Semantic Interpretation Based on the Multi-World Model

Katashi Nagao
IBM Research, Tokyo Research Laboratory
5 19 Sanbancho, Chiyoda-ku, Tokyo 102, Japan
E-mail: nagao@jpntscvm.bitnet

Abstract

This paper proposes the *Multi-World Model* as a new logical framework for semantic interpretation of natural language. The model represents *worlds* as first-order axiomatic systems linked through *interpretations between worlds,* which are mappings of theorems between these axiomatic systems. Results of semantic interpretation are assumed to be constructed on such a multi-world structure.

To understand ambiguous sentences, people make assumptions to complement information that is unknown to them. If an assumption is consistent with information that is known to them, then they will support that assumption and make a unique interpretation of the sentence. When a new sentence comes in whose interpretation conflicts with a previous interpretation, they will consider another assumption. When the context is fixed, the whole interpretation will be decided by selecting the most, preferable assumptions.

The *Multi-World Model* was designed as a formal description of facts, assumptions, background knowledge, and their relationships. This paper defines the *Multi-World Model* and presents its application to semantic interpretation with concrete examples.

1 Introduction

People use background knowledge to interpret sentences, to resolve anaphora, and to check the coherence of a text. If their background knowledge is not sufficient for this purpose, then they make assumptions in order to understand their meaning.

Generally, people interpret sentences by making assumptions based on partial information. The assumptions may be inconsistent with each other, and therefore the factual information derived from the sentences is related to each of the assumptions separately.

It is difficult to represent the meaning of sentences by logical expressions when assumptions must be taken into account. Reiter\s default logic [Reiter, 1980] is one way to overcome this problem. The model-theoretical basis of the *Multi-World Model* is a Kripke Model [Landman, 1986], and default reasoning can be expressed by using the modality of possibility.

Another problem is that semantic interpretation changes according to the growth of information obtained from the sentences read. The *Multi-World Model* can treat dynamic interpretation by using assumptions and knowledge represented by logical formulae.

The main ideas in this paper are as follows:

1. When a sentence is processed, a fact or assumptions are derived from it.
2. *Abductive worlds* are constituted of individual assumptions, and the *confronted world* is constituted of facts.
3. These worlds are considered as first-order axiomatic systems.
4. A relationship between a fact and an assumption is defined as a mapping from a confronted world to an abductive world.
5. When a new sentence is processed, the relationships between worlds will change as the occasion demands.
6. When the whole context is fixed, contextual references and ambiguities in the worlds are resolved.

Semantic interpretation of sentences consists of local relationships between facts and assumptions. These relationships are called *interpretations between worlds.* We call this model of *worlds* and *interpretations between ivorlds* the *Multi-World Model.*

In the next section of this paper, we define the *Multi-World Model.*

In the last section, we describe the semantic interpretation of natural language sentences by the *Multi-World Model,* and give concrete examples.

2 The Multi-World Model

The Multi-World Model is based on the idea that many logical domains are interconnected by mappings from formulae in one domain to those in other domains. This idea was originally developed by Deguchi and Oka, who called it a *Knowledge Network* [Deguchi and Oka, to appear].

We develop a model for semantic interpretation based on their formalization.

The Multi-World Model is a 2-tuple $< WS, IS >$, where:

- WS : Set of *worlds*.
- IS : Set of *interpretations between worlds*.

In the next paragraphs, we define worlds and interpretations between worlds.

2.1 Worlds

Worlds are represented by first-order axiomatic systems. Thus, a world is denoted by $W =< A_W, V_W >$.

1. Axioms A_W
   A_W : Axioms of world W.
2. Vocabularies V_W
   (a) V_{c_W} : Set of constants of world W.
   (b) $V^n_{f_W}$: Set of n-ary function symbols of world W.
   (c) $V^m_{P_W}$: Set of m-ary predicate symbols of world W.
   (d) $V_{f_W} = \cup\{V^i_{f_W} \mid i \in N\}$.
   (e) $V_{P_W} = \{V^j_{P_W} \mid j \in N\}$.
   (f) $V_W = V_{c_W} \cup V_{f_W} \cup V_{P_W}$: Vocabularies of world W.

2.2 Interpretations between Worlds

An interpretation between worlds is defined by mapping from a set of logical formulae in one world to a set in another world. It consists of an interpretation of vocabularies and an interpretation of formulae.

2.2.1 Interpretation of Vocabularies

Let W_1 and W_2 be worlds; J, the interpretation of vocabularies between W_1 and W_2, is then defined by the following injections:

1. For constants, $J_c : V_{c_{W_1}} \to V_{c_{W_2}}$.
2. For function symbols, $J^i_f : V^i_{f_{W_1}} \to V^i_{f_{W_2}} (i \in N)$.
3. For predicate symbols, $J^j_P : V^j_{P_{W_1}} \to V^j_{P_{W_2}} (j \in N)$.
4. $J =< J_c, J^i_f(i \in N), J^j_P(j \in N) >$: Interpretation of vocabularies between W_1 and W_2.

2.2.2 Interpretation of Formulae

Let W_1 and W_2 be worlds; K, the interpretation of formulae between W_1 and W_2, is then defined by the following functions, using J, the interpretation of vocabularies:

1. Interpretation of Terms
   (a) If x is a variable, then $K(x) = x$.
   (b) If a term t is a constant, then $K(t) = J_c(t)$.
   (c) If $t_1, \ldots, t_n$ are terms and f is a n-ary function symbol, then
   $K(f(t_1, \ldots, t_n)) = J_f(f)(K(t_1), \ldots, K(t_n))$.
2. Interpretation of Atomic Formulae
   (a) If t_1 and t_2 are terms and $t_1 = t_2$, then
   $K([t_1 = t_2]) = [K(t_1) = K(t_2)]$.
   (b) If $t_1, \ldots, t_m$ are terms and P is a m-ary predicate symbol, then
   $K(P(t_1, \ldots, t_m)) = J_P(P)(K(t_1), \ldots, K(t_m))$.
3. Interpretation of Formulae
   (a) If ϕ and ψ are formulae, then
   $K(\phi \wedge \psi) = K(\phi) \wedge K(\psi)$,
   $K(\phi \vee \psi) = K(\phi) \vee K(\psi)$, $K(\neg\phi) = \neg K(\phi)$,
   and $K(\phi \to \psi) = K(\phi) \to K(\psi)$.
   (b) If x is a variable and ϕ is a formula, then
   $K(\forall x\phi) = \forall x K(\phi)$ and $K(\exists x\phi) = \exists x K(\phi)$.

We denote the interpretation between worlds by the symbol I, consisting of the above J and K.

2.3 Conditions on Interpretations

We now specify some conditions for the interpretation between worlds.

2.3.1 Strong Interpretation

Let W_1 and W_2 be worlds, and let I be an interpretation between W_1 and W_2. If any axioms of W_1 are interpreted in W_2 as theorems by I (i.e., if $\forall\phi \in A_{W_1}$, then $A_{W_2} \vdash I(\phi)$), then interpretation I is called a *strong interpretation*.

If I is a strong interpretation, then the following theorem holds.

Theorem 1 *Let W_1 and W_2 be worlds, and let I be a strong interpretation from W_1 to W_2. If ϕ is a theorem in W_1 (i.e., $A_{W_1} \vdash \phi$, then $I(\phi)$, the interpretation of ϕ, is also a theorem in W_2 (i.e., $A_{W_2} \vdash I(\phi)$).*

We give a proof of this theorem is given in another paper [Nagao, 1988].

2.3.2 Weak Interpretation

Strong interpretation is inadequate for dealing with semantic interpretation of natural language. Therefore, we introduce another interpretation.

It seems rarely to happen that all the axioms of one world are interpreted in another world as theorems. However, there are many cases where reasoning in one world needs the axioms in another world. We therefore define *weak* interpretation as considering interpreted axioms as axioms in the target world if they are consistent with the original axioms in the target world.

A more precise description is as follows:

Let W_1 and W_2 be worlds, and let I be interpretation between W_1 and W_2. If the interpretation results of all axioms of W_1 are consistent in W_2 (i.e., if $\forall\phi \in A_{W_1}$, then $A_{W_2} \not\vdash \neg I(\phi)$), then interpretation I is called a *weak interpretation*. Weakly interpreted axioms can be regarded as axioms of W_2 (i.e., $I(\phi) \in A_{W_2}$).

If I is a weak interpretation, then the following theorem holds.

Theorem 2 *Let W_1 and W_2 be worlds, and let I be a weak interpretation from W_1 to W_2. If ϕ is a theorem in W_1 (i.e., $A_{W_1} \vdash \phi$), then $I(\phi)$, the interpretation of ϕ, is also a theorem in W_2 (i.e., $A_{W_2} \vdash I(\phi)$).*

Proof of this theorem is also given in another paper [Nagao, 1988].

2.4 Building a Kripke Model with the Multi-World Model

We now show that a Kripke Model [Landman, 1986] can be constructed by using the Multi-World Model.

2.4.1 Henkin-Extension of Worlds

We introduce into the first-order language a ***witnessing constant*** [Barwise ***et al.***, 1977], representing an existentially closed sentence such that $\exists x P(x)$ (P is a predicate symbol). It is denoted by $c_{P(x)}$.

The Henkin-Extension is an extension of first-order logic that includes witnessing constants of level n. They are defined recursively:

Let L_W be a language of world W.

1. Level 0 : a sentence $\exists x P(x)$ includes only constants of L_W.
2. Level n > 0 : a sentence $\exists x P(x)$ includes at least one witnessing constant of level n-1.

The Henkin-Extension of world W is the deductively closed world W_{ex}, which has witnessing constants of each level in addition to L_W and axioms $\exists x P(x) \rightarrow P(c_{P(x)})$ (called ***Henkin axioms***), corresponding to each witnessing constant, in addition to A_W.

2.4.2 Interpretation between Henkin-Extended Worlds

Let W_1 and W_2 be worlds, let W_{ex_1} and W_{ex_2} be their Henkin-Extensions, and let I be an interpretation between W_1 and W_2. An interpretation between W_{ex_1} and W_{ex_2}, I^* is then defined as follows:

1. For every constant c, $I^*(c) = I(c)$.
2. For every witnessing constant $c_{P(x)}$, $I^*(c_{P(x)}) = c_{I(P(x))}$.
3. For every function and predicate symbol, I^* is same as I.

If I is a strong/weak interpretation, then I^* is also a strong/weak interpretation.

2.4.3 Canonical Structure

The canonical structure of W_{ex}, $Cano(W_{ex})$, consists of:

1. Canonical base: A quotient set constructed by an equivalence relation $\cong$ such that $\forall t_1, t_2 \in CT_{ex}$, $t_1 \cong t_2 \rightarrow A_{W_{ex}} \vdash [t_1 = t_2]$, where CT_{ex} is a set of closed terms of W_{ex}. It is denoted by $Cbase(W_{ex})$.
2. Canonical relation: A relation on the canonical base constructed by functions and predicates of $L_{W_{ex}}$.

2.4.4 Possible Worlds

A set of possible worlds corresponding to Henkin-Extended world W_{ex} is defined as follows:

$S_{W_{ex}} = \cup^1\{W'_{ex} \mid I^* : W_{ex} \rightarrow W'_{ex}\}$.

Let s be a possible world corresponding to W_{ex} (i.e., $s \in S_{W_{ex}}$). We denote the canonical structure of s as $Cano(s)$, the canonical base of s as $Cbase(s)$ and the world which is extended to s as $W(s)$.

[1] This means that worlds into which functional compositions of intrepretations are mapped are also included.

2.4.5 Kripke Model Constructed by the Multi-World Model

The Kripke Model which is constructed by the Multi-World Model $< WS, IS >$ and the world $W (W \in WS)$ is a 5-tuple $< S, \leq, D, q, V >$, where:

- S is a non-empty set such that $S = S_{W_{ex}} = \cup\{W'_{ex} \mid I^* : W_{ex} \rightarrow W'_{ex}\}$.
- $\leq$ is a partial order on S such that $\forall s, s' \in S, s \leq s' \equiv \exists I \in IS, I : W(s) \rightarrow W(s')$, I is a strong/weak interpretation.
- D is a set of domains of variables such that $D = \cup\{Cbase(W'_{ex}) \mid I^* : W_{ex} \rightarrow W'_{ex}\}$.
- q is a function that selects from D a domain in each possible world such that $\forall s \in S, q(s) = Cbase(s)$, and if $s_1 \leq s_2$ then $I^*(q(s_1)) \subseteq q(s_2)$.
- V is a function assigning V_s to every possible world s. V_s is a partial function that maps from the set of formulae of L_W to the set of truth values, such that

  1. $V_s(\phi) = 1$ iff $Cano(s) \models I(\phi)$,
  2. $V_s(\phi) = 0$ iff $Cano(s) \models I(\neg\phi)$,
  3. Undefined otherwise,

  where ϕ is an atomic formula of L_W.

  Truth value assignments to arbitrary formulae are the same as for the Kripke Model.

2.4.G Modalities in the Multi-World Model

Modalities in a Kripke Model are defined in the Multi-World Model by introducing modal operators to the first-order language of worlds.

Let $M =< S, \leq, D, q, V >$ be a Kripke Model constructed by the Multi-World Model $< WS, IS >$ and the world $W (W \in WS)$, and let ϕ be a formula of W.

1. Necessity
   (a) $\Box\phi$ is true relative to s iff $\forall s'$ such that $s \leq s', V_{s'}(\phi) = 1$,
   (b) $\Box\phi$ is false relative to s iff $\exists s'$ such that $s \leq s', V_{s'}(\phi) = 0$.
2. Possibility
   (a) $\Diamond\phi$ is true relative to s iff $\exists s'$ such that $s \leq s', V_{s'}(\phi) = 1$,
   (b) $\Diamond\phi$ is false relative to s iff $\forall s'$ such that $s \leq s', V_{s'}(\phi) = 0$.

In any world W' that connected with W by a strong/weak interpretation I, $I(\phi)$ is a theorem (i.e., $\forall W', \exists I : W \rightarrow W', I$ is a strong/weak interpretation, $A_{W'} \vdash I(\phi)$). Therefore, a theorem ϕ of a world W (i.e., $A_W \vdash \phi$) satisfies the above definition of ***necessity***.

On the other hand, the axioms satisfying the condition of ***possibility*** play important roles in the Multi-World Model. These axioms may be inconsistent with a world, but consistent with another. We call such axioms ***possible-modal axioms***.

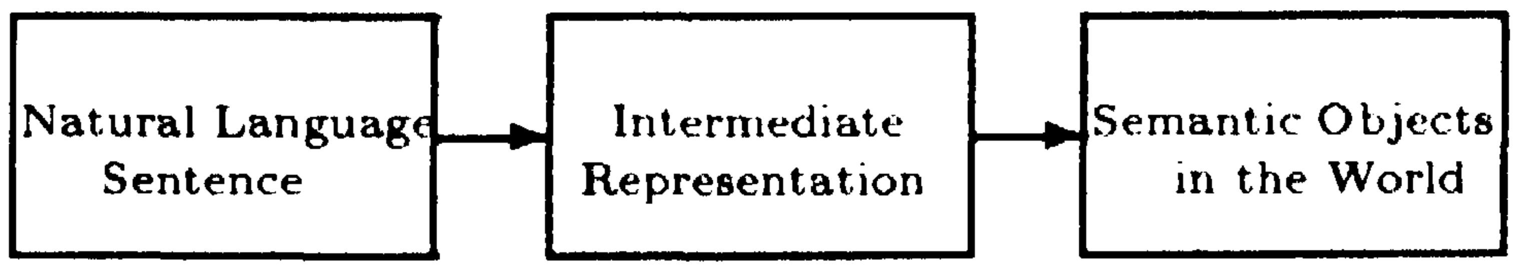

Figure 1: Semantic Interpretation Process

3 Semantic Interpretation

In general, the system for semantic interpretation of natural language sentences consists of the following process:

1. Translate the input sentence into the *intermediate representation.*

   The *intermediate representation* is a well-formed construction with no ambiguity (e.g., a first-order logical formula or a formula of *intensional logic* [Dowty *et al.*, 198]]).

2. Relate the *intermediate representation* to its denotation, the *semantic objects* in the world involved in the system (e.g., database entries).

The process mentioned above is illustrated in Figure 1.

However, an interpretation will generally be changed according to the context,. When a person tries to interpret a sentence, he constructs semantic objects based on background knowledge or knowledge obtained from preceding sentences. Rut if he fails to construct them, then he makes some assumptions and constructs semantic objects based on these assumptions. When he tries to interpret the next sentence and notices that a formerly constructed assumption is not. correct, or that, another assumption is better, he replaces the old assumption with a new one.

Thus, the system for semantic interpretation should be flexible enough to allow changes of interpretation.

3.1 Process of Semantic Interpretation

In our method, semantic interpretation is done according to the following process:

1. The sentence is translated into a logical formula. If it is translated into a higher-order logical formula, then it is reconstructed into first-order formulae and meta-operators. (E.g., "John knows that Mary is young" will be translated as $Know_{john}Young(Mary)$. Here, $Know_{john}$ is a meta-operator. The meaning of $Krow_{john}$ is defined in the Multi-World Model.)

   In this paper, we consider only first-order formulae.

2. One unambiguous sentence is translated into one logical formula, and is considered as a fact.

3. If a sentence is ambiguous and can be translated into several logical formulae, then these are considered as assumptions.

4. A fact, is regarded as an axiom of the confronted world, and each assumption is regarded as an axiom of the abductive world separately.

5. The confronted world is connected with the abductive worlds by weak interpretations.

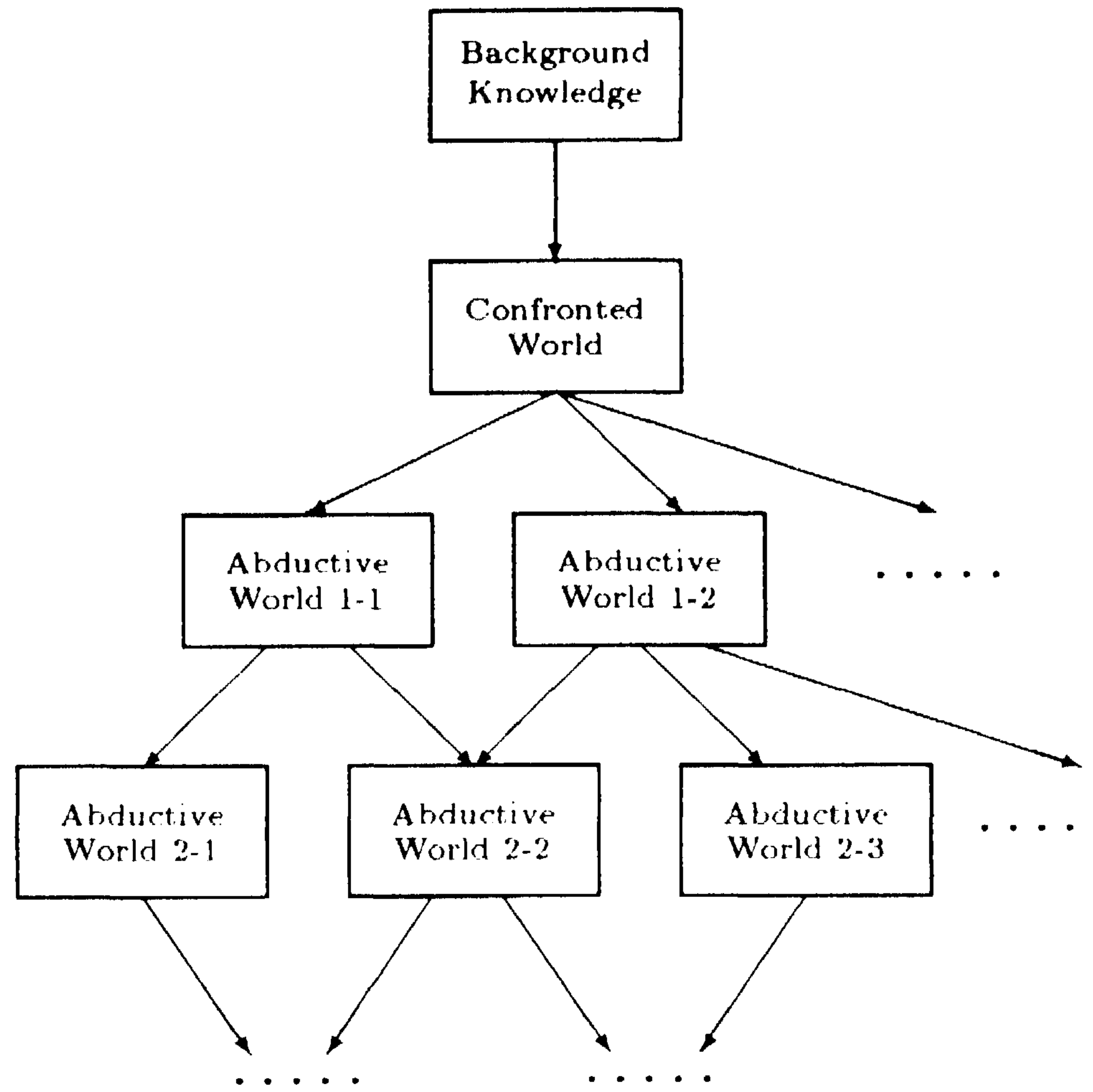

Figure 2: World Hierarchy for Semantic Interpretation

G. An interpretation from the confronted world to the abductive world determines the semantic interpretation of a sentence.

7. The abductive worlds construct a hierarchy if the following sentences produce other assumptions. The abductive worlds constructed bv one sentence are on the same level in the hierarchy of worlds.

   The hierarchy of worlds for semantic interpretation is represented in Figure 2.

   All arrows in this figure represent, weak interpretations.

8. As new sentences are processed by the system, the weakly interpreted axioms may be found to be inconsistent with the axioms of an abductive world. In such cases, the weak interpretation into this abductive world is eliminated.

9. When the context is fixed and there exist several abductive worlds that are consistent with the interpreted axioms of the confronted world, then we must select, the most, plausible abductive world. There is a heuristic rule for defining a preference relation between abductive worlds:

   "The abductive world with the latest and most information is the most preferable."

   The preferability of abductive worlds depends on the amount, of possible-modal axioms in the background knowledge that, are consistent with assumptions in these worlds and the recency of facts in the confronted world that are also consistent with assumptions.[2]

[2] This is an intuitive definition of preference.

3.2 Sentence Understanding Using Assumptions

As mentioned before, when people understand a sentence ("understanding a sentence" means assigning a proper interpretation to that sentence), they do so by using their background knowledge, and if their background knowledge is not sufficient for the purpose, then they build suitable assumptions in order to complement their understanding.

For example, take the ambiguous sentence

"John just saw a man with a telescope."

There arc two interpretations of "with a telescope":

1. *John saw a man by using a telescope.*

   This interpretation is based on the assumption that "."John" had a telescope and that he used it as an instrument to see "a man."

2. *John saw a man who had a telescope.*

   This interpretation is based on the assumption that "a man" had a telescope.

Suppose that the next sentence is:

".John bought it yesterday."

In this case, the first interpretation seems to be better than the second one.

However, when the third sentence is:

"Hut John gave it to him this morning,"

the second interpretation seems to be more adequate than the first.

In general, we cannot decide the complete interpretation of sentences until the whole context is fixed.

We now explain the process of making facts and assumptions for interpreting each sentence.

- Sentence 1. "John just saw a man with a telescope."

  This sentence is translated into the following two logical formulae:

  $F_{1.1}$: $\exists t_1, e_1, e_2, x, y\{Just(t_1) \wedge See(e_1, t_1) \wedge agent(e_1) = John \wedge patient(e_1) = x \wedge Man(x) \wedge instrument(e_1) = y \wedge Telescope(y) \wedge Have(e_2, t_1) \wedge agent(e_2) = John \wedge patient(e_2) = y\}$.

  Here, t_1 is a time variable, $e_i (i \in \{1, 2\})$ is an event variable, and $John$ is a constant corresponding to the individual "John."

  $F_{1.2}$: $\exists t_1, e_1, e_2, x, y\{Just(t_1) \wedge See(e_1, t_1) \wedge agent(e_1) = John \wedge patient(e_1) = x \wedge Man(x) \wedge Have(e_2, t_1) \wedge agent(e_2) = x \wedge patient(e_2) = y \wedge Telescope(y)\}$.

  1 he first formula includes the result of inference using the background knowledge: "if a person uses a telescope to see something, then he/she has the telescope."

  The two formulae are inconsistent, because the event *Have* of each formula is derived from "with a telescope," the event variable e_2 of each formula is equivalent, and *agent* is a function.

  Thus, since the two formulae cannot be placed in the same world, they are considered as assumptions and are placed in separate abductive worlds.

  The constant *John* and the witnessing constants $c_{Man(x)}, c_{Telescope(y)}$ are placed in the confronted

- Sentence 2. "John bought it yesterday."

  This sentence is translated into the following logical formula:

  F_2 : $\exists t_2, e_3, e_4\{Yesterday(t_2) \wedge Buy(e_3, t_2) \wedge agent(e_3) = John \wedge patient(e_3) = it_1 \wedge Have(e_4, t_2) \wedge agent(e_4) = John \wedge patient(e_4) = it_1\}$.

  This formula also includes the result of inference using the background knowledge that "if a person bought something, then he/she has it."

  This formula is placed in the confronted world.

  it_1 is a constant corresponding to the pronoun "it." The pronoun resolution procedure described in the next section returns the formula $it_1 = c_{Telescope(y)}$, and is placed in the confronted world.

  Consider the following possible-modal axiom in the background knowledge:

  K_1 : $\Diamond \forall t, e, x, y[\{Have(e, t) \wedge agent(e) = x \wedge patient(e) = y\} \rightarrow \forall t', e'\{After(t, t') \wedge Have(e', t') \wedge agent(e') = x \wedge patient(e') = y\}]$.

  The axiom means "it is possible that if a person has something at a moment then he/she has it after the moment."

  The result of inference by using F_2 and K_1 is consistent with $F_{1.1}$, but not with $F_{1.2}$. This makes the abductive world including $F_{1.1}$ more plausible than the one including $F_{1.2}$.

- Sentence 3. "But John gave it to him this morning."

  This sentence is translated into the following logical formula:

  F_3 : $\exists t_3, e_5, e_6\{Morning\_of\_Today(t_3) \wedge Give(e_5, t_3) \wedge agent(e_5) = John \wedge patient(e_5) = it_2 \wedge recipient(e_5) = he_1 \wedge Have(e_6, t_3) \wedge agent(e_6) = he_1 \wedge patient(e_6) = it_2\}$.

  This formula also includes the result of inference using the background knowledge that "if a person gave something to another person, then the person given it has it."

  This formula is placed in the confronted world. it_2 is a constant corresponding to the pronoun "it" and he_1 is a constant corresponding to the pronoun "him." The pronoun resolution procedure returns the formula $it_2 = c_{Telescope(y)} \wedge he_1 = c_{Man(x)}$, which is placed in the confronted world.

  The result of inference using F_3 and K_1 is consistent with $F_{1.2}$, but not with $F_{1.1}$. In addition, F_3 comes after F_2. Therefore the abductive world including $F_{1.2}$ is more plausible than the one including $F_{1.1}$.

We consider the interpretation of an ambiguous sentence as a problem of making assumptions to complement partial information and of determining which assumption is supported at the moment, when the context seems to be fixed. Dynamic interpretation involves the problem that a supported assumption will change according to the change in the context.

The M ulti-World Model is intended to provide a representational basis for these problems. But a more complicated problem to formalize the preference relation of worlds should be contemplated in order to sophisticate the model.

3.3 Anaphora Resolution Using the Multi-World Model

There are many discourse-oriented phenomena in natural language understanding, such as *focus, topic, coherence, cohesion* and so on.

One fundamental phenomenon that can be observed in many natural languages is called *anaphoric reference* or simply *anaphora.*

Pronouns are the typical, but not the only, markers for anaphoric reference. Once a term is introduced, it can be referred to in various ways, for example, by similarity (i.e., using the same word, a synonym, or a closely related word), by a definite noun phrase or paraphrase reference, and by zero anaphora (i.e., ellipsis).

According to the traditional approach, the anaphora resolution algorithm consists of the following steps: [Maruyama, to appear]

1. Select a definite expression from the current sentence. If all the definite expressions have been resolved, then terminate.
2. From the discourse structure, find the referent that satisfies certain syntactic and semantic constraints.

We assume that the semantic constraints are included in the background knowledge.

Here is an example of resolution of the pronoun "it." This example is a modified version of that in a forthcoming paper [Zadrozny and Jensen, to appear].

- *Sentences*

  Consider the following sentences:

  - S_1 : *A ship entered a port.*
  - S_2 : *It brought a disease.*
  - S_3 : *It struck rapidly.*

- *Logical Formulae*

  The above sentences are translated into the following formulae:

  - F_1 : $\exists t_1, e_1, x, y\{Past(t_1) \wedge Enter(e_1, t_1) \wedge agent(e_1) = x \wedge Ship(x) \wedge patient(e_1) = y \wedge Port(y)\}$.

    This formula is placed in the confronted world, along with the witnessing constants $c_{Ship(x)}, c_{Port(y)}$.
  - F_2 : $\exists t_2, e_2, z\{Past(t_2) \wedge Bring(e_2, t_2) \wedge agent(e_2) = it_1 \wedge patient(e_2) = z \wedge Disease(z)\}$.

    This formula is placed in the confronted world, along with the witnessing constant $c_{Disease(z)}$. The following equalities are produced for the pronoun resolution:

    * $E_{2.1}$: $it_1 = c_{Ship(x)}$.
    * $E_{2.2}$: $it_1 = c_{Port(y)}$.

    These two formulae are inconsistent with each other, since there is no entity that can be a ship and a port simultaneously. Consequently, these formulae are placed in the abductive worlds separately.
  - F_3 : $\exists t_3, e_3, x\{Past(t_3) \wedge Strike(e_3, t_3) \wedge agent(e_3) = it_2 \wedge patient(e_3) = x \wedge Rapidly(e_3)\}$.

    This formula is placed in the confronted world, and the following equalities are produced for the pronoun resolution:

    * $E_{3.1}$: $it_2 = c_{Ship(x)}$.
    * $E_{3.2}$: $it_2 = c_{Port(y)}$.
    * $E_{3.3}$: $it_2 = c_{Disease(z)}$.

    Since each formula is inconsistent with the others, these formulae are placed in different abductive worlds, which then construct a hierarchical structure.

- *Background Knowledge*

  The background knowledge includes the following formulae:

  - $Enter_1$: $\Diamond\forall t, e, x, y[\{Enter(e, t) \wedge agent(e) = x \wedge patient(e) = y\} \rightarrow \forall e'\{Come\_in(e', t) \wedge agent(e') = x \wedge patient(e') = y\}]$.
  - $Enter_2$: $\Diamond\forall t, e, x, y[\{Enter(e, t) \wedge agent(e) = x \wedge patient(e) = y\} \rightarrow \forall e'\{People(x) \wedge Join(e', t) \wedge agent(e') = x \wedge patient(e') = y \wedge Group(y)\}]$.
  - $Ship_1$: $\Diamond\forall x[Ship(x) \rightarrow \forall t, e, y\{Carry(e, t) \wedge agent(e) = x \wedge patient(e) = y \wedge (People(y) \vee Goods(y))\}]$.
  - $Ship_2$: $\Diamond\forall x[Ship(x) \rightarrow \{Aircraft(x) \vee Space\_vehicle(x)\}]$.

    $Port$: $\forall x\{Port(x) \rightarrow Harbour(x)\}$.
  - $Bring_1$: $\Diamond\forall t, e, x, y[\{Bring(e, t) \wedge agent(e) = x \wedge patient(e) = y\} \rightarrow \forall e'\{Carry(e', t) \wedge agent(e') = x \wedge patient(e') = y\}]$.
  - $Bring_2$: $\Diamond\forall t, e, x, y[\{Bring(e, t) \wedge agent(e) = x \wedge patient(e) = y\} \rightarrow \forall e'\{Cause(e', t) \wedge agent(e') = x \wedge patient(e') = y\}]$.
  - $Disease$: $\forall x\{Disease(x) \rightarrow Illness(x)\}$.

    $Strike_1$: $\Diamond\forall t, e, x, y[\{Strike(e, t) \wedge agent(e) = x \wedge patient(e) = y\} \rightarrow \forall e'\{Hit(e', t) \wedge agent(e') = x \wedge patient(e') = y\}]$.

    $Strike_2$: $\Diamond\forall t, e, x, y[\{Strike(e, t) \wedge agent(e) = x \wedge patient(e) = y\} \rightarrow \forall e'\{Harm(e', t) \wedge agent(e') = x \wedge Illness(x) \wedge patient(e') = y \wedge Suddenly(e')\}]$.

  The formula $E_{2.1}$ can be derived from the possible-modal axioms $Ship_1$ and $Bring_1$ in the background knowledge. But $E_{2.2}$ cannot be derived from any axioms, and thus $E_{2.1}$ is preferable to $E_{2.2}$.

  The formula $E_{3.3}$ can be derived from the axiom $Disease$ and the possible-modal axiom $Strike_2$ in the background knowledge, but the other formulae $E_{3.1}$ and $E_{3.2}$ cannot, thus $E_{3.3}$ is the most preferable.

We assume that the semantic interpretation needs background knowledge and that the assumptions are made in the process of reading the sentences.

The example of anaphora (pronoun) resolution mentioned above is evidence that the Multi-World Model is capable of handling discourse-oriented problems.

4 Relation to Other Work

Our approach is closely related to those that consider interpretation of a sentence as a problem of abduction [Ilobbs *ei al,* 1988] and decide the reference of anaphoric expressions by using assumptions that are consistent with the context [Charniak, 1988].

We intend to introduce a logical framework for interpretation of sentences, but we have not yet discussed the inference mechanism precisely, as the above works do.

Their ideas will be useful in enabling us to introduce a more formalized concept of preferential relationships between assumptions into our model. We will also study an efficient algorithm for abductive inference in order to implement our model.

Another logical framework for natural language understanding is proposed in a forthcoming paper [Zadrozny and Jensen, to appear]. This is aimed at building a computational model of a paragraph that represents interaction between sentences, the background knowledge to which these sentences refer, and metatheoretical operators that indicate the types of models permitted. The model is based on *-partial models* of logical formulae [Zadrozny, 1987].

The special emphasis of our model is its capability of dealing with dynamic interpretation, while Zadrozny's model is mainly concerned with topic and coherence of context.

5 Concluding Remarks

Logic plays an important role in constructing a theoretical basis for the semantics of natural language. However, we should not use higher-order logic, since no theorem provers of higher-order logic work efficiently. On the other hand, first-order logic is not sufficient for semantic interpretation of natural language sentences, since ambiguous sentences have several interpretations that may contradict each other, and contradictions cannot be handled properly in first-order logic. We assumed that a semantic interpretation is constructed by building assumptions and selecting the most likely assumption. When people interpret ambiguous sentences, they make several possible interpretations, but cannot always choose the correct one because their background knowledge is limited. In addition, we must take account of the fact that an interpretation changes because of incoming data. The Multi-World Model has the functions of a logical system for a formal treatment of natural language, and a flexible mechanism for dynamic interpretation.

We are planning to extend the model to describe preferential relationships between worlds more formally. We will also be investigating an efficient inference algorithm in order to implement a system based on our formalization.

Acknowledgements

We would like to thank members of the IBM Tokyo Research Laboratory, Wlodek Zadrozny of the IBM Thomas J. Watson Research Center, and the reviewers for their valuable comments on a draft of this paper, and Michael McDonald for his helpful advice on the wording.

References

[Barwise *et al,* 1977] Jon Barwise (editor), *Handbook of Mathematical Logic.* North-Holland, 1977.

[Charniak, 1988] Eugene Charniak, "Motivation Analysis, Abductive Unification and Nonmonotonic Equality," *Artificial Intelligence,* 34:275-295, 1988.

[Deguchi and Oka, to appear] Hiroshi Deguchi and Toshiyuki Oka, "Hierarchical Reasoning and Knowledge Networks" (in Japanese), to appear.

[Dowty *et ai,* 1981] David R. Dowty, Robert E. Wall, and Stanley Peters, *Introduction to Montague Semantics.* D. Reidel, Dordrecht, 1981.

[Hobbs *et ai,* 1988] Jerry R. Hobbs, Mark Stickel, Paul Martin, and Douglas Edwards, "Interpretation as Abduction," *Proceedings of the 26th Annual Meeting of the Association for Computational Linguistics,* 95-103, 1988.

[Landman, 1986] Fred Landman, *Towards a Theory of Information.* Foris, Dordrecht, 1986.

[Maruyama, to appear] Hiroshi Maruyama, "Maximal Join for Anaphora Resolution," in John Sowa, Norman Foo, and Anand Rao (editors), *Conceptual Graphs for Knowledge Systems,* Addison-Wesley, to appear.

[Nagao, 1988] Katashi Nagao, "Semantic Interpretation Based on the Multi-World Model," *TRL Research Report* TR87-1025, 1988.

[Reiter, 1980] Raymond Reiter, "A Logic for Default Reasoning," *Artificial Intelligence,* 13:81-132, 1980.

[Zadrozny, 1987] Wlodek Zadrozny, "A Theory of Default Reasoning," *Proceedings of the Sixth National Conference on Artificial Intelligence (AAAJ-87),* 385 390, 1987.

[Zadrozny and Jensen, to appear] Wlodek Zadrozny and Karen Jensen, "Semantics of Paragraphs," to appear.