Probabilistic Inferences in Dynamic Semantics

One of the major strengths of Dynamic Semantics is its ability to model both (i) the constraints governing the introduction, persistence, and accessibility of discourse referents; and (ii) the role played by contextual factors such as beliefs and the common ground in steering the interpretation of utterances. Both (i) and (ii) are crucially involved in the resolution of pronominal anaphoric reference: The interpreter's choice relies on a range of information sources, both linguistic (e.g., embedding structures; grammatical gender) and extralinguistic (e.g., knowledge about what is likely to be true). Many of the inferences involved, especially of the latter kind, are uncertain and heavily dependent on the content of the utterance in question. To model these inferences and their interaction with more genuinely linguistic constraints, one would want an integrated framework in which both uncertain beliefs about the world and the linguistic constraints traditionally modeled by Dynamic Semantics can be represented together.

In this paper we introduce such a framework, called *Probabilistic Dynamic Predicate Logic* (PDPL). Its ultimate purpose is to seamlessly integrate what is conventionally split into "world knowledge" on the one hand, and "linguistic knowledge," on the other. The guiding philosophical assumption is that this division is a theoretical artifact which does not correspond to any clear cognitive division that could be drawn on independent grounds. Linguistic knowledge *is* (part of) world knowledge, and both go hand in hand in the reasoning involved in communication.

With this vision in mind, in this paper we focus on a relatively small but illuminating application of PDPL: The resolution of pronominal anaphora in *bridging inferences*. Bridging is a way to augment the information literally conveyed by what was said, typically by fleshing out a "scenario" relative to which the utterance is less ambiguous and more coherent that it would be on its own. To the extent that the hearer can be counted on to draw bridging inferences in the utterance situation, they may become part of what is communicated. Often bridging involves establishing coreferential links for anaphoric (in)definites, as shown in (1) (Clark & Haviland 1977; Hawkins 1978; Asher & Lascarides 2000; Levinson 2000; Matsui 2001):

(1) A bridge collapsed. The strut was rotten.
   a. A bridge collapsed because a strut of the bridge was rotten.
   b. A bridge collapsed because a strut *[of some unmentioned antecedent]* was rotten.

(1a) is likely to be inferred to connect the sentences in (1). Importantly, though, the anaphoric link identifying the strut as part of the bridge is not required for the truth of (1), and the weaker inference in (1b) may be sufficient, even preferred, in certain contexts.

We argue that standard accounts of bridging fail to provide a satisfactory formal account of the way in which cross-sentential anaphora are resolved in bridging, particularly the role of world knowledge in adjudicating between multiple alternative resolutions of anaphoric coreference. PDPL fills this gap by modeling the information states of interlocuters in a given discourse and keeping track of the probabilities they assign to alternative resolutions.

For a further illustration, consider (2) from Clark and Haviland (1977) and imagine some speaker $\mathcal{A}$ utters (2a) to a hearer $\mathcal{B}$:

(2)
a. John unpacked the picnic. The beer was warm.
b. The beer was part of the picnic.
c. $\exists! x[\mathsf{unpack}(j, x)] \wedge \exists! y[\mathsf{beer}(y) \wedge \mathsf{warm}(y)]$
d. $\exists! x[\mathsf{unpack}(j, x)] \wedge \exists! y[\mathsf{beer\text{-}of}(y, z) \wedge \mathsf{warm}(y)] \wedge x = z$

We would not expect $\mathcal{A}$ to explicitly state (2b) in addition to (2a). Speakers can often assume that during a given discourse, hearers will extrapolate from underspecified utterances. In (2), the inference consists in finding an appropriate anaphoric antecedent for *the beer*. Standard pragmatic accounts of bridging inferences (Hawkins 1978; Levinson 2000) assume that in order for successful bridging to occur, $\mathcal{A}$ expects $\mathcal{B}$ to use her general knowledge of language and the world to draw stereotypical relationships between the objects and events referred to in the utterances, in such a way that that overall discourse coherence is maintained. Specifically in (2), $\mathcal{A}$ expects $\mathcal{B}$ to know that picnics usually involve beer drinking, whence $\mathcal{B}$ can resolve the anaphoric definite by reinterpreting *beer* as a relational noun of sorts. Thus the logical form of (2a) is (2d): Here, the beer is said to be “beer of” z, and the coreference is established by the assertion that $x = z$. (Needless to say, the variables would be free in standard predicate logic, but we are assuming a dynamic system here.)

For the purposes of this paper, we presuppose the mechanism behind these inferences. What we are concerned with is the following: While the coreference resolution in (2) was straightforward, the standard pragmatic story breaks down when faced with what we refer to as *ambiguous bridging inferences*:

(3) A car hit a truck. The windshield shattered.
a. A car hit a truck. The windshield of the car shattered.
b. A car hit a truck. The windshield of the truck shattered.
c. A car hit a truck. The windshield *[of some unmentioned antecedent]* shattered.

Given current theories of bridging, (3a) and (3b) are equally valid inferences. Since both cars and trucks have windshields, (3a) is no more more salient an inference than (3b), or vice versa. Nor is there a ‘stereotypical relationship’ between the objects and events mentioned in (3) that could be relied upon in building an unambiguous bridge. However, irrespective of context, we have intuitions that the relative likelihood of the (3a) is greater than that of (3c). It is asymmetries like this that we use PDPL to explain.

We begin with a non-classical logical language $\mathcal{L}$ whose syntax we take to be that of standard predicate logic. We define a *world model* to be a triple $\langle W, D, I\rangle$, where W is a non-empty set of possible worlds; D a domain of individuals common to all worlds in W; and I a function from worlds to interpretation functions assigning values to the non-logical constants of $\mathcal{L}$ in the usual way. We assume for simplicity that the sets W of worlds and D of individuals are finite. Based such a model, the information state of an agent can be represented by two ingredients: a prior probability distribution P over W; and a accessibility relation R in W^2 that is serial, transitive and euclidean, and such that for all $w, v \in W$, if wRv then $P(v) > 0$. The probability $P_w(\{v\})$ of the singleton proposition $\{v\}$ is written as $P_w(v)$. The agent's subjective probability at $w \in W$ of a world v is $P_w(v|R) := P(v)/\sum_{u:wRu} P(u)$ if wRv, 0 otherwise. The probability of a sentence $\varphi \in \mathcal{L}$ at world w is the expectation of its truth value: $P_w(\varphi|R) := \sum_{v:wRv} P_w(v|R) \times I_w(\varphi)$.

We are interested in modeling discourse information in addition to world knowledge. Let the set of *possibilities* be $\mathcal{I} = \{\langle w, g\rangle \mid w \in W, g \in D^X, X \subseteq Var\}$, i.e., pairs of worlds and partial variable assignments. For any possibility $i = \langle w, g\rangle$, we write '$i(\alpha)$' for the denotation of α at i, i.e., $I(w)(\alpha)$ if α is a constant and $g(\alpha)$ if α is a variable. An information state again involves two ingredients. The first is an accessibility relation c in $\mathcal{I}^2$ that is serial, transitive and, for our present purposes, euclidean. This relation is updated in response to incoming linguistic information. Def. 1 and 2 fix some auxiliary notions; Def. 3 gives the recursive definition.

Definition 1 (Referent activation)**.** For all variables $u \in VAR$, a relation $[u] \subseteq \mathcal{I}^2$ is defined as follows: $\langle w, g\rangle[u]\langle w', g'\rangle$ iff (i) $w = w'$; (ii) $u \notin dom(g)$; (iii) $dom(g') = dom(g) \cup \{u\}$; (iv) $g'(v) = g(v)$ for all $v \neq u$.

Definition 2 (Descendants)**.** A pair $\langle i', j'\rangle \in \mathcal{I}^2$ is a *descendant* of $\langle i, j\rangle \in \mathcal{I}^2$ iff for some sequence $x_1, \ldots, x_n$ with $0 \leq n$, $i[x_1]\ldots[x_n]i'$ and $j[x_1]\ldots[x_n]j'$.

Definition 3 (Update)**.**

$$\begin{aligned} c[Pt_1, \ldots, t_n] &:= \{\langle i, j\rangle \in c | \langle j(t_1), \ldots, j(t_n)\rangle \in j(P)\} \\ c[\neg\varphi] &:= \{\langle i, j\rangle \in c | \langle i, j\rangle \text{ has no descendants in } c[\varphi]\} \\ c[\varphi \wedge \psi] &:= c[\varphi][\psi] \\ c[\exists x] &:= \{\langle i', j'\rangle | \text{for some } \langle i, j\rangle \in c, i[x]i' \text{ and } j[x]j'\} \end{aligned}$$

The semantics of PDPL follow closely to that of Groenendijk *et al.* (1996), except for the fact we deal in sets of pairs of possibilities rather than sets of possibilities. This allows for easy extension to multi-agent settings and second-order beliefs.

The second ingredient is again a prior probability distribution Pr over W. Let $\Pr_{\langle w,\emptyset\rangle}(\langle v, \emptyset\rangle|c) := \Pr_w(v|R_c)$, where wR_cv iff $\langle w, \emptyset\rangle c\langle v, \emptyset\rangle$. According to Definition 3, all updates except those with $\exists$ are eliminative. In these cases the probability is updated by conditioning. In an update with $\exists$, the probabilities of individual possibilities are distributed uniformly over their descendants.

Definition 4 (Probability update)**.** Probabilities are updated as follows:

1. For eliminative update $[\varphi]$, $\Pr_i(j|c[\varphi]) := \begin{cases} \Pr_i(j|c)/\sum_{k:i(c[\varphi])k} \Pr_i(k|c) & \text{if } i(c[\varphi])j \\ 0 & \text{otherwise} \end{cases}$
2. If icj, $i[x]i'$ and $j[x]j'$, then $\Pr_{i'}(j'|c[\exists x]) := \Pr_i(j|c)/|D|$

To illustrate the workings of the theory, consider the model, $\mathbb{M}$, where $W = \{v, w\}, D = \{a_1, \ldots, a_n\}$ and the interpretation of basic terms are as follows:

	bridge	collapse	strut	rotten		x	y	z
v	$\{a_1, a_2\}$	$\{a_1\}$	$\{\langle a_3, a_1\rangle, \langle a_4, a_6\rangle\}$	$\{a_3, a_4\}$	g_v	a_{39}	a_{99}	a_1
w	$\{a_1, a_2\}$	$\{a_2\}$	$\{\langle a_3, a_1\rangle, \langle a_4, a_6\rangle\}$	$\{a_3\}$	g_w	a_{392}	a_{27}	a_2

Assume $\mathcal{B}$'s fixed possibility of departure is $\langle w, g_w\rangle$. We take the set of successors of this possibility to be $c_{\mathcal{B}}[\langle w, g_w\rangle] = \{\langle w, g_w\rangle, \langle v, g_v\rangle\}$, where the probabilities $\mathcal{B}$ ascribes each possibility are $\Pr_{\langle w,g_w\rangle}(\langle v, g_v\rangle) = .7$ and $\Pr_{\langle w,g_w\rangle}(\langle w, g_w\rangle) = .3$. Now imagine $\mathcal{A}$ utters (1a) to $\mathcal{B}$, the logical form of which is $\varphi := \exists x(\mathsf{bridge}(x) \wedge \mathsf{collapse}(x)) \wedge \exists y \exists z(\mathsf{strut}(y, z) \wedge \mathsf{rotten}(y))$. The fact that z is bound represents $\mathcal{B}$'s inference that the strut is part of some object, (s)he just doesn't know which. After $\mathcal{B}$'s update with φ we have that $c'_{\mathcal{B}}[\langle u, g'_u\rangle] = \{\langle v, g'_u\rangle, \langle v, g''_v\rangle, \langle w, g'_w\rangle\}$, where the assignment function of each possibility is as follows:

	x	y	z
g'_v	a_1	a_3	a_1
g''_v	a_1	a_4	a_6
g'_w	a_2	a_3	a_1

$\mathcal{B}$'s belief that $\mathcal{A}$ intends the strut to be interpreted as part of the bridge or some unmentioned antecedent, but not both, is an equivalence relation, $\sim$, over $c_{\mathcal{B}}[\varphi]$ such that $[c_{\mathcal{B}}[\varphi]]_\sim := \{\langle i, j\rangle \in c_{\mathcal{B}}[\varphi] \mid \text{for some } P, \text{for all } j, j(z) \in j(P)\}$. There are such two classes, $\{\langle\langle w, g'_w\rangle, \langle w, g'_w\rangle\rangle, \langle\langle w, g'_w\rangle, \langle v, g'_v\rangle\rangle\}$ and $\{\langle\langle w, g'_w\rangle, \langle w, g''_v\rangle\rangle\}$. The probability of each class is the sum total of its possibilities such that the probability of the former is .65 and the latter .35. Now we have that $\mathcal{B}$ interprets the strut to be part of the bridge and not some unmentioned antecedent. Because relevant contextual information can be encoded into $\mathcal{B}$'s initial probability distribution, it should be clear how this theory can be extended to account for (3).

Asher, Nicholas, Alex Lascarides: 2003, Logics of Conversation. Cambridge University Press; Groenendijk, Jereon, Martin Stokhof: 1991, Dynamic Predicate Logic. In: Linguistics and Philosophy,Vol. 14, no. 1, pp. 39-100. Groenendijk, Jeroen, Martin Stokhof, Frank Veltman: 1996, Coreference and Modality. In: S. Lappin (ed), Handbook of Contemporary Semantic Theory,Oxford, Blackwell, pp. 179-216; Hawkins, John A: 1978, Definiteness and Indefiniteness. Atlantic Highlands, NJ: Croom Helm London; Levinson, S: 2000, Presumptive Meaning. Cambridge, Mass.: The MIT Press; Matsui, Tomoko; 2001, Experimental pragamatics: towards testing relevance- based predictions about anaphoric bridging inferences. V. Akman et al. (Eds.): Context 20001, LNAI 2116, pp. 248-260.