

## 「概念を獲得する知能ネットワークシステム (INS) の提案」

中山功一<sup>(1)(2)</sup>, 下原勝憲<sup>(2)</sup>, 片井修<sup>(1)</sup>

(1)京都大学大学院情報学研究科, (2)ATR 人間情報科学研究所

連絡先氏名: 中山功一,

〒619-0288 京都府相楽郡精華町光台二丁目2番地2

Tel: 0774-95-1098, Fax: 0774-95-2647, E-mail: knakayama@atr.co.jp

概要: 本論文は, 人と違和感なくコミュニケーションできる人工知能 (AI) の実現に必要なメカニズムの一つとして, 概念の獲得に注目する. 人は, さまざまな言葉で表現される概念を先天的に持っているわけではなく, 後天的な学習により獲得している. 人と AI の円滑なコミュニケーションを目指したとき, この学習メカニズムの実現は重要であると考えられる.

本論文では, 2章において, 人の概念が言語的・文化的にどのように異なるのかを検証する. 文化により概念がさまざまに異なる例から, 概念が絶対的なものではなく, 環境や文化に大きく依存する可能性について述べる. 3章で, 人が概念をどのように獲得・学習していくのかを, 小児の発言を例に検証する. 概念を獲得する学習が, 概念を用いた表現の学習と同時並行的である可能性について述べる. 4章で, 概念を獲得できる手法として知能ネットワークシステム (INS: Intelligence Network System) を提案する. INS を用いて初期段階の知能から高度な知能を獲得していくシステムの創出を目指す. 5章で, INS の実現に向けた予備実験について述べる. 6章で, まとめと今後の課題を述べる.

### 1. 研究の背景と目的

人間の頭脳の働きをコンピュータ上で実現する AI 研究において, 人間の頭脳の働きはさまざまに分類され, さまざまな視点から研究されている. なかでも, 知識の表現方法や論理構造モデルを自己組織的に構築していく研究や, 概念や感情など表現が難しい問題に対する研究がひとつの流れをつくりつつある. 本論文では, 概念を人間の知能の基礎と捉え, 概念を生成・操作するための知能ネットワークシステム (INS: Intelligence Network System) を提案する. ここで概念とは, (1)複数のものを同類と認識することにより, 同類のものそれぞれのそれについてを表象から共通部分をぬき出して得た表象, 及び(2)複数のものを同類と認識する基準, または基準となる集団や集合, と定義する.

従来の学習手法の一つである Q 学習では, 予め状態空間を分割した Q テーブルを用いる. このとき, 環境からの入力, どの分割されたテーブルであるかによって入力を同類と認識する. 状態空間の分割を自律的に再構成する研究[Asada97]は少ない. また, 学習分類子システム (クラシファイアシステム) では, 予め入力区分される候補が設定されており, その組み合わせの変更はできる. しかし, 予め設定していない入力区分を新たに設定する研究[Wada02]は始まったばかりである. このように, 環境からの入力を同類と認識して新たな概念を生成し, その概念に基づき学習を進めていく手法の研究は少ない. 概念を生成できない手法では, 予め全ての概念やその候補を設定する必要がある. 高度な知能の実現を目指す上でボトルネックになる可能性がある. なぜなら, 人が用いる全ての概念を予め定義することは困難であるからである. 一方, 人は, 環境や文化に応じてさまざまな概念を学習により獲得していると考えられる. 人の知能をコンピュータ上で実現する AI 研究において, この概念を学習するという機能は重要であると考えられる. 特に, 人と AI のコミュニケーションの実現を目指す場合, 人とは全く異なる概念や概念学習・概念化のメカニズムを AI に用いると, 円滑なコミュニケーションが実現できない.

本論文では, 人が用いている概念の文化や環境による相違と, 人が概念を学習する過程に注目する. 人と円滑にコミュニケーションする手法の実現を目指し, 概念を生成・操作するための知能ネットワークシステム (INS) を提案する.

## 2. 人が用いる概念

人は、同じ事象や物体を見たときでも、環境や文化によりさまざまに異なった概念に区分（概念化）される。本章では、最も基礎的な概念化の例として、“色”と“空間”に関する概念に着目し、文化により概念がどのように異なるのかを検証する。

### 2.1 色の概念



図1: 二次元で示す色情報



図2: 色情報の概念化の例

一般に、色は赤（R）、緑（G）、青（B）や、色相（H）、彩度（S）、明度（V、B）といった三次元の情報として表現される。図1はその三次元空間を二次元で表現したものである。このとき、人はいくつのか色に見えるだろうか？ある人は図2のように三次元情報を同類の色と認識し、概念化するかもしれないし、他の人は異なる区分で概念化するかもしれない。つまり、三次元情報の概念化といった単純な例でも、その概念化のパターンは無限に存在する。実際に、虹は何色として表現されるか調査した文献[鈴木 90]によると、使用する言語（母国語）の違いにより表1のように異なる。また、別の資料によると国ごとに一般的な虹の色数と色名は、表2のようになる。

言語	色数
日本語	7色
Korean	7色
フランス語	7色
Shona語	3色
Bassa語	2色
英語(民衆レベル)	決められていない
英語(百科事典)	7色
ドイツ語	決められていない

表1: 言語ごとの虹の色数

国	色数	色名
日本	7色	赤, 橙, 黄, 緑, 青, 藍, 堇(紫)
イギリス	6色	赤, 橙, 黄, 緑, 青, 堇
アメリカ	6色	赤, 橙, 黄, 緑, 青, 堇
ドイツ	5色	赤, 黄, 緑, 青, 堇
フランス	7色	赤, 橙, 黄, 緑, 青, 藍, 堇(紫)
ソ連	4色	橙, 黄, 緑, 青

表2: 国ごとの一般的な虹の色数と色名

このように、同一の事象や物体を見た場合に、どのように概念化されるかは、文化や言語によりまちまちであるが、全く異なる概念化がなされるとコミュニケーションが困難となるため、同一の文化圏や言語圏では、ほぼ同一の概念化がなされていると考えられる。

すなわち、人と人工知能のコミュニケーションを考えた場合に、

(a)人は、文化や環境により、それぞれ異なる概念を用いる可能性がある。

(b)AIは、人が用いる概念に近い概念を用いる必要がある。

と考えらる。そのため、相手に応じた概念化が必要であると考えられる。概念を予め固定する手法では、相手に応じた概念化ができないのは明らかである。

## 2.2 別の次元の色

先述の通り、一般に、色は三次元情報の区分として概念化される。このような、予め定められた三次元入力に対し、どのように概念化するのかといった問題も、工学的に大変興味深い。しかし、“色”のなかには、これらの三次元情報という定義に当てはまらないものも存在する。



図3: 金色



図4: 玉虫色

三次元情報では表せない“色”の例として金色(図3)や玉虫色(図4)がある。金色は、本来、三次元色情報には含まれない「光沢がある(光の拡散係数が小さい)」や「輝きがある(光の反射率が高い)」といった、先述の色の三次元情報とは異なる情報に基づく条件が含まれる色である。また、玉虫色は、光の当たり具合や見る角度で変化するという条件を含んだ色である。これらの色の表現は、三次元情報の平面的な変化や時間的な変化から知覚されている。このように、概念を獲得するにあたり、さまざまな情報を元に概念が生成されていると考えられる。つまり、適切な概念化のためには、RGB系とHSV系のいずれを用いるかといった問題や、そのような予め定められた次元で示される状態空間を自律的に構成する方法[Asada97]だけでなく、どの入力を用いて概念化するのもも含めた自由度を持つメカニズムが重要であると考えられる。

## 2.3 空間の概念

無重力空間という環境では「上」や「下」という概念は存在しえないであろうし、イソギンチャクのような、ほぼ線対称(真上から平面的に見たら点対称)な生物にとって、「前」や「後」の概念は理解できないであろう。本節では、前後左右といった平面方向の概念化に注目する。

日本語には、東西南北といった話者に依存しない方向(絶対的方向)と、前後左右といった話者に依存する方向(相対的方向)が存在する。多くの言語は絶対的方向と相対的方向の両方を持っている。しかし、文献[井上 98]によると、自らが受けた刺激を他人にジェスチャーでどのよう

に伝えるかの実験（図 5）や、並べた動物を別の方向を向いた別の場所でどのように並べなおすかを調べた実験（図 6）などから、全世界に存在する言語の約 30%は、「右」や「左」という相対的な方向を持たないことが分かったという。

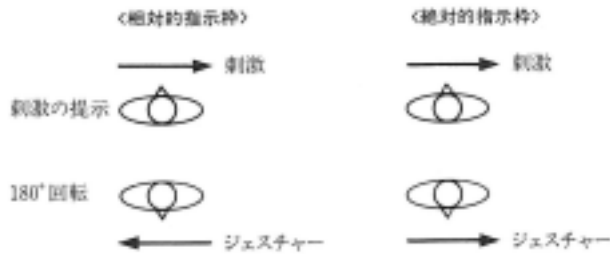


図5 ジェスチャーの指示枠

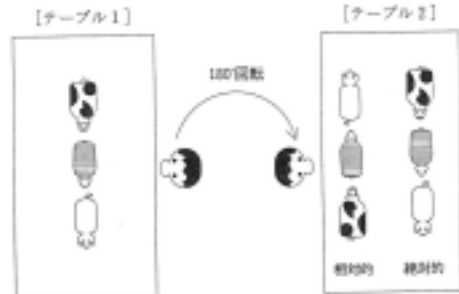


図6 「一列並びの動物たちゲーム」のセットアップ

相対的な方向を持たない言語を用いる文化圏の人間は、図 7 の[絵 A]と[絵 B]の違いを相手に伝える実験では、絶対的方向と相対的方向を共に用いる人々と大きな差は無い。しかし、[絵 B]と[絵 C]の違いを相手に伝える実験では、「右」や「左」という相対的な方向を持たない人々は、絵の説明により多くの時間を費やした。

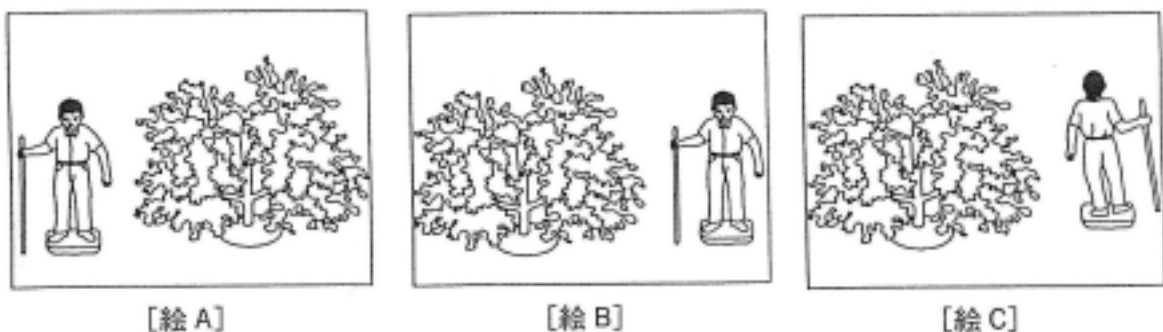


図7 位置の識別

実際には、絶対的方向や相対的方向だけではなく、「池の方向」や「林の方向」といった特定の物体との位置関係を示す場合もある。これらの例からも、連続空間を概念化する基準となる軸は、複数存在することがわかる。

## 2.4 さまざまな概念

ここで、三次元情報としての「色」と、平面（二次元）情報としての「空間」の概念化を包括的に考えてみる。「色」における「赤」や「青」といった概念は、状態空間上のある範囲を示すという意味で、「空間」における「日本」や「京都」といった場所を示す概念に対応すると考えられる。一方、「空間」における「前」や「右」といった概念は、状態空間上の方向を示すという意味で、「色」における「明度(明るい)」や「彩度(鮮やか)」といった概念に対応すると考えられる。すなわち、一部の言語圏の人が、「相対的方向」と「絶対的方向」の両方を用いないのと同様に、環境や文化によって状態空間を「位置」で概念化するのか「方向」で概念化するのか、あるいは両方を並列的に用いるのかも、環境に依存すると考えられる。そのため、概念化を実現する手法としては、このような自由度を持つメカニズムが重要であると考えられる。



### 3. 概念の学習

人間は、実際に成長過程において、概念をどのように学習しているのでしょうか？ 本章では、人の概念学習が、AIにおける概念の獲得メカニズムの参考になると考え、朝日新聞コラム「あのね こどものつぶやき」から、大人とは異なる概念を用いた小児の発言を例に検証を試みる。



図8 学習途中の概念を用いた例

ここでは、図8右上に示す春樹君の発言に注目する。日本の文化では、「元気」とは、体内で増減するものである。そのため、「元気が増える」「元気が出てくる」「元気が出る」と用いるべきである。しかし、春樹君の「切ったところから元気がすうっと出ていってしまっくん。元気が空気みたいなものやから」という発言から分かるように、「元気」を、体の内外を移動する概念として捉えている。つまり、「元気が出て行く」という発言は、「元気が体外に排出される」

「元気が出ていく」「元気が出る」と、「出る」という言葉を別の意味で用いている。このような、概念の相違による興味深い例として、「目に眠いが入った」や、「おこりんぼ、もらった」といった例がある(図8右下)。

このコラムの発言で示されるように、小児は「元気」、「眠い」といった言葉の概念を先に学習してからその概念を用いた表現を後から学習するのではなく、概念自体の学習と概念を用いた表現の学習を同時並行的に行っていると考えられる。すなわち、人が概念を獲得するメカニズムを実現する場合には、概念化や概念の再構成だけを先に学習するのではなく、概念自体を再構成しながらコミュニケーションの中において学習を進めるモデルが、より好ましいと考えられる。

なお、このコラムで取り上げられる特に多い表現として、「お空がないている(雨が降っている)」や「星が瞬きしている(星がきらめいている)」といった過度の擬人化がある。このような表現も、人間(特に自分)に関する概念を拡張して用いた学習過程の概念と捉える事ができる。

#### 4. 概念を学習する智能ネットワークシステム (INS)

本章では、2章、3章で検証した人の概念や概念化のメカニズムを参考に、概念の学習を可能にする智能ネットワークシステム(INS)提案する。本論文では、知能とは外部からの入力に対して外部への適切な出力を決定する機能と定義する。生命個体の場合、入力は知覚により得られた外界状態であり、出力は外界に影響を与える行動全般である。さらに、知能は、さまざまなエージェントの“協調”や“競合”などの相互作用から構成され、相互作用を通してタスクを実行していると考えられる[Minsky 90][戸田 92]。すなわち、ここでは、個体の知能を複数エージェントの相互作用からなるシステムとしてモデル化する。

具体的には、INSを図9のように設計し、このシステムを持つ個体によるシミュレーションを考える。各個体は、INSを持ち、各INSはエージェントの相互作用からなるものとする。INSにより、人の“概念”に相当するものが獲得できると考える。

システム構成法としては、知能及びその表現をあらかじめ設計する手法と、進化・学習機構などにより知能を獲得させる手法がある。本研究では、自己組織化的な知能の獲得を目指すため、進化・学習機構に基づくボトムアップな手法を導入する。

##### 4.1 モデルの概要

マルチエージェントシステムに基づくINSのモデル化にあたっては、エージェント間の協調と競合の制御性の点から、エージェント同士の相互作用をコロニー内に制限するコロニー離隔モデル[中山 02]を採用する。INSを用いてシステムが知能を持たない段階から順次知能を獲得していく段階は、INS自体が発達・進化していくプロセスそのものである。そこで、そのプロセスをメタモルフォーゼ [Metamorphose: 変態, 変形, 変身] と呼ぶこととする。すなわち、INSはメタモルフォーゼにより以下の4段階を経て知能を獲得していく。

本研究では、概念モジュールの生成とその組合せ、相互作用により、各段階を順に実現し、それぞれの機能が獲得されることを目指す。以下に各段階の概要とその意味を述べる。

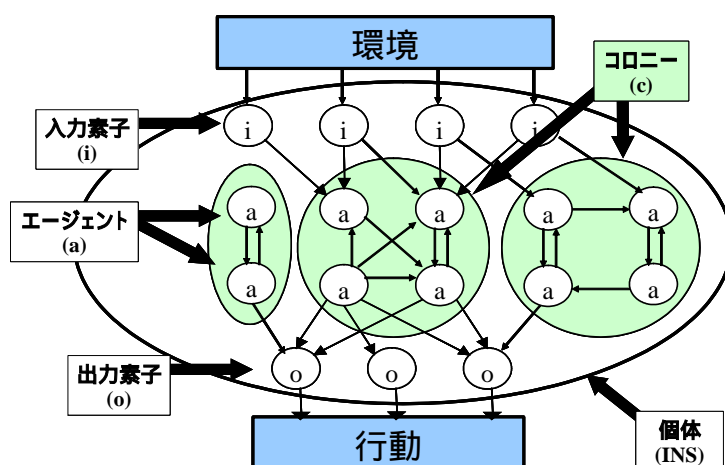


図9:INSの概念

**第一段階.:外部状態のクラスタ化による概念生成・・・外部からの入力から適した行動を出力する問題において、INSは適した行動別に外部入力をクラスタ化し、行動を導く概念モジュールを生成する段階である。この段階では、ある瞬間的な入力に対して適した出力を獲得する。本研究では、エージェント(ニューラルネットワークではニューロンに相当する)による自律的な学習を取り入れる。独立したエージェントが個別**

に学習することにより、ニューラルネットワークと同様のシステムが一つのINSとして得られることが期待される。

**第二段階：複数の概念から新しい概念の生成**・・・第一段階で得られた複数の概念モジュールをさらにクラスタ化することにより上層の概念モジュールを生成する段階である。

第一段階で得られた概念モジュールの出力を入力とする新たな概念モジュールを生成することにより、余分な情報の切り落とし（必要情報の抽出）と抽出された情報に基づく行動出力が得られる。また、エージェントが互いに出しあうことで、過去において概念に基づき得られた結果を利用することも可能となる。これは、短期記憶のメカニズムに相当する。これにより、外部入力を時系列的にとらえた行動出力が実現できる。また、過去の情報を得る入力素子を導入すべきか否かを設計時に決める必要がなくなり、必要に応じて過去の入力を用いることができるシステムが期待される。

**第三段階：複数の概念の並列処理と操作**・・・外部状態を並列な複数の概念モジュールに入力し、その結果を操作することにより、概念に基づいた思考を目指す段階である。

ある外部状態を複数の概念モジュールに入力することにより、多面的な異なる複数の概念が同時に発生する。その概念の選択方法や組合せ方法を操作する機構を INS 自体が構築することで、より高度な知能を獲得することが期待される。これにより、あらかじめ知識の表現や論理構造モデルが定義されているシステムとは異なり、新しい論理構造モデルの構築など、新たな知見が得られることも期待される。

**第四段階：概念の拡張**・・・第三段階までの知能をさらに発展させて、概念モジュールの組合せと相互作用により知能の実現を目指す。ここでは、概念を拡張し、外部入力なしにエージェントが自ら出力を生成することによる想像機構や、他の環境で学習した内容を別の概念モジュールに当てはめる推論機構、複数の INS の組合せによる論理的思考機能などの実現を検討する。

## 5. 予備実験

INS の学習における特徴は、ここのニューロンの学習によりシステム全体の最適性を実現することである。ここで、予備実験として、2層パーセプトロンの学習（図10）を例に、ネットワーク全体を一つの遺伝子としてコーディングした場合と、ニューロン毎に遺伝子をコーディングした場合の実験結果（図11）を示す。詳細[Nakayama02]は割愛するが、コーディング以外の条件を同様にした場合、ニューロンごとにコーディングした手法が、より正解率が高い事が分かる。

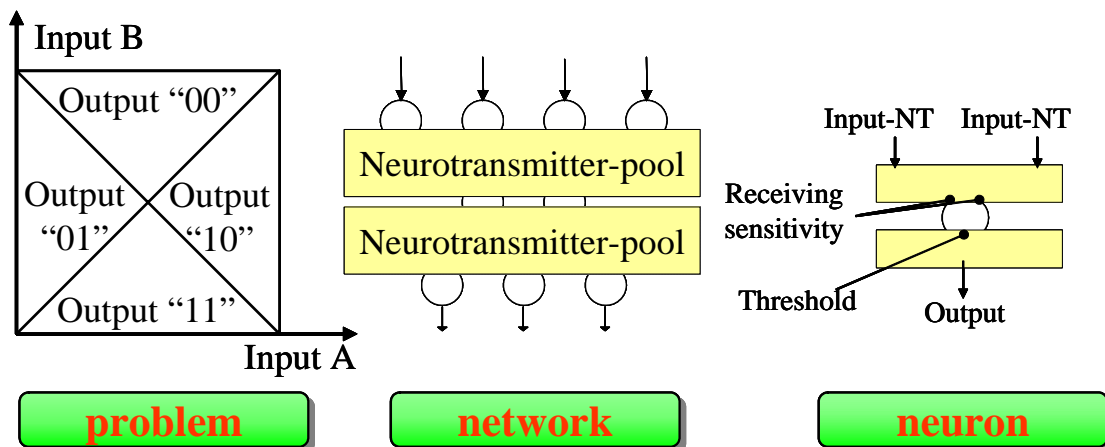


図10 ニューロンを選択単位とするネットワークの学習実験

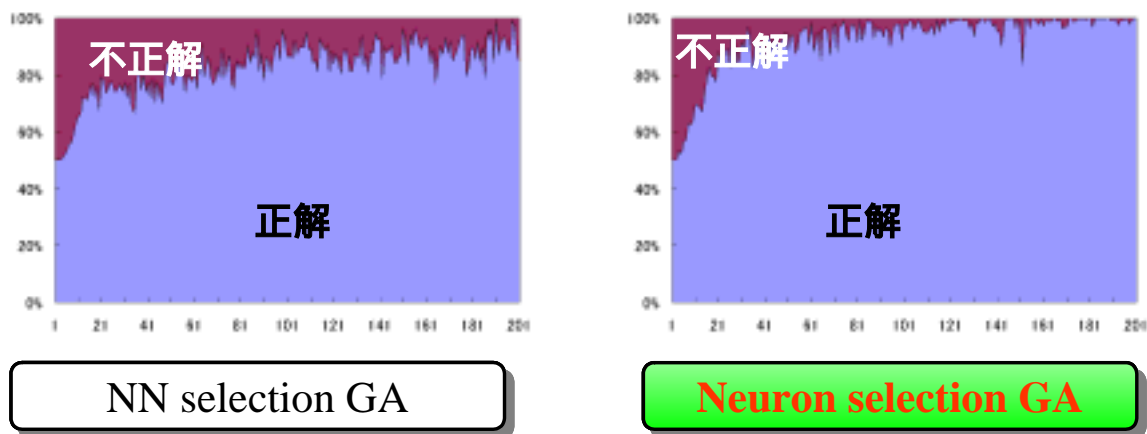


図11 実験結果

これらの結果から，従来のネットワーク全体を学習主体とする手法ではなく，ニューロンを学習主体とする手法の有効性が示唆される．

## 6. まとめ

本論文では，人と円滑にコミュニケーションできる AI の実現を目指したときに必要なメカニズムとして，学習による概念の獲得に注目した．人が用いる概念は，文化的にさまざまに異なっていることを例証した．また，小児が概念を学習する過程について例をあげ，概念学習に必要な条件について述べた．また，そのような概念学習が可能なメカニズムとして，知能ネットワークシステム(INS)を提案した．INS による学習の予備実験として，ニューロン一つ一つを選択単位としたニューラルネットの学習について述べた．

今回の論文では，INS については提案だけであり，INS の詳細な設計や実装，及び実験が不可欠である．今後の課題として，これらのシステムを実際に構成し，実験により有効性を検証していきたい．

### 参考文献

- [Asada97] 浅田稔, 野田彰一, 細田耕. ロボットの行動獲得のための状態空間の自律的構成. 日本ロボット学会誌, Vol. 15, No. 6, pp. 886–892(1997)
- [Wada02] Atsushi WADA, Keiki TAKADAMA, Katsunori SHIMOHARA, Osamu KATAI. Adaptive state space segmentation for LCS. PRIMA2002 SUMMER SCHOOL ON AGENTS AND MULTIAGENT SYSTEMS ( The 5th Pacific Rim International Workshop on Multi-Agents)
- [鈴木 90] 鈴木孝夫 「日本語と外国語」岩波新書(1990)
- [井上 98]井上京子 「もし「右」や「左」がなかったら」 大修館書店(1998)
- [Minsky 90] Marvin Minsky 「心の社会」 産業図書(1990)
- [戸田 92] 戸田正直 「感情」 東京大学出版会(1992)
- [中山 02] 中山 功一, 松井 博和, 野村 由司彦 “動的離隔型 GA(DS-GA)の提案,” 情報処理学会論文誌: 数理モデル化と応用 (進化的計算特集号), Vol.43, No. SIG 10(TOM7), pp.95-109(2002)
- [Nakayama02] Koichi NAKAYAMA, Katsunori SHIMOHARA, Osamu KATAI, "The Wireless NN Using the Neuron Selection GA" PRIMA2002 SUMMER SCHOOL ON AGENTS AND MULTIAGENT SYSTEMS.( The 5th Pacific Rim International Workshop on Multi-Agents)