

多重世界モデルを用いた対話における信念の翻意

長尾 確

(株) ソニーコンピュータサイエンス研究所

概要

対話において、我々はお互いに相互の信念があることを前提としているが、実際は、それは各個人にとっての仮説に過ぎず、対話の進行に伴ってその確信度(優先度)が変化していく。筆者は以前に、多義性を含む文の解釈に対するアブダクションに基づいたアプローチとして多重世界モデルを提案している。本論文では、このモデルを、対話において相互の信念が仮説として形成され、その優先度が動的に変化していくプロセスを説明するのに用いる。このプロセスをここでは信念の翻意と考える。信念の翻意は、これまでTMSのように真理値の反転、あるいはデフォルト間の階層性によって扱われてきたが、対話によって動的にもっともらしさが変化する信念の特性を表すのに適切でなかった。そこで、仮説としての信念に、文脈に依存した数的な優先度を考えることによって、信念の翻意に新しい視点を与える。

Belief Revision in a Dialogue Using the Multi-World Model

Katashi Nagao

Sony Computer Science Laboratory Inc.

3-14-13 Higashi-gotanda, Shinagawa-ku,

Tokyo, 141 Japan

E-mail: nagao@csl.sony.co.jp

Abstract

In natural language dialogue, people assume that each dialogue participant has mutual beliefs with another participant. In fact, they are just assumptions of each one and their plausibility will dynamically change according to a progress of dialogue. The author formerly proposed a model for interpreting ambiguous sentences by dynamic change of preferences among assumptions. In this paper, we applied the model to belief revision in a dialogue with a similar method that describes change of preferences among beliefs. Then, we can treat belief revision as dynamic assignments of preferences among beliefs constructed in a dialogue not as conventional static assignments of truth values or preferences.

1 はじめに

対話において、話者と聴者は、それぞれ相手との間に共有の信念や認識が存在することを仮定している。しかし、それはあくまでも仮定に過ぎず、自分の信念であり、かつ相手の信念であると思っている相互の信念は、「仮説」として扱われる。多くの場合、相手の発話に対して考えられる仮説は複数存在する。その顕著な例は、指示対象に関するものである。例えば、「○ ○さんは、～だそうですね。」と言われたとき、「○ ○さん」に相当する人物に関する信念に対して、複数の仮説が可能となる場合がある。この場合、可能な仮説を保持しておいて、対話の進行に伴って、各々の仮説の確信度(あるいは優先度)を動的に変化させることができる。先の例では、「○ ○さん」のその他の情報が与えられるにつれて特定の仮説の優先度が高くなり、それに伴ってその他の仮説の優先度が低くなる。そのようにして構成された優先度付きの仮説の集合が、相手の信念であると思っている自分の信念、すなわち相互の信念に相当するわけである。つまり、相手の信念のモデルを自分の信念に投影して対話を理解するのではなく、相互の信念だと思っている自分の信念を仮説の集合として保持して、その優先度を動的に変化させることによって対話を理解するのである。

我々は、以前に多義性を含む文の解釈において、複数の可能な解釈に相当する論理式を含む世界を構成し、背景知識や事実を含む別の世界と世界間の翻訳(以前は、世界間の解釈と呼んでいた)という関数で連結した多重世界を用いてモデル化を行った [7]。それは、人間は、多義性の含まれる文を解釈するとき、不完全な背景知識しか持たないがゆえに、とりあえず仮の解釈を与えておき、後になって文脈が明らかになったときに、以前に与えた解釈を修正するか、より適切な解釈を与えることができる、という解釈の動的な性質を説明するためであった。

本論文で議論する対話における信念の翻意も、基本的には同じ枠組みで説明できる。その

ため、多重世界モデルを再び導入する。次の章では、多重世界モデルの定義について簡単に述べ、それに続く章で仮説間の優先性に関する議論を行ない、そのあとで、多重世界モデルを用いた対話における信念の翻意について述べる。最後に、発話行為の扱いに関する議論を行なう。

2 多重世界モデル

多重世界モデルは次のように定義される。多重世界モデルは、世界の集合 W と世界間の翻訳の集合 T の 2 項組 $\langle W, T \rangle$ で表わされる。

2.1 世界

一階の述語論理の公理系を一つの世界とみなす。すなわち、語彙 V_w と公理の集合 A_w から構成される。

定義 1. 世界

世界 w は次のような一階の言語の語彙 V_w と公理の集合 A_w の 2 項組 $w = \langle V_w, A_w \rangle$ として定義される。

1. 語彙 V_w

- (a) V_{c_w} : 世界 w の定数記号の集合
- (b) $V_{f_w}^n$: 世界 w の n 項関数記号の集合
- (c) $V_{P_w}^m$: 世界 w の m 項述語記号の集合
- (d) $V_{f_w} = \cup \{V_{f_w}^i \mid i \in N\}$
 V_{f_w} には、合成によって得られる関数も含まれる。
- (e) $V_{P_w} = \{V_{P_w}^j \mid j \in N\}$
- (f) $V_w = V_{c_w} \cup V_{f_w} \cup V_{P_w}$: 世界 w の語彙

2. 公理の集合 A_w

A_w : 世界 w の語彙、変数、論理記号で構成される公理の集合

2.2 世界間の翻訳

ある世界の語彙から別の世界の語彙への関数 vt を次のように定義する。

定義2. 世界間の語彙の翻訳

w_1 と w_2 を世界とする。このとき、 vt を次のような単射として定義される世界 w_1 から w_2 への語彙の翻訳と呼ぶ。

1. 定数記号に関して、 $vt_c : V_{c_{w_1}} \Rightarrow V_{c_{w_2}}$.
2. 関数記号に関して、 $vt_f^i : V_{f_{w_1}}^i \Rightarrow V_{f_{w_2}}^i$ ($i \in N$).
3. 述語記号に関して、 $vt_P^j : V_{P_{w_1}}^j \Rightarrow V_{P_{w_2}}^j$ ($j \in N$).
4. $vt = \langle vt_c, vt_f^i (i \in N), vt_P^j (j \in N) \rangle$: 世界 w_1 から世界 w_2 への語彙の翻訳

この関数 vt を使って、ある世界の論理式の集合から別の世界の論理式の集合への写像 ft を次のように定義する。

定義3. 世界間の論理式の翻訳

w_1 と w_2 を世界とする。このとき、 ft を語彙の翻訳 vt を用いて次のように帰納的に定義される、 w_1 から w_2 への論理式の翻訳と呼ぶ。

1. 項の翻訳

- (a) x が変数ならば、 $ft(x) = x$.
- (b) 項 tm が定数記号ならば、 $ft(tm) = vt_c(tm)$.
- (c) $tm_1, \dots, tm_n$ が項であり、 f が n 項関数記号ならば、 $ft(f(tm_1, \dots, tm_n)) = vt_f(f)(ft(tm_1), \dots, ft(tm_n))$.

2. 原子式の翻訳

- (a) tm_1 と tm_2 が項であり、 $tm_1 = tm_2$ ならば、 $ft([tm_1 = tm_2]) = [ft(tm_1) = ft(tm_2)]$.

- (b) $tm_1, \dots, tm_m$ が項であり、 P が m 項述語記号ならば、 $ft(P(tm_1, \dots, tm_m)) = vt_P(P)(ft(tm_1), \dots, ft(tm_m))$.

3. 任意の論理式の翻訳

- (a) ϕ と ψ が論理式ならば、 $ft(\phi \wedge \psi) = ft(\phi) \wedge ft(\psi)$,
 $ft(\phi \vee \psi) = ft(\phi) \vee ft(\psi)$,
 $ft(\neg \phi) = \neg ft(\phi)$,
 $ft(\phi \supset \psi) = ft(\phi) \supset ft(\psi)$.
- (b) x が変数で、 ϕ が論理式ならば、 $ft(\forall x. \phi) = \forall x. ft(\phi)$,
 $ft(\exists x. \phi) = \exists x. ft(\phi)$.

ここで、以上の vt と ft をまとめて、記号 t で表わし、それを世界間の翻訳と呼ぶことにする。

定義4. 弱い翻訳

w_1 と w_2 を世界とし、 t を w_1 から w_2 への翻訳とする。 $\phi_i (i \in \{1, \dots, n\})$ を w_1 の公理とする。 t が次の条件を満たすとき、弱い翻訳と呼ばれる。

1. $t(\phi_1)$ が世界 w_2 の公理 A_{w_2} と無矛盾である (すなわち、 $A_{w_2} \not\vdash \neg t(\phi_1)$ である)。
2. $t(\phi_2)$ が $A_{w_2} \cup \{t(\phi_1)\}$ と無矛盾である。
3. $t(\phi_i)$ が $A_{w_2} \cup \{t(\phi_1), t(\phi_2), \dots, t(\phi_{i-1})\}$ と無矛盾である。 ($i \in \{3, \dots, n\}$)

定義5. 可能様相公理

w_1 と w_2 を世界とし、 t を w_1 から w_2 への弱い翻訳とする。このとき、以下の条件を満たす公理を可能様相公理 (possible-modal axiom) と呼ぶ。

$\Diamond(\phi \supset \psi) \in w_1$ iff $\exists w_2 \in W, \exists t \in T$ such that $t : w_1 \Rightarrow w_2, A_{w_2} \vdash t(\phi \supset \psi)$

このとき、 $t(\phi \supset \psi)$ は、 w_1 から弱い翻訳 t で写像された世界 w_2 において公理となることが出来る。ここでは、矛盾がなければ積極的に公理として用いることにする。この意味で、可能様相公理は、デフォルト論理 [9] の (正規) デフォルト・ルールと同じように扱われる。

対話参加者の心的状態を表す多重世界モデルは、共有知識(可能様相公理として表される常識的知識を含む)を含む世界と個別知識を含む世界と信念を含む複数の仮説世界が弱い翻訳で連結された階層構造を成している(図1参照)。個別知識の世界は対話参加者(例えば、話者と聴者)ごとに作られる。話者の心的状態における聴者の個別知識は「話者が知らないことで、かつ聴者が知っている」と話者が信じていることであり、例えば、聴者の誕生日などが挙げられる(この場合、話者は聴者の誕生日がいつであるか知らない)。弱い翻訳によって、仮説世界は共有知識や個別知識を継承して推論に用いることができる。また、図1のグレーの仮説世界は、相互信念を公理として持つ世界である。つまり、相互信念は、双方の個別知識と矛盾しない仮説であると考えられる。

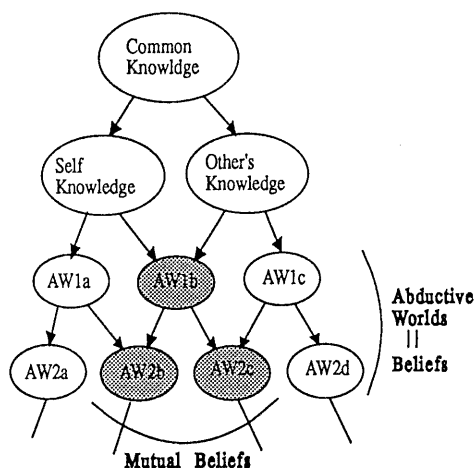


図1: 対話参加者の心的状態を表す多重世界

3 仮説間の優先性

多重世界モデルにおいて、特定の事象に関する信念(例えば、ある発話の解釈)は、同じ層の複数の仮説世界(例えば、図1の AW_{1a} , AW_{1b} , AW_{1c})における公理となる。同じ層の各々の仮説世界

の公理は同時には成り立たないものとする。このとき、複数の仮説世界間には優先性が考えられる。

このような優先性は文脈に応じて変化するが、それを完全に記号的に書き尽くすのは不可能である。一方、記号的な記述以外に、量的(数的)な記述も必要である、という指摘もある[3]。例えば、ある文の解釈の優先度[4]やアブダクションにおける仮説間の優先度[2]などにアナログ値を用いる例がある。

仮説世界の優先度は、その世界に弱い翻訳によって繋がっている上位の世界(上の層の知識あるいは仮説世界)との間の公理の依存関係や下位の仮説世界の優先度によって決まる。また、同じ層の異なる仮説世界の優先度が上がれば相対的にその世界の優先度が下がる。これは同じ層の仮説世界間の優先度の総和が一定であると考えればよい。

まず、世界間の公理の依存関係を次のように定義する。

定義6. 公理の依存関係

ある世界 w_1 と w_2 に関して、弱い翻訳 $t: w_1 \Rightarrow w_2$ が存在し、 w_1 の公理 ϕ と w_2 の公理 ψ の間で、次のような関係があるとき、 ϕ と ψ の間で依存関係があるという。

1. 引数に等価性がある。

$\phi = \dots \wedge P(a) \wedge \dots, \psi = \dots \wedge Q(b) \wedge \dots$ のとき、 $t(a) = b$ が成り立つ。

2. 述語に等価性がある。

$\phi = \dots \wedge P(a) \wedge \dots, \psi = \dots \wedge Q(b) \wedge \dots$ のとき、 $t(P) = Q$ が成り立つ。

3. 部分命題のどちらかがもう一方を含意している。

$\phi = \dots \wedge P(a) \wedge \dots, \psi = \dots \wedge Q(b) \wedge \dots$ のとき、 $t(P(a)) \supset Q(b)$ または $Q(b) \supset t(P(a))$ が成り立つ。

これらの依存関係には、依存の強さの度合いが考えられ、3の場合、2の場合、1の場合の順に高

い値となる。各々の値を d_i とする ($d_1 < d_2 < d_3$)。次の章では、多重世界モデルと仮説世界の優複数の場合があてはまるときは、各々の値を加 先度を用いた対話における信念の翻意について算した値となる。例えば、 $\phi = P_1(a_1) \wedge P_2(a_2)$, 述べる。

$\psi = Q_1(b_1) \wedge Q_2(b_2)$ のとき、 $t(a_1) = b_1, t(a_2) = b_2, t(P_1) = Q_1, t(P_2(a_2)) \supset Q_2(b_2)$ が成り立つならば、 ϕ と ψ の依存関係の度合は、 $2d_1 + d_2 + d_3$ となる。

この世界間の公理の依存関係の度合を用いて、仮説世界の優先度を次のように定義する。

定義 7. 仮説世界の優先度

仮説世界 w の優先度 $p(w)$ は次のように計算される。ここで、 i 番目の層の仮説世界の集合を W_i と書くことにする。このとき、0 番目の層は、個別知識の世界の層である。

1. $p(w) = 1$ iff $w \in W_0$
2. $p(w) = 0$
iff $w \in W_i \wedge \forall w' \in W_{i-1} \wedge \exists t : w' \Rightarrow w$
この場合 w は inconsistent な仮説世界である。
3. $p(w) = 0$
iff $w \in W_i \wedge W_{i+1} \neq \Phi \wedge \forall w' \in W_{i+1} \wedge \exists t : w \Rightarrow w'$
この場合 w は incoherent な仮説世界である。
4. $p(w) = \frac{d_w}{\sum_{w' \in W_i \setminus \{w\}} d_{w'}}$
iff $w \in W_i \wedge W_{i+1} = \Phi$
ここで、 d_w は w の上の層 ($i-1$ 層) の仮説世界または個別知識の世界との依存関係の度合の合計である。
5. $p(w) = \frac{p_w}{\sum_{w' \in W_{i+1} \setminus \{w\}} p_{w'}}$
iff $w \in W_i \wedge W_{i+1} \neq \Phi$
ここで、 p_w は w から弱い翻訳で繋がっている下の層 ($i+1$ 層) の仮説世界の優先度に w とその世界との依存関係の度合との積を合計したものである。
つまり、 $p_w = \sum_{w' \in W_{i+1}} d'_w \cdot p(w')$
such that $\exists t : w \Rightarrow w'$ である。

4 対話における信念の翻意

次のような S と H の対話を考えてみよう。

S: 「N さんは、最近結婚して名字が K になったんですよ。」

H: 「へえ、N さんは結婚したんですか。」

S: 「そう、彼はいま奥さんの名字を使ってるんですよ。」

H: 「えっ、そっちの N さんでしたか。」

このときの H の信念について、次のように考察する。

最初の S の発言に出てくる「N さん」が誰を指しているかに関して、H は、S の信念 (と 思っている H の信念) を仮説として生成する。このとき、複数の仮説が可能である場合がある。つまり、「N さん」が H の信念の中に存在する対象である n_1 と n_2 のどちらにも対応可能である場合である。このとき、最初の発言から $Married(n_1) \wedge ChangedName(n_1, k)$ と $Married(n_2) \wedge ChangedName(n_2, k)$ という 2 つの仮説が生成され、それぞれが異なる仮説世界 (それらを、 w_{1a}, w_{1b} とする) の公理となる (図 2 参照)。

このとき、共有知識の世界に、

$\Diamond \forall x, \exists y. Married(x) \wedge Female(x) \supset ChangedName(x, y)$ のような可能様相公理 (「女性が結婚すると名字が変わるだろう」という常識的知識) が含まれていて、また、H の個別知識として、 $Female(n_1)$ (つまり H は n_1 が女性であることを知っている) という知識があるとすると、可能様相公理はデフォルトとして用いられるから、H の個別知識の世界 w_{0a} と仮説世界 w_{1a} の間の公理間の依存度は、 $d_1 + d_3 (\exists x. Married(n_1) \wedge Female(n_1) \supset ChangedName(n_1, x))$ より) となる。ここで、世界間の語彙の翻訳は、同じものへの写像 (identification mapping) であるとしている。また、

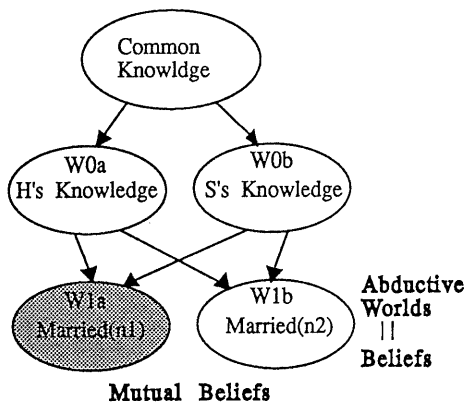


図 2: S の最初の発話後の H の心的状態を表す多重世界

同様に H の個別知識として、 $Male(n_2)$ があるとすると、H の個別知識の世界 w_{0a} と仮説世界 w_{1b} の間の公理間の依存度は、 d_1 となる。このとき、 w_{1b} より w_{1a} の優先度が高くなる。つまり、この時点では、H は「N さん」は n_1 であると思っている。

次の S の発話の後、H は S の発話の「彼」というキーワードから、 $He(n_1)$ と $He(n_2)$ という仮説を立て、それぞれを公理として含む仮説世界 w_{2a} と w_{2b} を作る (図 3 参照)。

このとき、 w_{1a} と w_{2b} は、同じ「N さん」に関する異なる指示対象を持つからコンフリクトすることになり、弱い翻訳が成り立たなくなる。同様に、 w_{1b} と w_{2a} の間の弱い翻訳も成立しない。また、共有知識の世界に、 $\Diamond \forall x. He(x) \supset Male(x)$ のような可能様相公理が含まれているとすると、 w_{0a} から弱い翻訳を通じて継承した $Male(n_2)$ と w_{2b} の公理 $He(n_2)$ の依存度は、 $d_1 + d_3$ であり、 w_{2a} の公理 $He(n_1)$ の依存度は d_1 であるから、 w_{2b} の優先度が高くなる。これらの、優先度は上の層の仮説世界 (w_{1a} と w_{1b}) に伝播される。このとき、明らかに仮説世界 w_{1b} の優先度が、 w_{1a} のそれより高くなる。つまり、この時点で、H は「N さん」が n_2 であると思うようになる。信念が翻意されたのである。

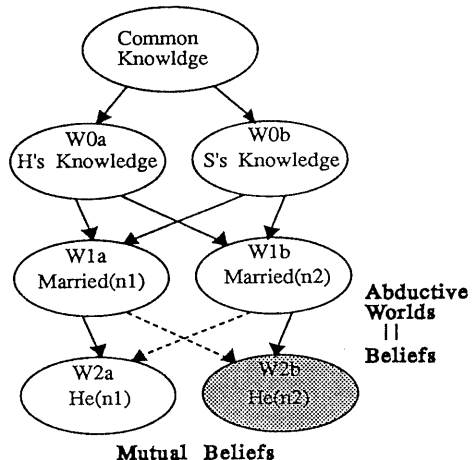


図 3: S の次の発話後の H の心的状態を表す多重世界

5 発話行為への応用

エージェントが発話行為を理解するモデルを作るためには、彼の信念や意図が、ある事象の結果としてどのように変化するかを定式化する必要がある。本論文での趣旨は、信念の翻意をどうモデル化するかということであるが、それは発話行為を扱う上で最も重要であると筆者は考えている。例えば、Perrault [8] は Reiter のデフォルト論理 [9] に基づいて発話行為の理論を構築した。しかし、発話行為の結果生じる信念の翻意をうまく扱えていない。信念の翻意を扱うモデルは、競合するデフォルト (仮説) 間の優先性を記述する必要がある。また、Appelt と Konolige [1] は、Moore の自己認識論理 [6] を拡張した階層的自己認識論理 [5] を用いて信念の翻意を扱っている。つまり、階層によってデフォルト (としての信念) 間の優先性を記述しているのである。しかし、対話の進行に従って、動的に優先性が変化する場合には対処できない。ここにおいて、我々の提案するモデルの有用性が主張できる。

また、もう一つ重要な問題として、相互信念の問題がある。相互信念とは、いったいどのよ

うに作られるのだろうか。我々の枠組みでは、相
互信念は一方の対話参加者の持つ仮説に過ぎな
い。それは、一方が、他方が知っているあるい
は信じていると思つてのことと矛盾せず、かつ
自分が知っていることあるいは信じていること
とも矛盾しない仮説の集合である。相互信念
とそうでない信念(一方だけの信念)との間には
(同じ層にあるものに関して)優先度が割り
当てられ、相互信念だと思つていたが実は違つ
ている(かも知れない)という状況を通常の信念
翻意と同様に扱える。

例えば、相手に嘘をつかれた場合のことを考
えてみよう。相手の言つたことから、いくつか
の仮説が立てられたとしよう(それを、ここで
は p_1, p_2, p_3 とする)。

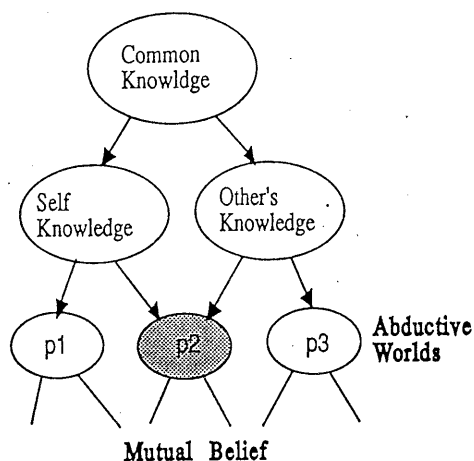


図 4: 心的状態を表す多重世界

このとき、自分の心的状態が図 4 のような多
重世界になっているとする。つまり、仮説 p_2 は
お互いの知識と矛盾しない信念(つまり相互信
念)であり、 p_1 は自分だけの信念(自分の知識
とは整合しているが、相手の知識と思つてい
ることとは矛盾してしまう信念)である。またこ
れとは逆に、 p_3 は相手だけの信念(自分の知識
とは矛盾しているため自分は信じていないが、
相手の知識と思つていこととは整合している

ので、相手は信じていると思つている信念)で
ある。

ここで、最初は p_2 を含む仮説世界の優先度
が、他に比べて高かったとする。つまり、相手
の言つたことを自分も信じているし、相手もそ
れを信じていると自分が信じているという可能
性が最も高いという状態である。しかし、その
後の発話などで、 p_1 を含む仮説世界の優先度
が変化して、他の仮説より優先度が高くなつた
とする。この場合、相手は知らないことを自分
に信じさせようとして言つた(可能性が高い)の
であるから、自分は嘘をつかれたと思つてしま
うのである。

6 おわりに

筆者の提案する多重世界モデルを対話におけ
る信念の翻意を説明する道具として応用した。
これによって、対話における信念は、階層的論
理システムにおいて、複数の仮説世界を形成し、
共有知識や個別知識の構成する世界との関係に
おいて、動的に優先度が変化していく、という
プロセスを記述できる。

この問題は一種のアブダクション [4, 2] であ
ると考えることもできる。つまり、相手の信念
(と思つていこと)は、相手の発話に対する説
明であり、それは相手の発話を真にするための
前提(あるいは、仮説)であると考えられるか
らである。

また、発話行為をより詳細に扱うためには意図
の考察が不可欠である。このため、さらに *Intend*
や *Goal* などの高階な述語を導入する必要がある。
高階な述語を扱う場合には、その意味論を
十分に限定して用いないと手に追えない問題に
なってしまう危険がある。今後さらに、発話行
為に関するより有用なモデルについて考察する
予定である。

References

- [1] 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 Nonmonotonic Reasoning*, 164–175, 1988.
- [2] Eugene Charniak and Solomon E. Shimony, "Probabilistic Semantics for Cost Based Abduction," In *Proceedings of the 8th National Conference on Artificial Intelligence*, 106–111, 1990.
- [3] Koiti Hasida, "Constraint with Potential Energy - A Theory of Situated Information," To appear in *Proceedings of the 3rd Conference on Situation Theory and its Applications*, 1991.
- [4] Jerry R. Hobbs, Mark Stickel, Paul Martin, and Douglas Edwards, "Interpretation as Abduction," In *Proceedings of the 26th Annual Meeting of the Association for Computational Linguistics*, 95–103, 1988.
- [5] Kurt Konolige, "Hierarchic Autoepistemic Theories for Nonmonotonic Reasoning: Preliminary Report," In *Proceedings of the 2nd International Workshop on Nonmonotonic Reasoning*, 42–59, 1988.
- [6] Robert C. Moore, "Semantical Considerations on Nonmonotonic Logic," *Artificial Intelligence*, 25:75–94, 1985.
- [7] Katashi Nagao, "Semantic Interpretation Based on the Multi-World Model," In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, 1467–1473, 1989.
- [8] C. Raymond Perrault, "An Application of Default Logic to Speech Act Theory," In P. R. Cohen et al. (eds) *Intentions in Communication*, MIT Press, 161–185, 1990.
- [9] Raymond Reiter, "A Logic for Default Reasoning," *Artificial Intelligence*, 13:81–132, 1985.