

# 博士論文審査結果報告書

## 論 文 題 目

Genetic Network Programming based  
Rule Accumulation for Agent Control

申 請 者  
**Lutao WANG**

情報生産システム工学専攻  
ニューロコンピューティング研究

2013 年 1 月

本論文は、2000年に平澤が開発した Genetic Network Programming(GNP)の拡張に関するものである。GNPでは進化で最適化された有向グラフ構造のノード遷移により解を求めている。具体的には、複数の判定ノードと処理ノードの遷移で表現される IF, THEN ルールにより、たとえば、エージェントが環境を認識し行動を決定することができる。しかし、GNPでは適合度が最も高い1個の個体のノード遷移により解を求めるため、信頼度の高い解を求めるのが困難な場合がある。

そこで本論文では、1個の個体のノード遷移からではなく、多くの良い個体から多くの優れた IF, THEN ルールを抽出し、これらをルールプールに蓄積して利用するメモリ機能を備えた Genetic Network Programming with Rule Accumulation(GNP-RA)を提案し評価している。従って、GNPの進化では最適な個体を求めるのに対して、GNP-RAの進化では最適な個体ではなく最適なルールプールを求めることになる。

本論文の各章では、GNPによりルールを抽出し、これと環境データとの平均マッチング度を計算することによりエージェントの行動を決定する基本的な GNP-RA、GNP-RAを多重ルールに拡張した GNP with Multi-Order Rule Accumulation(GNP-MRA)、良いルールプールと悪いルールプールを利用してルールのプルーニングを行う GNP with Rule Accumulation and Pruning(GNP-RAP)、蓄積したルールプールをテスト時にオンライン更新する GNP with Updating Rule Accumulation(GNP-URA)、および、有害なノードをスキップする機能を持つ Credit GNP を利用した Credit GNP with Rule Accumulation(CGNP-RA)を提案しその有効性を検証している。

第1章では、GNPの特徴と課題を整理し、有向グラフ遺伝子を持つ GNP からルールを抽出しこれを利用すると、エージェントの制御にとって有効であるという着想に至った経緯および期待できる効果を従来方式と比較しながら述べ、本論文の内容を要約している。

第2章では、GNPを使用してルールを抽出し、これを利用してエージェントの行動を決定する基本的な GNP-RA を提案し評価している。GNP-RAのルールの抽出では、1重 IF, THEN ルール(1個の IF と1個の THEN から構成されるルール)を進化のすべての世代の良い個体群から抽出するため、エージェントの過去の経験をルールとして蓄積することが可能になり汎化能力の向上が期待できる。一方、GNP-RAのルールの利用では、エージェントが環境データを認識するこのデータとエージェントの行動ルール群との平均マッチング度を計算し、平均マッチング度が一番高くなる行動をエージェントは実行する。これにより、GNP-RAでは GNP より適切な行動を実行することが期待できる。

シミュレーションでは、できるだけ多くのタイルをできるだけ短時間にホール

に落とすタイルワールドベンチマーク問題を取り上げ、エージェントの人工脳を GNP-RA および GNP を使用して実現し評価している。落としたタイル数やタイルをホールに近づけた距離などから構成される適合度により GNP-RA と GNP を評価した結果、テスト時のそれぞれの適合度は 97.9 および 79.6 となり、GNP-RA の適合度は GNP より約 23%向上することを明らかにしている。

第 3 章では、1 重 IF, THEN ルールでは処理が困難な非マルコフプロセスの **Perceptual Aliasing Problem** を解決するため、複数個の IF と複数個の THEN から構成される多重 IF, THEN ルールを使用した GNP-MRA を提案し評価している。多重 IF, THEN ルールの抽出は、1 重 IF, THEN ルールと基本的に同じであるが、多重 IF, THEN ルールの利用では、平均マッチング度を計算する際に、すべての多重ルールと環境データの平均マッチング度を計算する完全マッチング方式、および、過去の環境に適合する一部の多重ルールと環境データの平均マッチング度を計算する部分マッチング方式を提案し評価している。

シミュレーションでは、第 2 章と同じタイルワールドベンチマーク問題を取り上げ、GNP-MRA と GNP-RA を評価した結果、部分マッチング方式の 2 重 IF, THEN ルールを利用した GNP-MRA と GNP-RA のテスト時の適合度がそれぞれ 122.4 および 97.9 となり GNP-MRA が GNP-RA より約 25%優れていることを明らかにしている。また、部分マッチング方式が完全マッチング方式に比較し適合度が約 24%向上することを示している。

第 4 章では、適合度が高い個体より抽出したルールから構成される良いルールプールと適合度が低い個体より抽出したルールから構成される悪いルールプールを利用してルールのプルーニングを行う **GNP with Rule Accumulation and Pruning(GNP-RAP)**を提案し評価している。基本的な考え方は、良いルールプールのルールの中に悪いルールプールにも含まれるルールがあればこれをプルーニングし、エージェントがより適切な行動を実行することを目指している。なお、ルール抽出では **GNP with Sarsa Learning(GNP-SL)**を使用しルール抽出効率の更なる向上を図っている。

シミュレーションでは、前述の静的なタイルワールドとは異なり、タイルを穴に落とすと他のランダムな場所にタイルと穴が再び発生する動的で複雑なタイルワールドベンチマーク問題を取り上げ GNP-RAP を評価している。その結果、テスト時の GNP-RAP, GNP-RA, GNP の適合度がそれぞれ 7.38, 6.43 および 5.14 となり、GNP-RAP のルールプルーニング効果を明らかにしている。

第 5 章では、強化学習の 1 種である **Sarsa Learning** を利用して、訓練時に蓄積したルールプールをテスト時にオンライン更新する **GNP with Updating Rule Accumulation(GNP-URA)**を提案し評価している。ルールプールの更新では、テ

スト時に生成したルールがルールプールになければ、これをルールプールに追加し、ルールプールに存在すれば、生成したルールの強度でルールプール内の同一ルールの強度を更新している。

シミュレーションでは、動的で複雑なタイルワールドベンチマーク問題を使用して GNP-URA と GNP-RA を評価している。その結果、GNP-URA ではルールプールのオンライン更新中に適合度が 8.8 から 9.4 に上昇するが、GNP-RA では適合度が 7.0 から 5.5 に下降することを示し、Sarsa Learning によるルールプールのオンライン更新が有効であることを明らかにしている。

第 6 章では、Credit GNP を利用した Credit GNP with Rule Accumulation (CGNP-RA) を提案しその有効性を検証している。Credit GNP では Sarsa Learning を使用した GNP のすべてのノードに新たな Credit Branch を追加し、Credit Branch の Q 値とその他の従来 of ブランチの Q 値を比較し、その大小により Credit Branch を選択している。Credit Branch を選択した場合には Credit Branch の Q 値を更新せず、また、Credit Branch が選択されたノードを有害であるとしてスキップすることにより、悪いルールをブルーニングできることを論理的に明らかにしている。

シミュレーションでは、動的で複雑なタイルワールドベンチマーク問題を取り上げ CGNP-RA を評価している。その結果、テスト時の CGNP-RA, GNP-RA, GNP の適合度がそれぞれ 8.13, 6.70 および 4.81 となり、CGNP-RA が GNP-RA より約 21% 優れていることを明らかにしている。

第 7 章では、本論文で提案し評価を行った各種のメモリ機能を備えた GNP-RA の研究成果を総括している。

以上、本論文では、多くの良好な個体から多くの優れた IF, THEN ルールを抽出し、これらをルールプールに蓄積して利用するメモリ機能を備えた Genetic Network Programming with Rule Accumulation (GNP-RA) を提案し、動的で複雑なタイルワールドベンチマーク問題によりその有効性を検証している。その結果、進化論的計算の性能向上と応用拡大に大きく寄与している。よって、本論文は博士（工学）の学位論文として価値あるものと認める。

2012 年 12 月 21 日

主査	早稲田大学	教授	博士(情報工学)	(九州工業大学)	古月敬之
	早稲田大学	教授	工学博士	(早稲田大学)	吉江修
	早稲田大学	教授	博士(工学)	(早稲田大学)	藤村茂
	早稲田大学	名誉教授	工学博士	(九州大学)	平澤宏太郎