

Abduction and Dynamic Preference in Plan-Based Dialogue Understanding

Katashi Nagao

Sony Computer Science Laboratory Inc.

3-14-13 Higashi-gotanda, Shinagawa-ku, Tokyo 141, Japan

nagao@csl.sony.co.jp

Abstract

In this paper, we present a unified model of dialogue understanding involving various sorts of ambiguities, such as lexical, syntactic, semantic, and plan ambiguities. This model is able to estimate and revise the most preferable interpretation of utterances along with the progress of dialogues, the features of which successfully capture the dynamic nature of dialogue understanding. The model supports abductive reasoning and dynamic preference, which is comprised of two portions: (1) an extension of first-order logic for maintaining multiple interpretations of ambiguous utterances in a dialogue; (2) a device which estimates and revises the most preferable interpretation of these multiple interpretations. Since the model is logic-based, it provides a good basis for formulating rational justification of its current interpretation, which is one of the most desirable aspects in generating helpful responses.

1 Introduction

To properly understand natural language utterances, dialogue participants must deal with several sorts of ambiguity in utterances. Such ambiguity can involve lexical, syntactic, semantic, and pragmatic ambiguity. In this paper, we propose a model of plan recognition for utterances involving the above mentioned sorts of ambiguities. This model is able to estimate and revise the most preferable interpretation of the utterances along with the progress of dialogues. Our model successfully captures the dynamic nature of dialogue understanding. Our framework is comprised of two portions: (1) an extension of first-order logic for maintaining multiple interpretations of ambiguous utterances in a dialogue; (2) a device which estimates and revises the most preferable interpretation among these multiple interpretations. Since the model is logic-based, it provides a good basis for formulating rational justifications of its current interpretation. The ability to provide justification is one of the most desirable aspects in generating helpful responses. The features contained in our model are extremely useful for interactive dialogue management.

According to Pollack [1986], a plan represents an agent's beliefs and intentions about achieving a goal. Similarly, a recognized plan represents the recognizing agent's beliefs about the plan constructed by the planning agent. Since planning and recognizing agents are different, their knowledge and beliefs need not be the same, and the plan constructed by the planning agent may not be the plan that the recognizing agent would have constructed in the same circumstances. Based upon this observation, we can view the recognized plan as a set of assumptions which the recognizing agent inferred from utterances using his/her background knowledge and beliefs about the plan construction agent.

Because of the dynamic nature of dialogues, it is important to maintain sets of possible assumptions, as multiple interpretations, whenever ambiguous utterances are encountered. Also we need a device which estimates the most plausible interpretation among multiple competitive interpretations¹. A measure of the plausibility should also capture the dynamic nature of dialogues. After some information is provided by means of an utterance, the plausibility of a certain context² will increase, while plausibilities of other contexts will decrease.

Methods and devices to resolve ambiguities of utterances have been proposed in the literature. These methods include Small et al. [1988] for lexical disambiguation, Charniak and Goldman [1988] for lexical and semantic disambiguation, Hirst [1987] and Nagao [1990] for syntactic disambiguation, Nagao [1992] for syntactic and semantic disambiguation, and Carberry [1990a] and van Beek and Cohen [1991] for pragmatic (or recognized plan) disambiguation.

Simply combining these approaches to disambiguate all sorts of ambiguities quickly leads to combinatorial explosion. Because these methods have different frameworks, cooperation among disambiguation methods is difficult at best. While our method still has a combinatorial explosion, because it has a unified framework, it allows synergy among different types of ambiguity resolution.

¹Competitive interpretations are those which are mutually exclusive.

²Note that throughout this paper, we call a consistent set of facts and assumptions a *context*. If there are multiple consistent sets for a dialogue, we say there are *multiple contexts*.

Waltz and Pollack [1985] proposed a connectionist model for natural language interpretation which allows strong interaction among lexical, syntactic, and pragmatic processes. The model is able to give a proper interpretation for a semantically ambiguous sentence by considering the contextual influence on meaning interpretation. However, the model has two major disadvantages as a framework for natural language dialogue understanding systems, particularly for formulating helpful responses in cooperative dialogues. First, it doesn't address plan recognition problems. We think that their framework is not suitable for dealing with plan ambiguity. Second, the model is fairly weak in explaining and rationally justifying its current interpretation. This problem is inherent in methods which represent interpretation by means of the energy distribution of a network. The latter disadvantage is, for other reasons, also the case with probabilistic approaches to plan recognition [Charniak and Goldman, 1991].

Towards a unified framework for interpreting ambiguous utterances, we introduce a dialogue understanding model based on the *Multi-World Model* [Nagao, 1989], which is a formal model for maintaining multiple interpretations of ambiguous natural language sentences. The model represents worlds as first-order axiomatic systems linked through translations between worlds. These translations are mappings of theorems between axiomatic systems. The results of the semantic interpretation are assumed to be constructed on such a multi-world structure. In this model, a *context* is defined as a sequence of worlds which include a consistent set of facts and assumptions.

In this paper, we extend the model so that it overcomes the problems mentioned above. First, preferences among contexts are dynamically computed based on facts and assumptions in each context. The preference captures a measure of the plausibility of a context. The current most preferable context contains a currently recognized plan. Second, changing the most plausible context among mutually exclusive contexts in response to new information (utterances) is formally treated as belief revision of a plan recognizing agent. Most previously proposed models for resolving ambiguity in plan recognition [Carberry, 1990a; van Beek and Cohen, 1991] cannot revise their pragmatic decisions once they commit to one recognized plan.

2 The Multi-World Model Revisited

We previously developed a logical framework for semantic interpretation called the *Multi-World Model* [Nagao, 1989]. In this paper, we extend it so that it can provide a mechanism of abductive reasoning and of preferences among multiple contexts each of which consists of consistent sets of facts and assumptions. In Section 2.1 and Section 2.2, we revisit the model. In Section 2.3, we describe the extension to the model.

2.1 Worlds

The Multi-World Model is a 2-tuple $\langle W, T \rangle$, where:

- W : Set of *worlds*.

- T : Set of *translations between worlds*.

Worlds are represented by first-order axiomatic systems. Thus, a world is denoted by $w = \langle A_w, V_w \rangle$.

1. Axioms A_w
 A_w : Set of axioms of world w .
2. Vocabulary V_w
 - (a) V_{c_w} : Set of constants of world w .
 - (b) $V_{f_w}^n$: Set of n -ary function symbols of world w .
 - (c) $V_{P_w}^m$: Set of m -ary predicate symbols of world w .
 - (d) $V_{f_w} = \cup \{V_{f_w}^i \mid i \in N\}$.
 - (e) $V_{P_w} = \{V_{P_w}^j \mid j \in N\}$.
 - (f) $V_w = V_{c_w} \cup V_{f_w} \cup V_{P_w}$: Vocabulary of world w .

2.2 Translation between Worlds

A translation between worlds is defined by a mapping of a set of logical formulae in one world to a set in another world. It consists of a translation of vocabulary and a translation of formulae.

2.2.1 Translation of Vocabulary

Let w_1 and w_2 be worlds; vt , the translation of vocabulary between w_1 and w_2 , is then defined by the following injections:

1. For constants, $vt_c : V_{c_{w_1}} \Rightarrow V_{c_{w_2}}$.
2. For function symbols, $vt_f^i : V_{f_{w_1}}^i \Rightarrow V_{f_{w_2}}^i (i \in N)$.
3. For predicate symbols, $vt_P^j : V_{P_{w_1}}^j \Rightarrow V_{P_{w_2}}^j (j \in N)$.
4. $vt = \langle vt_c, vt_f^i (i \in N), vt_P^j (j \in N) \rangle$: Translation of vocabulary between w_1 and w_2 .

2.2.2 Translation of Formulae

Let w_1 and w_2 be worlds; ft , the translation of formulae between w_1 and w_2 , is then defined by the following functions, using vt , the translation of vocabulary:

1. Translation of terms
 - (a) If x is a variable, then $ft(x) = x$.
 - (b) If a term t is a constant, then $ft(t) = vt_c(t)$.
 - (c) If $t_1, \dots, t_n$ are terms and f is a n -ary function symbol, then $ft(f(t_1, \dots, t_n)) = vt_f(f)(ft(t_1), \dots, ft(t_n))$.
2. Translation of atomic formulae
 - (a) If t_1 and t_2 are terms and $t_1 = t_2$, then $ft([t_1 = t_2]) = [ft(t_1) = ft(t_2)]$.
 - (b) If $t_1, \dots, t_m$ are terms and P is an m -ary predicate symbol, then $ft(P(t_1, \dots, t_m)) = vt_P(P)(ft(t_1), \dots, ft(t_m))$.
3. Translation of formulae
 - (a) If ϕ and ψ are formulae, then $ft(\phi \wedge \psi) = ft(\phi) \wedge ft(\psi)$, $ft(\phi \vee \psi) = ft(\phi) \vee ft(\psi)$, $ft(\neg \phi) = \neg ft(\phi)$, and $ft(\phi \rightarrow \psi) = ft(\phi) \rightarrow ft(\psi)$.
 - (b) If x is a variable and ϕ is a formula, then $ft(\forall x.\phi) = \forall x.ft(\phi)$ and $ft(\exists x.\phi) = \exists x.ft(\phi)$.

We denote the translation between worlds by the symbol t , consisting of the above vt and ft .

2.2.3 Consistent Translation

There are many cases where reasoning in one world needs the axioms in another world. We therefore define a *consistent translation* as a translation whose resultant axioms are consistent with original axioms in the target world.

A more precise description is as follows:

Let w_1 and w_2 be worlds, and let t be translation between w_1 and w_2 . If the translation results of all axioms of w_1 are consistent in w_2 (i.e., $\neg\exists\phi \in t(A_{w_1}) \equiv \{t(\psi) \mid \psi \in A_{w_1}\} \cup A_{w_2} \text{ s.t. } t(A_{w_1}) \cup A_{w_2} \vdash \phi \wedge \neg\phi$), then translation t is called a *consistent translation*.

Consistently translated axioms can be regarded as axioms of w_2 (i.e., $\forall\phi \in A_{w_1}, t(\phi) \in A_{w_2}$).

2.2.4 Possible-Modal Axiom

Let w be a world, and ϕ be a formula defined in w . If there is a world w' that is connected with w by a consistent translation t and $t(\phi)$ is a theorem (i.e., $\exists w', \exists t : w \Rightarrow w', t$ is a consistent translation, $A_{w'} \vdash t(\phi)$), then the world w has a *possible-modal axiom* that is described as $\Diamond\phi$ ($\in A_w$). Possible-modal axioms are distinguished from (normal) axioms. So possible-modal axioms do not affect consistent translations.

The formulae satisfying the condition of modality of *possibility* play important roles in the Multi-World Model. These axioms may be inconsistent with one world, but consistent with some other world.

2.3 Abductive Reasoning Based on the Multi-World Model

Abduction is an inference schema, such that from $\forall x.\phi(x) \rightarrow \psi(x)$ and $\psi(a)$, one concludes $\phi(a)$. One can think of $\psi(a)$ as the observable evidence, of $\forall x.\phi(x) \rightarrow \psi(x)$ as a general principle that explains $\psi(a)$'s occurrence, and of $\phi(a)$ as the underlying cause or explanation of $\psi(a)$. This type of inference is not valid. There may be alternative explanations, such as $\phi(a)$'s, which are considered as assumptions. Therefore, other criteria are required to choose among the possible assumptions. In the Multi-World Model, abductive reasoning is described by means of possible-modal axioms and preferences among contexts (consistent sets of facts and assumptions). For example, a principle used in abduction $\forall x.\phi(x) \rightarrow \psi(x)$ is represented as a possible-modal axiom $\Diamond\forall x.\psi(x) \rightarrow \phi(x)$. From this axiom and a factual belief $\psi(a)$, one can conclude $\phi(a)$ in a certain world. However, this conclusion is treated as an assumption. Each possible assumption is placed on an (abductive) world as an axiom. A *context* that includes such an abductive world is assigned a preference according to degrees of dependencies of axioms in the context. The preference can change as information grows. When we need to select among multiple mutually exclusive contexts, a preference operator is used to choose the 'best' context. This context, along with its assumptions, then becomes the preferred context.

2.3.1 Multiple Contexts

In the Multi-World Model, a world sequence in which worlds are connected through consistent translations is called a *context*. Abductively inferred beliefs (recognized

plans and interpretations of utterances) are considered as axioms of special worlds called *abductive worlds*. If these beliefs are mutually exclusive (i.e., alternatives of the same phenomenon), then each one is placed on different abductive worlds in the same layer. The Multi-World Model that is structured as a hierarchy (shown in Figure 1) describes *multiple contexts*. For example, a

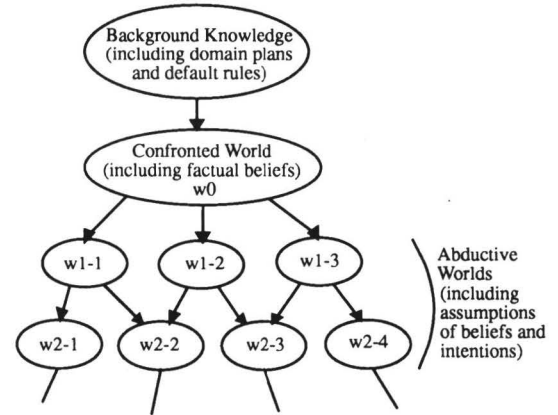


Figure 1: World Hierarchy that Represents Multiple Contexts

world sequence, w_0, w_{1-1}, w_{2-1} , represents one context, while the sequence w_0, w_{1-1}, w_{2-2} describes a different context.

Multiple contexts are constructed as follows:

1. When an utterance is processed, facts or assumptions are derived from it.
2. *Abductive worlds* are constituted of individual assumptions, and the *confronted world* is constituted of facts.
3. A relationship between facts and assumptions is defined as a consistent translation from the confronted world to an abductive world.
4. When a new utterance is processed, relationships between worlds change as the occasion demands.
5. Multiple contexts are constructed from several relationships among facts and assumptions (i.e., the confronted world and abductive worlds).
6. Preferences among multiple contexts are calculated according to degrees of dependency between worlds, which can change as information grows.
7. A preferred context includes preferred assumptions.

Since the preferences can dynamically change in response to the situation, we have to design a device for computing preferences that sufficiently captures the dynamic nature of preferences. Numerical descriptions of preferences can sufficiently capture information dynamics. Actually, the numerical plausibility of abductive inference is widely utilized in other works [Hobbs *et al.*, 1988; Charniak and Shimony, 1990].

The preferences we use are defined by degrees of dependency between worlds connected via consistent translations.

2.3.2 Dependency between Worlds

Let w_1 and w_2 be worlds, let t be consistent translation between w_1 and w_2 , and let $\phi = \dots \wedge P(a) \wedge \dots$ be an axiom of w_1 , and $\psi = \dots \wedge Q(b) \wedge \dots$ be an axiom of w_2 (except the axioms translated by t from w_1). If at least one of the following relations between ϕ and ψ hold (i.e., provable in w_2), then we say that there is a *dependency* between w_1 and w_2 by a consistent translation t :

1. **Argument equality**
If $t(a) = b$ holds, then there is a dependency between ϕ and ψ called *argument equality*.
2. **Predicate equality**
If $t(P) = Q$ holds, then there is a dependency between ϕ and ψ called *predicate equality*.
3. **Positive entailment**
If $t(P(a)) \rightarrow Q(b)$ or $Q(b) \rightarrow t(P(a))$ holds (except for the case where $t(P(a)) = Q(b)$ holds), then there is a dependency between ϕ and ψ called *positive entailment*.
4. **Negative entailment**
If $t(P(a)) \rightarrow \neg Q(b)$ or $\neg Q(b) \rightarrow t(P(a))$ holds, then there is a dependency between ϕ and ψ called *negative entailment*.
In this case, at least one (translation of) possible-modal axiom must be included in a proof sequence of $t(P(a)) \rightarrow \neg Q(b)$ or $\neg Q(b) \rightarrow t(P(a))$.

Each dependency, except negative entailment, has a positive degree of dependency (let each be u_i , then $u_1, u_2, u_3 > 0$). While negative entailment has a negative degree of dependency $u_4 < 0$.

If several dependencies hold, the degree of dependency is defined as the sum of each value, which is called the *degree of dependency between worlds* and is denoted by d_{w_1, w_2} .

For example, if $\phi = P_1(a_1) \wedge P_2(a_2)$ is in w_1 , $\psi = Q_1(b_1) \wedge Q_2(b_2)$ is in w_2 , and $t(a_1) = b_1, t(a_2) = b_2, t(P_1) = Q_1, t(P_2(a_2)) \rightarrow Q_2(b_2)$ hold, then the degree of dependency between w_1 and w_2 becomes $d_{w_1, w_2} = 2u_1 + u_2 + u_3$.

2.3.3 Preferences among Multiple Contexts

A *context* is represented by a world sequence like $w_0.w_1\dots$. Let W_i be a set of abductive worlds in the i -th layer and w_i be its element (i.e., $w_i \in W_i$). W_0 means the layer of the confronted world which includes factual information.

A preference of a context $w_0.w_1\dots, p(w_0.w_1\dots)$ is calculated as follows:

1. $p(w_0) = 1$.
2. $p(w_0.w_1) = \frac{d_{w_0, w_1}}{\sum_{w' \in W_0, w'' \in W_1} d_{w', w''}}$.
3. If $\exists i, \neg \exists t : w_i \Rightarrow w_{i+1}$, then $p(w_0\dots w_i.w_{i+1}\dots) = 0$.
In this case, $w_0\dots w_i.w_{i+1}\dots$ is inconsistent.
4. In the case where $i > 1$, $p(w_0\dots w_{i-1}.w_i) = \frac{p(w_0\dots w_{i-1}) + d_{w_{i-1}, w_i}}{\sum_{w' \in W_{i-1}, w'' \in W_i} p(w_0\dots w') + d_{w', w''}}$.

In this case, the preference of $w_0\dots w_i$ is calculated by adding the preference of $w_0\dots w_{i-1}$ and the degree of dependency between w_{i-1} and w_i , and normalizing by the sum of all the contexts including a world in W_i .

3 Plan Recognition and Belief Revision in Dialogue Understanding

We apply the Multi-World Model to plan recognition and belief revision in utterance comprehension with a method that describes the changing of preferences among contexts of beliefs. We treat belief revision as dynamic assignments of preferences among contexts constructed from utterances, not as conventional static assignments of truth values or preferences.

Recognizing the speaker's beliefs and plans is important for understanding utterances. If there are alternative beliefs or plans inferred from an utterance, then we should make assumptions which correspond to each alternative. Perrault [1990] constructed a formal theory of speech acts based on Reiter's default logic [Reiter, 1980], but he could not properly deal with belief revisions that result from speech acts. The model for belief revision should describe preferences among competitive assumptions of beliefs. Appelt and Konolige [1989] proposed a formal treatment of belief revision based on hierarchic autoepistemic logic [Konolige, 1989], an extension of Moore's autoepistemic logic [Moore, 1985]. In this framework, a hierarchy of defaults represents preferences among beliefs (i.e., the result of a default rule in one level is preferred to the result of a default rule in a higher level). However, it cannot deal with preferences among consistent beliefs that dynamically change as the volume of observed information increases.

A heuristic method has been proposed to assess ambiguous contexts and choose a preferred alternative [Carberry, 1990a]³. However, once a context is selected by the method, even if a more appropriate context (including alternative plans) is later found, it cannot switch the current selected context to that context. Our framework allows the context model to dynamically switch to a more preferable context by maintaining multiple contexts and updating preferences among contexts.

3.1 Example

In many cases, there can be mutually exclusive hypothetical interpretations of an utterance. This multiplicity is caused by the uncertainty of the recognizing agent's beliefs about the planning agent's beliefs and ambiguity in utterances. As an example of ambiguous utterances which yield multiple interpretations, consider the following dialogue between two agents, **A** and **B**. We assume that agents **A** and **B** live in the same town, and they are acquaintances of each other. One day agent **A** met agent **B** on his way to a river bank.

³Carberry [1990b] discusses the problem of building and revising the context model based on evidential reasoning. Her motivation is similar to ours, but her mechanism is still unclear about basic probability assignments and preference rules.

A1: Where are you headed?

B1: I'm going to the bank.

See, here I have a terrific gun which I bought yesterday.

A2: Gosh! Are you going to rob it?

B2: No, I'm going to Tama river.

Across the river there are woods.

A3: Oh, I see. You are going hunting.

We assume that the plan library of agent **A** in this example is the same as that of Kautz's example [Kautz, 1990] (Figure 2). Additionally, we assume that there is

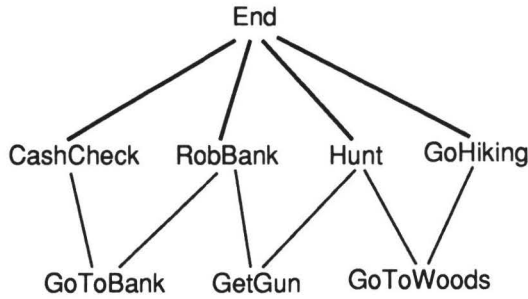


Figure 2: Plan Library for Cashing a Check, Robbing a Bank, and Hunting

a river called Tama in the town. Also, in the town there is only one bank (a place in which money is kept).

Agent **B**'s goal is to go to the woods across the river to hunt. After agent **A** heard utterance **B1** by **B**, **A** might have assumed that agent **B** is going to rob a bank. This recognition is entailed by Kautz's method of plan recognition. However, after **B** uttered **B2**, **A** would have recognized that **B** actually intended to go to the woods across the river to hunt. Since a recognized plan is based on the beliefs of a recognizing agent, it may not be the plan that the planning agent constructed. We can view the recognized plan as a set of the most plausible assumptions held by the recognizing agent. As it is a set of assumptions, the recognizing agent may revise his/her assumptions after additional information is provided to him/her by means of utterances. Therefore, we have to design the plan recognition process so that recognizing agents can revise their assumptions appropriately based upon information provided.

In our framework, semantic interpretations of ambiguous utterances and intentions, which are abductively inferred by using ambiguous plans, are uniformly treated as assumptions. These assumptions are represented by logical formulae and placed on different abductive worlds as axioms. In addition, we assume that (domain) plans are represented as possible-modal axioms asserted in the background knowledge.

For example, the domain plans in the Kautz's example [Kautz, 1990] are represented as follows:

$P_1 : \Diamond \forall x. GoToBank_1(x) \rightarrow \exists y. CashCheck(y)$

$P_2 : \Diamond \forall x, y. GoToBank_1(x) \wedge GetGun(y) \rightarrow \exists z. RobBank_1(z)$

$P_3 : \Diamond \forall x, y. GoToWoods(x) \wedge GetGun(y) \rightarrow \exists z. Hunt(z)$

$P_4 : \Diamond \forall x. GoToWoods(x) \rightarrow \exists y. GoHiking(y)$

We assume that they have *a priori* knowledge about the above plans and the following default rule:

$D_1 : \Diamond \forall x. GoToBank_2(x) \rightarrow \exists y. GoToRiver(y)$

When agent **B** makes the utterance **B1**, **A** infers **B**'s beliefs and intentions as follows:

The word "bank" has two meanings which are denoted by $Bank_1$ (a place in which money is kept) and $Bank_2$ (land along the side of a river), so the assumptions $GoToBank_1(e_1)$ and $GoToBank_2(e_2)$ are constructed. The abductive worlds w_{1-1} and w_{1-2} corresponding to each assumption are created as shown in Figure 3, where the constants e_1 and e_2 represent instances of events⁴.

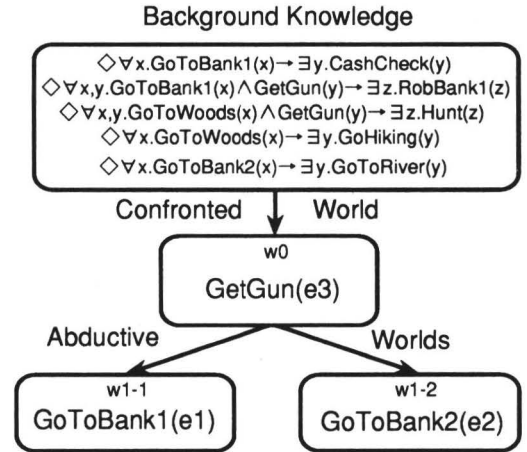


Figure 3: Multi-World of Semantic Interpretation of **B1**

In addition, the factual belief $GetGun(e_3)$ is placed on the confronted world w_0 . Consistent translations of w_0 -to- w_{1-1} and w_0 -to- w_{1-2} are successfully constructed. Translations of vocabulary are identical mappings. Multiple contexts $w_0.w_{1-1}$ and $w_0.w_{1-2}$ are then considered for calculating preferences. In this case, the preferences of contexts $w_0.w_{1-1}$ and $w_0.w_{1-2}$ are the same since no dependency between worlds is found.

Then, by using plans P_1 and P_2 , the two inferred intentions, $CashCheck(e_4)$ and $RobBank_1(e_5)$, are generated and placed on the lower layer abductive worlds w_{2-1} and w_{2-2} , respectively (Figure 4).

Consistent translations are constructed in order to connect worlds⁵. Multiple contexts are then considered for computing preferences. The preferences of $w_0.w_{1-1}.w_{2-1}$ and $w_0.w_{1-1}.w_{2-2}$ are calculated as follows. The degree of dependency between w_{1-1} and w_{2-1} becomes $u_2 + u_3$ because there is a predicate equality of $GoToBank_1$ and $GoToBank_1(e_1)$ entails

⁴We already have a syntactic/semantic analysis device which can recognize utterances and extract facts and assumptions from them. For the sake of simplicity, we don't describe this device in this paper. See Nagao [1992].

⁵If we represent logical semantic interpretations more strictly, then w_{1-2} will be inconsistent with w_{2-2} (and perhaps with w_{2-1}) because these worlds refer to the different semantic objects of "bank."

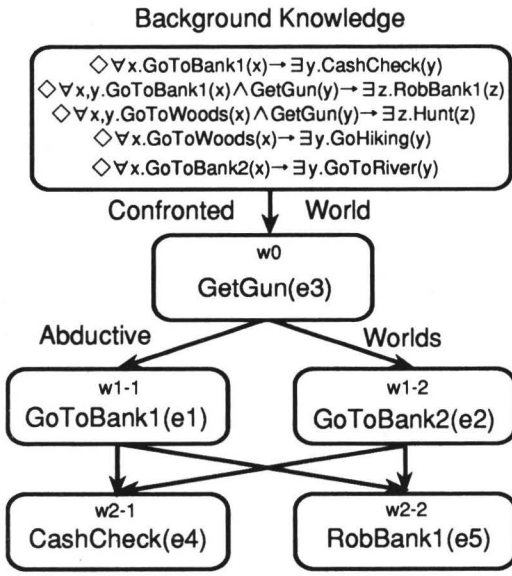


Figure 4: Multi-World of Plan Recognition of **B1**

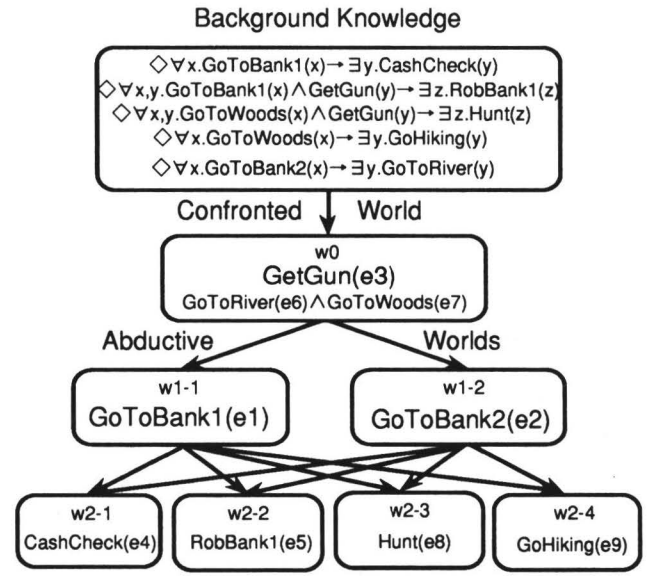


Figure 5: Multi-World of Semantic Interpretation and Plan Recognition of **B2**

CashCheck(e_4) by using the plan P_1 (P_1 is inherited from the background knowledge to w_{2-1})⁶. While the degree of dependency between w_{1-1} and w_{2-2} becomes $2u_2 + u_3$ because there are two predicate equalities of *GoToBank1* and *GetGun*, and *GoToBank1*(e_1) entails *RobBank1*(e_5) by using the plan P_2 (inherited from the background knowledge to w_{2-2}) and the fact *GetGun*(e_3) exists in w_{1-1} (inherited from w_0 to w_{1-1}). The context $w_0.w_{1-1}.w_{2-2}$ is the most preferable because the preference value is the highest among the alternative contexts.

After **B** gives the utterance **B2**, the factual belief *GoToRiver*(e_6) $\wedge$ *GoToWoods*(e_7) is generated and placed on the confronted world. Plans P_3 and P_4 are activated when the fact *GoToWoods*(e_7) is encountered. Assumptions of **B**'s intention *Hunt*(e_8) and *GoHiking*(e_9) are then inferred. Abductive worlds w_{2-3} and w_{2-4} , which have each assumption, are generated in the same layer of w_{2-1} and w_{2-2} . This is done because these four worlds are mutually exclusive. The multi-world is constructed as shown in Figure 5.

Now, we consider the preferences of the contexts including w_{1-2} . The preferences of $w_0.w_{1-2}.w_{2-3}$ and $w_0.w_{1-2}.w_{2-4}$ are calculated as follows. The degree of dependency between w_0 and w_{1-2} becomes $u_2 + u_3$, because *GoToBank2*(e_2) entails *GoToRiver*(e_6) by using the default rule D_1 which is inherited from the background knowledge to w_{1-2} . The degree of dependency between w_{1-2} and w_{2-3} becomes $2u_2 + u_3$, since *GoToWoods*(e_7) and *GetGun*(e_3) (inherited from w_0 to w_{1-2}) entails *Hunt*(e_8) by using the plan P_3 (inherited

from the background knowledge to w_{2-3}). While the degree of dependency between w_{1-2} and w_{2-4} becomes $u_2 + u_3$, because *GoToWoods*(e_7) (inherited from w_0 to w_{1-2}) entails *GoHiking*(e_9) by using the plan P_4 (inherited from the background knowledge to w_{2-4}). Now, the context $w_0.w_{1-2}.w_{2-3}$ is the most preferable since the contexts including w_{1-1} have lower preference values.

This example shows that beliefs and intentions inferred by the hearer can be revised as observed (heard) information increases. Our model captures the phenomena by maintaining multiple contexts and dynamically assigning a preference value to each context.

4 Concluding Remarks and Further Work

We have developed a new framework for plan recognition and belief revision which properly treats ambiguities in dialogue understanding. These ambiguities include lexical, syntactic, semantic, and plan ambiguities. This model is able to estimate and revise the most preferable interpretation of utterances as the dialogue progresses. The model successfully captures the dynamic nature of real world conversation. The framework is comprised of two portions: (1) an extension of first-order logic for maintaining multiple interpretations of ambiguous utterances in a dialogue; (2) a device which estimates and revises the most preferable interpretation among these multiple interpretations. Since the model is logic-based, it provides a good basis for formulating rational justifications of its current interpretation, which is one of the most desirable aspects in generating helpful responses.

Also, the model deals with the dynamic nature of dialogues by means of the following two mechanisms. First, preferences among contexts are dynamically computed based on facts and assumptions within each context.

⁶The composition of consistent translations are also considered as consistent translations. Mappings of constants to skolemized variables are also constructed in translations. See Nagao [1989]. Unification between a skolemized variable and a constant is based on Charniak's *abductive unification* [Charniak, 1988].

The preference captures a measure for the plausibility of a context. The currently most preferable context contains a currently recognized plan. Secondly, changing the most plausible context among mutually exclusive contexts along a dialogue is formally treated as a belief revision of a plan recognizing agent. Most previously proposed models for resolving ambiguity in plan recognition cannot revise their pragmatic decisions once they commit to a recognized plan.

We are currently working on the following extensions: (1) providing a mechanism for detecting a critical dialogue situation where several competitive interpretations have closer preference measures, and entering a clarification dialogue; (2) extending the model to entail temporal reasoning and planning.

Our model regards the most preferable interpretation as the recognized plan (or the most plausible interpretation). However, in some dialogues, many alternatives may have nearly equal preference values. In this critical situation, one may wish to query for additional information in order to commit to the preferable context with more certainty. A criterion for detecting the critical situation using preference measures of mutually exclusive contexts is being carried out.

Acknowledgments

The author would like to give great thanks to the other members of Sony CSL, especially Ei-Ichi Osawa, for their helpful comments and support.

References

- [Appelt and Konolige, 1989] Douglas E. Appelt and Kurt Konolige. A Nonmonotonic Logic for Reasoning about Speech Act and Belief Revision. In *Proceedings of the 2nd International Workshop on Non-Monotonic Reasoning*, Lecture Notes in Artificial Intelligence, Vol.346, pp.164–175, Springer-Verlag, 1989.
- [Carberry, 1990a] Sandra Carberry. *Plan Recognition in Natural Language Dialogue*. MIT Press, 1990.
- [Carberry, 1990b] Sandra Carberry. Incorporating Default Inferences into Plan Recognition. In *Proceedings of AAAI-90*, pp.471–478, 1990.
- [Charniak, 1988] Eugene Charniak. Motivation Analysis, Abductive Unification and Nonmonotonic Equality. *Artificial Intelligence*, Vol.34, pp.275–295, 1988.
- [Charniak and Goldman, 1988] Eugene Charniak and Robert Goldman. A Logic for Semantic Interpretation. In *Proceedings of the 26th Annual Meeting of ACL*, pp.87–94, 1988.
- [Charniak and Goldman, 1991] Eugene Charniak and Robert Goldman. A Probabilistic Model of Plan Recognition. In *Proceedings of AAAI-91*, pp.160–165, 1991.
- [Charniak and Shimony, 1990] Eugene Charniak and Solomon E. Shimony. Probabilistic Semantics for Cost Based Abduction. In *Proceedings of AAAI-90*, pp.106–111, 1990.
- [Hirst, 1987] Graeme Hirst. *Semantic Interpretation and the Resolution of Ambiguity*. Cambridge University Press, 1987.
- [Hobbs et al., 1988] Jerry R. Hobbs, Mark Stickel, Paul Martin, and Douglas Edwards. Interpretation as Abduction. In *Proceedings of the 26th Annual Meeting of ACL*, pp.95–103, 1988.
- [Kautz, 1990] Henry Kautz. A Circumscriptive Theory of Plan Recognition. In Philip R. Cohen, Jerry Morgan, and Martha E. Pollack, editors, *Intentions in Communication*, pp.105–133. MIT Press, 1990.
- [Konolige, 1989] Kurt Konolige. Hierarchic Autoepistemic Theories for Nonmonotonic Reasoning: Preliminary Report. In *Proceedings of the 2nd International Workshop on Non-Monotonic Reasoning*, Lecture Notes in Artificial Intelligence, Vol.346, pp.42–59, Springer-Verlag, 1989.
- [Moore, 1985] Robert C. Moore. Semantical Considerations on Nonmonotonic Logic. *Artificial Intelligence*, Vol.25, pp.75–94, 1985.
- [Nagao, 1989] Katashi Nagao. Semantic Interpretation Based on the Multi-World Model. In *Proceedings of IJCAI-89*, pp.1467–1473, 1989.
- [Nagao, 1990] Katashi Nagao. Dependency Analyzer: A Knowledge-Based Approach to Structural Disambiguation. In *Proceedings of COLING-90*, Vol.2, pp.282–287, 1990.
- [Nagao, 1992] Katashi Nagao. A Preferential Constraint Satisfaction Technique for Natural Language Analysis. In *Proceedings of ECAI-92*, pp.523–527, 1992.
- [Perrault, 1990] C. Raymond Perrault. An Application of Default Logic to Speech Act Theory. In Philip R. Cohen, Jerry Morgan, and Martha E. Pollack, editors, *Intentions in Communication*, pp.161–185. MIT Press, 1990.
- [Pollack, 1986] Martha Pollack. *Inferring Domain Plans in Question-Answering*. Ph.D. Thesis, University of Pennsylvania, 1986.
- [Reiter, 1980] Raymond Reiter. A Logic for Default Reasoning. *Artificial Intelligence*, Vol.13, pp.81–132, 1980.
- [Small et al., 1988] Steven I. Small, Garrison W. Cottrell, and Michael K. Tanenhaus, editors. *Lexical Ambiguity Resolution*. Morgan Kaufmann Publishers, Inc., 1988.
- [van Beek and Cohen, 1991] Peter van Beek and Robin Cohen. Resolving Plan Ambiguity for Cooperating Response Generation. In *Proceedings of IJCAI-91*, pp.938–944, 1991.
- [Waltz and Pollack, 1985] David L. Waltz and Jordan B. Pollack. Massively Parallel Parsing: A Strongly Interactive Model of Natural Language Interpretation. *Cognitive Science*, Vol.9, pp.51–74, 1985.