

否定推論と肯定推論を同時に実行できる コネクショニスト推論システム

麻生川 稔[†]

本研究は、「ニューラルネットワークを用いた知識獲得」に関するものである。目的は、知られているほぼ正しい知識を、知られているほぼ正しい例題とニューラルネットワークの学習機能を用いて洗練することである。これは、知られているほぼ正しい知識は、不完全であったり矛盾を含んでいたりするので、ニューラルネットワークの学習によって足りない部分を補い矛盾している部分を訂正するものである。このニューラルネットワークの特徴としては：(1) 肯定推論と否定推論を同時に行うことができる。(2) 例題が正しく解釈できるように、知識が完全化かつ完備化される。(3) 真と偽以外にこれらの中間状態や、矛盾等が表現可能で、非単調性推論の機能を有する。(4) ニューラルネットワークでは入力と出力の区別がないので、知られている事実を入力とすると、知られていないことについての推論が行われる。(5) 知られている事実が不十分であるときは、真偽不明と答えることもできる。これまでのエキスパートシステムでは上記した機能の一部は実現可能であるが、すべてを同時に実現したものはない。正しい知識をニューラルネットワークに展開する方法と、それを用いた推論方法について説明し、推論時のシステムの動作がつねに平衡状態に到達することをリヤプノフ関数の存在を示すことにより証明した。さらに、例題からの学習方法等について説明し、実装例を用いて実現方法の妥当性について検討する。

A Connectionist Production System Which can Perform Both Modus Ponens and Tollens Simultaneously

MINORU ASOGAWA[†]

Modus Ponens is used in forward inference and backward inference where the truth of the conclusion is inferred from the truth of the premise. On the other hand, in Modus Tollens, the falseness of the premise is inferred from the falseness of the conclusion. Although Modus Ponens is used in general connectionist production systems, Modus Tollens is rarely used. The authors have developed a novel connectionist production system, which can perform both Modus Ponens and Tollens simultaneously. One of the advantages of this system is that it has a learning mechanism. Often, the rules and examples given as external knowledge are erroneous and incomplete. Rules are refined with supervised learning, which is performed presenting both positive and negative examples to this system on which external rules and observations are mapped. The system stability is proved by using a Lyapunov function. Moreover, this system implementations of an implication, a conjunction, a disjunction and a negation are intuitively consistent with boolean logic and discussed by presenting a simple example implementation.

1. はじめに

本研究は、「ニューラルネットワークを用いた知識獲得」に関するものである。目的は、知られているほぼ正しい知識を、知られているほぼ正しい例題とニューラルネットワークの学習機能を用いて洗練することである。これは、知られているほぼ正しい知識は、不完全であったり矛盾を含んでいたりするので、ニューラ

ルネットワークの学習によって足りない部分を補い矛盾している部分を訂正するものである。このような目的に対し、コネクショニスト的なエキスパートシステムが多用されている^{1)~10)}。コネクショニスト的なエキスパートシステムでは、ルールの前件部や後件部の各命題をセルに対応付け、命題間の関係であるルールをセル間のリンクに対応付ける。たとえば、「りんごが赤いならば、熟れている」というルールが与えられたとき、赤いと熟れているとがそれぞれ1つのセルに対応付け、ルール自体は赤いセルより熟れているセルへのリンクに対応付ける。もし、前件部の赤いセルに真

[†] 日本電気株式会社 C&C メディア研究所
C&C Media Research Laboratories, NEC Corporation

の状態が割り当てられたら、その状態はリンクを通じて、後件部の熟れているセルに真の状態を伝搬する。つまり、前件部が真であるときに後件部が真である肯定推論（モダス・ポネンス）を行う。

この推論の他に、否定推論（モダス・トレンス）がある。これは、「熟れていない（熟れているの否定）」という事実を知ったとき、「赤くない（赤いの否定）」であることを推論することである。否定推論は、従来のエキスパートシステムで使われている後向き推論とは異なる。後向き推論は肯定推論の1つの実現方法である。つまり、後向き推論では、「熟れている」（後件部が真）ということ推論するためには、「赤い」（前件部が真）という事実があればよいことをルールから知り、前件部が真である事実があるかどうかを調べ、その結果、前件部が真であることを示せたら、後件部も真であると推論する。コネクショニスト的なエキスパートシステムを含め従来のエキスパートシステムでは、肯定推論である前向き推論または後向き推論を用いており、否定推論（モダス・トレンス）は積極的に活用されていない。つまり、知識の一部のみしか利用していないことになる。

本研究で提案するシステムは、与えられた知識を用いて肯定推論と否定推論を同時に行うという機能以外に、以下のような機能を有する。

- 命題に対し、真と偽以外にこれらの中間状態や、矛盾等が表現可能で、非単調性推論の機能を有する。
- 各命題に知られている事実を入力とすると、真偽値が未知の命題の値を前向き推論と後向き推論を用いて推論できる。
- 知られている事実が不十分であるときは、真偽不明と答えることもできる。
- 与えられた例題を用いた学習により、知識が完全化かつ完備化される。

これまでの推論システムでは上記した機能の一部は実現可能であるが、すべてを同時に実現したものはない。

2. 従来法

従来より推論に用いられてきたコネクショニスト的な方式には、各命題を1つのセルに対応させたコネクショニスト的なエキスパートシステムやその拡張、ボルツマンマシン等がある。これらの方式においては、否定推論に対して適用はなされておらず、命題に対して真偽不明と矛盾との区別をつけることができない等の問題点があった。同様の問題が、Fuzzy理論を用い

たシステムにもあった。また、デンプスターシェファ理論¹¹⁾を含むいくつかのシステムでは、真と偽以外にこれらの中間状態や、真偽不明を表現することができ、非単調性推論の機能を有するが、例題からの学習に対する機能が欠けていた。

以下、各システムについて概観する。

コネクショニスト的なエキスパートシステムでは、学習に各セルを命題に対応させ、ルールをセル間のリンクに対応させており、学習にはバックワードプロパゲーション^{1)~7)}、ヘップ則に基づくアルゴリズム^{8),9)}、ジェネティックアルゴリズム¹⁰⁾を用いる。これらのシステムでは、前件部が偽であるときに、後件部を偽と推論してしまう。しかしながら、論理的には、後件部は真でも偽でも成り立ち、真か偽かのどちらかに一意に決めることはできないはずである。それ以外に、否定推論を行うことができず、真偽不明と矛盾との区別をつけることができない等の問題点がある。各命題の否定を用いた補完表現を適用したコネクショニスト的なエキスパートシステムにおいても（補完表現以外に本システムで述べるような機構を有しないと）、否定推論を行うことができない。

μ LKONE¹²⁾は、コネクショニスト的なエキスパートシステムであるが、ボルツマンマシンを用いている。推論では、観測データを正しいものとし対応するセルに設定した後にアニーリングすることにより行う。ルールはセル間のリンクの重みとして表現されており、観測データを入力とすると、観測されていない命題について推論が行われる。DCPS¹³⁾も同様に、ボルツマンマシンを用いた、1階の述語論理の表現を可能としたコネクショニスト的なエキスパートシステムである。しかしながら、これらのシステムでは、否定推論を行うことができず、真偽不明と矛盾との区別をつけることができない等の問題点が残る。

Fuzzy理論を用いたシステムでは^{14)~17)}真と偽以外にこれらの中間状態を表現できるが、否定推論を行うことができず、真偽不明と矛盾との区別をつけることができない等の問題点がある。

デンプスターシェファ理論¹¹⁾やInferno¹⁸⁾やGinsberg¹⁹⁾が提案しているシステムでは、真と偽以外にこれらの中間状態や、真偽不明を表現することができ、非単調性推論の機能を有する。また、Inferno¹⁸⁾やGinsberg¹⁹⁾が提案しているシステムでは、肯定推論と否定推論を同時に行うことができるが、学習機能に欠けるため、例題が正しく解釈できるように知識を変更する手段はない。

3. 肯定推論と否定推論を同時に実現する方法

3.1 内部解釈の状態を2つのセルの活性状態で表現

命題の内部解釈の状態を表現する体系として、Inferno¹⁸⁾やデンプスターシェファ理論¹¹⁾のように、真であることを示すセルと偽であることを示すセルの2つを用いる。真セル（以下、真であることを示すセルをこのように省略する）と偽セル（偽であることを示すセル）はその活性度として $[0.0, 1.0]$ の範囲の実数をとる。この2つの実数値の解釈は、極端な場合として、真セルの活性度が 1.0 で偽セルが 0.0 のときは真と解釈し、反対に真セルの活性度が 0.0 で偽セルが 1.0 のときは偽と解釈する。上記以外の活性度の組合せとして、中間的な値の組合せも可能である。特に、真セルと偽セルの活性度がともに 0.0 の場合は、真であるとも偽であるとも断言できない場合に対応するので、直観的には真偽不明と解釈し、ともに 1.0 の場合は真であるとも偽であるともいえるので矛盾と解釈する（デンプスターシェファ理論においては矛盾の状態を許さず、真と偽と真偽不明の総和が 1.0 になるように保たれる）。

3.2 肯定推論と否定推論の実現

まず図1に示すように、ルールの前件部と後件部の各命題をそれぞれ、真セルと偽セルで表現する、つまり都合4つのセルを使って表現する。ルールは、2つの一方向リンクを用いて表現する。1つめの一方向リンクは、前件部の真セルから後件部の真セルへの一方向リンクである。この一方向リンクは2つの命題の真

セルを接続する。2つめの一方向リンクは、後件部の偽セルから前件部の偽セルへの一方向リンクであり、2つの命題の偽セルをつないでいる。注意していただきたいのは、真セル間の一方向リンクの方向と偽セル間の一方向リンクの方向は逆であることを除いては、一方向にしか情報を通さない機能は同じということである。一方向リンクは、リンクの出ているセルの活性度はリンクが入ってくるセルの活性度よりつねに大きいという状態に保つもので、この関係が成立しない場合は、その関係を満たすようにリンクの入ってくるセルの活性度が変化する（この実現方法については、6.2節に述べる）。

3.2.1 肯定推論

肯定推論は、前件部が真であるときに後件部が真と推論することである。まず、前件部が真であることは、前件部の真セルの活性度が 1.0 で偽セルが 0.0 として表現される。また、推論を始める前は後件部に対して何も分かっていないので真偽不明、すなわち真セルと偽セルの活性度が 0.0 であるとする（図2参照：図中ではセルの活性度を棒グラフで示した）。

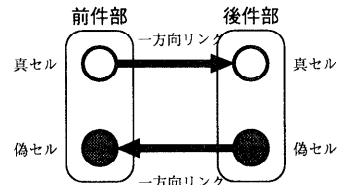


図1 ルールの表現方法
Fig. 1 Rule representation.

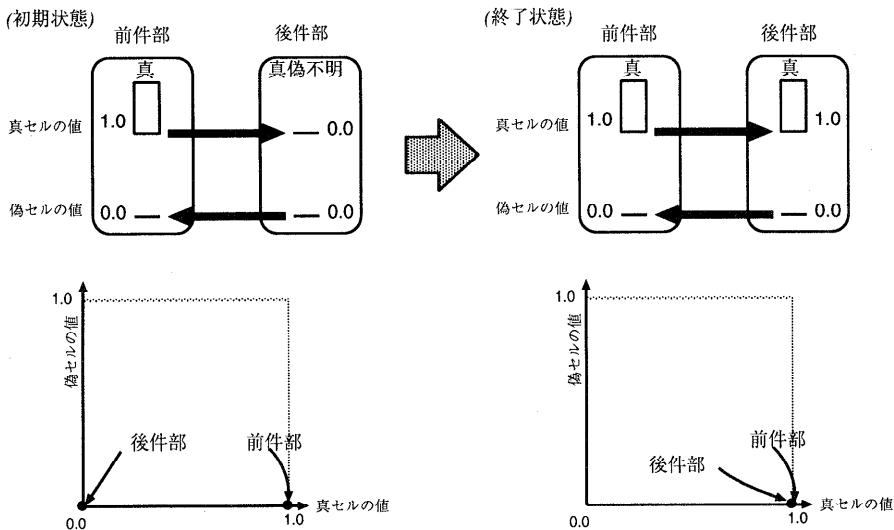


図2 肯定推論の実現

Fig. 2 Modus ponens realization.

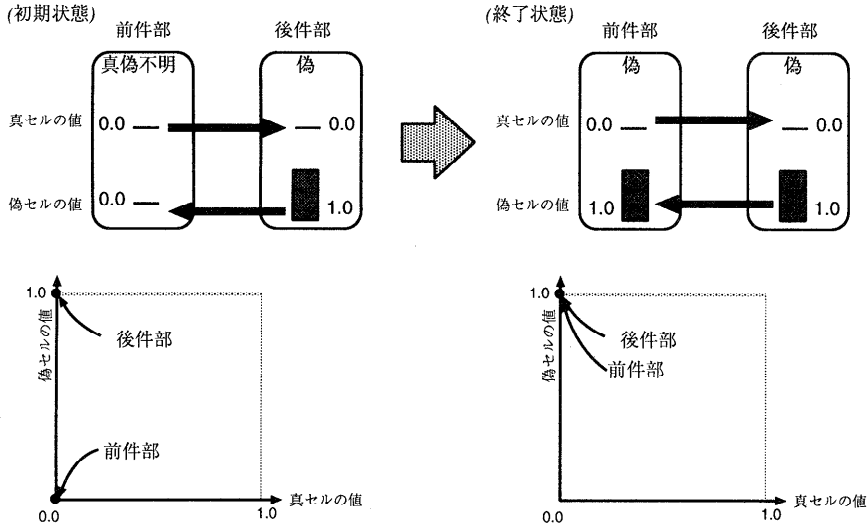


図3 否定推論の実現
Fig. 3 Modus tollens realization.

前件部の値		後件部の値			
		真 (1,0)	偽 (0,1)	真偽不明 (0,0)	矛盾 (1,1)
真 (1,0)	安定	不安定	不安定	不安定	
偽 (0,1)	安定	安定	安定	安定	
真偽不明 (0,0)	安定	不安定	安定	不安定	
矛盾 (1,1)	安定	不安定	不安定	安定	

図4 推論の値の組合せの安定性
Fig. 4 Stability of implication.

前件部の真セルの活性度が一方向リンクを通して後件部の真セルに伝搬し、その結果として活性度が0.0より1.0に変化する。また、後件部の偽セルから一方向リンクは出ていく方向だけなので、どのセルからも影響を受けないため活性度は0.0より変化しない。これらの結果、後件部の真セルの活性度は1.0になり、偽セルは0.0となり、後件部は真と解釈される。すなわち、前件部が真であるときに、後件部も真であると肯定推論できる。

3.2.2 否定推論

肯定推論は、後件部が偽であるときにその前件部が偽と推論することである。

後件部が偽であることは、後件部の真セルの活性度が0.0で偽セルが1.0として表現される(図3参照)。後件部の偽セルの活性度が一方向リンクを通して、前件部の偽セルに伝搬し、活性度が0.0より1.0に変化する。同時に、前件部の真セルから一方向リンクは出ている方向だけなので、どのセルからも影響を受けないので、活性度は0.0より変化をしない。この結果、前件部の真セルと偽セルの活性度は0.0と1.0になり、偽と解釈される。すなわち、後件部が偽であると

きに、前件部が偽であると否定推論できる。

3.2.3 2値論理との関係

一方向リンクを用いたとき、推論において前件部と後件部の値のうち、どの組合せが安定して成立しうるかを図4に示した。不安定となる場合は、前件部または後件部の値が一意に変化し安定な状態に落ち着く。

2値の論理では、含意 $A \rightarrow B$ の真偽値は、 $\neg A \vee B$ と等しく、その真偽値は図5のようになる。ここで、図4の安定と不安定をそれぞれ、TとFと解釈しなおせば、2値の論理真偽値と一致する。

4. 連言と選言の実現

連言「AかつB」の演算は、図6に示すように定義する。すなわち、「AかつB」の真セルの値は、Aの真セルとBの真セルの最小値とし、「AかつB」の偽セルの値は、Aの偽セルとBの偽セルの最大値とする(以下では説明を簡単にするために、最大値と最小値関数を用いているが、実際には6.3節に示す関数を用いる。この関数は、その極限においては最大値最小値関数と同じであるために、これらを用いて説明する)。

これによって、連言の命題で1つでも偽があれば、

		B	
		T	F
A	T	T	F
	F	T	T

図5 A→Bの論理表

Fig. 5 Boolean value of A→B.

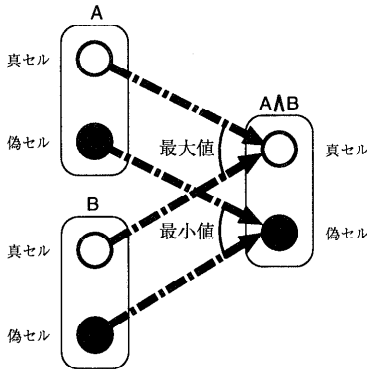


図6 連言の展開方法

Fig. 6 Conjunction realization.

	真 (1,0)	偽 (0,1)	真偽不明 (0,0)	矛盾 (1,1)
真 (1,0)	真 (1,0)	偽 (0,1)	真偽不明 (0,0)	矛盾 (1,1)
偽 (0,1)	偽 (0,1)	偽 (0,1)	偽 (0,1)	偽 (0,1)
真偽不明 (0,0)	真偽不明 (0,0)	偽 (0,1)	真偽不明 (0,0)	偽 (0,1)
矛盾 (1,1)	矛盾 (1,1)	偽 (0,1)	偽 (0,1)	矛盾 (1,1)

図7 2つの命題による連言の値

Fig. 7 Conjunction value with two propositions.

連言自体の偽の値は 1.0 となる。また、連言の命題ですべて真であるときだけ、連言自体の真セルの値は 1.0 となる。このことは、連言の定義に直観的に矛盾しないし、2 値の論理にも合致する。この連言の演算の性格を示すために、2 つの命題の値によって連言がどのような値をとるかを、図 7 に示した。

選言「A または B」の演算では、連言の最大値と最小値の演算を入れ替えたものを用いる。

5. 否定の実現方法

A の否定 ($\neg A$) の真セルと偽セルの活性度は、A の偽セルと真セルの活性度と等しいとする。これは、A に対する真の度合は、A の否定の偽の度合であると解釈することによって理解できる。また、A に対する偽の度合についても同様である。

6. 推論時のネットワークの動作

簡単のために、前章までは各セルの値は入力情報が決まると即座にすべてのセルの値が決まるかのように説明した。実際には、各セルの値は動的に更新した結果決まる。また、例題からの学習が可能となるように、肯定推論と否定推論の一方方向リンクと連言と選言の定義を重みつきに拡張する。学習は、例題を提示し、重みを調整することによって行う。

以下、連言の推論知識の例を用いて、ニューラルネットワークの動作を説明する。

ルールは、

「もし B かつ C ならば、X である」

である。これは、図 8 に示すように展開する。ここで、外界の観測として、B と C が真であると分かっていたとする。ここで、本モデルでは外界からの事実知識とそれを解釈する状態は分けて考える。これは、外界からの事実知識を基に推論を進めると、「特定の事実知識が間違っていると解釈しないと全体との辻褃が残すためである。初期状態においては、図 8 に示すように展開する。図中に示す例では、命題 X に対する観測値は存在しないとした。観測以外の内部解釈の状態がすべて真偽不明 (0,0) となっている。これは、推論を開始する前は、観測された値以外は真とも偽とも決定できない状態であるためである。

各内部解釈の値は、観測値の影響によって初期状態から変化する。この例では、ルールの後件部に対して 3 種類の影響がある。

(1) 観測値から内部解釈への影響

これは、観測値が内部解釈におよぼす影響で、できる限り内部解釈が観測値に等しくなるように、内部解釈を変化させる。

(2) 肯定推論と否定推論の一方方向リンクによる影響

これは、2 つの一方方向性リンクを通して内部解釈の影響が伝搬する。

(3) 連言の各命題から連言のセルに対する影響

連言の真セルと偽セルの活性度は、連言の各命題の真セルと偽セルの活性度によって決められる。

(4) 内部解釈の値をもっともらしい範囲に保持する機構からの影響

各内部解釈に対して、なるべく矛盾に陥らないように保持する機構がある。具体的には、真セルと偽セルの値の和が 1.0 を超えないようにする。

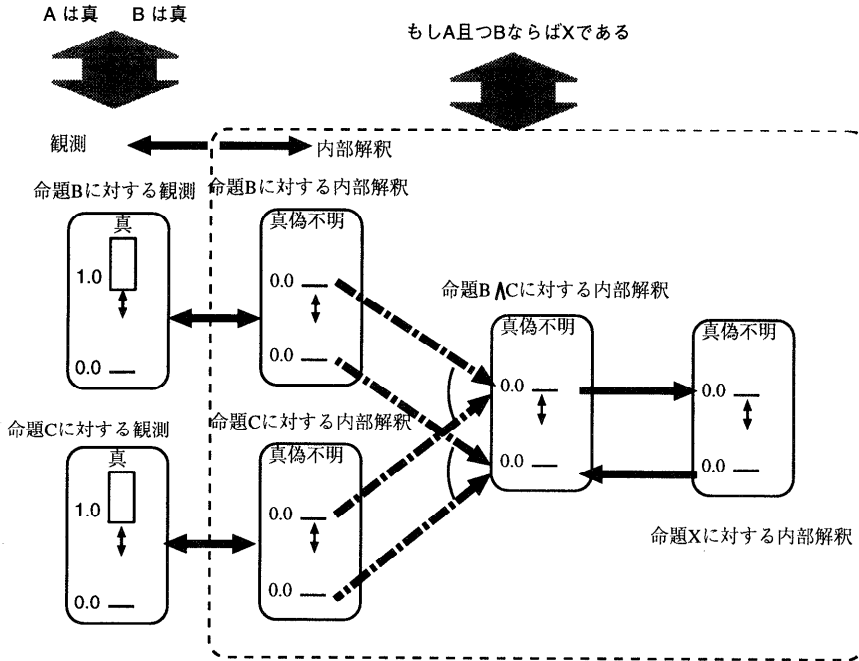


図8 観測と連言を含む知識の展開方法
Fig. 8 Representation of observations and rule with conjunction.

以下、各影響について述べる。

6.1 観測値から内部解釈への影響

これは、観測値が内部解釈に及ぼす影響で、できる限り内部解釈が観測値に等しくなるように、内部解釈を変化させる。これを実現するため、内部解釈と観測値の差によるエネルギーを定義して、それが少なくなるように、Gauss-Newton法を用いて、内部解釈の値を変化させる。注意しなくてはならないのは、内部解釈が変化しても、観測値は変化しないことである。

具体的に **B** に対しては、内部解釈と観測値の差によるエネルギー F_1 を以下のように定義する。

$$F_1 = \frac{W_{1,B}}{2} (B_{true}^{Obs} - B_{true}^{Belief})^2 + \frac{W_{1,B}}{2} (B_{false}^{Obs} - B_{false}^{Belief})^2 \quad (1)$$

ここで、 B_{true}^{Obs} と B_{false}^{Obs} は観測値の真セルと偽セルの値。 B_{true}^{Belief} と B_{false}^{Belief} は内部解釈の真セルと偽セルの値である。 $W_{1,B}$ は、**B** の観測値がどのくらい信頼できるかを示す重みである。特に、 $W_{1,B} = 0.0$ のときは、**B** に対する観測値がないことを示しており、 F_1 は恒等的に 0.0 になる。

内部解釈の時間変化量は、 F_1 に対しての内部解釈による偏微分係数に -1 を乗じたものを用いる。すなわち、

$$\frac{dB_{true}^{Belief}}{dt} \propto - \frac{\partial F_1}{\partial B_{true}^{Belief}} = W_{1,B} (B_{true}^{Obs} - B_{true}^{Belief}) \quad (2)$$

$$\frac{dB_{false}^{Belief}}{dt} \propto - \frac{\partial F_1}{\partial B_{false}^{Belief}} = W_{1,B} (B_{false}^{Obs} - B_{false}^{Belief}) \quad (3)$$

これによって、内部解釈の真セルと偽セルの活性度は、対応する観測値のそれに等しくなるまで増大し、この変化量は内部解釈と観測値の差 F_1 に比例する。したがって、 B_{true}^{Belief} の時間変化は Γ 関数に従い、時間が経過するにつれて、 B_{true}^{Belief} は B_{true}^{Obs} に限りなく近づく。 F_1 が 0 になると、式 (2), (3) の右辺は 0 となり、変化量が 0 となり変化が止まる。

6.2 肯定推論と否定推論の一方方向リンクによる影響

肯定推論と否定推論の一方方向リンクによる影響は、3章に述べたように、2つの一方方向性リンクを通して影響が伝搬する。学習を可能にするために、重み付きの肯定推論と否定推論を以下のように定義する。

確信度 W_r で「**A** が真ならば **X** は真」であるという知識が与えられたときに、図9に示すように、重み付き一方方向リンクを用いて展開する。重み付き一方方向性リンクは、重み付き一方方向性リンクの両端のセルの活性度による差 F_3 を以下のように定義し、この偏微分を用いることにより実現する。ここで、 A_{true}^{Belief} と

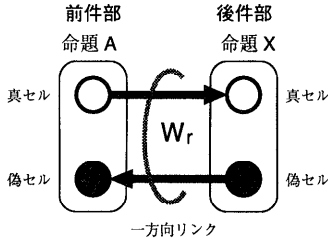


図9 重み付きルールの表現方法
Fig. 9 Weighted rule representation.

A_{false}^{Belief} は命題 A の真セルと偽セルの活性度, X_{true}^{Belief} と X_{false}^{Belief} は命題 X の真セルと偽セルの活性度である。

$$F_3 = \frac{1}{2}m(W_r A_{true}^{Belief} - X_{true}^{Belief})^2 + \frac{1}{2}m(W_r X_{false}^{Belief} - A_{false}^{Belief})^2 \quad (4)$$

ここで、関数 $m(\cdot)$ は以下のような関数。

$$m(x) = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

このように定義すると、リンクの出ているセルの活性度がリンクが入ってくるセルのよりも大きければ F_3 は 0 であるが、そうでない場合は正の値を持つ。内部解釈の時間変化量は、 F_3 に対しての各内部解釈による偏微分係数を用いる。ここで、 A_{true}^{Belief} と X_{false}^{Belief} に対しては、 F_3 の影響によって変化させない。

$$\frac{dX_{true}^{Belief}}{dt} \propto -\frac{\partial F_3}{\partial X_{true}^{Belief}} = m(W_r A_{true}^{Belief} - X_{true}^{Belief}) \quad (6)$$

$$\frac{dA_{false}^{Belief}}{dt} \propto -\frac{\partial F_3}{\partial A_{false}^{Belief}} = m(W_r X_{false}^{Belief} - A_{false}^{Belief}) \quad (7)$$

ルールの確信度 $W_r \neq 0.0$ のとき、重み付き一方向リンクを用いて肯定推論と否定推論が可能であることは、前に述べた。

ルールの確信度 $W_r = 0.0$ のとき、 $F_3 = 0.0$ となり、すべてのセルに関して F_3 は影響を与えない。したがって、前件部と後件部の各セルの値は、 F_3 によって変化しない。これはルールがまったくくないのと同じであり、直観と矛盾しない。

6.3 連言の各命題から連言のセルに対する影響

連言の真セルと偽セルの活性度は、連言の各命題の真セルと偽セルの活性度によって決まる。重み付き連

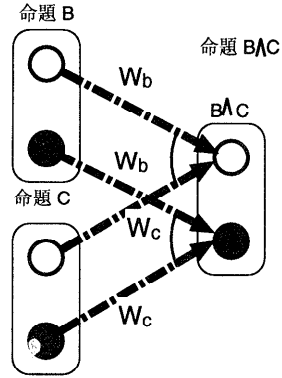


図10 重み付き連言
Fig. 10 Weighted conjunction.

言は、図10に示すように展開する。 W_B は命題 B に対する重み付けで、 W_C は命題 C に対する重み付けであり、それぞれ、 $[0.0, 1.0]$ の範囲の実数である。命題 B と命題 C の活性度で決まる命題 $B \wedge C$ の活性度は、具体的には $And_{true}(\cdot)$ 関数と、 $And_{false}(\cdot)$ 関数により与えられる。これらの定義は、Inferno で用いられて定義とよく似ているが、重み付きである点が異なる。

$$And_{true}(B^{Belief}, C^{Belief}; W_B, W_C) = \{1 - W_B(1 - B_{true}^{Belief})\} \times \{1 - W_C(1 - C_{true}^{Belief})\} - (1 - W_B)(1 - W_C) \quad (8)$$

$$And_{false}(B^{Belief}, C^{Belief}; W_B, W_C) = 1 - (1 - W_B B_{false}^{Belief})(1 - W_C C_{false}^{Belief}) \quad (9)$$

特に、 $W_B = W_C = 1.0$ の場合、式(8)、(9)はそれぞれ以下ようになる。

$$And_{true}(B^{Belief}, C^{Belief}; 1.0, 1.0) = B_{true}^{Belief} C_{true}^{Belief} \quad (10)$$

$$And_{false}(B^{Belief}, C^{Belief}; 1.0, 1.0) = 1 - (1 - B_{false}^{Belief})(1 - C_{false}^{Belief}) \quad (11)$$

これによって、連言の命題に1つでも偽があれば、連言自体の偽の値は 1.0 となる。また、連言の命題がすべて真であるときだけ、連言自体の真セルの値は 1.0 となる。このことは、連言の定義に直観的に矛盾しないし、2値の論理にも合致する。

また、 $W_B = 0.0$ の場合、式(8)、(9)はそれぞれ、以下ようになる。

$$\begin{aligned} & And_{true}(B^{Belief}, C^{Belief}; 0.0, W_C) \\ & = W_C C_{true}^{Belief} \end{aligned} \quad (12)$$

$$\begin{aligned} & And_{false}(B^{Belief}, C^{Belief}; 0.0, W_C) \\ & = W_C C_{false}^{Belief} \end{aligned} \quad (13)$$

これによって、連言の値は、命題 **C** の真セルの活性度に W_C を重み付けした値になり、直観に矛盾しない。特に、 $W_B = W_C = 0.0$ とすると、連言の値は真偽不明となり、直観に矛盾しない。

一般的に連言が n の命題で構成されているときは、式 (8), (9) は以下ようになる。

$$\begin{aligned} & And_{true}(I^{Belief}; W_i)^{i=1\dots n} \\ & = \prod_{i=1}^n \{1 - W_i(1 - I_{true}^{Belief})\} - \prod_{i=1}^n (1 - W_i) \end{aligned} \quad (14)$$

$$\begin{aligned} & And_{false}(I^{Belief}; W_i)^{i=1\dots n} \\ & = 1 - \prod_{i=1}^n (1 - W_i I_{false}^{Belief}) \end{aligned} \quad (15)$$

ここで、 I^{Belief} は命題 i の内部解釈の値で、 W_B は命題に対する重み付けである。

式 (14), (15) を変形すると、以下ようになる。

$$\begin{aligned} & And_{true}(I^{Belief}; W_i)^{i=1\dots n} \\ & = And_{true}(I_{true}^{Belief}; W_i)^{i=1\dots n-1} \\ & \quad \times \{1 - W_n(1 - N_{true}^{Belief})\} \\ & \quad - W_n N_{true}^{Belief} \prod_{i=1}^{n-1} (1 - W_i) \end{aligned} \quad (16)$$

$$\begin{aligned} & And_{false}(I^{Belief}; W_i)^{i=1\dots n} \\ & = And_{false}(I_{false}^{Belief}; W_i)^{i=1\dots n-1} \\ & \quad \times (1 - W_n N_{false}^{Belief}) + W_n N_{false}^{Belief} \end{aligned} \quad (17)$$

ここで、 $W_n = 0.0$ とすると、以下のようになり、第 n 番めの命題がないときの And 関数と一致し、直観に矛盾しない。

$$\begin{aligned} & And_{true}(I^{Belief}; W_i)^{i=1\dots n} \\ & = And_{true}(I_{true}^{Belief}; W_i)^{i=1\dots n-1} \end{aligned} \quad (18)$$

$$\begin{aligned} & And_{false}(I^{Belief}; W_i)^{i=1\dots n} \\ & = And_{false}(I_{false}^{Belief}; W_i)^{i=1\dots n-1} \end{aligned} \quad (19)$$

さらに、 $W_1 = \dots = W_{n-1} = 0.0$ とすると、以下のようになり、第 n 番めだけの命題があるときの And 関数と一致し、直観に矛盾しない。

$$\begin{aligned} & And_{true}(I^{Belief}; W_i)^{i=1\dots n} \\ & = W_n N_{true}^{Belief} \end{aligned} \quad (20)$$

$$\begin{aligned} & And_{false}(I^{Belief}; W_i)^{i=1\dots n} \\ & = W_n N_{false}^{Belief} \end{aligned} \quad (21)$$

システムを動作させるとき、必ずしも、命題 **B** \wedge **C** の真セルの活性度は、命題 **B** と命題 **C** の真セルの And 関数の値と一致していないことがあるので、各セルの値を変更することが必要となる。このため、連言のセルの値と各命題のセルの差によるエネルギー F_2 を定義し、これが少なくなるように連言のセルと各命題の値を変化させる。

$$\begin{aligned} F_2 = & \frac{1}{2} \{B \wedge C_{true}^{Belief} \\ & - And_{true}(B^{Belief}, C^{Belief}; W_B, W_C)\}^2 \\ & + \frac{1}{2} \{B \wedge C_{false}^{Belief} \\ & - And_{false}(A^{Belief}, B^{Belief}; W_B, W_C)\}^2 \end{aligned} \quad (22)$$

ここで、 $B \wedge C_{true}^{Belief}$ は命題 **B** \wedge **C** の真セルの活性度、 B_{true}^{Belief} は命題 **B** の真セルの活性度、 C_{true}^{Belief} は命題 **C** の真セルの活性度である。内部解釈の時間変化量は、 F_2 に対しての各内部解釈による偏微分係数を用いる。すなわち、各真セルの活性度に対しては、

$$\begin{aligned} & \frac{dB \wedge C_{true}^{Belief}}{dt} \propto - \frac{\partial F_2}{\partial B \wedge C_{true}^{Belief}} \\ & = And_{true}(B^{Belief}, C^{Belief}; W_B, W_C) \\ & \quad - B \wedge C_{true}^{Belief} \end{aligned} \quad (23)$$

$$\begin{aligned} & \frac{dB_{true}^{Belief}}{dt} \propto - \frac{\partial F_2}{\partial B_{true}^{Belief}} \\ & = (B \wedge C_{true}^{Belief} \\ & \quad - And_{true}(B^{Belief}, C^{Belief}; W_B, W_C)) \\ & \quad \times \frac{\partial And_{true}(B^{Belief}, C^{Belief}; W_B, W_C)}{\partial B_{true}^{Belief}} \end{aligned} \quad (24)$$

$$\begin{aligned} & \frac{dC_{true}^{Belief}}{dt} \propto - \frac{\partial F_2}{\partial C_{true}^{Belief}} \\ & = (B \wedge C_{true}^{Belief} \\ & \quad - And_{true}(B^{Belief}, C^{Belief}; W_B, W_C)) \\ & \quad \times \frac{\partial And_{true}(B^{Belief}, C^{Belief}; W_B, W_C)}{\partial C_{true}^{Belief}} \end{aligned} \quad (25)$$

実際に動作させると、各セルの活性度が $[0.0, 1.0]$ の範囲を超えてしまうことがあるので、セルの活性度が範囲内に収まるようにエネルギー F_4 を以下のように定義することによって防ぐ。

$$F_4 = \frac{1}{2}l(A_{true}^{Belief})^2 + \frac{1}{2}l(A_{false}^{Belief})^2 \quad (26)$$

ここで、関数 $l(\cdot)$ は以下のような関数。

$$l(x) = \begin{cases} -x & \text{if } x < 0 \\ (x-1) & \text{if } x > 1 \\ 0 & \text{otherwise} \end{cases} \quad (27)$$

上記の説明では、連言の内部解釈セルの値は *And* 関数の値に収束するので、時間変化量を求め、動的な更新は無意味なように思えるが、連言の内部解釈が他の要因によって影響を受ける場合は、連言の内部解釈セルの値は *And* 関数の値に収束しない (図 15 参照)。

6.4 選言の各命題から選言のセルに対する影響

選言の真セルと偽セルの活性度は、選言の各命題の真セルと偽セルの活性度によって決まる。重み付き選言は、命題 **B** と命題 **C** の活性度で決まる、命題 **BVC** の活性度は、 $Or_{true}(\cdot)$ 関数と、 $Or_{false}(\cdot)$ 関数により与えられる。

$$Or_{true}(B^{Belief}, C^{Belief}; W_B, W_C) = 1 - (1 - W_B B_{true}^{Belief})(1 - W_C C_{true}^{Belief}) \quad (28)$$

$$Or_{false}(B^{Belief}, C^{Belief}; W_B, W_C) = \{1 - W_B(1 - B_{false}^{Belief})\} \times \{1 - W_C(1 - C_{false}^{Belief})\} - (1 - W_B)(1 - W_C) \quad (29)$$

これ以外は、重み付き連言と同様である。セルの時間変化量も重み付き連言と同様であり、 $And_{true}(\cdot)$ と $And_{false}(\cdot)$ が入れ替わった形式になる。さらに、一般形についても同様である。

$$Or_{true}(I^{Belief}; W_i)^{i=1\dots n} = 1 - \prod_{i=1}^n (1 - W_i I_{true}^{Belief}) \quad (30)$$

$$Or_{false}(I^{Belief}; W_i)^{i=1\dots n} = \prod_{i=1}^n \{1 - W_i(1 - I_{false}^{Belief})\} - \prod_{i=1}^n (1 - W_i) \quad (31)$$

ここで、 I^{Belief} は命題 i の内部解釈の値で、 W_B は命題に対する重み付けである。

6.5 内部解釈の一貫性を保持

内部解釈の一貫性の保持は、真セルと偽セルの活性度が排他的であるように保持する。これによって、たとえば内部解釈 **A** に対しては、 A_{true}^{Belief} と A_{false}^{Belief} の和が 1.0 以下となり、矛盾をさけることができ、直観的に理解しやすくなる。

このために、内部解釈の一貫性によるエネルギー F_5 を以下のように定義する。

$$F_5 = m(A_{true}^{Belief} + A_{false}^{Belief} - 1.0) \quad (32)$$

内部解釈の真セルと偽セルの時間変化量は、 F_5 に対しての各内部解釈による偏微分係数を用いる。

$$\begin{aligned} \frac{dA_{true}^{Belief}}{dt} &\propto -\frac{\partial F_5}{\partial A_{true}^{Belief}} \\ &= \begin{cases} -1 & \text{if } A_{true}^{Belief} + A_{false}^{Belief} - 1.0 > 0 \\ 0 & \text{otherwise} \end{cases} \end{aligned} \quad (33)$$

6.6 推論時のシステムの動作と安定性

推論時のシステム全体的な動作は、以下のように行う。

- (1) 観測値を設定する。
- (2) システム全体のエネルギー総和 $F (= \sum_i F_i)$ を減少させるように各内部解釈セルの値を同期的に変更し、全体的に落ち着くところまでシステムを動作させる

このように推論時のシステムを動作させると、必ず各セルの活性度の時間微分がゼロとなる平衡状態に到達する。このことを、システム全体のエネルギー総和 F がリヤプノフ関数であることを示す。そのため、 F に下限値があり、かつ、各セルの値を更新すると F が単調に減少することを示す。

エネルギーの定義式 (1), (4), (22), (26) より、 $F_1 \sim F_4$ は 2 乗項の和であり、式 (32) より、 F_5 は下限値がある。したがって、システム全体のエネルギー総和 F は下限値がある。

また、各セルの活性度の時間微分は、エネルギー関数に対する偏微分の負で定義されている。したがって、システム全体のエネルギー総和 F の時間微分は、以下に示すようになる。

$$\begin{aligned} \frac{dF}{dt} &= \sum_j \frac{\partial F}{\partial x_j} \frac{dx_j}{dt} \\ &= \sum_j -\epsilon \left(\frac{\partial F}{\partial x_j} \right)^2 \\ &\leq 0 \end{aligned} \quad (34)$$

ここで、 x_j は各セルの活性度であり、 $-\epsilon$ はセルの活性度の時間微分のパラメータである。上式よりシステム全体のエネルギー総和 F は単調に減少することが示された。

以上より、システム全体のエネルギー総和 F がリヤプノフ関数であることが示され、各セルの活性度を

Cup :- Stable AND OpenV AND Lifiable
 Stable :- Flat
 Lifiable :- Grasp AND Light
 OpenV :- Concavity AND Pointsup
 Grasp :- Handle AND Sideh

領域知識

Case	1	2	3	4	5	6	7	8	9	10	11	12	13
Handle	T	F	F	T	T	F	T	F	F	F	T	T	T
TopHandle	F	F	F	F	F	F	T	F	F	F	T	T	F
SideHandel	T	F	F	T	T	F	F	F	F	F	F	F	T
Flat	T	T	T	T	T	T	T	F	T	T	T	T	F
Concavity	T	T	T	T	T	F	T	T	T	T	T	T	T
Pointsup	T	T	T	T	T	F	T	T	T	F	T	T	F
Light	T	T	T	T	T	T	T	F	T	F	T	T	T
Ceramic	T	F	F	F	F	F	F	T	F	F	T	F	F
Styrofoam	F	T	T	F	F	T	F	F	F	T	F	F	F
Paper	F	F	F	T	T	F	T	F	T	F	F	F	T
Expensive	T	F	T	F	T	F	T	F	T	F	F	T	T
Fragile	T	T	F	F	T	T	F	T	F	T	T	F	F
Cup	T	T	T	T	T	F	F	F	F	F	F	F	F

例

図 11 Cup の例題

Fig. 11 Cup example.

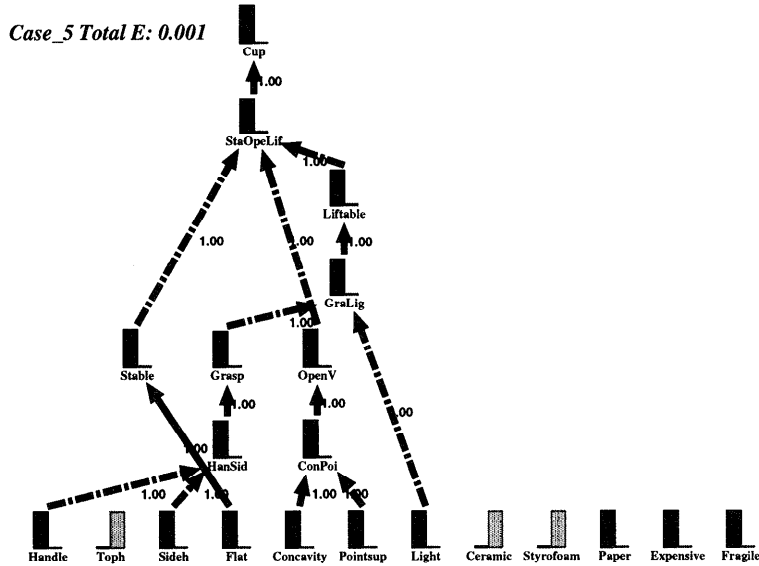


図 12 推論の例 1: 最終状態

Fig. 12 Final state of inference (example 1).

更新することにより、平衡状態に収束することが証明された。

7. 推論の実行例

例題として、Shavlik ら¹⁾が用いたコップの例を用いた。この例では、図 11 に示すような、領域知識と Cup の命題に対して正と負の例題が与えられている。

図 12 に領域知識をニューラルネットワークにマッ

プし Case 5 を提示したときの推論の最終状態を示した。図には、内部解釈の状態を示し、観測値の値は省略した。図中、実線の矢印は重み付きルールを示しており、点線は、連言の展開を示している。矢印の線の太さは、重み付きルールまたは重み付き連言の重みに対応させている。StaOpeLif は、連言 (Stable AND OpenV AND Lifiable) を示すものである。同様に GrasLig は連言 (Grasp AND Light) を、ConPoi は

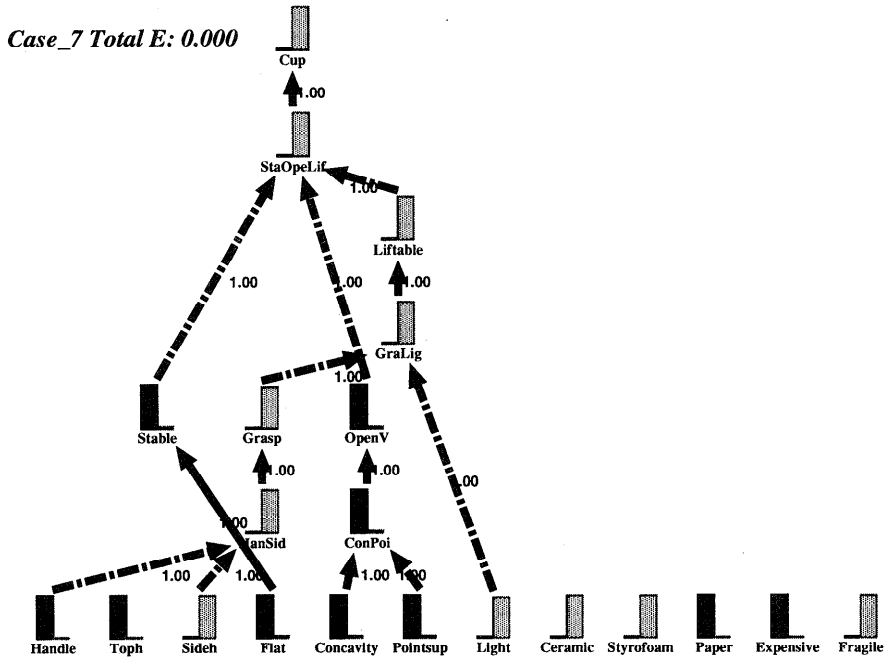


図 13 推論の例 2:最終状態
Fig. 13 Final state of inference (example 2).

連言 (Concavity AND Pointsup) を, HanSid は連言 (Handle AND Sideh) を示すものである。

本実験では, セルの活性度の時間微分のパラメータである ϵ を 0.20 として, 1,000 回同期的に変更した。

Cup の Case 7 において, Light に関する観測値を与えなかったときの推論の最終状態を, 図 13 に示した。Light に関しては観測値が与えられていないので, 他の観測値により偽と推論される。これは, 以下のように解釈できる。

Cup に関する観測値は偽であるので, 否定推論を用いて StaOpeLif が偽と推論される。連言の要素である, Stable と OpenV については, 与えられた観測値より真と推論されているので, Liftable が偽と推論される。これを受けて, 否定推論を用いて GraLig が偽と推論され, Grasp は与えられた観測値より真と推論されているので, Light が偽と推論される。

これらの例に示されるように, 肯定推論と否定推論を同時に行うことにより, 従来に比べて柔軟な推論が可能となっている。

8. 学習時のシステムの動作

学習時のシステム全体的な動作は,

- (1) 観測値の値を対応する各セルに設定する。
- (2) システム全体のエネルギー総和を減少させるよ

うに各セルの値を同期的に変更し, 全体的に落ち着くところまでシステムを動作させる。

- (3) 十分に落ち着いたら, エネルギーの総和に対し各重みで偏微分した値を用いて, 重みを変更する。
- (4) 内部解釈を初期値に戻し, (1) へ戻る。

ルールの確信度 W_r の学習は, システムを動作し落ち着いた状態での $F_{3'}$ を用いて行う。前出の $F_{3'}$ と少々異なる。これは, 知識の完備化の要請より生じるものである。

$$F_{3'} = \frac{1}{2} (W_r A_{true}^{Belief} - X_{true}^{Belief})^2 + \frac{1}{2} (W_r X_{false}^{Belief} - A_{false}^{Belief})^2 \quad (35)$$

重みの学習には, 全体的に落ち着くところまでシステムを動作させたときの, エネルギーの総和 F に対し各重みで偏微分した値を用いる。実装では, 非線形の関数が介在するので偏微分を計算で求めるのではなく, 以下のような近似計算を用い計算量を軽減している。

$$\frac{\partial F}{\partial W_i} \approx \frac{F(\mathbf{W} + \Delta \mathbf{W}) - F(\mathbf{W} - \Delta \mathbf{W})}{2W_i} \quad (36)$$

ここで, $F(\mathbf{W} + \Delta \mathbf{W})$ は, 各重みに $\Delta \mathbf{W}$ をさせ, 推論を実行したときのエネルギーの総和 F である。 $\Delta \mathbf{W}$ には $\pm[0.001, 0.01]$ の一様乱数を用いた。乱数の $[-0.001, 0.001]$ の範囲を除くのは, 近似式の分母

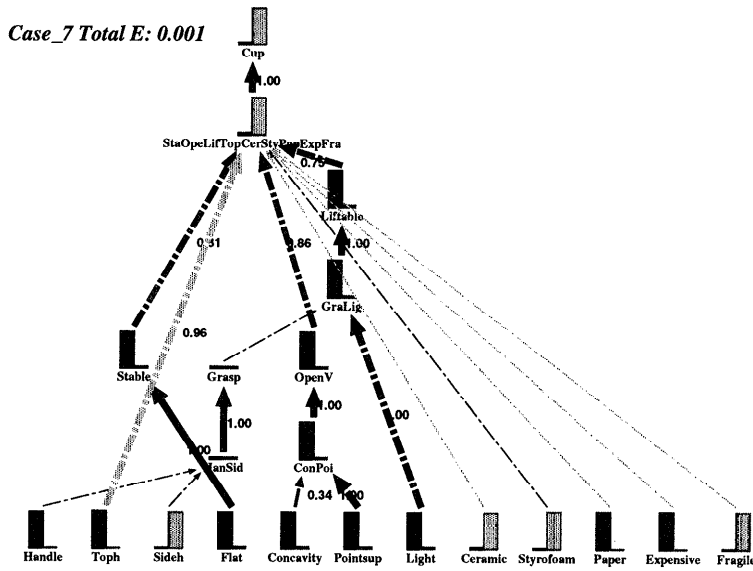


図 14 学習を含んだ推論の例 1: 学習後の最終状態
Fig. 14 Learned state of inference (learning example 1)

が 0.0 に近い数であると、すべての重みを同時に変更しているため近似値が非常に大きな値になってしまう可能性があるからである。

8.1 学習の実行例

7章で示した例題においては、領域知識が完全でないために、Case 7, Case 9, Case 11, Case 12 に対して推論が正しく行われなない。

そこで、学習のために領域知識に用いられていなかった命題を用いて、

Cup :- Stable AND OpenV AND Liftable
AND not Toph AND not Ceramic
AND Styrofoam AND not Paper
AND not Expensive AND not Fragile

とした。ただし、新たに付け加えた命題に対する連言の重みは 0.0 とした。新たに付け加えた命題のいくつかが否定となっているのは、例題中の命題 Cup に対しての偏りを考慮して決めた。

すべての例題を 500 回提示し学習したときの Case 7 に対する推論結果を図 14, 図 15 に示す。図中、内部解釈 Toph より内部解釈 StaOpeLifTopCerStyPapExpFra へのハッチングの矢印は、内部解釈 Toph の否定を用いることを示している。学習前は、内部解釈 Toph の否定に対する連言の重みが 0.0 であったが、学習の結果、連言の重みが 0.96 となり、Case 7 に対して正しく推論できている。同様に、Case 11, Case 12 に対して推論が正しく行われるようになる。得られた領域知識は、大雑把に、

Cup :- Stable AND OpenV AND Liftable
Stable :- Flat
Liftable :- Light
OpenV :- Concavity AND Pointsup

と表すことができ、Case 9 を除いて正しく推論が行えるようになった。Case 9 を正しく推論するためには、領域知識を新たに付け加えて学習する必要がある。

図 16 に、学習中のシステム全体のエネルギー総和 F と Toph 連言の重みの変化を示す。重みの学習のために、偏微分の近似を用いるが、うまく近似できていないときがあるため、全体のエネルギー総和 F と滑らかではないが、学習にともない減少している。また、50 回付近で Toph 連言の重みが負になっているが、これは、偏微分に従って連言の重みを変更すると本質的に起こりうる現象である。意味的には連言の重みは $[0.0, 1.0]$ の範囲にあるべきなので、以下のように、式 (27) で定義した関数 $l(\cdot)$ を用いてエネルギー F_6 を定義し、重みが範囲内に収まるようにしている。

$$F_6 = \sum l(W_i)^2 \tag{37}$$

9. 結 論

知られているほぼ正しい知識を、知られているほぼ正しい例題とニューラルネットワークの学習機能を用いて洗練する方法について提案した。知られているほぼ正しい知識は、不完全であったり矛盾を含んでいたりでするので、ニューラルネットワークの学習によって

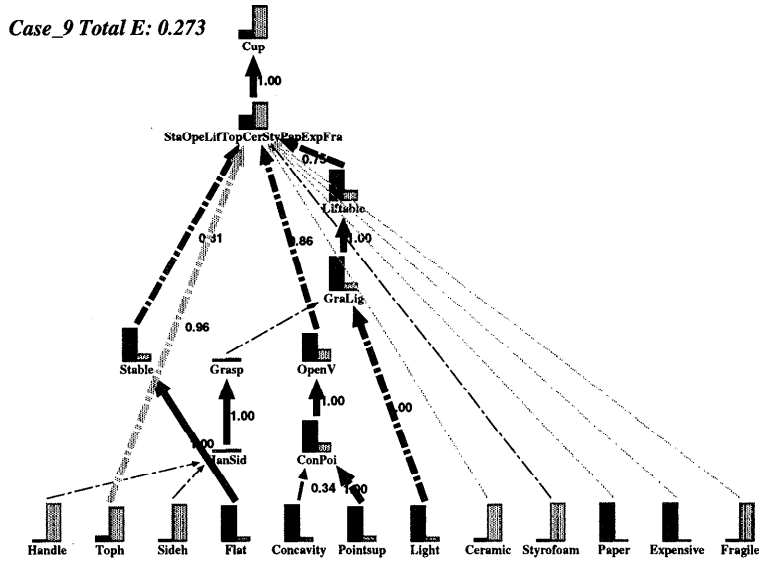


図 15 学習を含んだ推論の例 2: 学習後の最終状態
 Fig. 15 Learned state of inference (learning example 2)

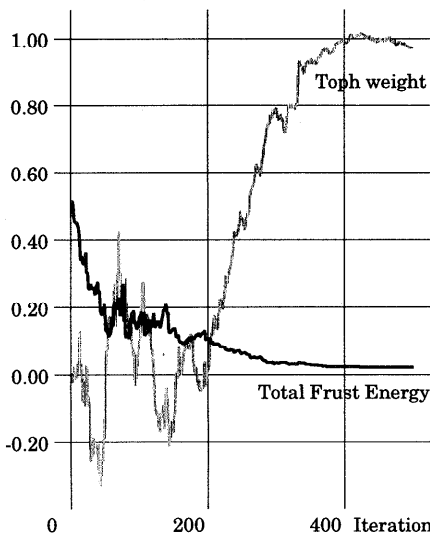


図 16 学習中のシステム全体のエネルギー総和 F と Toph 連言の重みの変化
 Fig. 16 Changes of total energy and weighted conjunction for toph.

足りない部分を補い矛盾している部分を訂正するものである。本システムのニューラルネットワークの特徴としては：(1) 肯定推論と否定推論を同時に行うことができる。(2) 例題が正しく解釈できるように、知識が完全化かつ完備化される。(3) 真と偽以外にこれらの中間状態や、矛盾等が表現可能で、非単調性推論の機能を有する。(4) ニューラルネットワークでは入力

と出力の区別がないので、知られている事実を人力とすると、知られていないことについての推論が行われる。(5) 知られている事実が不十分であるときは、真偽不明と答えることもできる。これまでのエキスパートシステムでは上記した機能の一部は実現可能であるが、すべてを同時に実現したものはない。正しい知識をニューラルネットワークに展開する方法と、それを用いた推論方法、学習方法等について説明し、実現方法の妥当性について検討した。

参考文献

- 1) Shavlik, J. and Towell, G.: An Approach to Combining Explanation-based and Neural Learning Algorithms, *Connection Science*, Vol.1, No.3, pp.231-251 (1989).
- 2) 石淵久生：ニューラルネットによる近似識別, 数理科学, No.331, pp.45-49 (1991).
- 3) 石川真澄：忘却を用いたコネクショニストモデルの構造学習アルゴリズム, 人工知能学会誌, Vol.5, No.5, pp.595-603 (1990).
- 4) 下川勝千：ファジィニューラルネットワークの一手法, ニューロコンピューティング研究会予稿集, Vol.NC 90-82, pp.83-88 (1990).
- 5) 林 陽一, 中井正人：ニューラルネットワークを用いたファジー IF-THEN ルールの自動抽出, 電気学会論文誌, Vol.110-C, No.3, pp.198-206 (1990).
- 6) 堀川慎一, 古橋 武, 大熊 繁, 内川嘉樹：ファジィニューラルネットワークの構成法とその評価, 第6回ファジィシンポジウム講演論文集, pp.57-

- 62 (1990).
- 7) 川村 旭, 渡部信雄, 大和田有理, 益岡竜介, 浅川和雄: ニューロ・ファジィ融合システムの試作, 第5回生体・生理工学シンポジウム講演論文集, pp.197-200 (1990).
 - 8) Gallant, S.: Connectionist Expert Systems, *Comm. ACM*, Vol.31, No.2, pp.152-169 (1988).
 - 9) Gallant, S.: A Connectionist Learning Algorithm with Provable Generalization and Scaling Bounds, *Neural Networks*, Vol.3, pp.191-201 (1990).
 - 10) Opitz, D.W. and Shavlik, J.W.: Connectionist Theory Refinement: Genetically Searching the Space of Network Topologies, *Journal of Artificial Intelligence Research*, Vol.6, pp.177-209 (1997).
 - 11) Shafer, G.: *Mathematical Theory of Evidence*, Princeton University Press (1976).
 - 12) Derthick, M.: Mundane Reasoning by Parallel Constraint Satisfaction, Technical Report, Carnegie Mellon University (1988). CMU-CS-88-182.
 - 13) Touretzky, D.: A Distributed Connectionist Production System, *Cognitive Science*, Vol.12, pp.423-466 (1988).
 - 14) 向殿政男: ファジィ推論 1, 人工知能学会誌, Vol.6, No.6, pp.843-850 (1991).
 - 15) 向殿政男: ファジィ推論 2, 人工知能学会誌, Vol.7, No.1, pp.60-68 (1992).
 - 16) 田崎栄一郎: ファジィ・エキスパートシステム, 数理科学, No.284, pp.46-54 (1987).
 - 17) Kosko, B.: *Fuzziness versus Probability*, chapter 7, pp.263-298, Prentice-Hall (1992).
 - 18) Quinlan, J.: Inferno: A Cautious Approach to Uncertain Inference, *Computer Journal*, Vol.26, No.3, pp.255-269 (1983).
 - 19) Ginsberg, M.L.: Non-monotonic Reasoning using Dempster's Rule, *Proc. 4th AAAI '84* (1984).

(平成 10 年 6 月 23 日受付)

(平成 10 年 10 月 2 日採録)

麻生川 稔 (正会員)



1984年南カリフォルニア大学より Ms.Sc. (Computer Science). 1986年東京工業大学理工学研究科電子物理工学専攻修士課程修了。同年日本電気(株)入社。1992年から1993年にかけてCMU客員研究員。遺伝子情報処理, ニューラルネットワーク, エキスパートシステム等の研究に従事。1992年情報処理学会全国大会奨励賞受賞。