

知識に基づいた進化的行動によるエージェントの生成

－大災害時におけるの消火活動シミュレーション－

日大生産工(院) ○渡邊 良太

日大生産工 松田 聖

1 まえがき

大都市に直下型の大地震が起きたと仮定しよう。地震が起きた時による災害は様々であるが、その中で最も大きな被害をもたらす災害は、地震直後に各地で起こる火災であると考えられている。その火災によって想定される被害を可能な限り減らすための一つの方法として、消防士（以下エージェントと呼ぶ）同士が上手く協力し合い、効率良く消火活動を行うことが挙げられる。本研究は、まず、コンピュータ上に上記の環境の条件を満たしたモデルを作成する。そのモデルを作成する際に、マルチエージェント・シミュレーション（以下MASと呼ぶ）という手法を用いる。次に、作成したモデルから何度もシミュレーションを重ねることで、その環境におけるエージェントの最も評価の高い行動パターンを獲得する。その行動パターンを獲得する際に、遺伝的プログラミング（以下GPと呼ぶ）という手法を用いる。そして、シミュレーションから得た結果が最適解、もしくは準最適解であるか考察を行い、本研究の結論を導くとする。そうすることで、いざ現実で大震災が起きようとも、本研究の導き出した結論から模範的な防災対策を発見できると筆者は考える。

2 MAS

現実社会において調査を行うことが大変困難である問題が、数多く存在する。それらの問題を、コンピュータ上に作成した仮想社会でシミュレーションし、認識・分析・理解することができる。このコンピュータ・シミュレーションの一手法を、MASと呼ぶ。MASの特徴に、環境とエージェントとの相互依存的な関係性が挙げられる（Fig.1）。環境がエージェント達に影響を与え、エージェント達が

環境に影響を与える。さらに、エージェントが他のエージェントにも影響を与える。このように、ひとつの要素が全てと密接に関係し、相互作用を生むことが解る。

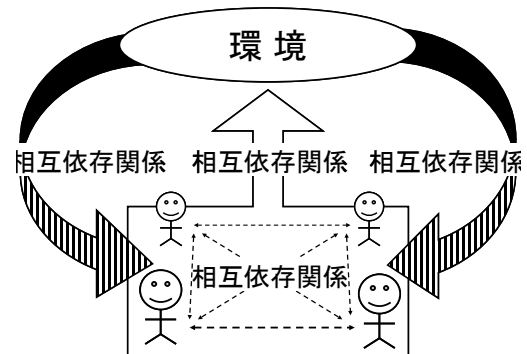


Fig.1 環境とエージェントの相互依存関係

本研究では、地震直後に各地で起こる火災という環境と、それに対するエージェント達の消火活動をMASで表現して、モデル化を行う。そうすることで、火災とエージェントの相互依存関係や、エージェント同士の自律性・協調性がどのようなものであるか認識・分析・理解でき、新たな発見が期待できると考える。

3 モデル化

地震直後に各地で起こる火災という社会現象をモデル化するために、エージェントが学習を行う訓練モデル、エージェントが訓練モデルで学習した知識をテストする実験モデルという二つのモデルを作成した（Fig.2）。この二つのモデルに共通する構造として、道路情報、エージェント情報、火災情報という三つの情報を格納したウェイ・ポイントが存在する。

The Generation of the Agents by Evolutionary Behavior based on Knowledge

－ The Fire Fighting Simulation at the Great Disaster －

Ryota WATANABE, Satoshi MATSUDA

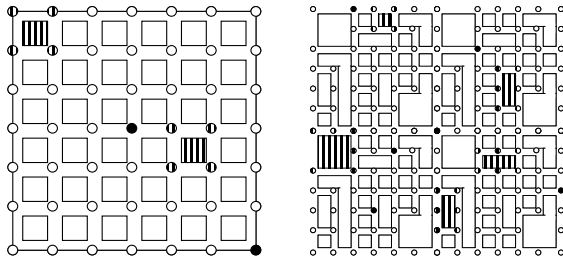


Fig.2 訓練モデルと実験モデル

4 ウェイ・ポイント

エージェントが意思決定をする際に、エージェントに座標の知識表現を与える必要がある。座標の知識表現を与えることによって、その環境に置かれたエージェントが、現在の状況を理解して、最適な行動をとると考えられる。その座標の知識表現を与える一手段が、ウェイ・ポイントである。本研究でのウェイ・ポイントは、二つの性質を持つ。一つ目は、一つのウェイ・ポイントに、道路情報、エージェント情報、火災情報が格納されていることである (Fig.3)。二つ目は、エージェントが最も隣接するウェイ・ポイントの情報に依存することである (Fig.4)。

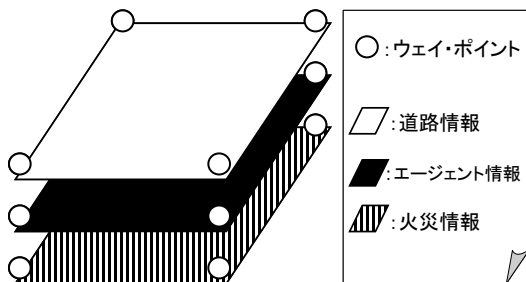


Fig.3 ウェイ・ポイントの構造

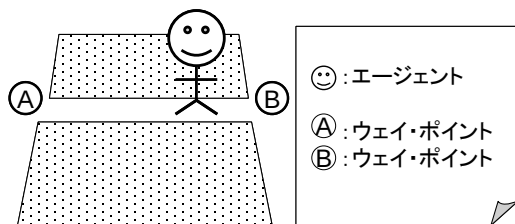


Fig.4 Bの情報に依存するエージェント

5 訓練モデル

複雑な環境は、シミュレーション毎に著しく状況が変化する。この苛酷な環境に何も学習していないエージェントを置いた時、エージェントは賢くなるのであろうか。同様な状況に訪れることは、大変多くのシミュレーションを必要とする。つまり、無知なエージェ

ントを複雑な環境で学ばせることは、効率が悪いと考えられる。エージェントを効率良く賢くするためには、ある程度の知識表現を与えればよいと思われる。そこで、エージェントを学習させるための環境である訓練モデルを作成した。

訓練モデルの環境を以下に説明する (Fig.5)。

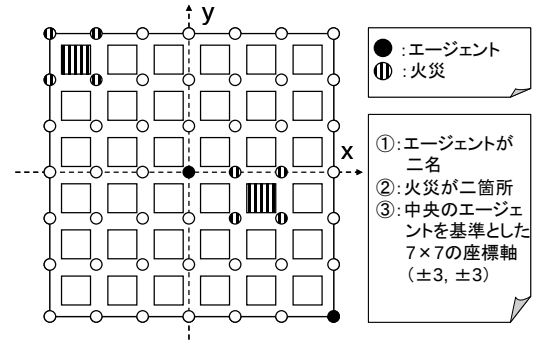


Fig.5 訓練モデルの環境

訓練モデルの環境には、三つの特徴がある。一つ目は、エージェントが中央の基準を含み二名存在すること。二つ目は、火災が二箇所発生していること。三つ目は、中央のエージェントを基準とした7x7のx, y座標軸 ($\pm 3, \pm 3$) があること。この環境の中でエージェントが自律的に消火活動を行うものとする。

6 エージェントの知識表現

上記の三つの特徴から、エージェントに三つの知識表現を与える。一つ目は、中央のエージェントを基準としたある範囲内のウェイ・ポイントに、火災が一箇所あるかということ。二つ目は、中央のエージェントを基準としたある範囲内のウェイ・ポイントに、残りの火災があるかということ。三つ目は、中央のエージェントを基準としたある範囲内のウェイ・ポイントに、もう一方のエージェントがいるかということ。Fig.5の環境を例として、この二つの知識表現である中央のエージェントを基準としたある範囲内を、以下の四つに定める。

- A) 中央のエージェントを基準とした座標 (0, 0) (Fig.6)
- B) 中央のエージェントを基準とした座標 ($\pm 1, \pm 1$) (Fig.7)
- C) 中央のエージェントを基準とした座標 ($\pm 2, \pm 2$) (Fig.8)
- D) 中央のエージェントを基準とした座標 ($\pm 3, \pm 3$) (Fig.9)

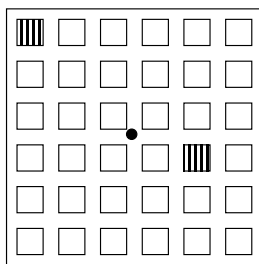


Fig.6 座標(0, 0)

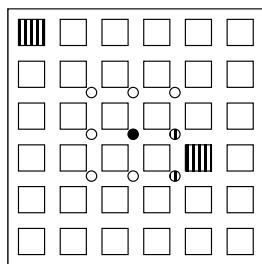


Fig.7 座標(±1, ±1)

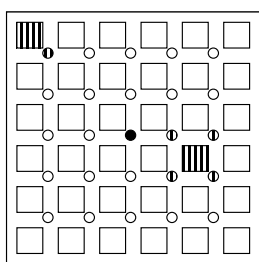


Fig.8 座標(±2, ±2)

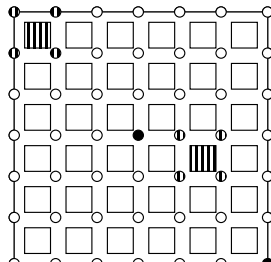


Fig.9 座標(±3, ±3)

この環境は、火災がBとCの範囲内にあり、エージェントがDの範囲内にいることがわかる。つまり、この範囲内に二箇所の火災とエージェントが存在するという初期状況を、エージェントに知識として与えることで、効率の良い学習が行われると考えられる。A、B、C、Dをそれぞれ組み合わせることによって、計40通りの初期状況が存在する。そして、それぞれの初期状況からシミュレーションを重ねることで、エージェントが賢くなっていくと考えられる。

7 訓練モデルのサイクル

訓練モデルのサイクルを以下に説明する (Fig.10)。

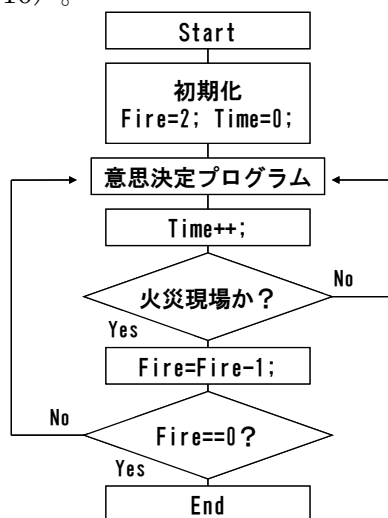


Fig.10 訓練モデルのサイクル

知識表現として与えた初期状況が16通りあるために、40通りの別のサイクルが存在する。このサイクルでは、意思決定プログラムによってエージェントが消火活動をする。火災が全て消えた時の時間によって評価値が決まり、シミュレーションが終了する。エージェントを賢くするためには、意思決定プログラムの良し悪しで決まる。エージェントの行動群は、「一番目に近い火災へ向かい消火すること」と、「二番目に近い火災へ向かい消火こと」の二種類のみである。この二つの行動を状況によって使い分けるためには、複雑な数式の意思決定プログラムが必要不可欠である。その意思決定プログラムを獲得するために、シミュレーションから得た評価値を基にして、意思決定プログラムを進化させる必要がある。ここで、GPという手法を用いることで、進化した意思決定プログラムを獲得できると考えられる。

8 GP

自然界の生物は、食うか食われるかの生存競争の中に置かれている。少しでも優れているものは長く生きる可能性が高く、子孫を残す確率も高い。その子孫は、親の特徴を遺伝子として受け継ぎ、時にはその遺伝子が突然変異を起こす。この進化のメカニズムを模範して、遺伝子を工学的に応用したものを遺伝的アルゴリズム (以下GAと呼ぶ) と呼ぶ。そして、GAの遺伝子型と呼ばれるデータのコードを木構造で扱えるように拡張したものを、GPと呼ぶ。

GPのサイクルを以下に説明する (Fig.10)。まず始めに、ランダムに集団数の個体分だけ、木構造の遺伝子型を生成する。そのコードに沿ってシミュレーションを行い、個体の評価値をそれぞれ求める。各集団の中から、評価値の高い個体を親として選択して、交叉を行う。交叉をすることで、子が生成され、評価値の低い親は集団の中から淘汰される。新たな集団の中の個体が低い確率で突然変異を起こし、遺伝子型が変化することがある。このサイクルを繰り返すことで、最適解、もしくは準最適解を獲得できる。

GPの遺伝子型とは、あらかじめ用意したデータである終端集合 $T=\{T_1, T_2, T_3, \dots\}$ と、データ間の関係や演算である関数集合 $F=\{F_1, F_2, F_3, \dots\}$ を、任意に組み合わせた木構造のデータコードである。

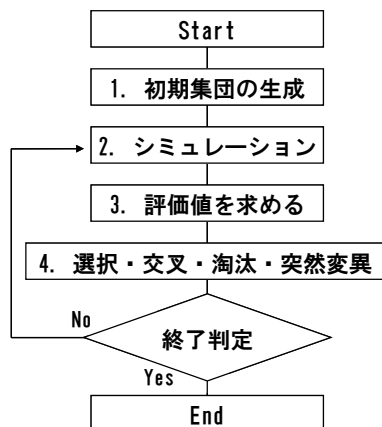


Fig.11 GPのサイクル

9 意思決定プログラム

終端集合Tと関数集合Fを組み合わせた木構造のデータコードから、意思決定プログラムは生成される。本研究で用いる終端集合Tと関数集合Fを、以下に説明する (Fig.12)。

終端集合T	関数集合F
Get1stFireX	GoFireFight(2)
Get1stFireY	AndGate(2)
Get2ndFireX	OrGate(2)
Get2ndFireY	IFLTE(2)
GetAgentX	
GetAgentY	

Fig.12 終端集合Tと関数集合F

終端集合Tは、基本的に座標のデータを受け取るものである。Get1stFireXは、一番目に近い火災のx座標を示す。Get1stFireYは、一番目に近い火災のy座標を示す。Get2ndFireXは、二番目に近い火災のx座標を示す。Get2ndFireYは、二番目に近い火災のy座標を示す。GetAgentXは、エージェントのx座標を示す。GetAgentYは、エージェントのy座標を示す。

関数集合Fは、それぞれ引数を二つずつ持つ。GoFireFight(2)は、一つ目の引数と二つ目の引数を論理積で計算し、1ならば一番目に近い火災へ向かう。それ以外ならば二番目に近い火災へ向かう機能を持つ。AndGate(2)は、ANDゲートの入力を引数1と引数2で行い、0もしくは1を出力する機能を持つ。OrGate(2)は、ORゲートの入力を引数1と引数2で行い、0もしくは1を出力する機能を持つ。

つ。IFLTE(2)は、引数1と引数2を比べ、引数1 $\leq$ 引数2の場合は1を出力する。それ以外の場合は0を出力する機能を持つ。

これらを組み合わせることで、訓練モデルにおいて最適な意思決定プログラムができると考えられる。

10 実験モデル

訓練モデルで学習したエージェントの意思決定プログラムを、より大局的な環境である実験モデルで普遍的に活用できるかテストを行う。訓練モデルでは、40通りの状況における意思決定プログラムを作成した。この40通りのパターンをより広大な環境で利用するために、エージェントの知識表現である座標軸 (± 3 , ± 3) の間隔を 3^n 倍して、ウェイ・ポイントも 3^n 倍の間隔にする (Fig.13)。そうすることで、座標の範囲を拡大、もしくは縮小することができる。

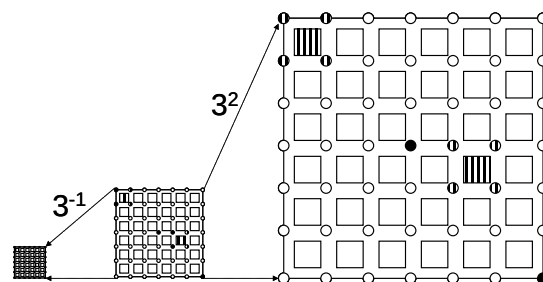


Fig.13 知識表現の拡大・縮小

11 まとめ

実験モデルでは、訓練モデルで生成した意思決定プログラムを利用して、大局的なシミュレーションを行う。よって、環境が変わることにより、プログラムやシステムの普遍性が問われる。もし、どの環境にも適応できる普遍性を持つものならば、さらに火災について詳しい情報を実装することによって、現実での模範的な防災対策として進化することができる、筆者は考える。

「参考文献」

- 1) 山影進, 「人工社会構築指南書」 artisocによるマルチエージェント・シミュレーション入門, (2007), pp.8~12, pp.300~302, pp.421~429.
- 2) 山影進, 服部正太, 「コンピュータのなかの人工社会」マルチエージェントシミュレーションモデルと複雑系, (2002), pp.2~20.