【土木計画学研究・論文集　No. 15　1998年9月】

多様性を前提とした多層パーセプトロンに基づく現象の構造化手法に関する研究*

A Study on Structured Analysis Based on the Multi-Layered Perceptrons

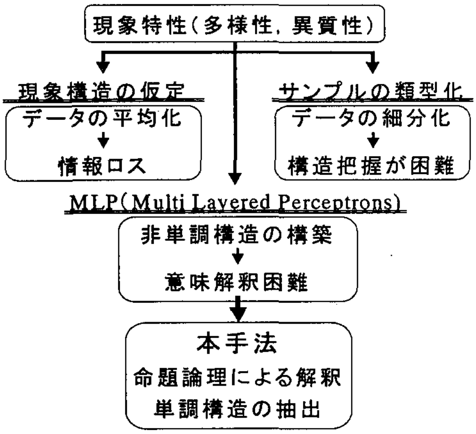
紀伊雅敦**, 土井健司***
Masanobu KII,　Kenji DOI

1. はじめに

市民ニーズの多様性や高度化への対応が求められる地域、地区計画等の各種のプラニングプロセスにおいては、より表現力の豊かな分析手法や支援手法が必要とされてきている. これに呼応した一つの方向性はGISに代表される情報の視覚化と操作化であり、もう一つはAIや知識ベース等に見られる適応性の高いモデル化あるいは構造化手法の開発である.

多様性や異質性を有するデータから現象の構造化を試みる場合, 従来の方法においてはモデル推定においてサンプル・セグメンテーションを取り入れる方法が開発されている. しかし, この場合においてもモデルをアプリオリに設定する必要があり, パラメータ決定に際して数式に適合するようにデータの平均化が行われるため, 現象の背景にある多様性が無視され、情報の損失を余儀なくされる. また, 現象をデータの特性に従って類型化する方法としては、判別分析, クラスター分析および主成分分析等の手法が用意されているが, その結果から構造的意味や法則性を理解することは一般に容易ではない(図—1).

多様性を前提として構造を模倣する方法論として, パターン認識的アプローチがあり, その代表的な手法として, 誤差逆伝播学習を行う階層型ニューラルネットワーク(以下 NN)である多層パーセプトロン(MLP：Multi-Layered Perceptrons)がある. MLP はデータより自律的に現象構造を構築し, 非単調な現象についても高精度な近似を行うことが可能である. ただし, MLP は複数の単調な構造の結合として全サンプルに対する非単調な



図—1　本手法の位置づけ

構造を構築しており, その非単調構造がモデルとしての意味解釈を困難としている. この困難さが MLP を用いる際の問題点の 1 つとされてきたことは言うまでもない.

本研究では MLP が単調な構造の結合であることに着目し, 単調構造の記述を簡便に行える命題論理を用い, 学習済み MLP の解釈をシステマティックに行う方法の開発を試みる. さらに, 得られた命題論理による解釈に対応する構造を MLP より分離, 抽出する方法について提案する.

以下では, まず MLP の概略とそれを用いた既往の現象知識の獲得手法を示す. 次に 3. において学習された MLP から命題論理式を導出する方法を示し, 4. ではその解釈に基づく構造の分離抽出とそれに対応した具体的な数式の導出方法を示す.

2. MLP による現象知識の獲得

2-1. MLP

MLP とは学習といわれる内部パラメータの調整により高精度なデータの近似が可能な手法である. これは次式のように表される.

*　キーワード：情報処理
**　学生員　工修　東京工業大学大学院土木工学専攻
***正会員　工博　東京工業大学大学院情報環境学専攻
(〒152　東京都目黒区大岡山 2-12-1
TEL:03-5734-2695　FAX:03-3726-2201)

$$x_{ij} = \sum_k w_{ijk}\, y_{i-1,k} + \beta_{ij} \qquad (1)$$

$$\mathbf{y}_i = f(\mathbf{x}_i) = \frac{1}{1+\exp(-\mathbf{x}_i)} = \frac{1}{1+\exp(-\mathbf{W}_i \mathbf{y}_{i-1})} \qquad (2)$$

ここで, x_{ij}は第i層のj番目ユニットの入力値, $y_{i-1,k}$は第i-1層のk番目ユニットの出力値, w_{ijk}は第i-1層のユニットkと第i層のユニットjの結びつきの強さを表わす重み, β_{ij}は第i-1層の常に1を出力するユニットの重みであり, 定数項に相当する. また$f(x)$はシグモイド関数, $\mathbf{x}_i$は第i層の入力ベクトル(ユニット番号は省略), $\mathbf{y}_i$, $\mathbf{y}_{i-1}$はそれぞれ第i層, i-1層からの出力ベクトル, $\mathbf{W}_i$はw_{ijk}とβ_{ij}を要素とする重み行列である. 式(1), (2)で表わされる出力の受け渡しと変換は, 入力層から最終の出力層へと順次伝達され, MLP における学習は教師データに基づく重み行列 $\mathbf{W}$ の調整によって行われる.

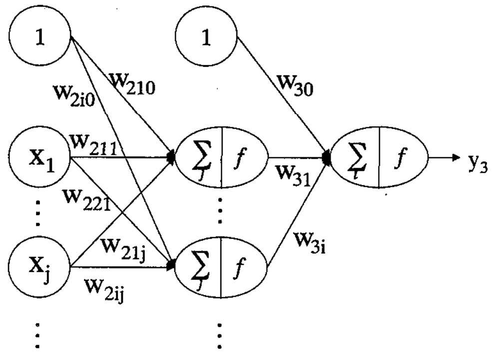
また本稿で扱う3層の MLP は以下のように与えられ, ネットワークを用いて表すと図－2 のようになる.

$$y3 = f\left(\beta_3 + \sum_i w_{3i}\, f\left(\beta_{2i} + \sum_j w_{2ij}\, x_j\right)\right) \qquad (3)$$

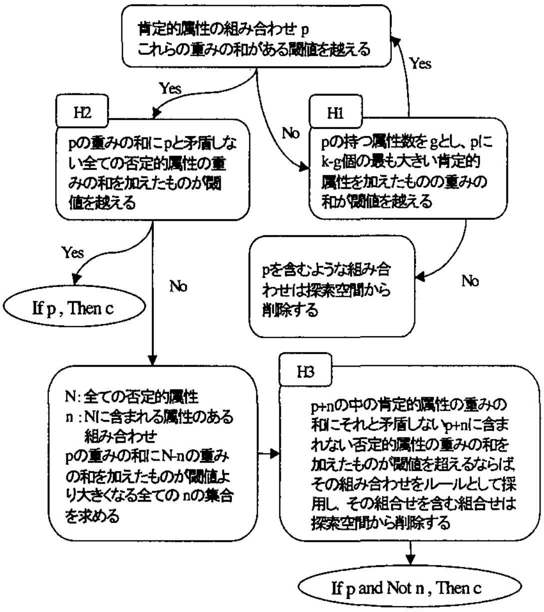
2-2. KT アルゴリズム

知識ベース分野においては, Towell and Shavlik[1)]による M-of-N アルゴリズム, Horikawa ら[2)]によるファジィルールの抽出手法など, 言語的な知識や命題を用いた記号的な知識, あるいはルールを NN の学習結果から導出する手法が数多く提案されている[3)]. ただし上記手法において, 前者は得られるルールが Boolean タイプといわれる特殊な形式であり, 後者は学習に用いるネットワークが特殊な構造である, といった制約を有する. 一方, Fu による KT (Knowledgetron)アルゴリズム[4)]は, ネットワーク構造に制約をおかず, 得られるルールが命題論理と整合する含意形式であり, NN からの知識獲得手法としてもっとも基本的なものであると言える.

この手法は学習済みの MLP において2つの層ごとに上位層のユニットと下位層のユニットの関係を命題論理の含意で表わす. 図－3は KT アルゴリズムの基本的流れを著者らが整理したものである. ここでは説明される上位層のユニットのもつ命題を概念 c とし, 下位層のユニットのもつ命題を属性としている. また, ルールに用いられ



図—2　3 層パーセプトロン



図－3　KT アルゴリズム

る属性に対応する重みの和と, 説明されるユニットの設定した閾値との大小比較により適切なルールが決定される. KT はこれらの組み合わせを探索木により求めるが, それらを網羅的に探索すると探索空間は属性数に対し指数的に増大する. そこで, 図に示す3つのヒューリスティックス, H1, H2, H3 を用い探索空間を減少させる. ただし, この手法では1つのルールの表現に用いられる属性数が k として予め定められる. この探索により得られる結果は属性(下位層ユニット)の集合であり, その属性をすべて論理積で結んだものが得られる含意の前件となり, 上位層のユニットの持つ概念が後件となる. また, それぞれの組み合わせは別個のルールを形成し, 同一のルールでは表わされない.

3. 現象構造の導出方法—論理式の導出—

ここではまず，先に述べたKTアルゴリズムによるルールの導出方法を命題論理式の導出方法へとより一般化し，データの入出力間の論理的意味付けにおいて必要十分性を満たす知識の獲得方法を提案する．

本研究ではFuによる方法と同様に，着目するユニットをu_i，その前階層のユニットをu_jとし，u_iとu_jのもつ命題をp_i, p_jと表す．

ここで，ユニットu_iの出力y_iは，u_iが持つ命題の真理値であるとみなす．ここで，真理値とはその命題の真偽を決定する値であり，値が1ならば真，0ならば偽を表す．ここでは，古典命題論理により解釈を行うことから，次のように閾値qを用いて，y_iを0又は1に2値化する．

$$z_i = \begin{cases} 1, & \text{if} \quad y_i = f(x_i) \geq q \\ 0, & \text{if} \quad y_i = f(x_i) < q \end{cases} \tag{4}$$

ここにz_iは命題p_iの2値化された真理値であり，y_iが閾値qを越えるならば，p_iは十分に真であるとしてz_i=1とする．逆にy_iがqを越えない時には十分に真ではないものとしてz_i=0とする．次に，式(1)の右辺を以下のように変形することにより，ユニットの内部関数が入力に対して単調増加関数であるとする．

$$\sum_j w_{ij} y_j = \sum_{\forall j \in (w_{ij} < 0)} \left\{ -w_{ij} \left(1 - y_j\right) \right\} + \sum_{\forall j \notin (w_{ij} < 0)} w_{ij} y_j + \beta_{ij} + \sum_{\forall j \in (w_{ij} < 0)} w_{ij} \tag{5}$$

この式において，第1項目の$-w_{ij}$, $1-y_j$をそれぞれ，w_{ij}', y_j'と置き換え，第3項目，第4項目をまとめてβ_{ij}'とする．これにより2層間の入出力関数の単調増加を保証する．また，置き換えられたy_j'に対応する命題は元の命題の否定($\neg p_j$)という意味になる．これによりKTにおける探索プロセスが単純化され，ヒューリスティックスは一元化される．置きかえられた重みを用いると式(3)におけるシグモイド関数$f(x)$はxに関して単調増加となることから，式(1)を用いて次の関係が得られる．

$$z_i = \begin{cases} 1, & \text{if} \quad \sum_j w_{ij} y_j + \beta_i \geq f^{-1}(q) \\ 0, & \text{if} \quad \sum_j w_{ij} y_j + \beta_i < f^{-1}(q) \end{cases} \tag{6}$$

また，上式を命題論理における真理値に対応させるために出力y_jを2値化されたz_jで置き換える．

$$z_i = \begin{cases} 1, & \text{if} \quad \sum_j w_{ij} z_j + \beta_i \geq f^{-1}(q) \\ 0, & \text{if} \quad \sum_j w_{ij} z_j + \beta_i < f^{-1}(q) \end{cases} \tag{7}$$

次に，上式に対応する命題論理式を考える．そのために，各ユニットに対して具体的に2値の真理値を当てはめ，ユニットの真理値に応じて集合Jとその補集合$\bar{J}$を次式のように定義する．

$$\forall j \in J \ , \ z_j = 1 \tag{8}$$

$$\forall j \in \bar{J} \ , \ z_j = 0 \tag{9}$$

式(8)は命題が真であるユニットの集合を表わし，式(9)は命題が偽であるユニットの集合を表わす．

個々の命題p_jを'and'で結んだ論理式(連言，conjunction)を$\wedge p_j$と表わすならば，式(8)より$\bigwedge_{j \in J} p_j$は明らかに真である．ここで，仮にユニットiに影響を及ぼす前階層のユニットの集合が式(8)で表わされる集合Jであるとすれば，u_iの論理式は論理積と含意によって次のように構成される．

$$\bigwedge_{j \in J} p_j \rightarrow p_i \tag{10}$$

上式は，「$\bigwedge_{j \in J} p_j$が真であるならばp_iも真である」という論理表現である．

次に，式(7)の条件において式(9)を満たすユニットu_jは関与しないため，これを以下のように書き直す．

$$z_i = \begin{cases} 1, & \text{if} \quad \sum_{j \in J} w_{ij} z_j + \beta_i \geq f^{-1}(q) & (11a) \\ 0, & \text{if} \quad \sum_{j \in J} w_{ij} z_j + \beta_i < f^{-1}(q) & (11b) \end{cases}$$

このとき，z_i=1が成り立つこと，すなわち式(11.a)の条件式は式(10)が成り立つための必要条件である．なぜなら，式(8)の下で式(11.a)が成り立つ時には，必ずp_iが真となるためである．

一方，式(10)の十分条件を考えるならば，'and'の性質より，集合Jにおける任意の命題p_jが偽の場合にはp_iが偽となることが条件となる．この十分条件を示すために，式(10)を満たす集合Jの部分集合J^-を考える．このとき，$^\forall J^-$に対して式

(11)において z_i=0 が成り立つこと，すなわち 式(11.b)の条件式は式(10)が成り立つための十分条件となり，以下のように示される．

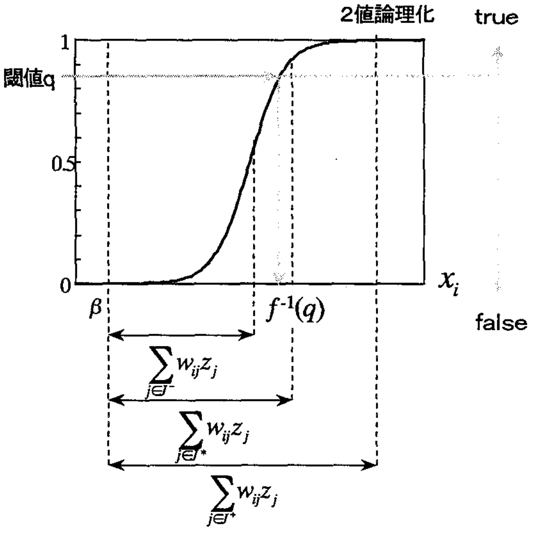
$$\sum_{j\in J^-} w_{ij}z_j + \beta_i < f^{-1}(q) \tag{12}$$

結局，集合 J が式(11.a)の条件を満たし，なおかつその全ての部分集合 $^\forall J^- \subset J$ が式(12)を満たすならば，J は式(10)を満たすための必要十分な大きさの集合であるとし，これを J^*と表記する．また，式(10)の含意（→）は隣接する i, j の 2 層間の命題の一方向の関係を示すものであるが，同値関係が導かれるならば任意の層のユニットの命題を入力層の命題を用いて表わすことができる．そこで，式(10)を満たす集合 J^*の群を$\{J_s^*$, s=1,…,S$\}$とすれば，$\{J_s^*\}$のみが式(11.a)を満たすものであるため，第 i 層の命題 p_i が真であるならば前階層の第 j 層においては$\{J_s^*\}$のうちのいずれかは真なる命題を有するユニットの集合であることが要請される．逆に，既に述べたように，ユニットの集合が J_s^*であれば p_i は真となる．したがって，以下の同値関係が導かれる．

$$\bigvee_{s=1}^{S}\left(\bigwedge_{j\in J_s^*} p_j\right) \leftrightarrow \ p_i \tag{13}$$

ただし，‘∨’は論理和‘or’を表わす記号であり，左辺は括弧中の連言の 1～S までの論理和を表す．上式は，「少なくとも 1 つの s について J_s^*に含まれる全ての j について p_j が同時に真であるならば，p_i は真である」ことを意味する．

図―4 は，3 つの集合 J^*, J^-, J^+における それぞれの u_i の出力 y_i を見たものである．ただし，J^+は J^*を部分集合として含む任意の集合であり，$J^* \subset J^+$と仮定する．このとき，J^-については $y_i<q$ であることから u_i において $\bigwedge_{j\in J^-} p_j$ は真とはなり得ない．次に，J^+については $y_i \geq q$ であるが，J^*を部分集合として含むことから必要十分な集合とは言えない．よって，$\bigwedge_{j\in J^+} p_j$ は u_i の持つ命題 p_i を表わす論理式とはならない．



図―4　重みパラメータによる論理式の導出

4. 構造の分離抽出

ここでは前節で得られた論理式に従い MLP を単調な構造に分離し対応する数式を抽出する方法について述べる．

まず，3. で述べた方法は連続する2層間の関係を表したものであり，ここでは互いに矛盾する表現は含まれない．例えば，「$p_j \wedge \neg p_j$:p_j が真であり，なおかつ p_j は偽である」といった常に成立しない論理式はここでは取り除かれている．常に成立しない論理式を矛盾を含む論理式ということにする．

しかしながら，入力層と中間層，中間層と出力層のそれぞれの関係を表す論理式から，入力層と出力層の関係を表す論理式を導くと，そこには先に述べたような矛盾が含まれる可能性がある．そこで，このような矛盾を取り除いた，入力層と出力層の関係を表す論理式を現象の構造を表す論理式とし，その式に基づき MLP の分離を行う．

入力層と出力層との関係を表す論理式は，論理計算の代入により式(13)の形式で導かれる．これは論理和∨で結ばれた式のいずれかが真であるときに，出力を表す命題が真となる，という意味である．したがって，式(13)において s が 2 以上であれば，この命題が成立するための構造は複数あると言える．すなわち∨で結ばれた各々の論理式が 1 つの構造を表わす論理式であるということになる．本章では具体的にその論理式に対応する関数形を MLP より分離する方法について述べる．

ここで，入力層により出力層を説明する 1 つの構造を次式で表す．ただし $p_1(j)$は第 1 層のユニット j のもつ命題を表す．

$$p_3 \leftarrow \bigwedge_{j\in J_s^*} p_1(j) \tag{14}$$

これは J_s^*に含まれる全ての j について $u_1(j)$の命題 $p_1(j)$ が真ならば，u_3 の命題 p_3 も真であるという意味の論理式であ

る．したがって，入力層ユニットに関しては，J_s*に属するもののみを考えれば良いことになる．ここで関数形を分離するために用いる重みを選択する．ネットワーク上で考えると，中間ユニットとリンクを選択する問題となる．ネットワーク上では入力層と出力層は中間層を通じて結ばれているので，中間層の第kユニット$u_2(k)$について，その論理式が次式で表される場合を考える．

$$p_2(k)=\bigvee_{t_k=1}^{T_k}\left(\bigwedge_{j'\in J_{t_k}^*} p_1(j')\right) \tag{15}$$

ここで，$\bigwedge_{j'\in J_{t_k}^*} p_1(j')$ は，J_{tk}*に含まれる全てのj'について$p_1(j')$が真である場合に真となる論理式を表す．いま，式(14)におけるJ_s*に属するj以外は抽出される構造において影響を与えないことから，J_s*に含まれないユニットは考えなくて良いことになる．次に入力層の命題$p_1(j)$の符号がjとj'で逆になるとき，すなわち出力層において$p_1(j)$，中間層においてその否定 $\neg p_1(j)$が論理式に含まれる場合は，このリンクは導出すべき構造と逆の意味を与えていることになるため，求める構造から取り除く．結局，入力層と出力層との関係に矛盾しない論理式を持つ中間ユニットを構造に採用すれば良いということになる．このアルゴリズムを示したものが**表-1** である．

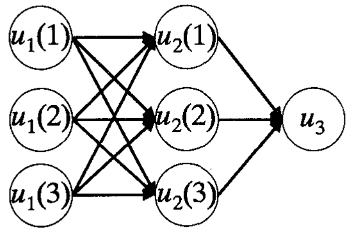
簡単な例によりこのことを示す．まず，学習済み MLP の解釈が 3. の方法により，以下のような論理式として抽出されるものとする．

$$p_1(1)\wedge p_1(2)\wedge p_1(3)\rightarrow p_3 \tag{16}$$
$$p_2(1)\leftrightarrow\left(p_1(1)\wedge\neg p_1(2)\right)\vee\left(p_1(1)\wedge p_1(3)\right) \tag{17}$$
$$p_2(2)\leftrightarrow\left(\neg p_1(1)\wedge p_1(2)\right)\vee\left(\neg p_1(1)\wedge p_1(3)\right) \tag{18}$$
$$p_2(3)\leftrightarrow\left(p_1(1)\wedge p_1(2)\right) \tag{19}$$

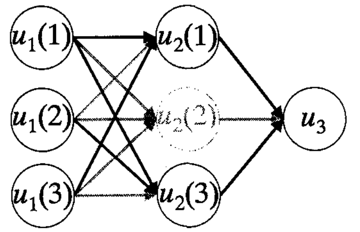
ただし，式(16)は入力層と出力層の関係を表す論理式の一部分であり，矢印が一方向であるのは右辺が左辺の必要十分条件ではないということを表す．また，式(16)の意味は，出力p_3は入力$p_1(1)$, $p_1(2)$, $p_1(3)$すべてが真である場合には真となる，という意味である．式(17)は，中間層の第1ユニットの命題 $p_2(1)$は，$p_1(1)$が真でありかつ $p_1(2)$が偽であるとき，もしくは$p_1(1)$が真でありかつ$p_1(3)$が真であるときに真となる，という意味を表す．以下式(18), (19)についても同様に意味を与えられる．

表—1 構造の分離抽出アルゴリズム

1.	ある $k\in K_t$*とある$j\in J_s$*について，$p_1(k)=\neg p_1(j)$ならば $u_1(k)$と$u_2(T)$を結ぶリンクを削除する．
2.	全ての t についてある $k\in K_t$*とある$j\in J_s$*において $p_1(k)=\neg p_1(j)$ならば $u_2(T)$ならびに関係するリンクをすべて削除する．
3.	全ての$k, j,$ について繰り返す．



図—5 分離前のネットワーク



図—6 分離されたネットワーク

*灰色は削除されたユニットとリンク

抽出前の MLP により与えられる構造をネットワークにより表すと，**図—5**のように表すことができる．先述の方法に従い、MLP より式(16)に対応する構造を分離することを考える．まず，式(17)の$p_2(1)$に対応する中間層の第1ユニット$u_2(1)$は右辺の第 1 項が $\neg p_1(2)$を含んでいることから，この項は矛盾を含むため偽となることがわかる．しかし，第2項については矛盾する表現が含まれていないことから，$u_1(2)$から$u_2(1)$へのリンクを削除し，$u_2(1)$への他のリンクは残す．次に式(18)の中間層第 2 ユニット $u_2(2)$に対応する命題 $p_2(2)$は全ての項に$\neg p_1(1)$という式(16)と矛盾する命題を含んでおり，抽出する構造に対して逆の影響を与えていることになるため，$u_2(2)$は削除される．また，式(19)の第3ユニット$u_2(3)$は矛盾する表現は含まれないことから，求める構造にそのまま残す．このようにリンクあるいはユニットの削除を行うと，式(16)に対応する構造は**図—6** のように抽出され，対応する数式を導くことが可能となる．

5. おわりに

本稿では，MLP により現象を近似させた際の重みパラメータを情報として，現象の背後にある論理的構造を知識として獲得する手法を示した．ここで得られる知識は，KT の IF-THEN ルールよりも一般的な，双方向の含意形式の命題論理式である．次に MLP は本質的に並列処理機構であることに着目し，非単調な MLP の構造を複数の単調な構造に分離抽出する方法を提示した．

本手法は，多様性の解釈に役立つとともに、従来の統計的分析手法に比べてより柔軟な表現力を有する．この特徴は地域分析や空間データの特性把握に有効であり，個別性の高い局地レベルの土地利用現象や交通現象のメカニズム解明に資するものと期待される．

なお、本稿で提案した方法論においては，MLP の解釈に古典命題論理を用いたが，これは関数の値を 0 又は 1 の 2 値についてのみ扱ったものである．より精緻な分析を指向するためには，多値論理，無限値論理を用いた意味解釈方法が必要と思われる．また，2 値論理に依る場合でも、結果として導かれる論理的意味解釈は MLP のネットワーク構造や内部関数の設定にも依存する．これらについては、改めてコネクショリスト・アプローチというより広い枠組みでの検討を試みたい．

参考文献

1) Towell, G. and Shavlik, J. : The extraction of refined rules from knowledge based neural networks, Machine Learning, Vol.131, pp.71-101, 1993.
2) Horikawa, S. , Furuhashi, T. and Uchikawa, Y. : On fuzzy modeling using fuzzy neural networks with the back-propagation algorithm, IEEE Transactions on Neural Networks, Vol.3, No.5, pp.801-806, 1992.
3) Giles, C.L., Miller , C.B., Chen, D., Chen, H., Sun, G.Z. and Lee, Y.C. : Learning and extracting finite state automata with second-order recurrnet neural networks, Neural Computation, Vol.4, pp.406-414, 1992.
4) Li Min Fu: Rule Generation from Neural Networks, IEEE Transactions on Systems, Man, and Cybernetics, Vol.24, No.8, pp.1114-1124, 1994.

多様性を前提とした多層パーセプトロンに基づく現象の構造化手法に関する研究

紀伊雅敦，土井健司

局地レベルでの地域分析においては，その地域の個別性やミクロな要因の影響等が顕在化し，その現象構造を先験的に仮定することは困難である．本研究では，多様性や個別性の存在に起因して構造のアプリオリな設定が困難な現象に対して，MLP の学習により自律的に導出された現象構造を論理式の導出に基づき記述する方法を開発した．また得られた論理式に基づき，本来非単調な構造を有する MLP を複数の並列的な単調構造に分離し、抽出する方法の提案を行った．

A Study on Structured Analysis Based on the Multi-Layered Perceptrons

by Masanobu KII and Kenji DOI

Recently, the importance of analyzing relatively small district is increasing in the field of regional planning. At the regional analysis of small district, because of the emergence of the individuality of the districts and the effects of the micro factors, it is difficult to assume the structures of the phenomenon beforehand. In this paper, we developed methodology that shows the way to describe the explanations of the structures of phenomenon that has been derived by the MLP with propositional formula, and to divide the MLP into some monotonous functions based on the formula.