiguous, updating a context $\sigma$ with $[P x \varphi]$ does not have to return a single update state for each $s \in \sigma$, but a set of them.

7 Definition

$$s[[P x \varphi]] = \{s[[\varphi[x/y]]] \mid y \in D(s), s[[\varphi[x/y]]] \neq \emptyset\}, \text{ where } P \in \{\text{he, she, it}\}$$

The restriction $y \in D(s)$ ensures that only accessible antecedents are considered.

To make the return value of an update to be consumable by the next update function we have to generalize our notion of an update: an update function is a partial function from sets of states to sets of states. We will call sets of states also ambiguous contexts.

8 Definition (Pronoun Resolution) Let $\sigma$ be an ambiguous state.

$$\sigma[[P x \varphi]] = \bigcup_{s \in \sigma} s[[P x \varphi]], \text{ where } P \in \{\text{he, she, it}\}$$

It is obvious that this way of resolving pronouns is very inefficient, but we want to give a monotone semantics for ambiguity, and of course it is possible to augment our theory with some heuristics that allow to restrict the set of readings or ways of resolving pronouns, see for instance [JK96].

3.3 Quantificational Ambiguities

Next we consider quantificational ambiguities, for which a compact representation was given earlier. The semantics of underspecified representations is defined in terms of total disambiguations see [Rey95], so avoiding redundancy on the the syntactic side is possible, but it returns on the semantic side. Often this is not a problem, because one can consider the syntactic side, only. For instance given a complete and sound calculus, reasoning in this system amounts to syntactic manipulation. But in an ambiguous setting things are a bit different: underspecified representations are compact representations of several formulas, i.e. they are syntactically ambiguous, and at the moment it is not obvious how to reason with ambiguous formulas without disambiguating them at least partially, cf. [Jas97, MdR98].

The question is how can we restrict the massive branching of contexts when updating with ambiguous information. If we do not update with total disambiguations but instead interleave updating and disambiguation we might gain a decrease in complexity because further disambiguation steps depend on earlier successful updates. Updates of complex formulas are computed
in a top-down manner. The first step is to identify the possible outermost operators in a quantificationally ambiguous expression and start computing the updates until we face a hole. Then we make a choice again and check what are the operators which may have the next wider scope respectively, continue updating, and so on. Possibly some of these first steps exclude some disambiguations because they are not compatible with the given input context.

Partially updating with quantificationally ambiguous material changes also the set of ordering constraints. This makes it necessary to make contexts more complex. We will say that an ambiguous context $\sigma$ is a set of triples of the form $\langle s, C, LHF \rangle$. Updating $\sigma$ proceeds in a top-down manner:

9 Definition Let $UR = \langle LHF, L, C \rangle$ be an underspecified representation. $top(UR)$ returns the formula of the top node and $dt(h)$ are the possible immediate subformulas of a hole $h$.

$$\sigma[UR] = \{ \langle s, C, LHF \rangle | \exists C', LHF' : \langle s, C', LHF' \rangle \in \sigma \}$$

$$\sigma[h] = \bigcup_{l_j \varphi_j \in dt(h)} \{ \langle s, C \cup \{ l_k \leq h \}, LHF \rangle | \langle s, C, LHF \rangle \in \sigma, l_k : \psi \in dt(h) \}$$

$$\sigma[h] = \sigma[\varphi_j], \text{ if } dt(h) = \{ l_j : \varphi_j \} \text{ and } \varphi_j \text{ does not contain a hole}$$

Having defined the update function of underspecified representations and expressions containing pronouns, we adapt the update conditions of the kind of expressions already considered in EDPL. First we have to adapt the update definition of unambiguous contexts from simple states to triples of the form $\langle s, C, LHF \rangle$.

10 Definition if $\varphi$ does not contain holes or underspecified representations:

$$\langle s, C, LHF \rangle[\varphi] = \begin{cases} 
\langle s[\varphi], C, LHF \rangle, & \text{if } s[\varphi] \neq \emptyset \text{ and defined and } C \not\vdash \bot \\
\text{undefined}, & \text{otherwise}
\end{cases}$$

11 Definition (Semantics of DPLA) The update potentials of the logical connectives are those given in Definition 3 plus the following ones, where $s^e$ is of the form $\langle s, C, LHF \rangle$.

$$\sigma[R(x_1 \ldots x_n)] = \{ s^e[R(x_1 \ldots x_n)] | s^e \in \sigma, s^e[R(x_1 \ldots x_n)] \text{ is defined} \}$$

$$\sigma[x = y] = \{ s^e[x = y] | s^e \in \sigma, s^e[x = y] \text{ is defined} \}$$

$$\sigma[\neg \varphi] = \{ s^e - t^e | s^e \in \sigma, t^e \in \{ s^e[\varphi] \} \}$$

$$\sigma[\exists x \varphi] = \{ s^e[\exists x \varphi] | s^e \in \sigma, s^e[\exists x \varphi] \text{ is defined} \}$$

$$\sigma[\varphi \land \psi] = \sigma[\varphi][\psi]$$

$$\sigma[P x \varphi] = \bigcup_{s^e \in \sigma} s^e[P x \varphi], \text{ if defined}$$

$$\sigma[UR] = \{ \langle s, C, LHF \rangle | \exists C', LHF' : \langle s, C', LHF' \rangle \in \sigma \}$$
\[ \sigma[h] = \bigcup_{l_j \in \text{dt}(t)} \{ (s, C \cup \{l_k \leq h\}, LHF) \mid (s, C, LHF) \in \sigma, l_k : \psi \in \text{dt}(h)\} [\phi_j[h']] \]

\[ \sigma[h] = \sigma[\phi_j], \text{if dt}(h) = \{l_j : \phi_j\} \text{ and } \phi_j \text{ does not contain a hole} \]

where \((s, C, LHF) - (t, C, LHF) = (s - t, C, LHF)\)

The next example illustrates the advantage of partial updating over regular updating with total disambiguations.

**12 Example** Consider a situation \(\sigma\), where nothing about a male person was said and somebody utters *every boy didn't see a movie he knows*. First of all we have to give an underspecified representation \(UR\) of the sentence:

\[
\begin{array}{ccc}
  l_0 : h_0 & l_1 : \exists y(m(y) \land \text{hcz}(k(z,y) \land h_1)) & l_2 : \forall x(b(x) \rightarrow h_2) \\
  & l_3 : \neg h_3 & l_4 : \text{see}(x,y)
\end{array}
\]

Updating \(\sigma\) with \(UR\) means updating \(\sigma\) with the top node of \(UR\):

1. \(\sigma[UR] = \sigma[h_0]^C\)

Then we have to consider the three operators which can have widest scope:

2. \(\{ (s, C \cup \{l_2, l_3 \leq h_1\}, LHF) \mid (s, C, LHF) \in \sigma \} [\exists y(m(y) \land \text{hcz}(k(z,y)) \land h_2)] \)
   \(\bigcup\ \{ (s, C \cup \{l_1, l_3 \leq h_2\}, LHF) \mid (s, C, LHF) \in \sigma \} [\forall x(b(x) \rightarrow h_1)] \)
   \(\bigcup\ \{ (s, C \cup \{l_1, l_2 \leq h_3\}, LHF) \mid (s, C, LHF) \in \sigma \} [\neg h_3] \)

If we focus on updating the first one, we see that the pronoun cannot be resolved properly.

3. \(\{ (s[y], C \cup \{l_2, l_3 \leq h_1\}, LHF) \mid (s, C, LHF) \in \sigma \}[m(y) \land \text{hcz}(k(z,y)) \land h_2] \)
   \(= \{ (j \mid \exists i \in s : i \leq [y] j), C \cup \{l_2, l_3 \leq h_1\}, LHF) \mid (s, C, LHF) \in \sigma \} = \sigma_1 \)
   \(= \sigma_1[m(y) \land \text{hcz}(k(z,y)) \land h_2] = \sigma_1[m(y)] [\text{hcz}(k(z,y)) \land h_2] \)
   \(= \{ (j \mid \exists i \in s : i \leq [y] j, j(y) \in F(m)), C_1, LHF) \mid (s, C_1, LHF) \in \sigma_1 \} = \sigma_2 \)
   \(= \sigma_2[\text{hcz}(k(z,y)) \land h_2] \)

Because \(F(m) \cap \text{he} = \emptyset\) the next update results in

4. \(= \{ (\emptyset, C_1, LHF) \mid (s, C_1, LHF) \in \sigma_2 \} \)

Therefore the update of the whole formula reduces to

219
5. \( \{\langle s, C \cup \{l_1, l_3 \leq h_2\}, LHF\} \mid \langle s, C, LHF \rangle \in \sigma \} \forall x(b(x) \rightarrow h_1) \)
   \( \cup \{\langle s, C \cup \{l_1, l_2 \leq h_3\}, LHF\} \mid \langle s, C, LHF \rangle \in \sigma \} \neg h_3 \)

Note, that we are able to rule out the wide scope reading of the existential quantifier by leaving the scoping relation of the universal quantifier and the negation still unspecified.

### 3.4 Some Properties of DPLA

EDPL is distributive and eliminative; but what about DPLA? It is easy to see that distributivity holds, because in none of our definitions the update depends on the whole context. Combining this with the fact that EDPL is distributive we get:

**13 Observation** \( \sigma[\varphi] = \bigcup_{\varphi' \in \sigma} \sigma[\varphi'] \), if defined

Although updates of ambiguous formulas possibly contain more readings than the original context, each of these reading eliminates some (possibly null) assignment functions.

**14 Observation** DPLA is eliminative.

For the fragment which does not involve ambiguities the properties of EDPL generally hold for DPLA, too.

**15 Definition (Entailment)** Following [Dek93], where \( s \preceq s' \) iff \( \forall i \in s \exists j \in s' : i \subseteq j \) we say that \( \sigma \preceq \sigma' \) iff \( \forall \langle s, C, LHF \rangle \forall \langle s', C', LHF' \rangle : s \preceq s' \).

The entailment relation can now be defined as:

\[ \varphi_1 \ldots \varphi_n \models \psi \text{ iff } \sigma[\varphi_1] \ldots [\varphi_n] \preceq \sigma[\psi] \]

This is just one of many possible definitions of the entailment relation. Two things should be pointed out: first, the dynamic force of the existential quantifier does not exceed the entailment relation. So anaphora occurring in the conclusion cannot be bound by an existential quantifier in the premises, cf. [Ben96, Gro95] for a comparison of dynamic entailment relations. Second, our notion of entailment is very weak, because it presupposes that all readings in the premises entail all readings in the conclusion, but for dealing with natural language semantics it seems to be an appropriate definition, see [vdD96, MdR98, Rey95] for further discussion and refinements.
4 Conclusion

Our formalism is monotone and it inherits the properties of eliminativivity and distributivity from EDPL. It is more general than DPL based approaches, because it gives an account for pronoun resolution without relying on coindexation and it does so by considering different possibilities in parallel. Contrary to [Bos95, Rey93] the update potential of ambiguous sentences is defined and it is possible to compute the update potential of a sequence of ambiguous sentences.

DPLA allows us to give a general picture of ambiguous updating, which can be extended in several directions by adding some heuristics to yield a more efficient formalism. We would like to point out, that we did not try to model how human speakers deal with ambiguous information, but we did try to give an account that is general enough to be specified in a way to do so.

It might be interesting to see how our framework interacts with approaches on deduction in an ambiguous setting, cf. [Jas97, MdR98]. Jaspar noticed that the notion of relative ambiguity is necessary in order to get a well behaved calculus for ambiguity. This notion is part of our framework and of dynamic semantics in general, because updates depend on the context to which they are applied, and so do ambiguous updates. The next step will be to see how a calculus for reasoning in ambiguous dynamic semantics can be developed.

Acknowledgments

The research in this paper was supported by the Spinoza project ‘Logic in Action’ at the University of Amsterdam.

References


