

明示的な意味共有の不要な自己組織型言語発生モデル

林 兼大[†]

有田 隆也[‡]

[†] 名古屋大学大学院人間情報学研究科

[‡] 名古屋大学大学院情報科学研究科

A self-organization model of language emergence without explicit meaning transmission

Tomohiro Hayashi[†] Takaya Arita[‡]

[†] Graduate School of Human Informatics, Nagoya University

[‡] Graduate School of Information Science, Nagoya University

Abstract: The origin of language has been a widely researched topic recently. One interesting research problem is to understand how humans are able to associate objects to words, without using any specific measures for conveying the meaning between them. This paper has focused on the establishment of communication among agents without using any explicit feedback. We have constructed a model where each agent classifies concepts by using self-organizing maps, and associates words with meanings by using neural networks. Simulations have shown that learning with an ambiguous supervisor exhibits higher communication success rates. Moreover, it has also shown that agents are able to establish communication, despite having different input properties. We expect to use this model for generating autonomous communication between agents.

1 はじめに

言語というシステムがどのようにして誕生し、進化してきたのか、これに関する様々な議論がこれまでなされてきた。本稿で扱う言語とは、人間の生み出す自然言語のことであり、特にその中でも、言語の最もプリミティブな側面だと考えられる音声言語という側面に焦点を当てる。

こうした言語の起源を探る研究として、近年、人工生命アプローチによる研究が盛んに行われている。これは、言語の数理的側面をモデル化し、コンピュータシミュレーションを行うことで、言語が成り立つ原理についての何らかの知見を得ようとするものである。また、そこで得られた知見をもとに、分散システムなどにおける通信手段の獲得といった工学的応用を探ることも研究の目的の一つである。

言語の起源や進化に関する人工生命研究では、言語の持つ様々な側面を対象としてきた。例えば、言語の多様性に着目し、言葉と意味の対応を進化させ語彙の多様性を考察した Arita らのモデル [1] がある。また、文法的構造を

創発させようとした試みとして、ニューラルネットワークを使って両者の語彙を一致させていく Batali のモデル [2] や Arita らのモデル [3] などがある。しかし、こういった個体間での会話を模擬した人工生命モデルの多くが、個体間で共通の意味を共有している状況を想定し、交わされる言葉をその共通の意味に合致させていくものである。したがって、こういったモデルでは、意思疎通の手段を持たない個体間で如何にしてコミュニケーションを成り立たせることができたのかといった問いに答えることは難しい。言葉の指す対象がどのようにして一致していくのか、そしてそのメカニズムに関してどのような学習理論が成り立つのかを検証することは、言語の起源を探る上で避けては通れない重要な課題である。

本稿では、他の情報交換手段を持たない個体同士が如何にして共通の言語体系をつくりあげてきたのかという謎に対して、モデルを立てシミュレーションを行い、明示的な意味共有のない個体間での言語成立の可能性について議論する。さらに、認識自体が異なる個体間における、言語に

よる意志疎通の可能性についても詳しく議論する。

2 先行研究

言語が生まれたその起源において、ある言葉の指す意味は明示的に共有されることはないと考えられる。これは、言葉の指示対象が認知によって初めて確立するためである。このような状況においても、果たして人間は言葉の指す対象を他人と一致させていくことができるのだろうか。

この問いに対して人工生命アプローチで迫ったモデルとして、Smith の提案するモデル [4, 5, 6] がある。このモデルは、エージェントとオブジェクトから構成されている。各オブジェクトは複数の特徴値 (例えば、色、大きさ、形、香りなどに相当する値。以下、特徴ベクトルと呼ぶ) を持っている。各エージェントは、対応した知覚チャンネルを通して特徴値を知覚することでオブジェクトを認知し、オブジェクトを指す言葉を伝達しあう。会話は以下の手順で行われる。オブジェクト群からランダムに選ばれたオブジェクト (コンテキストと呼ぶ) を提示された 2 エージェントは、交代で 1 回ずつ話し手と聞き手になる。このとき話し手は、コンテキストの中からどれか 1 つのオブジェクトをトピックとして、これを聞き手に伝えようと言葉を発する。聞き手は聞き取った言葉を解釈し、コンテキストの中から話し手が伝えたかったトピックはどれなのかを推測する。

各知覚チャンネルはバイナリツリーで表現される。エージェントはこれを深化させるために、ディスクリミネーションゲームと呼ばれるアルゴリズムを実行する。これは、会話に必要な意味を見出していくことで認知の手続きを具現化するものである。さらに、言葉の使用頻度をデータベース化した辞書を使って認知された意味を指示する言葉を学習し、話し手の立場における言葉の選択、および聞き手の立場における言葉の解釈を行う。

上記のようなモデル設定に基づくシミュレーションの結果、会話を繰り返すうちに会話が成功するようになった。ここでいう会話の成功とは、話し手の意図したトピックと、聞き手の推測したトピックが一致することである。つまり、明示的な意味共有がない環境でも言語体系が確立していることを示している。

3 提案モデル

3.1 設計方針

Smith のモデルでは、エージェントの持つ概念空間の独自性は分割する知覚チャンネル選択のランダム性に依存するため、ディスクリミネーションゲームを繰り返すことで、必然的に似た概念空間を持つようになる。そのため、会話成功率が上昇するという現象が生じることは、当然の結果といえよう。また、Smith のモデルでは、対象の持つ全ての特徴を同時に使って認知することができないため、伝えたい対象そのものではなく、対象が持つ特徴の 1 つに対して言葉を割り当てていた。したがって、部分的な特徴を指す言葉から対象を推測しなければならず、こうした状況では、言語が必ずしも信頼性の高いコミュニケーションシステムになるとはいいがたかった。

そこで本稿では、Smith の提案したモデルのエッセンスを抽出し、ニューラルネットワークを使った新たなモデルを提案する。このモデルでは、認知を行い概念構造を形成していく過程を、教師なし学習の代表的なネットワークとして知られる自己組織化マップを用いて実現する。これにより、エージェントが完全に独自に概念空間の形成を行い、対象そのものに言葉を割り当てておくことを可能にする。このモデル設定に基づくシミュレーションを行い、明示的なフィードバックのない個体間における、より普遍的な言語確立の可能性を探る。

提案モデルにおける問題設定は、Smith のモデルと同様である。つまり、全オブジェクト n_o 個から抜き出された n_c 個のオブジェクトがコンテキストとして 2 エージェントに対して提示され、話し手がトピックを意図した言葉を発言し、聞き手が言葉を聞いてトピックを推測するという会話を繰り返す。ここで重要なことは、エージェント間での会話の成否に関するフィードバックはないということである。なお、本稿では、会話を聞き手と話し手が交代で 1 回ずつ行うことを 1 エピソードと定義する。

3.2 自己組織化マップ

Kohonen の自己組織化マップ (Self-Organizing Map; SOM) とは、教師を必要としない学習を行い、自律的に入力パターン群をその類似度に応じて分類していくニューラルネットワークである [7, 8]。SOM は入力層と、2 次元格子状にニューロンが配置されるマップ層の 2 層で構成され、入力層ニューロンとマップ層ニューロンは全結合している。

SOM に、ベクトル $x = (x_1, x_2, \dots, x_n)$ を入力すると、

その入力に最も近い結合荷重 $w_j = (w_{j1}, w_{j2}, \dots, w_{jn})$ を持つマップ層ニューロン (勝者ニューロン) とその近傍のニューロンの結合荷重を入力ベクトルに近づけるように、勝者ニューロンおよびその近傍の全ての結合荷重を変化させる。つまり、各入力データに対して選択的にマップ層ニューロンが反応するネットワークを構成できる。

SOM の学習手順を以下に示す。

1. 入力ベクトルとの距離の計算

マップ層の j 番目のニューロンと入力ベクトルとの距離 d_j は次式で求める。

$$d_j = \sqrt{\sum_{i=1}^n (x_i(t) - w_{ji}(t))^2} \quad (1)$$

2. 勝者ニューロンの決定

d_j が最小となるマップ層ニューロンを入力ベクトル x に対する勝者ニューロン j^* と定義する。

3. 結合荷重の更新

勝者ニューロン j^* および近傍のニューロン j と入力層ニューロン i の結合荷重 w_{ji} を以下の式に基づいて更新する。勝者ニューロンを中心とした 1 辺が $2r+1$ の正方形を大きさ r の近傍領域と定義し、そこに含まれるニューロンを勝者ニューロンの近傍と定義する。なお、この領域は時間の経過に比例して狭くなっていく。

$$w_{ji}(t+1) = w_{ji}(t) + \Delta w_{ji} \quad (2)$$

$$\Delta w_{ji} = \eta(t) h(j, j^*) (x_i(t) - w_{ji}(t)) \quad (3)$$

$$\eta(t) = \alpha \left(1 - \frac{t}{T}\right) \quad (4)$$

ここで、近傍関数 $h(j, j^*)$ は以下の式で表される。 $|j - j^*|$ は勝者ニューロン j^* とマップ層 j 番目とのニューロンの平面距離である。

$$h(j, j^*) = \exp\left(\frac{-|j - j^*|^2}{\sigma^2(t)}\right) \quad (5)$$

$$\sigma(t) = \beta \left(1 - \frac{t}{T}\right) \quad (6)$$

最初に入力層とマップ層間の結合荷重をランダムな値で初期化した後、全入力ベクトルに対して上記の操作を繰り返すことで、入力ベクトルの空間を反映した 2 次元マップをつくることことができる。

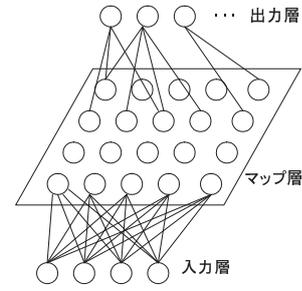


図 1: エージェントの持つネットワーク

3.3 自己組織化マップの会話のための拡張

提案モデルでは、各エージェントは、SOM に出力層を追加した 3 層構造のネットワークを持つ (図 1)。このネットワークの出力層ニューロン 1 つを 1 つの言葉として取り扱う。エージェントはこのネットワークを使ってオブジェクトを認知し、言葉を出力する。エージェントは、まず全てのオブジェクトを SOM に対して学習させ、概念構造を形成した後、会話を行う。

勝者ニューロン j^* および近傍のニューロン j と、出力層ニューロン k の結合荷重 w_{kj} を以下の式に基づいて更新する。ただし γ は、学習の時間的な減衰を表す値で、 $0 < \gamma \leq 1$ の定数とする。

$$w_{kj}(t+1) = \gamma w_{kj}(t) + \Delta w_{kj} \quad (7)$$

$$\Delta w_{kj} = 1 - \frac{|d_j - d_{j^*}|}{\sqrt{n}} \quad (8)$$

つまり、出力層ニューロンと勝者ニューロン間の結合荷重には 1 加算され、出力層ニューロンとその近傍のニューロンとの結合荷重には入力ベクトルとの距離に応じて 1 以下の値が加算される。

3.4 言葉の解釈

エージェントがある言葉からコンテキスト中のオブジェクトを特定する作業を言葉の解釈と呼び、その手順を以下に説明する。まず、エージェントは、ある言葉と最大の荷重で結合しているマップ層のニューロン (言葉解釈ニューロン) を選ぶ。もし、複数ある場合はその中からランダムに 1 つ選ぶ。次に、提示された n_c 個のコンテキストに対して反応した n_c 個の勝者ニューロン (コンテキスト反応ニューロン) の中から、最も平面距離が解釈ニューロンに近い勝者ニューロンに対応したオブジェクトを選ぶ。こうして、ある言葉を解釈して 1 つのオブジェクトが特定される。

3.5 会話におけるエージェントの処理

3.5.1 話し手の処理

話し手は、トピック (提示されたコンテキストの中から1つ選んだオブジェクト) をネットワークに入力し、選出された勝者ニューロンに結合されている出力層ニューロン群のうち、自身がその言葉 (出力層ニューロン) をそのトピックだと解釈できる出力層ニューロン (言葉) を発言する。これに該当する言葉が無い場合、未聴の新たな言葉を発する。続いて、言葉の学習、つまり出力層-マップ層ニューロン間の結合荷重の更新を行う。この更新方法として、4.1節で3つの手法を検討する。

3.5.2 聞き手の処理

聞き手は、聞いた言葉から解釈されるオブジェクトをトピックだと推測する。続いて、出力層-マップ層ニューロン間の結合荷重の更新を行う。コンテキストに含まれる全オブジェクトの勝者ニューロンおよびそれらの近傍にあるニューロンと聞いた言葉の出力層ニューロンとの結合を更新式(7)を用いて更新する。ただし、複数の近傍領域に属するマップ層ニューロンの重複学習を避けるため、複数のオブジェクト反応ニューロンの近傍に相当するニューロンは、入力ベクトルとの距離が最小となるオブジェクト反応ニューロンにしたがってのみ結合荷重が更新される。

4 シミュレーション実験

4.1 会話における学習方法の違いの影響

会話における言葉の学習方法の違いの影響を探るために、以下の3種類の設定で実験を行った。

- 1) 自分の伝えようとするトピックの勝者ニューロンとその近傍のみを学習する。
- 2) 話し手の立場のときは学習を行わない。
- 3) コンテキストに含まれる全てのオブジェクトの勝者ニューロンとその近傍を学習する (聞き手と同じ学習方法)。

エージェント数を2, オブジェクト持つ特徴値ベクトルの次元数 $n = 5$, 全オブジェクト数 $n_o = 10$, 初期近傍領域の大きさを8, $\alpha = 0.5$, $\beta = 10$, 学習回数 $T = 300$ として各エージェントが自分の 10×10 の SOM に全オブジェクトを学習させた後, $\gamma = 1$, 会話におけるコンテキスト数 $n_c = 5$, 言葉学習時の近傍領域の大きさを1, 実行

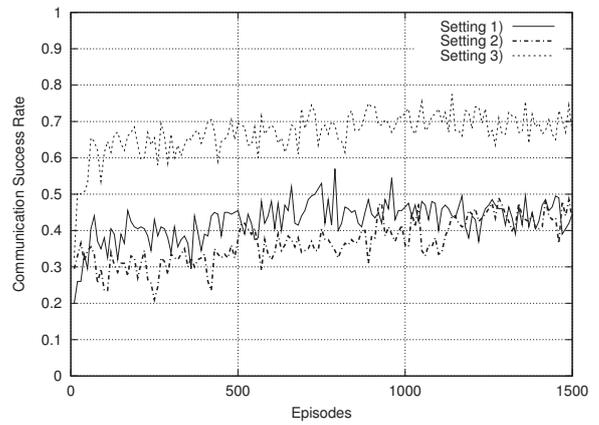


図 2: 会話成功率の推移

エピソード数を1500とし、話し手の学習方法として1), 2), 3)のそれぞれの設定で会話を繰り返す実験を10試行を行った。そのときの会話成功率の平均を図2に示す。なお、グラフの横軸は経過エピソード数を、縦軸は会話成功率を表している。

同図より、3つの設定全てにおいて、ほぼランダムにトピックを推測している初期エピソードを経て、会話成功率が上昇していることが分かる。しかし、1)および2)では、会話成功率は初期エピソードに比べると上昇してはいるものの、0.5を超えることは稀である。これは、フィードバックがないため、言葉へのオブジェクトの対応付けがエージェント間で異なったまま安定してしまっているためだと考えられる。Smithのモデルでは、ディスクリミネーションゲームによる学習ルールの存在によって、特に学習の初期段階において言葉の意味候補が絞られていた。それに対して、提案モデルでは、最初に誤った対応付けをしてしまうと、それを修正する力が働かないため、結果としてこのようなパフォーマンスになってしまったと考えられる。

1)においては、新しい言葉が生まれたとき、話し手側が言葉に対応する概念を聞き手側とは無関係に固定してしまっている。そのため、聞き手がその概念以外で捉えてしまったら、今度はその聞き手が話し手となって誤った概念として発言してしまい、それを修正することが困難になってしまうと考えられる。こうした話し手が言葉の指す概念を一方向的に固定する学習方法は、片方のエージェントを話し手と固定し、話し手の教を忠実に聞き手が学ぶ状況で用いられるならば、言語は容易に生まれるであろう (図3)。しかし、お互いが交代で話し手 (発信者) と聞き手 (解釈者) になるような状況では言語は生まれにくい。

そこで、あえて話し手も提示されたコンテキスト全てを等価に学習する仕組み、つまり、話し手も聞き手も同じよ

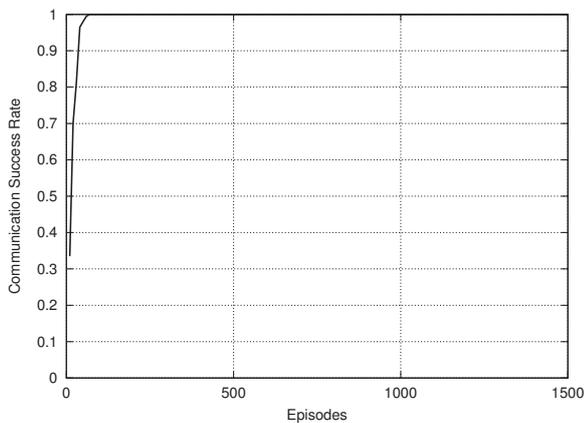


図 3: 1) で話し手エージェントを固定した場合の会話成功率の推移

うにあいまいに言葉の意味を学習していく設定が有効なのではないかと考えられる。つまり、3) の設定がそれに相当する。その設定での実験結果が図 2 の 3) である。グラフより、1), 2) と比べて会話成功率が上昇していることが分かる。これは、言葉を交わす過程で、お互いに徐々に探りあいながら共通の意味を捜し求めていく方法が有効に機能しているためだと考えられる。これらの結果より、話し手があいまいな言葉の学習を行うときのほうが、トピックのピンポイントな学習を行うときよりも、言語の起源における言葉の学習方法として有効であるといえる。

4.2 会話における認知特性の違いの影響

人が認知を行う際、他人と全く同じ情報が認知機構に到達しているとは限らない。そこで、認知の過程にフィルタがかかっている場合を想定する。これは、認知する情報が個体によって異なっている場合を想定するものである。こういった状況においても言語によるコミュニケーションが成立可能なのかを検証する。ここでは、提案モデルと Smith のモデルで、会話時に片方のエージェントがオブジェクトの特徴値をフィルタを通して認知するという設定で実験を行う。

オブジェクトを認知させる際にフィルタを通すエージェントと、オブジェクトをそのまま認知するエージェントの 2 個体用意する。ここでは、片方のエージェントに対してオブジェクトの特徴値を与えるときに関数を通すことによってこれを実現する。ここで述べた、認知特性が相違する設定の影響を探るため、本稿では、フィルタとして以下の 2 つの関数を使って実験を行った。

フィルタ A(図 4(a)) sin 関数: $f(x) = \frac{1}{2} \{\sin(2\pi x) + 1\}$

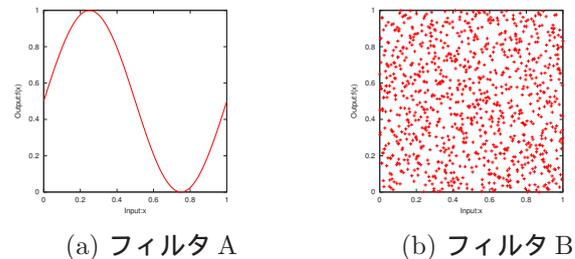
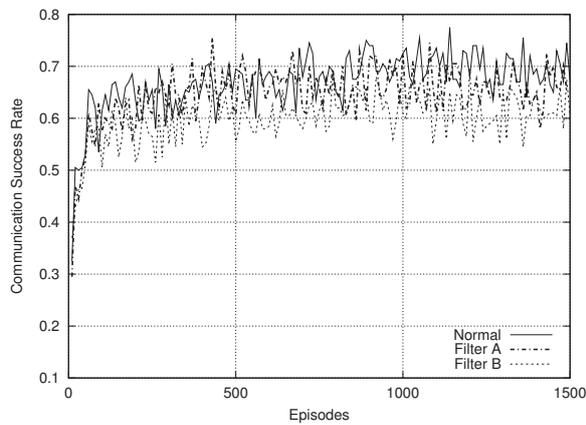


図 4: フィルタ

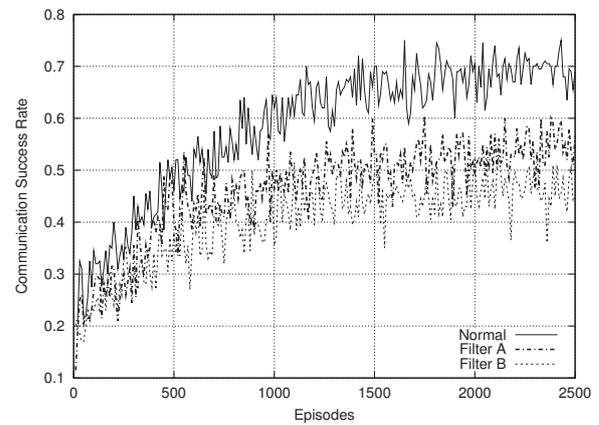
フィルタ B(図 4(b)) 非連続関数: 入力に対する出力をランダムに定義した関数

各エージェントが SOM に全オブジェクトを学習し概念空間を形成した後、話し手の学習方法として 4.1 節の 3) の条件で 10 試行の実験を行った。用いたパラメータは 4.1 節の実験と同様である。提案モデルの会話成功率を図 5(a) に示す。Normal が示すのは、両エージェントにフィルタを通さない場合 (認知特性が同一である設定) の結果である。また、提案モデルの特質を理解するために、同様の設定で Smith のモデルでも実験を行った。このときの会話成功率を図 5(b) に示す。

図 5(a) から、認知特性が同一である設定と比べて、認知特性が相違する設定での会話成功率は若干低いものの、エピソードを重ねるにつれ会話成功率が上昇していく様子が分かる。これは、フィルタを通して認知するエージェントの存在する状況でも、エージェントが独自につくりあげた概念空間とリンクした言語体系が確立し、その結果、言葉による情報のやりとりが可能になっていることを示している。また、図 5(b) から、提案モデルと同様に、エピソードを重ねるにつれ会話成功率が上昇していることが分かる。しかし、Smith のモデルでの認知特性が相違する設定と認知特性が同一である設定との終盤 (2000 ~ 2500 エピソード付近) の会話成功率の差は、提案モデルにおけるその差と比べて大きい。つまり、このことは、認知特性の相違する設定における提案モデルのロバスト性を示しているといえる。しかしながら、認知特性が相違する設定においては、片方のエージェントが認知する際に通すフィルタが複雑な関数になるほど、認知特性が同一である設定と比べて会話成功率が低い。これは、提案モデルにおいては、片方のエージェントが認知する際に通すフィルタが複雑になるほど、形成される SOM の構造が異なってくるため、コンテキストにないオブジェクトに解釈された言葉を聞いたときに最も概念的に近いオブジェクトをトピックとして推測できる可能性が下がるためだと考えられる。



(a) 提案モデル



(b) Smith モデル

図 5: 片方のエージェントがフィルタを通して認知したときの会話成功率の推移

5 おわりに

本稿では、言葉の指示する意味に関して明示的なフィードバックのないエージェント間での言語体系確立に焦点を合わせ、Smith の研究をふまえて、自己組織化マップを使い恣意的な手続きを極力排除した新たなモデルを提案した。この提案モデルで実験を行い、Smith のモデルの持つ言語確立に関する拘束条件を緩和した条件下においても、言語体系が確立しうることを示した。また、この実験から、話し手があいまいな言葉の学習を行うときのほうが、トピックのピンポイントな学習を行うときよりも、言語の起源における言葉の学習方法として有効でありうるという知見が得られた。さらに、認知特性の異なるエージェント間でも言語体系が確立しうることも示し、言語が成立するための拘束条件をさらに緩和することができたといえる。モデルの特質を比べた結果、認知特性の異なるエージェントの存在する状況では、Smith のモデルと比べて提案モデルの方が、言語体系確立のロバスト性が高いことが示された。

ここで示した言語成立のメカニズムを、自律分散ロボット間での通信手段を自動的に確立させる仕組みへと応用することなどが、今後の展開として考えられる。例えば、提案モデルにおける入力層ニューロン 1 つ 1 つをロボット等のエージェントにおけるセンサだと考えると、センサの入力特性の異なるロボット間でも言葉を使ったコミュニケーションを確立できる可能性がある。

参考文献

[1] Arita, T. and Koyama, Y.: Evolution of Linguistic Diversity in a Simple Communication System, *Artificial Life*, Vol. 4, No. 1, pp. 109–124 (1998).

- [2] Batali, J.: Computational simulations of the emergence of grammar, in *Approaches to the Evolution of Language - Social and Cognitive Bases*, pp. 405–426, Cambridge University Press (1998).
- [3] Arita, T. and Taylor, C. E.: A Simple Model for the Evolution of Communication, in *The Fifth Annual Conference On Evolutionary Programming*, pp. 405–410, MIT Press (1996).
- [4] Smith, A. D. M.: Establishing Communication Systems without Explicit Meaning Transmission, in *Advances in Artificial Life: 6th European Conference, ECAL 2001, Prague, Czech Republic, September 10–14, 2001, Proceedings*, pp. 381–390 (2001).
- [5] Smith, A. D. M.: Intelligent Meaning Creation in A Clumpy World Helps Communication, *Artificial Life*, Vol. 9, No. 2, pp. 559–574 (2003).
- [6] Smith, A. D. M.: Semantic Generalisation and the Inference of Meaning, in *Advances in Artificial Life: 7th European Conference, ECAL 2003, Dortmund, Germany, September 14–17, 2003 Proceedings*, pp. 499–506 (2003).
- [7] Kohonen, T., Schroeder, M. R. and Huang, T. S.: *Self-Organizing Maps (3rd ed.)*, Springer-Verlag New York, Inc. (2000).
- [8] Ritter, H. and Kohonen, T.: Self-Organizing Semantic Maps, *Biological Cybernetics*, Vol. 61, No. 4, pp. 241–254 (1989).