ICOT Technical Memorandom: TM-969

TM-0969

非単調推論

佐藤　健

November, 1990

ICOT

Mita Kokusai Bldg. 21F
4-28 Mita 1-Chome
Minato-ku Tokyo 108 Japan

(03)3456-3191 ~ 5
Telex ICOT J32964

Institute for New Generation Computer Technology

非単調推論

佐藤 健
(財) 新世代コンピュータ技術開発機構

1 はじめに

人間が何かを行動するときに完全な情報から何かをすることは稀である．たとえば，列車に乗って旅行に行くとしよう．われわれは列車の時刻表を見てそれに間に合うように駅に行く計画をたてる．時刻表の情報は完全な情報のように見えるがそうではない．なぜなら，列車が時刻表どうりにくるかどうかは定かではないからである．したがって，列車が時刻表どおりにくるというのはもっともらしい仮説（常識）であり，人間はそうした常識に基づいて行動しているといえる．

こうした不完全な状況下における常識に基づいた人間の合理的な推論を定式化することが非単調推論の目的である．

さらに，この常識による推論の結果は新たな情報の出現によってくつがえされる可能性を含んでいる．旅行の例でいえば，家を出発する前にテレビで列車が遅れているという情報を得たならば，駅に問い合わせて時刻表から決めた計画を変更するであろう．つまり，前に時刻表の情報に基づいて作った計画が新たな情報によって変更されたのである．

結論がくつがされる可能性があるような推論の定式化は通常の論理の考え方から反している．なぜなら，論理の目的はある情報が得られたときにそこからいつでも正しい結論を導くことにある．このため情報が付け加わったからといって結論がくつがえされることはない．実は，この結論がくつがえされない性質を「単調である」といい，その逆，つまり結論がくつがえされる性質を「非単調である」という．非単調推論という名前は，結論がくつがえされる性質をもつ推論という意味から来ている．

論理の一つの目的は，数学的思考の正しさを定式化したものであり，その目的では，単調性が重要である．しかし，人間の日常的な推論を定式化するためには非単調性を扱わなければならない．

非単調推論の枠組としていろいろなものが提案されているが，本稿ではそれらを大きく二つに分ける．一つは無矛盾性に基づくアプローチであり，もう一つは極小モデルによるアプローチである．

無矛盾性に基づくアプローチでは，ある常識を仮定しても何の矛盾も生じないときにはその常識を使って推論するという考え方をする．さきほどの時刻表の例に照らしていえば，列車が遅れるという情報がない時点では時刻表どおりに列車が来ると仮定しても矛盾が生じないために時刻表に基づいて列車が来ると推論するものである．

これに対し，極小モデルに基づくアプローチではいまわかっている知識が成り立つ可能世界（モデル）を考え，その可能世界間に起こりやすさの順序を導入することによりいちばん起こりやすい可能世界（極小モデル）において成り立つものを推論の結果とするものである．さきほどの時刻表の例でいえば時刻表から得られる情報に沿った世界，すなわち時刻表どおりに列車が来る世界をいちばん起こりやすい世界と考え，その世界を念頭において行動するということである．

本稿では無矛盾性に基づくアプローチとしてReiterのデフォルト論理とMooreの自己認識論理を紹介し,極小モデルに基づくアプローチとしてMcCarthyの極小限定,佐藤の解釈の順序に基づいた枠組を紹介する.

2　無矛盾性に基づく非単調推論

2.1　デフォルト論理

デフォルト論理 [Reiter80] では通常の一階述語論理にデフォルトと呼ばれる新しい推論規則を導入することによって定義される. たとえば,「時刻表の列車は普通,時刻表どおり発車する」という常識を表現するためには以下のようなデフォルトを用いればよい.

$$\frac{Timetable(train, time) : Depart(train, time)}{Depart(train, time)}$$

ここで, $Timetable(train, time)$ は「時刻表に列車 $train$ が時刻 $time$ で発車すると記載してある」ことを表し, $Depart(train, time)$ は「列車 $train$ が時刻 $time$ で発車する」ことを表す. この推論規則の意味は $Timetable(train, time)$ が成り立ち, $Depart(train, time)$ を仮定しても無矛盾であるならば, $Depart(train, time)$ が成り立つと推論するというものである. このようなデフォルトを使えば, 時刻表に記載がある列車は, 遅れるという情報がなければ時刻表どおりに列車が発車することを導くことができ, またもし列車が遅れて時刻表どおりに列車が発車できないことがわかれば,上のデフォルトによる推論はできないようになる.

この論理をもう少し詳細に調べてみよう. デフォルトの厳密な定義は以下のとおりである.

$\mathbf{x}$ を自由変数の組, $\mathbf{x}$ を自由変数としてもつ一階述語論理式を $\alpha(\mathbf{x}), \beta_1(\mathbf{x}), ..., \beta_m(\mathbf{x}), w(\mathbf{x})$ とする. デフォルトとは以下の形の推論規則のことである.

$$\frac{\alpha(\mathbf{x}) : \beta_1(\mathbf{x}), ..., \beta_m(\mathbf{x})}{w(\mathbf{x})}$$

$\alpha(\mathbf{x})$ の部分を前提部と呼び, $\beta_1(\mathbf{x}), ..., \beta_m(\mathbf{x})$ の部分を根拠部と呼び $w(\mathbf{x})$ の部分を結論部と呼ぶ. また, デフォルトの集合を D, 公理の集合を W とすれば (D, W) をデフォルト理論という. さて, このようなデフォルト理論に対して望ましい推論結果 E(拡張と呼ぶ) はどのようになるのであろうか. それは, E を決めたときに以下の条件を満たす最小の論理式の集合 $\Gamma(E)$ が E と一致することである.

1. 正しいとわかっている情報は $\Gamma(E)$ に入っている. 正しいとわかっている情報は公理の集合 W であるから $W \subseteq \Gamma(E)$.

2. $\Gamma(E)$ は論理的に閉じている. $Th(\Gamma(E)) = \Gamma(E)$[1].

3. 適用できるデフォルトルールがある場合, その結論部は $\Gamma(E)$ に含まれる. デフォルトを $\dfrac{\alpha : \beta_1, ..., \beta_m}{w}$ とする. $\alpha \in \Gamma(E)$ かつ $\neg\beta_1 \notin E$,...,$\neg\beta_m \notin E$ ならば $w \in \Gamma(E)$. ここ

[1] $Th(X)$ は論理式の集合 X から一階述語論理の推論規則によって導かれるすべての論理式の集合を表す.

で，$\neg\beta_i \notin E (i = 1, ..., m)$ ということは，$E \cup \{\beta_i\} \not\vdash \mathbf{F}$ [2]ということで，β_i を仮定しても E と矛盾しないということを表している．

この $\Gamma(E)$ の定義は公理 W を含み，仮定した推論結果 E に矛盾しないようなデフォルトを適用してできる最小の論理的帰結の集合である．この $\Gamma(E)$ が最初仮定した E と一致するということは，E を推論結果としても妥当であることを表している．しかるに，この定義ではまず何らかの論理的帰結の集合 E を仮定してそれから得られる $\Gamma(E)$ と一致するかどうかを調べてみないと拡張を得られない．したがって，今わかっている公理から始めて論理的帰結を得るようなことができない．さいわいにも最近の研究の結果から，今わかっている公理から始めて拡張を得る非決定的手続きが存在する [Etherington87].

拡張を計算する非決定的手続き

Step 0: $E_0 = Th(W), i := 0$.

Step 1: E_i に関して適用可能なデフォルト[3] $\delta_i \in D$ でその結論部が E_i に入っていないものを一つ選ぶ．もしそのようなデフォルトがなければ E_i を出力する．

Step 2: $E_{i+1} = Th(E_i \cup \{w_i\})$. もし，ある $k(k = 0, ..., i)$ が存在して δ_k が E_{i+1} に関して適用可能でなくなれば，この手続きは失敗する．さもなければ，i を 1 増やして Step 1 に行く．

この手続きの出力はすべて拡張であり，また，全解探索を行えばすべての拡張を求めることができる．しかしながら，この手続きはデフォルトの適用可能性を調べるために無矛盾性検査を必要とするので一般の一階述語論理式に対しては計算可能ではない．命題論理式や無矛盾性検査が決定可能な一階述語論理式のクラスに対しては計算可能である．

この手続きを使って以下のデフォルト理論 (D, W_1) の拡張を求めてみよう．

$$D = \{\frac{Bird(x) : Fly(x)}{Fly(x)}\}$$
$$W_1 = \{Bird(Tweety), \forall x(Penguin(x) \supset Bird(x)), \forall x(Penguin(x) \supset \neg Fly(x))\}$$

ここで，$Bird(x)$ は「x が鳥である」ことを表し，$Penguin(x)$ は「x がペンギンである」ことを表し，$Fly(x)$ は「x が飛ぶ」ことを表す．また，D のデフォルトは，x に変数を含まない項を代入した以下のような無数のデフォルトを表していると考える．

$$\frac{Bird(A) : Fly(A)}{Fly(A)}, \frac{Bird(B) : Fly(B)}{Fly(B)}, ..., \frac{Bird(Tweety) : Fly(Tweety)}{Fly(Tweety)}, ...$$

このデフォルト理論に対して上の手続きを適用してみる．

0. $E_0 = Th(W_1)$.

---

[2] $\mathbf{F}$ は矛盾を表す

[3] デフォルト δ を $\frac{\alpha : \beta_1, ..., \beta_m}{w}$ としたとき $\alpha \in E$ かつ $\neg\beta_1 \notin E, ..., \neg\beta_m \notin E$ ならば，δ は E に関して適用可能であるという．

1. $Bird(Tweety)$ かつ $\neg Fly(Tweety) \notin E_0$ であるからデフォルト

$$\delta_0 = \frac{Bird(Tweety) : Fly(Tweety)}{Fly(Tweety)}$$

は E_0 に対して適用可能であり $Fly(Tweety) \notin E_0$ であるから

$$E_1 = Th(W_1 \cup \{Fly(Tweety)\})$$

となる.

2. E_1 に対して適用可能なデフォルトは存在しないので E_1 が出力される.

したがって, デフォルトから $Fly(Tweety)$ が導かれたことになる. しかしながら, この結論は新たな情報の付加によって無効になることがある.

たとえば, $W_2 = W_1 \cup \{Penguin(Tweety)\}$ としてみる. つまり, Tweety がペンギンであることがわかったとする. W_2 からは, $\neg Fly(Tweety)$ が導かれるので W_1 のときに適用可能であったデフォルト $\delta_0 = \dfrac{Bird(Tweety) : Fly(Tweety)}{Fly(Tweety)}$ は, W_2 では, 適用可能でなくなったことに注意されたい. (D, W_2) に関する拡張は以下のように計算される.

0. $E_0 = Th(W_2)$.

1. E_0 に対して適用可能なデフォルトは存在しないので E_0 が出力される.

したがって, さきほどは $Fly(Tweety)$ が導かれたのに対し W_2 ではもはや $Fly(Tweety)$ は導かれない.

さて, この拡張は通常の論理の論理的帰結とは異なる性質を持つ. たとえば, $D = \{\dfrac{: \neg P}{Q}, \dfrac{: \neg Q}{P}\}$ かつ $W = \{\}$ とすると, 拡張は2つ存在する.

Case 1:

0. $E_0 = Th(\{\})$.
1. $P \notin E_0$ より $\delta_0 = \dfrac{: \neg P}{Q}$ は E_0 に関して適用可能であり, かつ $Q \notin E_0$ より $E_1 = Th(\{Q\})$ となる.
2. E_1 に対して適用可能なデフォルトは存在しないので E_1 が出力される.

Case 2:

0. $E_0 = Th(\{\})$.
1. $Q \notin E_0$ より $\delta_0 = \dfrac{: \neg Q}{P}$ は E_0 に関して適用可能であり, かつ $P \notin E_0$ より $E_1 = Th(\{P\})$ となる.
2. E_1 に対して適用可能なデフォルトは存在しないので E_1 が出力される.

推論結果が2つ存在するということは通常の論理からみると奇異な感じがするかもしれないが、人間の推論ではこのようなことが起こりうる．たとえば，「人を見たら泥棒と思え」と「渡る世間に鬼はない」のように常識の中には矛盾するものもあり，どちらを選ぶかによって自分の信じることが変わるというようなことである．

また，$D = \{\frac{:\neg P}{P}\}$ かつ $W = \{\}$ とすると，拡張は存在しない．

0. $E_0 = Th(\{\})$.

1. $P \notin E_0$ より $\delta_0 = \frac{:\neg P}{P}$ は E_0 に関して適用可能であるので $E_1 = Th(\{P\})$ となる．しかし，δ_0 はもはや E_1 に対しては適用可能でなくなるのでこの手続きは失敗する．なおかつ，これ以外のデフォルトの選択は存在しないので拡張は存在しない．

上のデフォルトの意味を考えると $\neg P$ を仮定しても推論結果と矛盾しないときに P が推論結果に入っているという矛盾したものになっているので拡張が存在しないのが直感にあっているといえよう．

これに関連して以下の形をしたデフォルトだけからなるデフォルト理論の場合には，いつでも拡張が存在することが知られている．

$$\frac{\alpha : w}{w}$$

2.2 自己認識論理

自己認識論理 [Moore85] は通常の命題論理に様相記号 $\mathbf{L}$ を導入した様相命題論理式を扱う[4]．$\mathbf{L}\alpha$ の意味は「α を信じる」ということである．たとえば，「β ならば普通 α である」という常識を表現するためには以下のような論理式を用いればよい．

$$\mathbf{L}\beta \wedge \neg\mathbf{L}\neg\alpha \supset \alpha$$

この式の意味は β を信じていて，$\neg\alpha$ を信じていなければ（α を信じても信念が矛盾しないならば）α が成り立つということである．つまり，デフォルト論理で推論規則であったものを論理式として直接記述できるようにしたものが自己認識論理であると考えられる．ここで，デフォルト論理でのデフォルトの前提部，根拠部，結論部は，それぞれ上の自己認識論理式の $\mathbf{L}\beta$, $\neg\mathbf{L}\neg\alpha$, α の部分に対応している．しかしながら，自己認識論理では，望ましい推論結果は以下のように定義されているので若干の相違がある．

自己認識論理式の集合 A が与えられたとき，自己認識論理式の集合 E は以下の条件をみたすとき拡張となる．

1. わかっている情報は E に入っている．つまり，$A \subseteq E$.

2. E は論理的に閉じている．つまり，$E = \{\alpha | E \vdash \alpha\}$．ここで $\vdash$ は，各 $\mathbf{L}\alpha$ を一つの命題記号としてみたときの命題論理の論理的帰結を表す．

---

[4] 通常の一階述語論理に拡張したときの定義は，Levesque [Levesque90] にあるがまだ決定版があるとはいえない．

3. α を信じていたときかつそのときにかぎり $\mathbf{L}\alpha$ を信じている. つまり, E に α が含まれている場合かつそのときにかぎり $\mathbf{L}\alpha$ は E に含まれる.

以上をまとめると, E は以下の式の不動点となる.

$$E = \{\alpha | A \cup \mathbf{L}E \cup \neg\mathbf{L}\overline{E} \vdash \alpha\}$$

ここで, $\mathbf{L}E$ は, $\{\mathbf{L}\alpha | \alpha \in E\}$ なる自己認識論理式の集合を表し $\neg\mathbf{L}\overline{E}$ は, $\{\neg\mathbf{L}\alpha | \alpha \notin E\}$ なる自己認識論理式の集合を表す. この定義の直感的意味は E に含まれていればそれを信じるとし, E に含まれていないければそれを信じないとしたときに, そこから A によって推論できるものが E に一致するということである. デフォルト論理と同じようにこの定義ではまず何らかの論理的帰結の集合 E を仮定してそれから得られるものと一致するかどうかを調べてみないと拡張を得られないので今わかっている公理から始めて論理的帰結を得るようなことができない. 残念ながら, デフォルト論理と違ってこの場合の拡張を得る構成的な手続きは知られていない. そのかわりに, 拡張のモデル論が与えられており、それを用いて拡張の形を知ることができる [Moore84].

今, 公理 A の中にある命題記号に対して解釈を考える. 可能な命題解釈は, 命題記号の数を n 個とすれば 2^n 個ある. その中の m 個の命題解釈の集合 K と 1 個の命題解釈 V(m 個の中に入っていても入っていなくてもよい) をとったとき, (K, V) を A の可能世界解釈と呼ぶ. そして, 「可能世界解釈が論理式 α を満たす」ということを以下のように定義する.

1. α が $\mathbf{L}$ を含まない式 (通常式と呼ぶ) のとき $(K, V) \models \alpha$ とは $V \models \alpha$ のことである.

2. α が $\mathbf{L}\beta$ のとき $(K, V) \models \alpha$ とはすべての $W \in K$ に対して $W \models \beta$ のことである.

3. それ以外のとき, 通常の命題論理の再帰的な定義にしたがう.

もし, (K, V) が A 内のすべての論理式を満たせば, (K, V) を A の可能世界モデルと呼ぶ. 以下の事実が知られている.

> A の命題解釈の集合から K を選んだとき, 以下の条件を満たすときかつそのときにかぎり K から生成される自己認識論理式の集合は A の拡張になっている.
> 条件: $V \in K$ のときかつそのときにかぎり (K, V) が A の可能世界モデルになっている.

ここで K から生成される自己認識論理式の集合とは, 以下の集合を表す.

$$\{\alpha | \forall V \in K, (K, V) \models \alpha\}$$

この定義を使って拡張を求めてみる.

$A = \{\neg\mathbf{L}P \supset Q, \neg\mathbf{L}Q \supset P\}$ とすると, 拡張は 2 つ存在する. A に含まれる命題記号は P,Q であるから, 命題解釈は 4 個ある. その中から m 個選んで K とし, 上の条件を満たすかどうか調べてみる.

Case 1: K として $\{\{P, Q\}, \{P, \neg Q\}\}$ [5]を選ぶと, 以下に示すように上の条件を満たしている.

---

[5]ここでは解釈をその解釈で真となるリテラルの集合で表すとする

1. $V_1 = \{P, Q\}$ のとき:
   $\forall V \in K, V \models P$ より, $(K, V_1) \models \mathbf{L}P$ であるから, $(K, V_1) \models \neg\mathbf{L}P \supset Q$ である. また, $V_1 \models P$ より $(K, V_1) \models \neg\mathbf{L}Q \supset P$ である. したがって, (K, V_1) は A の可能世界モデルである.

2. $V_2 = \{P, \neg Q\}$ のとき:
   上と同様にして, $(K, V_2) \models \neg\mathbf{L}P \supset Q$ かつ $(K, V_2) \models \neg\mathbf{L}Q \supset P$ であるから (K, V_2) は A の可能世界モデルである.

3. $V_3 = \{\neg P, Q\}$ のとき:
   $\exists V \in K, V \not\models Q$ より, $(K, V_3) \models \neg\mathbf{L}Q$ である. しかしながら, $V_3 \models \neg P$ より $(K, V_3) \not\models P$ である. したがって, $(K, V_3) \not\models \neg\mathbf{L}Q \supset P$ となり, (K, V_3) は A の可能世界モデルではない.

4. $V_4 = \{\neg P, \neg Q\}$ のとき:
   上と同様にして $(K, V_4) \not\models \neg\mathbf{L}Q \supset P$ より, (K, V_4) は A の可能世界モデルではない.

したがって, K から生成される自己認識論理式の集合は A の拡張になっている. この場合, P が拡張の中に入るが, Q は入らない.

Case 2: K として $\{\{P, Q\}, \{\neg P, Q\}\}$ を選ぶと, 上の条件を満たしている. したがって, K から生成される自己認識論理式の集合は A の拡張になっている. この場合, Q が拡張の中に入るが, P は入らない.

$A = \{\neg\mathbf{L}P \supset P\}$ とすると, 拡張は存在しない. A に含まれる命題記号は P であるから, 命題解釈は 2 個ある. その中から m 個選んで K とし, 上の条件を満たすかどうか調べてみると条件を満たす K は存在しない. したがって拡張は存在しない.

また, $A = \{\mathbf{L}P \supset P\}$ とすると, 拡張は 2 つ存在する. A に含まれる命題記号は P であるから, 命題解釈は 2 個ある. その中から m 個選んで K とし, 上の条件を満たすかどうか調べてみる.

Case 1: K として $\{\{P\}\}$ を選ぶと上の条件を満たしている. したがって, K から生成される自己認識論理式の集合は A の拡張になっている. この場合, P が拡張の中に入る.

Case 2: K として $\{\{P\}, \{\neg P\}\}$ を選ぶと, 上の条件を満たしている. したがって, K から生成される自己認識論理式の集合は A の拡張になっている. この場合, P が拡張の中に入らない.

一番最初に述べたデフォルト論理との対応だと $A = \{\mathbf{L}P \supset P\}$ は $D = \{\frac{P:}{P}\}, W = \{\}$ なるデフォルト理論 (D,W) に対応するが, このデフォルト理論では, P を含まない拡張だけが得られる. これが, デフォルト論理と自己認識論理の拡張の定義の違いを表しているといえる. すなわち, 自己認識論理の場合「P を信じていたら P が成り立つ」ということから P が成り立つとする循環論法が使えるということである. デフォルト論理と自己認識論理の関係については [Konolige88, Marek89] を参照していただきたい.

3 極小モデルに基づく非単調推論

3.1 極小限定

極小限定 [McCarthy80, Lifschitz84] では, 2階述語論理を用いて「P という性質を満たすものは今わかっているものだけでそれ以外のものは P という性質を満たさない.」ということを表す.

このことを簡単な例で示す. 述語 P を含む公理を $A(P)$ で表すとする. すると, P に関する極小限定の定義 (P は極小化述語と呼ぶ) は以下のようになる.

$$A(P) \wedge \neg\exists p(A(p) \wedge p < P)$$

ここで, $A(p)$ は $A(P)$ 内のすべての P を述語変数 p で置き換えたものであり, $p < P$ は以下の論理式の略である.

$$\forall x(p(x) \supset P(x)) \wedge \neg\forall x(P(x) \supset p(x))$$

この式の直感的な意味は以下のようである. $A(P)$ という式が P という性質の定義を表していると考える. $A(p)$ も p が P と同じ性質を持っていることを表している. $p < P$ は p を満たす要素の集合 (外延と呼ぶ) が P の外延よりも小さいことを表している. $\neg\exists p(A(p) \wedge p < P)$ はそのように P の外延より小さい外延を持つ p が存在しないことを表している. したがって, 極小限定の定義は P が $A(P)$ を満たす外延の集合の中で一番小さいもの (正確には極小) であることを表している.

たとえば $A(P) = P(B)$ としてみよう. もし $P(C)$ なる $C(B \neq C)$ があると考えると p として $p(B)$ だけが成り立つとすれば $p < P$ なるものが存在するということになり P が極小であることに反する. したがって, P として考えられるのは $P(B)$ のみが成り立ち B 以外は成り立たないものすなわち $P = \lambda x(x = B)$ [6] になる.

また, たとえば $A(P) = P(B) \vee P(C)$ としてみよう. このときは, $P(B)$ だけが成り立つものと $P(C)$ だけが成り立つものの 2 つが出てくる. なぜなら, 両方とも集合の包含関係において極小なものになっているからである.

常識推論ではパラメータを含む極小限定が必要となる. パラメータの述語の働きは, パラメータの述語の定義を変化させたとき極小化する述語を極小化できるならばその定義を変化できるというものである. このときの定義は極小化する述語を P として, パラメータの述語を Z とすれば以下のように表される.

$$A(P,Z) \wedge \neg\exists p\exists z(A(p,z) \wedge p < P)$$

この定義だと, とにかく z として $A(p,z)$ を満たすものをもってきても $p < P$ なるものがないということを表し, Z としてどんなものを持ってきても P が極小であることを示している.

---

[6] t を項とすれば $P(t)$ は $\lambda x(x = B)(t)$ となる. この式はラムダ式の本体の中の x を t で置き換えたもの, すなわち, $t = B$ となる. したがって, P の真偽値は t が B のとき真であり t が B に等しくないものは偽となる.

このパラメータつきの極小限定の式を使うと常識は以下のように表される．たとえば，時刻表の例では

$$\forall train \forall time((Timetable(train, time) \wedge \neg Ab(train, time)) \supset Depart(train, time))$$

として，$Timetable, Depart$ の述語をパラメータとして，Ab をできるだけ小さくすればよい．ここで，Ab は例外述語と呼ばれ，常識の例外を表すために使われている．つまり，上の式は，$train, time$ が例外でないならば常識を用いることができるということを表している．そして，ここでの極小限定の使われ方は例外を今わかっているものだけに限定することで，例外とわからないものに関してはこの常識が成り立つようにするということである．

推論の例を示そう．たとえば，

$$\begin{aligned}&A(Ab, Bird, Penguin, Fly) = \\ &\quad Bird(Tweety) \wedge \forall x(Penguin(x) \supset Bird(x)) \wedge \forall x(Penguin(x) \supset \neg Fly(x)) \wedge \\ &\quad \forall x((Bird(x) \wedge \neg Ab(x)) \supset Fly(x))\end{aligned}$$

として，$Bird, Penguin, Fly$ をパラメータとして Ab を極小化してみる．ここで Ab は「鳥は普通飛ぶ」という常識に対する例外を表す．

$$A(Ab, Bird, Penguin, Fly) \wedge \neg\exists ab \exists b \exists p \exists f(A(ab, b, p, f) \wedge ab < Ab)$$

この式は以下の式に変換される．

$$\begin{aligned}&A(Ab, Bird, Penguin, Fly) \wedge \\ &\quad \forall ab \forall b \forall p \forall f((A(ab, b, p, f) \wedge \forall x(ab(x) \supset Ab(x))) \supset \forall x(Ab(x) \supset ab(x)))\end{aligned}$$

この式は以下の式と同値である[7]．

$$\begin{aligned}&A(Ab, Bird, Penguin, Fly) \wedge \\ &\forall ab \forall b \forall p \forall f((A(ab, b, p, f) \wedge \forall x(ab(x) \supset Ab(x))) \supset \forall x(Ab(x) \equiv ab(x)))\end{aligned}$$

この式の意味は $A(ab, b, p, f)$ を満たす述語の組 (ab, b, p, f) があったときに $\forall x(ab(x) \supset Ab(x))$ を満たせば ab の定義は Ab と同じであることを表している．さて，ここで，以下のような述語の組を考えてみる．

1. $ab = \lambda x \mathbf{F}$
2. $b = \lambda x(x = Tweety)$
3. $p = \lambda x \mathbf{F}$
4. $f = \lambda x(x = Tweety)$

としてみる．すると

$$A(ab, b, p, f) = b(Tweety) \wedge \forall x(p(x) \supset b(x)) \wedge \forall x(p(x) \supset \neg f(x)) \wedge \forall x((b(x) \wedge \neg ab(x)) \supset f(x))$$

---

[7] 一般に $(\alpha \wedge \beta) \supset \gamma$ は $(\alpha \wedge \beta) \supset (\beta \wedge \gamma)$ と同値であることに注意．

において, 上の式を代入すると $b(Tweety)$ は $\lambda x(x = Tweety)(Tweety)$ より $Tweety = Tweety$ となり, $p(x)$ は $\lambda x\mathbf{F}(x)$ より $\mathbf{F}$ となり, $b(x)$ は $\lambda x(x = Tweety)(x)$ より $x = Tweety$ となり, $f(x)$ は $\lambda x(x = Tweety)(x)$ より $x = Tweety$ となり, $ab(x)$ は $\lambda x\mathbf{F}(x)$ より $\mathbf{F}$ となるので, $A(ab, b, p, f)$ は以下の式になる.

$$
\begin{aligned}
&(Tweety = Tweety) \wedge \forall x(\mathbf{F} \supset (x = Tweety)) \wedge \\
&\forall x(\mathbf{F} \supset \neg(x = Tweety)) \wedge \forall x(((x = Tweety) \wedge \neg\mathbf{F}) \supset (x = Tweety))
\end{aligned}
$$

上の式は恒真となる.

また $\forall x(ab(x) \supset Ab(x))$ は $\forall x(\mathbf{F} \supset Ab(x))$ となり, いつでも満たされるので, ab の定義, すなわち $\lambda x\mathbf{F}$ が Ab の定義となる. したがって, 極小限定の式から $Ab = \lambda x\mathbf{F}$ が導かれる. これは, $A(Ab, Bird, Penguin, Fly)$ には, 例外とわかっているものがないことを表す. したがって, $\forall x((Bird(x) \wedge \neg Ab(x)) \supset Fly(x))$ は $\forall x(Bird(x) \supset Fly(x))$ と同値となりその結果 $Fly(Tweety)$ が導かれる.

さて, ここで $Penguin(Tweety)$ がわかったとしよう. つまり,

$$
\begin{aligned}
&A(Ab, Bird, Penguin, Fly) = \\
&\quad Bird(Tweety) \wedge Penguin(Tweety) \wedge \forall x(Penguin(x) \supset Bird(x)) \wedge \\
&\quad \forall x(Penguin(x) \supset \neg Fly(x)) \wedge \forall x((Bird(x) \wedge \neg Ab(x)) \supset Fly(x))
\end{aligned}
$$

となったときについて考える. ここで, 以下のような述語の組を考えてみる.

1. $ab = \lambda x(x = Tweety)$
2. $b = \lambda x(x = Tweety)$
3. $p = \lambda x(x = Tweety)$
4. $f = \lambda x\mathbf{F}$

としてみる. すると前と同様にして $A(ab, b, p, f)$ は以下のようになる.

$$
\begin{aligned}
&(Tweety = Tweety) \wedge (Tweety = Tweety) \wedge \forall x((x = Tweety) \supset (x = Tweety)) \wedge \\
&\forall x((x = Tweety) \supset \neg\mathbf{F}) \wedge \forall x(((x = Tweety) \wedge \neg(x = Tweety)) \supset \mathbf{F})
\end{aligned}
$$

この式は恒真である.

また $\forall x(ab(x) \supset Ab(x))$ は $\forall x(x = Tweety \supset Ab(x))$, すなわち, $Ab(Tweety)$ となり, $A(Ab, Bird, Penguin, Fly)$ から $Ab(Tweety)$ が導かれるので[8], $\forall x(ab(x) \supset Ab(x))$ は満たされることがわかり, ab の定義, すなわち $\lambda x(x = Tweety)$ が Ab の定義となる. したがって, 極小限定の式から $Ab = \lambda x(x = Tweety)$ が導かれる. これは, $A(Ab, Bird, Penguin, Fly)$ には, 例外とわかっているものが $Tweety$ だけであることを表す. したがって, $\forall x((Bird(x) \wedge \neg Ab(x)) \supset Fly(x))$ は $\forall x((Bird(x) \wedge (x \neq Tweety)) \supset Fly(x))$ と同値となりその結果もはや $Fly(Tweety)$ が導かれなくなる.

[8] $Penguin(Tweety)$ と $\forall x(Penguin(x) \supset \neg Fly(x))$ より $\neg Fly(Tweety)$ が導かれ, $Bird(Tweety)$ と $\neg Fly(Tweety)$ と $\forall x((Bird(x) \wedge \neg Ab(x)) \supset \neg Fly(x))$ とから $Ab(Tweety)$ が導かれる.

このような計算方法ではいろいろな定義を述語変数に当てはめてみて条件にあうかどうか調べることになるが，この方法のほかにいろいろな計算方法がある．[中川 89] を参照していただきたい．ただし，あとで述べるように正しい論理式をすべて導く計算方法は一般には存在しない．これは，極小限定の定義が 2 階述語論理で表されていることに起因する．

さて，この極小限定には極小モデルに基づく意味論が存在する．P を極小化する述語とし，Z をパラメータ述語としたときに以下の順序をドメイン D をもつ解釈の間に導入する．2 つの解釈を I_1, I_2 とする．$I_1 \leq I_2$ を以下のように定義する．

1. I_1, I_2 は P と Z 以外の記号に対して同じ解釈を持つ．

2. $I_1[P] \subseteq I_2[P]$.

2 番目の条件は I_1 において P の外延は I_2 の外延に含まれる，すなわち I_1 において満たされる D の要素が I_2 において満たされる要素より集合の包含関係の意味で小さいことを表している．

解釈 I が論理式 A の P, Z に関する極小モデルであるとは，I が A のモデルであって，$I' \leq I$ かつ $I \leq I'$ でない A のモデル I' が存在しないことである．

つまり，極小モデルは A という条件を満たしつつできるだけ P を満たす D の要素を少なくしたものである．ここで注意してほしいのは，極小限定でのモデル間の順序は極小化する述語を満たす外延に関する包含関係に基づいているということである．より一般的な順序に基づいた極小モデルの定義は次章の佐藤の解釈の順序に基づいた枠組でのべる．

2 階述語論理式で表された極小限定の定義のモデルは極小モデルと一致することが示されている．すなわち，極小限定の定義から得られる推論結果はすべての極小モデルで成り立つ． しかしながら，すべての極小モデルで成り立つすべての論理式を極小限定の定義から導き出すことは一般には不可能であることが示されている．したがって，すべての極小モデルで成り立つすべての論理式を導くためには最初に与える公理の形を制限しなければならない．

また，最初に与える公理にモデルがあっても極小モデルが存在しない場合がある．この場合には極小限定の結果は矛盾を引き起こす． こういった場合を避けるためにも，最初に与える公理の形を制限しなければならない．

さて，この基本的な極小限定からいろいろな極小限定の形式が派生している．ここでは，その中で最も重要と思われる優先順位付き限定 [McCarthy86] について述べる．

優先順位付き極小限定では，極小化する述語に対して何を先に極小化するかを指定できるものである．この指定によって競合する常識に対して優先順位を与えることができる．たとえば，「生き物は普通飛ばない」という常識と「鳥は普通飛ぶ」という常識 Tweety が鳥であることがわかったときには，両方適用可能であるように見えるが，われわれはよりくわしい常識である鳥の常識の方を優先する．しかしながら，これを通常の極小限定でやろうとするとうまく行かない．たとえば，

$$
\begin{aligned}
&A(Ab_1, Ab_2, Animal, Bird, Fly) = \\
&\quad Bird(Tweety) \land \forall x(Bird(x) \supset Animal(x)) \land \\
&\quad \forall x((Animal(x) \land \neg Ab_1(x)) \supset \neg Fly(x)) \land \forall x((Bird(x) \land \neg Ab_2(x)) \supset Fly(x))
\end{aligned}
$$

として，Ab_1, Ab_2 を極小化しても $Fly(Tweety)$ は導かれない．これは，2つの常識に対する優先順位が極小化する際に使われないためである．この問題を解決するために極小化する述語に極小化の優先順位を与えたものが優先順位付き極小限定である．

たとえば，上の例に関する優先順位付き極小限定の定義は以下のようになる．

$$A(Ab_1, Ab_2, Animal, Bird, Fly) \wedge$$
$$\exists ab_1, \exists ab_2, \exists a, \exists b, \exists f(A(ab_1, ab_2, a, b, f) \wedge (ab_2; ab_1) \leq (Ab_1; Ab_2))$$

ここで $(ab_2; ab_1) \leq (Ab_2; Ab_1))$ は以下の式の略である．

$$\forall x(ab_2(x) \supset Ab_2(x)) \wedge (\forall x(ab_2(x) \equiv Ab_2(x)) \supset \forall x(ab_1(x) \supset Ab_1(x))) \wedge$$
$$\neg(\forall x(Ab_2(x) \supset ab_2(x)) \wedge (\forall x(Ab_2(x) \equiv ab_2(x)) \supset \forall x(Ab_1(x) \supset ab_1(x))))$$

この式の意味は Ab_2 を先に極小化し，そのあとで Ab_1 を極小化するというものである．Ab_1 は生き物に関する常識の例外を表し Ab_2 は鳥に関する常識の例外を表すので，Ab_2 のほうを先に極小化することによって鳥に関する常識のほうが優先されるのである．

3.2 解釈の順序に基づく非単調推論

極小限定では，例外を満たすものを小さくするという考えであったがこれをもっと一般化すると直接常識を扱うことができる [Satoh90].

基本的な考え方は以下のようである．今わかっていることに矛盾しない可能世界の中に起こりやすさの順序を導入する．そして，我々が実際にいる世界はその順序のなかで一番起こりやすい世界であると考える．たとえば，時刻表の例を考えてみよう．すると，列車が来るまでは列車が時刻表どおりに来るかどうかわからないので可能世界としては「列車が時刻表どおりに来る」世界 w_1 と「列車が時刻表どおりに来ない」世界 w_2 の2つが考えられる．ここで，可能世界の起こりやすさの順序を考えると，「列車が普通時刻表どおりに来る」という常識から w_1 のほうが w_2 よりも起こりやすいということになる．このことから，我々の実際いる世界は列車が時刻表どおりに来る世界であると考えるのである．

この考え方を論理的には以下のように考えることができる．すなわち，可能世界を考えるということは，今わかっていることを公理として表したときにそれを成り立たせる解釈(モデル)を考えることになる．そして，起こりやすい可能世界を考えるということは，その解釈の間に順序をつけて一番起こりやすい解釈を選ぶということに対応する．この枠組では，解釈の順序を表すメタ言語を定義し，それによって世界の起こりやすさを表す．また，このメタ言語を変換して一番起こりやすい解釈の形式的定義を2階述語論理式で表すことができる．

簡単のため命題論理で説明する．生き物と鳥と飛ぶことに関する例を考える．いつでも正しい規則「鳥は生き物である.」と2つの常識「生き物はふつう飛ばない.」,「鳥はふつう飛ぶ.」から得られる解釈の順序は図1のようになる．ここで A は「Tweety が生き物であること」を表し，B は「Tweety が鳥であること」を表し，F は「Tweety が飛ぶこと」を表す．図では，下の方の解釈がより起こりやすいとする．2つの常識を $A \supset \neg F$, $B \supset F$ で表し，できるだけこの論理式が成り立っている世界のほうが起こりやすいようになっている．一番下の3つの解釈は，2つの常識を両方とも満たす解釈であり，真ん中の2つの解釈は，「鳥はふつう飛ぶ.」という常識が成り立っていて「生き物はふつう飛ばない.」という常識が成り

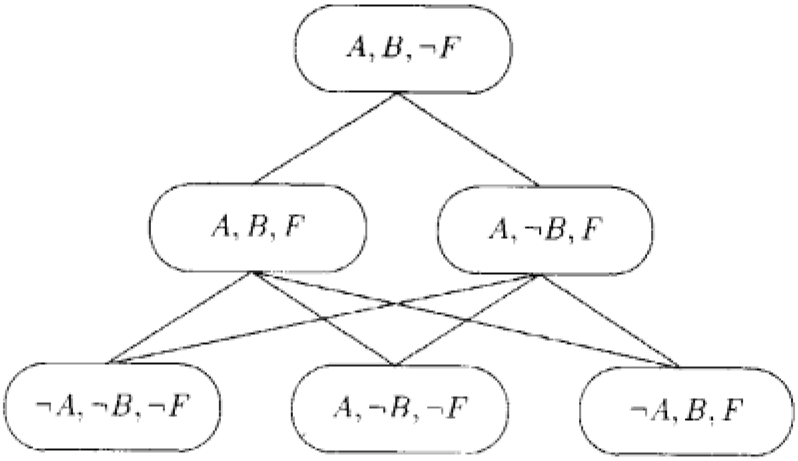

図 1: 生き物と鳥と飛ぶことに関する解釈の順序

立っていない解釈である．これは，鳥に関する常識のほうが生き物に関する常識よりも優先されることを表している．

ここで，Tweety が生き物であることがわかったとする．すると生き物が真でない世界はもはやモデルにはならなくなるので解釈の順序の様子は図 2のようになる．この状況では，一番起こりやすい解釈に対応するのは Tweety は鳥でなく飛ばない世界となる．

さらに，Tweety が鳥であることがわかったとする．すると鳥が真でない解釈はもはやモデルにはならなくなるので解釈の順序の様子は図 3のようになる．この状況では，一番起こりやすい解釈に対応するのはこの場合，Tweety が飛ぶ世界となる．

極小限定ではこの解釈の順序を極小化の順序に置き換えるために例外述語を導入し，それを極小化している．

図 1の順序をメタ言語で表すと以下のようになる．

$$I_2 < I_1 \stackrel{\mathrm{def}}{=}$$

$$((I_1 \models B \supset F) \boldsymbol{\supset} (I_2 \models B \supset F)) \mathbb{A}$$

$$(((I_1 \models B \supset F) \equiv (I_2 \models B \supset F)) \boldsymbol{\supset} ((I_1 \models A \supset \neg F) \boldsymbol{\supset} (I_2 \models A \supset \neg F))) \mathbb{A} \tag{1}$$

$$\neg(((I_2 \models B \supset F) \boldsymbol{\supset} (I_1 \models B \supset F)) \mathbb{A}$$

$$(((I_2 \models B \supset F) \equiv (I_1 \models B \supset F)) \boldsymbol{\supset} ((I_2 \models A \supset \neg F) \boldsymbol{\supset} (I_1 \models A \supset \neg F)))) \tag{2}$$

ここで，$\boldsymbol{\supset}$ はメタ言語における「ならば」を意味し $\mathbb{A}$ はメタ言語における「かつ」を意味し $\equiv$ はメタ言語における同値を意味する．

$I_2 < I_1$ は I_2 の方が I_1 よりも起こりやすいことを表す．上の式の意味は，$B \supset F$ を満たす世界が起こりやすく，2 つの世界が両方とも $B \supset F$ を満たせば，$A \supset \neg F$ を満たす世界の方が起こりやすいことを表している．

たとえば，I_1 として $\{A, B, F\}$，I_2 として $\{A, \neg B, \neg F\}$ を選び，$I_2 < I_1$ かどうか調べてみる．$I_1 \models B \supset F$ かつ $I_2 \models B \supset F$ より $(I_1 \models B \supset F) \boldsymbol{\supset} (I_2 \models B \supset F)$ は真となり $(I_1 \models B \supset F) \equiv (I_2 \models B \supset F)$ も真となる．また，$I_1 \not\models A \supset \neg F$ かつ $I_2 \models A \supset \neg F$ より $(I_1 \models A \supset \neg F) \boldsymbol{\supset} (I_2 \models A \supset \neg F)$ も真であり，(1) の部分は真となる．しかしながら，$(I_2 \models A \supset \neg F) \boldsymbol{\supset} (I_1 \models A \supset \neg F)$ は真でないので，(2) の一番外側の $\neg$ は偽となり (2) 全体

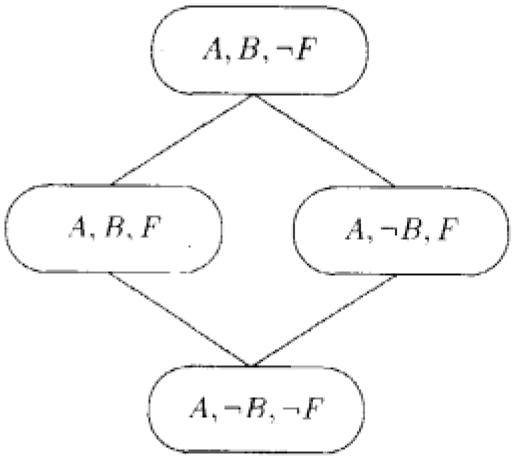

図 2: 生き物であることがわかったときの解釈の順序

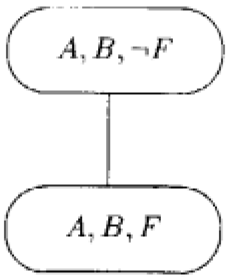

図 3: 鳥であることがわかったときの解釈の順序

は真となる, したがって, $I_2 < I_1$ であることがわかる. これを解釈のすべての2つの組に対して行うと図1が得られる.

このメタ言語で表された順序から得られる最も起こりやすい解釈を形式的に表現することができる. 公理を $A(A,B,F)$ とすると以下の論理式のモデルは最も起こりやすい $A(A,B,F)$ のモデルとなっている.

$$A(A,B,F) \wedge \neg\exists a\exists b\exists f(A(a,b,f)\wedge$$
$$((B\supset F)\supset(b\supset f))\wedge(((B\supset F)\equiv(b\supset f))\supset((A\supset\neg F)\supset(a\supset\neg f)))\wedge \quad (3)$$
$$\neg(((b\supset f)\supset(B\supset F))\wedge(((b\supset f)\equiv(B\supset F))\supset((a\supset\neg f)\supset(A\supset\neg F))))) \quad (4)$$

(3),(4) の部分はそれぞれ (1),(2) に対応している. この変換はじつは機械的に行うことができる. くわしい議論は [Satoh90] を参照していただきたい.

ここで $A(A,B,F)$ として, $(B\supset A)\wedge A$, すなわち, Tweety が生き物であることがわかったとする. すると, 上の式は, 以下のようになる.

$$(B\supset A)\wedge A\wedge\neg\exists a\exists b\exists f((b\supset a)\wedge a\wedge$$
$$(B\supset F)\supset(b\supset f)\wedge(((B\supset F)\equiv(b\supset f))\supset((A\supset\neg F)\supset(a\supset\neg f)))\wedge$$
$$\neg((b\supset f)\supset(B\supset F)\wedge(((b\supset f)\equiv(B\supset F))\supset((a\supset\neg f)\supset(A\supset\neg F)))))$$

命題論理の場合、命題論理変数に対する存在限量された式 $\exists p\Phi(p)$ は $\Phi(\mathbf{F})\vee\Phi(\mathbf{T})$ と同値であるから[9], 上の式の各 a,b,f に対してその変換を施せば, 上の式は通常の命題論理式に変換される. 変換した論理式を簡単化すると結局 $A\wedge\neg B\wedge\neg F$ が得られる. この結果は, 図2における最も起こりやすいモデルに対応している.

もしここで Tweety が鳥であることがわかったとする. すると, $A(A,B,F)$ は, $(B\supset A)\wedge A\wedge B$ すなわち、$A\wedge B$ となる. 上と同様の処理を施すと $A\wedge B\wedge F$ が得られる. すなわち, 飛ぶことに関して結論が変化した. また, この結果は, 図3における最も起こりやすいモデルに対応している.

4　おわりに

本稿では, 非単調推論を2つのアプローチにわけて基本メカニズムの解説を行った. これらのアプローチはどちらも計算不可能性の問題を持つため論理式を制限して役に立つ計算可能なクラスを見つけなければならない. このことに関係して論理プログラミングの意味論と非単調推論の定式化には密接な関係があることが最近わかってきた [Przymusinski87, Gelfond88]. このことは非単調推論の側からみると論理プログラミングの計算がそのまま非単調推論の計算として使えることを意味するので重要である.

この2つのアプローチのうち極小モデルに基づくアプローチの方が優先順位付き極小限定のように常識の優先順位を自然に表現できるので[10], 常識推論の表現の観点からみるとすぐれているように思われる.

[9] $\mathbf{F}$ は偽を表し, $\mathbf{T}$ は真を表す.

[10] もちろん無矛盾性に基づくアプローチでもある程度, 優先順位をつけることができる. [Etherington87] を見よ.

非単調推論の応用については述べなかったが，[Reiter80] や [McCarthy86] にはいろいろな応用が述べられている．また，[Nonmon87] には非単調推論における著名な論文が載せられているのでより興味のある方はそちらを参照していただきたい．

謝辞

本稿に対して貴重な御助言を頂いた ICOT の太田，岩山氏に感謝致します．

参考文献

[Etherington87] Etherington, D. W., Formalizing Nonmonotonic Reasoning Systems, *Artificial Intelligence*, Vol. 31, pp. 41 – 85 (1987).

[Gelfond88] Gelfond, M., Lifschitz, V., The Stable Model Semantics for Logic Programming, *Proc. of LP'88*, pp. 1070 – 1080 (1988).

[Nonmon87] Ginsberg, M. L., ed, Readings in Nonmonotonic Reasoning, Morgan Kaufmann Publishers (1987).

[Konolige88] Konolige, K., On the Relation between Default and Autoepistemic Logic, *Artificial Intelligence*, Vol. 35, pp. 343 – 382 (1988), この論文には間違いがあり訂正が *Artificial Intelligence*, Vol. 41, pp. 115 にある．

[Levesque90] Levesque, H., All I Know: A Study in Autoepistemic Logic, *Artificial Intelligence*, Vol. 42, pp. 263 – 309 (1990).

[Lifschitz84] Lifschitz, V., Some Results on Circumscription, *Proc. of Workshop on Nonmonotonoic Reasoning*, pp. 151 – 164 (1984).

[Marek89] Marek, W., Truszczynski, M., Relating Autoepistemic and Default Logics, Proc. of KR'89, pp. 276 – 288 (1989).

[McCarthy80] McCarthy, J.: Circumscription - a Form of Non-monotonic Reasoning, *Artificial Intelligence*, Vol. 13, pp. 27 – 39 (1980).

[McCarthy86] McCarthy, J.: Applications of Circumscription to Formalizing Common-sense Knowledge, *Artificial Intelligence*, Vol. 28, pp. 89 – 116 (1986).

[Moore84] Moore, R. C.: Possible-World Semantics for Autoepistemic Logic, *Proc. of Workshop on Non-monotonic Reasoning*, pp. 344 – 354 (1984).

[Moore85] Moore, R. C.: Semantical Considerations on Nonmonotonic Logic, *Artificial Intelligence*, Vol. 25, pp. 75 – 94 (1985).

[中川 89] 中川 裕志: 極小限定の研究動向，情報処理，Vol. 30, No. 6, pp. 684 – 693 (1989).

[Przymusinski87] Przymusinski, T. C.: On the Declarative Semantics of Deductive Databases and Logic Programs, *Foundations of Deductive Databases and Logic Programming*, Morgan Kaufmann, pp. 193 – 216 (1987).

[Reiter80] Reiter, R.: A Logic for Default Reasoning, *Artificial Intelligence*, Vol. 13, pp. 81 – 132 (1980).

[Satoh90] Satoh, K.: Formalizing Nonmonotonic Reasoning by Preference Order, *Proc. of InfoJapan'90*, pp. 155 – 162 (1990).