

分類学習を用いた参加型モデリング手法の拡張

Extension of Companion Modeling Using Classification Learning

鳥居 大祐
Daisuke Torii

京都大学情報学研究科社会情報学専攻
Department of Social Informatics, Kyoto University
torii@kuis.kyoto-u.ac.jp, <http://www.lab7.kuis.kyoto-u.ac.jp/~torii/>

ブスケ
フランソワ
François Bousquet

フランス国際農業研究所 (CIRAD)
Centre de coopération internationale en recherche agronomique pour le développement
bousquet@cirad.fr

石田 亨
Toru Ishida

京都大学情報学研究科社会情報学専攻
Department of Social Informatics, Kyoto University
ishida@i.kyoto-u.ac.jp, <http://www.lab7.kuis.kyoto-u.ac.jp/~ishida/>

keywords: multiagent-based simulation, modeling, machine learning, classification learning, feature selection

Summary

Companion Modeling is a methodology of refining initial models for understanding reality through a role-playing game (RPG) and a multiagent simulation. In this research, we propose a novel agent model construction methodology in which classification learning is applied to the RPG log data in Companion Modeling. This methodology enables a systematic model construction that handles multi-parameters, independent of the modelers ability. There are three problems in applying classification learning to the RPG log data: 1) It is difficult to gather enough data for the number of features because the cost of gathering data is high. 2) Noise data can affect the learning results because the amount of data may be insufficient. 3) The learning results should be explained as a human decision making model and should be recognized by the expert as being the result that reflects reality. We realized an agent model construction system using the following two approaches: 1) Using a feature selection method, the feature subset that has the best prediction accuracy is identified. In this process, the important features chosen by the expert are always included. 2) The expert eliminates irrelevant features from the learning results after evaluating the learning model through a visualization of the results. Finally, using the RPG log data from the Companion Modeling of agricultural economics in northeastern Thailand, we confirm the capability of this methodology.

1. はじめに

マルチエージェントシミュレーションは、社会システムを理解・設計するための手段として活用され、多くの研究がなされている。モデル構築・精錬は既存の文献、調査データ、ビデオなどを利用して行うことが一般的であるが [Gilbert 99, Helbing 00, Murakami 03], 現実を十分理解しモデルに反映するために、ロールプレイングゲーム (RPG) やユーザ参加型のシミュレーションを積極的に利用して、初期モデルを精錬する参加型モデリングという手法が研究されている [Bousquet 99, Gilbert 02]。

しかしながら、RPG やシミュレーションから得たデータを分析する場合、多くは、専門家の仮説を元に、統計分析などを利用して仮説を構成する要素を個別に検証するというものである。このような手法では、分析自体が個人の能力に依存するだけでなく、仮説に含まれないが関連する可能性のある要素の分析は、その数が多くなる

と手間がかかりすぎるため、網羅的に行われぬのが現実である。本研究では、このような現実に対処するために、参加型モデリングのモデリングプロセスに対し分類学習を組み込むことで、RPG ログから体系的にモデルを獲得する方法論を提案する。得られたモデルにより専門家による仮説を客観的に検証することが可能となる。分類学習を利用する利点は、人手では扱いきれないすべての要素を機械により同時に扱うことができ、そこから必要な要素をバイアスという客観的指標により選別しモデルを構築できる点である。ゲーミングのログに対し機械学習を利用する研究としては、U-MART [Kita 03] がある。これは、人工市場を対象とした具体的なテストベッドシステムであり、人間とエージェントが同時に参加するシミュレーション環境でエージェントが最適な取引戦略を学習するものである。本研究は、RPG ログデータからステイクホルダーの現実を反映したモデルを機械学習により得ようとする点で、手法の目的が異なる。

本手法の分類学習問題としてのタスクは、RPG ログデータから RPG 参加者の意志決定を一般的に説明する分類知識を獲得することである。多くの分類学習の適用研究では大量のデータから意味のある分類知識を獲得することを目的とするのに対し、本研究では、決して十分とは言えない量のデータから専門家が人間の意志決定として認める分類知識を獲得することが目的となる。単に RPG ログデータから生成した事例集合に対し学習アルゴリズムを適用するだけでは目的を達成するのは難しく、参加型モデリングのプロセスに対し、専門家知識を利用した学習プロセスを適宜組み込むことが必要である。特に、参加型モデリングの最初の過程で生成される専門家の仮説（初期モデル）を学習プロセスにうまく取り入れることが鍵となる。このような問題に対し、本研究では、分類学習 [Quinlan 86, Quinlan 93] に与える属性を専門家知識を反映して絞り込むことで、数が十分でない事例集合からでも一般性が高く、かつ現実を適切に説明するモデルを獲得する接近法を用いて、本手法を実現するエージェントモデル構築システムを実現した。専門家知識を利用して属性を絞り込む分類学習の手法としては、[石野 95] がある。これは、大量の属性・事例を持つアンケートデータを対象に、必要な属性集合を機械学習と専門家のインタラクションにより得るものである。専門家は学習前に重要な属性に関する知識を持っていない。本研究では、専門家が事前知識として持っている重要な属性の候補を利用し属性を絞り込む別の接近法を提案している。

以下、本論文では、まず参加型モデリングのプロセスについて述べる。次に、参加型モデリングにおける機械学習を用いたモデル構築手法と具体的にそれを実現するシステムを論じる。最後に、実際に、タイ東北部における農業経済を扱った参加型モデリング [Veipas 04] に対し、本手法を適用した結果とその考察を示す。

2. 参加型モデリング

2.1 参加型モデリングのモデリングプロセス

参加型モデリング (Participatory Modeling あるいは Companion Modeling) [Bousquit 99, Bousquit 02] は、ロールプレイングゲーム (RPG) やマルチエージェントシミュレーションをステイクホルダーと共に繰り返し行うことで、現実をより深く理解し、文献などから構築した初期モデルを精練していく手法である。最終的に、本手法で獲得したモデルを利用したマルチエージェントシミュレーションにより、将来起こりうるべき現象を提示することで、ステイクホルダー間にある問題の解決や、政府の施策策定に役立てられる。

図 1 に参加型モデリングにおけるモデリングプロセスを示す。始めに、文献やステイクホルダーに対する調査を元に初期エージェントモデルが構築される。次に、ステイクホルダーを取り巻く環境を再現したボード (図 2

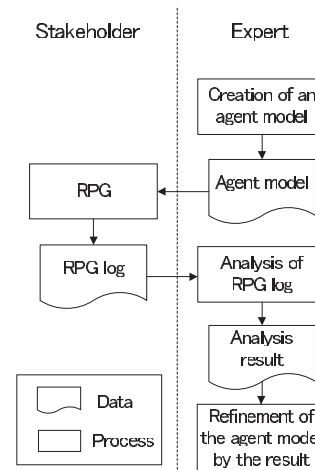


図 1 参加型モデリングにおける RPG を用いたモデリングプロセス [Bousquit 99]

左) を使用した RPG が実施される。RPG は、初期モデルを検査・改良するために行われる。RPG では、ステイクホルダーの意志決定の動機を明確にするためのインタビューが行われ、これらの情報と RPG データの解析により元のエージェントモデルの改良が行われる。RPG が行われた後に、RPG で使用されたボードと同じ特徴を持つインタフェースを用いたマルチエージェントシミュレーションが作成され (図 2 右)、ステイクホルダーにより評価が行われる。これは、図 1 における RPG のプロセスをシミュレーションにより置き換えたプロセスとなる。ステイクホルダーは RPG の体験により、シミュレーションを容易に理解可能であり、モデルの修正を示唆することが可能である。多くの場合、1 回以上 RPG によるモデル構築が行われた後、そのモデルを使用してマルチエージェントシミュレーションによるモデルプロセスが行われる。このように、RPG やシミュレーションを用いたプロセスを繰り返し行うことで、現実をより深く理解していくことができる。

2.2 参加型モデリングの特徴

参加型モデリングの特徴は次の 3 つにまとめられ、現実を反映したモデルの構築とモデルの妥当性評価に対する新しい解決法を与えている。

- (1) RPG によりステイクホルダーが無意識に持つ問題への理解を抽出でき、より現実を反映したモデル構築が行えること
- (2) RPG と同じ特徴を持つシミュレーションが実現されることにより、そのモデル評価を専門的知識のないステイクホルダー自身が行うことができること
- (3) 最終的に行われる参加型モデリングにより獲得されたモデルを利用したマルチエージェントシミュレーションは、ステイクホルダーがモデリングの過程を共にしたことで、現状から起こりえる教訓として捉え

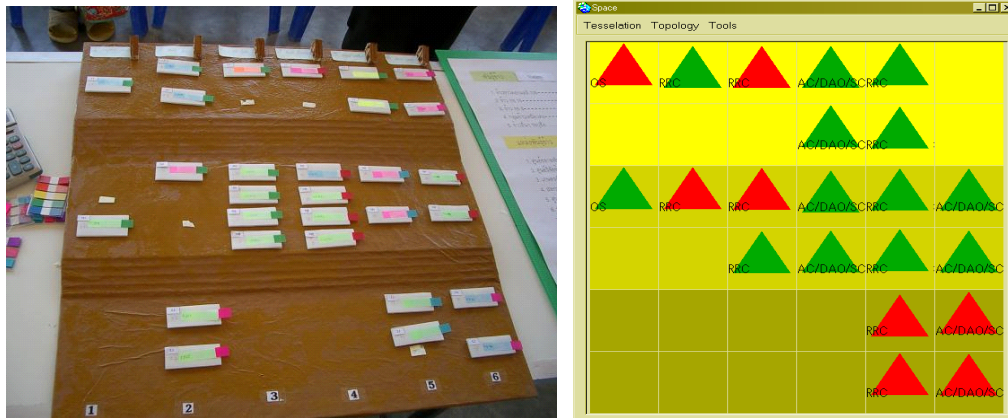


図 2 RPG ボード (左) とシミュレーションインタフェース (右) [Veipas 04]

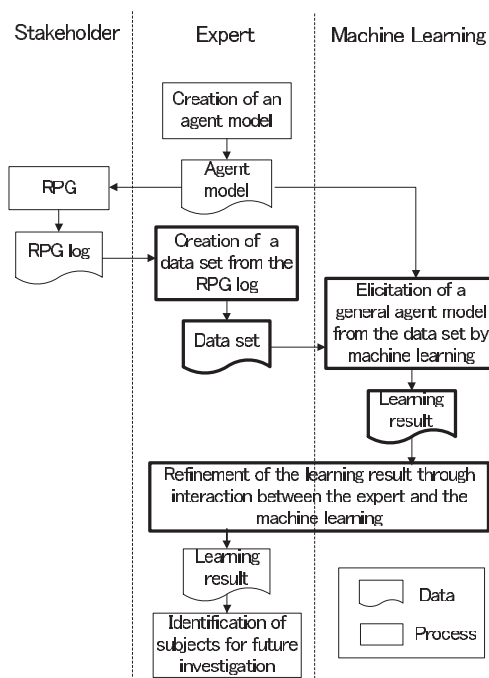


図 3 参加型モデリングの機械学習による拡張モデリングプロセス

ることができ、問題の解決に役立てることができる。特に、東南アジアやアフリカなどの農村などで、普段研究のプロセスに慣れていないステークホルダーを対象としたとしても、本方法は現実を抽出するための効果的な方法と言える。

3. 機械学習を用いた参加型モデリングの拡張

本研究では、RPG ログデータを元に機械学習によりモデルを獲得し、仮説として専門家が立てた意思決定モデルを検証し、検討課題を得ることを目的とする。検証のためには、仮説をある程度正しい前提知識として扱う演繹学習ではなく、前提知識を必要とせず、データから分類知識を抽出する分類学習が適切と言える。

本手法のタスクは、RPG ログデータからの分類学習

によって、RPG 参加者の意志決定を一般的に説明する分類知識を獲得することである。そのために、図 3 に示す機械学習（分類学習）を用いたエージェントモデル構築プロセスを提案する。提案手法は、図 1 に示す RPG を用いたモデリングプロセスに適用・拡張したもので、人手による RPG ログ分析部分を機械学習を用いたプロセスで置き換えたものとなっている（図 3 中、太枠部）。本部分は、次の 3 つの部分から構成される。

- (1) 初期設定
RPG ログデータを機械学習で使用できる形式に変換し、学習データを生成する。
- (2) 機械学習による一般モデルの導出
専門家により生成された初期モデルと RPG ログデータから変換された学習データから学習モデルを導出する。ここで重要なのは、初期モデルを利用して専門家知識を学習に対して反映することと、導いた学習モデルができるだけ高い一般性を持つことである。
- (3) 機械学習と専門家のインタラクションによる学習モデルの精練
上記手順により、専門家知識を反映した一般のモデルが出力されるものの、専門家が十分納得し、現実を反映したモデルであるためには、専門家と機械学習の間で適切なインタラクションを設計し、学習モデルをさらに精練できることが必要となる。

4. 分類学習によるエージェントモデル構築システム

本章では、前章で述べた手法を実現する具体的なシステムの設計を示す。

4.1 問題設定

分類学習問題としての特徴は次の 3 点である。

- (1) 専門家は、文献やサーベイを元に仮説を構築するものの、仮説に含まれる意思決定の要因以外にも重要な要因が存在すると考えている。事例データに含

まれる属性は、そのような専門家が分類に有効であると考えらるすべて含ませるため、その数が大きくなる傾向にある。しかしながら、属性数が増えるほど学習の性能を確保するのに必要な事例数が大きくなる傾向にある [Almuallim 94, Kohavi 97] のに対し、データ収集コストが高いために、学習で利用する属性数に対し十分な学習データを得ることが困難である。

- (2) データ量の不十分さゆえにノイズが学習結果に影響を及ぼす可能性が高い。
- (3) 学習から得た分類知識は、単に新しい事例を正しく分類する性能も必要であるが、エージェントの意思決定モデルとして説明がつくものである必要がある。このことは、学習結果を構成する分類条件は、現実を説明する適切な文脈を持ち、専門家を納得させるものでなくてはならないということである。

4.2 接 近 法

以上の問題への我々の接近法は、分類学習 (C4.5) [Quinlan 86, Quinlan 93] に与える属性を専門家知識を反映して絞り込み、数が十分でない事例集合からでも一般性が高く、かつ現実を適切に説明するモデルを獲得するものである。属性の絞り込みは次の 2 つの方法を取った。

- (1) 機械学習の側よりできるだけ不要な属性を削除し、より一般性の高いモデルを専門家に提供するために、Wrapper アプローチ^{*1} [Kohavi 97] を利用する。今回は、交差検定^{*2}から導かれる予測精度 (Prediction Accuracy) を最大にする属性組を初期の属性組から山登り法 (後向き探索) により除去していくことにより同定する。性能推定に交差検定を利用することで、未知のデータに対して分類精度の高い (一般性の高い) モデルを得ることができる。ただし、本手法では、事例数が十分でない場合、ノイズがあるために、現実を正しく説明するための必須属性も除外される可能性がある。そこで、専門家により分類学習適用前に作成される仮説 (決定木モデル) に含まれる属性は、現実を正しく説明するために必要な要素の候補として、この過程で除外しないようにする。
- (2) 学習結果を意思決定モデルとして説明できるものとするために、不適切な属性を専門家により除去することで学習結果の精錬を行う。学習結果が十分吟味されるためには、専門家が直感的に理解できる提

示法が必要である。そこで、実施した RPG と同じ環境と参加者をコンピュータ上でグラフィカルインタフェースとエージェントとして再現し、その条件下でエージェントに学習モデルを与え、エージェントの意志決定結果として示すことで理解促進を図る。その後、コンピュータ上の現象と決定木を照らし合わせることで不適切な属性を同定する。この方法の利点は、最初から決定木を評価した場合、専門家の初期仮説を正しいものとして単に学習モデルを批判しがちであるのに対し、エージェントの意思決定結果として見せることにより中立的に評価を行える点である、さらに、この評価をもとに決定木を再評価し、十分吟味を尽くすことができる。尚、この点を追求するために、学習モデルに加え、実際の RPG ログをそのままエージェントに与えた場合も提示する (どちらが学習モデルによるものかは専門家に伝えない)。また、もう一つの利点は、可視化手法は明らかにおかしな現象を容易に把握でき、不適切な属性の発見が簡単に行える点である。

4.3 分類学習を用いたモデル構築システム

上記に述べた接近法を組み込んだモデル構築システムを図 4 に示す。以下に、3 章で述べた手法に対応した 3 つの手順における各処理を示す。

(1) 初期設定

決定木学習に用いる属性と目標概念を決定後、RPG ログから決定された属性に対応したインスタンスを抽出し、事例集合を生成する。また、初期エージェントモデルを決定木モデルに変換し、そこに含まれる属性を重要な属性として抽出する。

(2) 属性選択手法による不要な属性の除外

初期設定で作成された属性と事例集合を用いて予測精度が最良となる属性組を同定する。属性選択探索法は山登り法 (後向き探索)、性能推定値は交差検定による予測精度を利用する。各段階では、現属性集合から一つ属性を除外して得られるすべてのパターンの属性集合を生成し、それぞれで交差検定から予測精度を得る。そのうち最も予測精度が改善するものがあれば、それを次の属性集合とする。予測精度の向上が得られなくなった時点で探索を終了し、そのときの属性集合を出力する。ただし、各段階で、専門家により選ばれた重要な属性は除外されない。

(3) 可視化による専門家への理解促進と精錬

最初に生成された事例集合と属性選択手法から得た属性集合から学習結果を得る (3 章との対応としては、ここで 2 番目の手順が終了する)。次に、実施した RPG と同じ環境と参加者をコンピュータ上でグラフィカルインタフェースとエージェントとして再現する。さらに、エージェントに学習モデルを与えた場合と RPG ログを与えた場合の 2 通りを、ど

*1 帰納学習により得られるモデルの性能が最良になるような属性組を探索により求める手法。多くの場合、属性組の良し悪しを定義する性能推定値を交差検定から求め、山登り法を探索手法として用い性能推定値を最良にする属性組を求める。山登り法は、すべての属性組から初めて推定値が上がる方向の一つずつ属性を除外していく後向き探索と、空集合から始めて推定値が上がる方向に属性を追加していく前向き探索がある。

*2 重複を許さずに、訓練集合を n グループにほぼ等分割し、そのうち $n-1$ グループを利用して学習を行い、残りの 1 グループでテストする。テスト事例を適用した予測精度を性能判定の基準とする。テスト事例として選ぶグループを変えながら本作業を n 回行う。今回、 n は一般的に用いられている 10 とした。

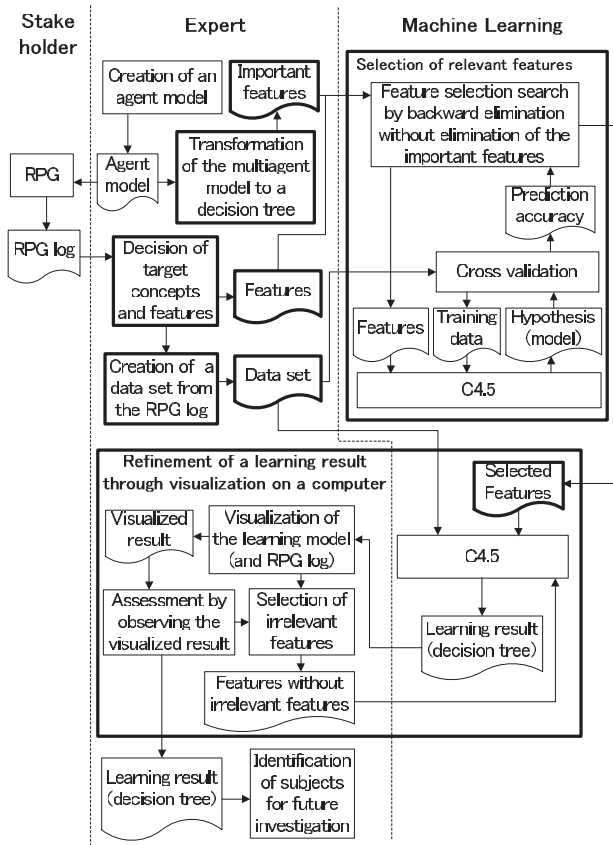


図 4 分類学習によるモデル構築システム

ちらか明かさずにエージェントの意志決定結果として専門家に提示しそれぞれの評価を得る。その後、専門家はどちらが学習モデルによるものであったかを知った上で、その評価を元に直接決定木を吟味する。もし不適切な属性があれば、それが除外され再び C4.5 が適用される。本プロセスは専門家が満足するモデルを得るまで繰り返えされる。本プロセス後に、学習結果と仮説を検証し新たな検討対象を得る。

5. 実データを利用した本手法の適用と評価

5.1 本研究における参加型モデリングプロジェクト

本研究では、IRRI (国際稲研究所)^{*3}と CIRAD (国際農業開発センター)^{*4}のタイ東北部における農業経済を扱った共同プロジェクトに参加した [Vejpas 04]。

本参加型モデリングの背景は次の通りである。タイ東北部において、タイ政府は数種類の米種を重点的に流通させることに成功したが、一方で、一部ではこれまで使用されてきた別の米種も未だに使用されている。また、重点的な米種を流通させるために政府の補助により新たな米種提供業者の参入を容認している。本参加型モデリングでは、1) 米種に関する農民のニーズと米種選択のモデル獲得の把握すること、2) 農民の米種提供業者の選択

表 1 属性候補とその選択過程：5.2 節の 3 つの各手順で選択された属性(表中, “s” にて示す)と除外された属性(表中, “d” にて示す)を示す。尚, 数字は手順を順番に表す。手順 1 では専門家により重要と見なされた属性(図 5 参照), 手順 2 では Wrapper 手法により除外された属性と C4.5 が出力した結果に含まれる属性(図 6 左参照), 手順 3 では専門家により除外された属性と最終的に C4.5 が出力した結果に含まれる属性(図 6 右参照)を示す。

属性名	説明	1	2	3
Change Variety	前年度に作付けしなかった米種を作付けするか	s		
Original Supplier is RRC	現在保持の米種が RRC から仕入れたものか	s		
Previous Seed Class	前年度使用の米種の等級	s	s	s
Close to Ubon	居住地域が Ubon Ratchathani に近いか	s	s	s
Accessibility of RRC	RRC から仕入れ可能か	s	s	s
Farmer Name	参加者 ID		d	
Step	ステップ			
Ethnicity	参加者が属する民族			
Family Number	家族数			
Accessibility of SC	SC から仕入れ可能か			
Total Size of Land	農地の合計			
Size of Upper Land	高地の面積			
Size of Middle Land	中地の面積		d	
Size of Lower Land	低地の面積			
Previous Supplier	前ステップで選択の業者		d	
Original Supplier	現在保持する米種の仕入れ先業者		s	d

モデルと米種提供業者間の米種流通モデルを得ることで、現在の米種提供システムの問題点を同定すること、の 2 点を目的としている。

我々は、複数の品種のそれぞれに、農民の米種選択の選択モデルと米種提供業者の選択モデルを対象として本手法を適用した。本論文では、商業用に広く流通している米種に関する米種提供業者の選択モデルを例に示す。

5.2 分類学習によるモデル構築システムの適用

4.3 節の図 4 で示したシステムを本例題に適用する。初期の全属性候補と各手順で専門家や機械学習により選択 / 削除された属性を表 1 に示す。また、モデルの説明は図 5, 図 6 中に付与された説明を参照のこと。

§1 初期設定

専門家により属性と目標概念(米種業者)の決定が行われ、RPG ログ(2 回分。それぞれ都市部と農村部で行われた)より事例集合が生成された。事例は、RPG における各ステップ(一年に相当)における選択を表す。属性数は 16, 事例数は 80 となった^{*5}。また、専門家の仮説であるエージェントモデルが決定木モデルに変換され(図 5), そこから重要な属性が 5 つ抽出された。この決定木で本事例集合を分類した際の精度は、58.0%であった。

本仮説の概要は次の通りである。まず、Sub Tree(Own Stock (OS) を利用しない場合の純粋な業者の選択モデルと見なせる)では、都市部(Ubon Ratchathani 近郊)においては、最良の品質を提供する RRC から仕入れる

*3 <http://www.irri.org/>

*4 <http://www.cirad.fr/>

*5 最初のステップは全員が業者から調達するという特殊な状況であり事例集合からは除外されている。

ためにはその仕入れルートを持特別に持っている必要がある。尚、都市より離れている場合、専門家は業者がどのように決定されるかは不明である。次に、メインの決定木では、まず、前年度に作付けしなかった品種を利用する場合は、その品種を他から仕入れる必要がある (OS 選択不可)、“Change Variety” が分類条件となる。それ以外の場合は、Own Stock を使うか、業者から仕入れるかを決定する。これには基本的に、前年使用した米の等級を表す “Previous Seed Class” の値 (0 が最もよい等級である) が影響するが、元々自分が仕入れた米種が品質の良い RRC であれば、より OS を長く使い続ける (RRC から仕入れた場合は 3 年間)。

§ 2 属性選択手法による不要な属性の除外

事例集合と属性を利用して、4・3 節 (2) で述べた属性選択を行った。即ち、16 個すべての属性組から始め、山登り法 (後向き探索) により、交差検定から得た予測精度がこれ以上上がらなくなるまで属性を一つずつ削除していった。その際に、5・2・1 節で選定された重要な属性 5 つは削除対象外となり、残りの 11 個の属性から削除が行われた。その結果、予測精度は初期の 66.2% から、3 つの属性が除外された結果 71.2% となった。

§ 3 可視化による専門家への理解促進と精練

最初に生成された事例集合と属性選択手法から得た 11 個の属性から C4.5 により学習モデルを導出した (図 6 左)。次に、可視化手法のために、RPG で使用したポートの特徴をグラフィカルインタフェースとして、RPG 参加者をエージェントとして、Cormas^{*6}[Bousquit 98] 上で実装した (図 2 右)。さらに、エージェントに対して学習モデル (図 6 左) と RPG ログを与えたそれぞれの場合に対して、コンピュータ上でエージェントの意志決定結果として専門家に示していった。専門家はステップ (一年分に相当) ごとに、各エージェントの意志決定結果に対して見解を記録しながら注意深く観察していった。

その後、どちらが学習モデルによるものであったか明かした上で、図 6 (左) の決定木を直接専門家に見せることで評価が行われ、現実に即さない不適切な属性の選出が行われた。専門家は、自分たちが述べた見解と決定木を照らし合わせることで、コンピュータ上で起こった現象の原因を理解することができた。今回は、農村部の農民に関しては、第一ステップに一部が “ST” を選び、それ以降のステップではすべて “OS” が選択されたという特徴的な状況が観察され、その原因である “Original Supplier” が不適切な属性と判断された。なぜならば、住む場所がどこであれ、自分が所持する米種の元々の仕入れ先のみが業者の選択を決定するものとはなりえないからである (この属性は、Own Stock を使用し続ける年数「Seed Class」のみに影響する以外ほとんど考えられないというのが専門家の一致した見解であった)。したがって、この属性を除去し、再び、C4.5 の適用を行った。そ

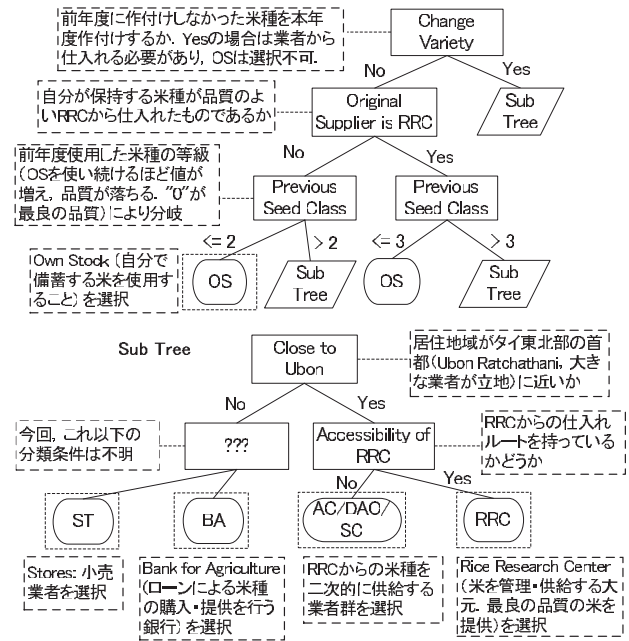


図 5 専門家により作成された決定木 (初期仮説)

の結果を図 6 (右) に示す。再度専門家の評価により問題がなかったため、これを最終的なモデルとした。尚、予測精度は 68.8%、分類精度は 75.0% となった。

5・3 仮説の検証

最終的なモデルは、RPG の結果を反映しステイクホルダーの意志決定の一部を一般に説明するものであるという評価であった。なぜならば、“Close to Ubon = yes” かつ “Previous Seed Class” の値がより大きい場合は、“Accessibility of RRC” が業者決定の要素になり、都市部でよりアクセスしやすい AC/DAO/SC か RRC が選択されていることや、“Close to Ubon = yes” と限定付きではあるが “Previous Seed Class” の値がより小さい場合は、Own Stock (OS) を選択するという専門家の仮説の本質的部分が反映された形となったからである。基本的な構成は問題なかったが、“Previous Seed Class” の閾値が 0 であることは仮説での値 (2 あるいは 3) とは違っており問題となった。しかし、これは RPG ログをエージェントに与え可視化した場合にも同様の現象を専門家が確認しており、RPG でも実際に起こった現象として理解されることとなった。

仮説 (図 5) と学習モデル (図 6 右) の違いから次の検討課題が得られた。まず、学習モデルと仮説との目立った違いは、“Previous Seed Class” の閾値である。この現象は、RPG という環境の中のみで起こるものであるのか (RPG では金のやり取りは仮想的であったため容易に米が購入されていた可能性がある)、一般的に起こるものであるのか検討する必要があると示唆された。また、仮説に含まれた 2 つの属性 (Change Variety, Original Supplier is RRC) は結果に含まれなかった。その原因としてデー

*6 <http://cormas.cirad.fr/indexeng.htm>

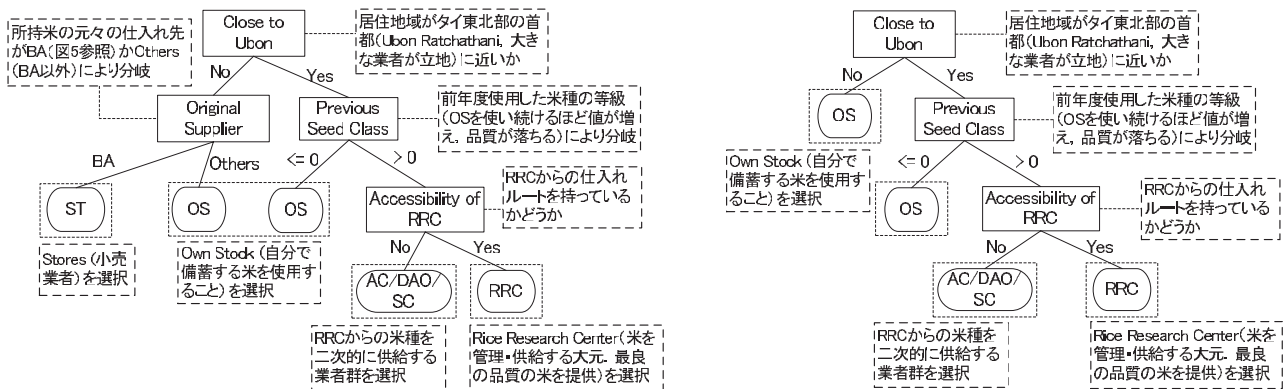


図 6 属性選択手法適用後出力された学習結果（左）と専門家の評価後改良された学習結果（右）

タが十分でなかったことも考えられ、その重要性について検討する必要が示唆された。また、結果の中には十分にステイクホルダーの行動を説明していると言い切れないものもあり（例えば、“Close to Ubon = no -> OS”では、都市から離れているだけで必ず OS を選ぶとは考えにくい）、今回データに含めた属性以外で専門家が気付いていない重要な属性が存在する可能性があり、これも検討課題となった。これらの課題は、ステイクホルダーへの追跡調査、専門家間の議論、及び、今後の RPG により検討される必要がある。

5.4 本手法の考察

今回採用した接近法の有効性を次のように確認した。まず、属性選択手法の利用に関しては、専門家による学習結果の精練の手間を軽減されるものとなった。今回の例では、属性選択手法を利用しない場合、即ち、専門家による属性の絞り込みだけを行う場合、属性の削除と C4.5 を適用する修正プロセスを 3 回要したのに対し、使用した場合は 1 回で済んだ（図 6 左から “Original Supplier” を削除したのみ）。今回の例題以外においても、属性選択手法が専門家への負担を軽減するために効果を挙げた。今回のように可視化手法を用いる場合はその準備コストを考えると、その回数がより少ない方が望ましく、本プロセスは不可欠と考えられる。また、属性選択の過程で、仮説である決定木から抽出した重要な属性は除外しないこととしたが、これを行わない場合、専門家が重要と考える属性も除外されてしまい、結果的に専門家が納得する結果を得ることが難しくなった場合があった。

次に、学習されたモデルの可視化による評価法は、モデルの持つ特性をより理解し、より中立的にモデルを評価するために役立つと考えられる。RPG と同じ状況を再現しているために、専門家は各ステップでエージェントの意志決定結果について様々な意見を述べる事ができた。これは、当初に決定木だけ示していたときよりも多くの意見を引き出す事ができた。今回の例では、農村部の農民に関してはほとんど “OS” が選択されるのに対し、最初のステップの一部のみが “ST” を選択したため、

その特異な状況を簡単に把握することができ、続いてそれに対応する属性の除外が容易に行われた。また、結果的に属性の除外はされなかったものの、“Previous Seed Class” の閾値が 0 だったために、一部のエージェントにおいて OS とその他の業者が交代で表示される結果となり、その状況について特に多くの発言を得た。専門家の仮説と相違があったにもかかわらずこの現象が受け入れられたのは、RPG ログをエージェントに与えた場合に示された結果も似た現象を示されたからであり、専門家がより中立的に評価できたという点でこの方法の有効性を示している。また、以上については、決定木を直接見せるよりも達成が容易と考えられ、視覚的に結果を見せることの利点を示している。

本手法により得られたモデルは、RPG ログから抽出できる最も一般性の高く唯一のものと言い切ることはできない。しかしながら、属性選択手法ができるだけ予測精度を向上させることでモデルの一般性を高め、さらに可視化によって専門家が学習モデルをより深く理解し精練することで、専門家が納得できるモデルを提供できた。したがって、専門家は仮説との比較より得られた検討課題を真摯に受け止め、今後その解決のため取り組むことができ、この点が特に重要である。

今後、その他の例題を取り上げるによりさらに本手法自体の評価・精練を行う必要があるが、機械学習により人手で同時に扱うには困難な数の属性組から、専門家の仮説を検証するための価値のあるモデルを提供する体系的な手法を提供できたと考えられる。

6. おわりに

本研究では、参加型モデリングにおいて、RPG ログデータに対し分類学習を適用する新しいエージェントモデル構築法を提案した。本手法により、モデル構築者の能力に依らず、多くの要因を一度に扱うことのできる体系的なモデル構築を行うことができる。分類学習を適用する際の問題としては、1) データ収集コストが大きいため、属性の数に比べ十分な事例が得られない。2) データ量の不十分さゆえにノイズが学習結果に影響を及ぼす可

能性が高い。3) 学習結果を構成する分類条件は、人間の意志決定モデルとして説明でき、専門家を納得させるものでなくてはならない、の3点である。これに対し、我々は、1) 専門家が重要と判断する属性は残しつつ、属性選択手法により初期属性集合から予測精度が最も高くなる属性集合を同定し、2) 実施したRPGと同じ環境と参加者をコンピュータ上でグラフィカルインタフェースとエージェントとして再現し、その条件下でエージェントに学習モデルを与え、意志決定結果を示すことで理解促進を図り、機械学習から得た結果に含まれる不適切な属性をさらに除外する、という属性選択の手法によりエージェントモデル構築システムを実現し、問題の解決を行った。最後に、実際に、タイ東北部の農業経済の参加型モデリングに参加し、当該の専門家と共に評価することで、本手法の有効性を確認した。

本論文の意義は、新たなエージェントモデル構築法の可能性を議論・提案し、その有用性を実際例により示したことである。これまでにRPGログから機械学習を用いて現実を反映するモデルを得る体系的な手法の研究はなかったと言える。東南アジア、アフリカ、ヨーロッパで様々な事例 [Bousquet 02] に適用され、研究が進んでいる参加型モデリングは、本研究により工学的な支援を得ることでさらに方法論として充実したものとなる。今後、その他の属性選択手法や演繹学習を含め他の学習手法を適用することで、機械学習によるエージェントモデル構築手法の更なる充実が期待される。

謝 辞

本研究の手法に対しご助言いただいた CIRAD の Guy Trébuil 氏、Ubon Ratchatani 大学の Chirawat Vejjapas 氏の両氏に心より感謝の意を表す。また、本研究は、21世紀 COE「知識社会基盤構築のための情報学拠点形成」若手研究者育成プログラム、日本学術振興会科学研究費基盤研究 (A)15200012 の助成金により行われた。

◇ 参 考 文 献 ◇

- [Almuallim 94] Almuallim, H. and Dietterich, T. G. Learning Boolean Concepts in the Presence of Many Irrelevant Features. *Artificial Intelligence*, Vol. 69, pp. 279-305, 1994.
- [Bousquet 98] Bousquet, F., Bakam, I., Proton, H. and Le Page, C. Cormas: common-pool resources and multi-agent Systems. *Lecture Notes in Artificial Intelligence*, Vol. 1416, pp. 826-838, 1998.
- [Bousquet 99] Bousquet, F., Barreteau, O., Le Page, C., Mullon, C. and Weber, J.: An environmental modelling approach. The use of multi-agent simulations. Blasco F and Weill A (Eds.) *Advances in environmental and ecological modeling*, Elsevier, pp. 113-122, 1999.
- [Bousquet 02] Bousquet F., Barreteau O., d'Aquino P., Etienne M., Boissau S., Aubert S., Le Page C., Babin D., Castella J.C.: Multi-agent systems and role games: an approach for ecosystem co-management. In: Janssen M. (eds) *Complexity and Ecosystem Management: The Theory and Practice of Multi-agent Approaches*. Edward Elgar Publishers, pp. 248-285, 2002.
- [Gilbert 99] Gilbert, N. and Troitzsch, K. G.: Simulation for the social scientist. *Milton Keynes: Open University Press*. 1999.
- [Gilbert 02] Gilbert, N., Maltby, S. and Asakawa, T.: Participatory simulations for developing scenarios in environmental resource management. In *Proceedings of Third Workshop on Agent-Based Simulation*, pp. 67-72, 2002.
- [Helbing 00] Helbing, D., Farkas, I.J. and Vicsek, T.: Simulating Dynamical Features of Escape Panic. *Nature*, 407(6803), pp. 487-490, 2000.
- [石野 95] 石野洋子, 寺野隆雄: 模擬育種法と帰納学習を適用したマーケティング情報分析, 人工知能学会誌, Vol. 12, No. 1, pp. 121-131, 1997.
- [Kita 03] Kita, H., Sato, H., Mori, N. and Ono, I.: U-Mart System, Software for Open Experiments of Artificial Market, In *Proceedings of IEEE International Symposium on Computational Intelligence in Robotics and Automation*, pp. 1328-1333, 2003.
- [Kohavi 97] Kohavi, R. and John, G.: Wrappers for feature selection. *Artificial Intelligence*, Vol. 97, pp. 273-324, 1997.
- [Murakami 03] Murakami, Y., Ishida, T., Kawasoe, T. and Hishiyama, R.: Scenario Description for Multi-Agent Simulation. In *Proceedings of International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-03)*, pp. 369-376, 2003.
- [Quinlan 86] Quinlan, J. R.: Induction for Decision Trees, *Machine Learning*, Vol. 1, No. 1, pp. 81-106, 1986.
- [Quinlan 93] Quinlan, J. R.: C4.5: Programs for Machine Learning, *Morgan-Kaufmann*, 1993.
- [Vejjapas 04] Vejjapas, C., Bousquet, F., Naivinit, W., Trébuil, G., Srisombat, N. Participatory modeling for managing rainfed lowland rice variety and seed systems in lower north-east Thailand, *Mekong Rice Conference*, OS3.4, 2004.

〔担当委員：柏原 昭博〕

2005 年 5 月 2 日 受理

著 者 紹 介



鳥居 大祐 (学生会員)

2001 年京都大学工学部情報学科卒。2003 年同大学院社会情報学専攻修士課程修了。現在、同大学院社会情報学専攻博士課程に在学中。機械学習、マルチエージェントシミュレーション、P2P システムに興味を持つ。



ブスケ フランソワ

1994 年リヨン第一大学にて生物数学の博士号を取得。1993 年よりフランスの国際農業研究所 (CIRAD) に所属し、現在、チュラロンコン大学との共同プロジェクトにてタイを活動の中心に置く。再利用可能な資源管理のシミュレーションのためにマルチエージェントシステムの開発や利用に興味を持つ。



石田 亨 (正会員)

1976 年京都大学工学部情報学工学科卒業、1978 年同大学院修士課程修了。京都大学大学院情報学研究所社会情報学専攻教授。工学博士。IEEE Fellow。情報処理学会フェロー。人工知能、コミュニケーション、社会情報システムに興味を持つ。