

# 環境問題の解決法のマルチエージェントシミュレーションによる模索

近藤良久\* 藤井晴行† 谷本潤‡

\*東京工業大学工学部建築学科 〒152-8550 東京都目黒区大岡山 2-12-1

†東京工業大学大学院理工学研究科 〒152-8550 東京都目黒区大岡山 2-12-1

‡九州大学大学院総合理工学研究院 〒816-8580 福岡県春日市春日公園 6-1

E-mail: \*ykondo@post.arch.titech.ac.jp †hfujii@arch.titech.ac.jp ‡tanimoto@cm.kyusyu-u.ac.jp

あらまし 人間、環境、社会を相互に影響し合うひとつの系としてマルチエージェントシミュレーションを行う。環境との共生を可能にする環境デザインの方法の提案を最終目的とする。本稿はエージェントの態度と系の振舞いに注目する。

キーワード マルチエージェントシミュレーション、人間-環境-社会系、Q学習

## Multi-Agent Simulation towards Solutions of the Environmental Problem

Yoshihisa KONDO\* Haruyuki HUIJII† Jun TANIMOTO‡

\*Department of Architecture, Tokyo Institute of Technology 2-12-1 O-okayama, Meguro-ku, Tokyo, 152-8550 Japan

† Graduate School of Science and Engineering, Tokyo Institute of Technology 2-12-1 O-okayama, Meguro-ku, Tokyo, 152-8550 Japan

‡ Interdisciplinary Graduate School of Engineering Sciences, Kyusyu University

6-1 Kasugakoen, Kasuga-shi, Fukuoka, 816-8580 Japan

E-mail: \*ykondo@post.arch.titech.ac.jp †hfujii@arch.titech.ac.jp ‡tanimoto@cm.kyusyu-u.ac.jp

**Abstract** Some case studies using multi-agent simulation to observe the relations between people's attitude towards environment and thermal comfort are described. The final goal is to find the crews to design an environmentally-symbiotic and sustainable society.

**Keyword** Multi-Agent Simulation, Man-Environment-Society System, Q-learning

### 1. はじめに

環境問題は社会の問題でもある。しかし、その発生は社会を構成する個々の人間の行動に起因する。一方、個人個人の行動は環境や社会の動きに影響される。即ち人間と環境と社会との間には transactional (交互浸透的) な関係がある。本研究は、人間と環境と社会を一つの系として捉え、環境問題の発生メカニズムや解決の糸口を探ろうとするものである。

環境問題に関する考察は、理学・工学的な観点及び社会学的な観点からなされている。しかし、これらは必ずしも緊密に結びついているわけではない。前者は環境の物理学的な振舞いに着目し、多くの場合、人間や社会の振舞いは、その内容よりもむしろ帰結となる状態値が予め設定された計算条件として、半ば固定的に与えられる。建築環境学においては、例えば村上<sup>[1]</sup>、奥山<sup>[2]</sup>、谷本<sup>[3]</sup>らが建物や都市の温熱環境のシミュレーション手法を構築しているが、人間の営為は伝熱・流体解析に用いられる物性値として与えられることが多い。後者は、前者に対して、人間の行動やその総体

である社会の振舞いの特徴に着目するが、環境の物理学的な挙動は、筆者らの知る限り、比喩的に表現されるにとどまっている。例えば、本研究でも用いるマルチエージェントシミュレーションにおいては出口<sup>[4]</sup>や大内<sup>[5]</sup>らが環境問題を放草地のジレンマとして捉え、社会的ジレンマの下で社会が持続するような役割、グループなどの創発について論じている。本研究は両観点を融合し、環境に着目するシミュレーションと人間の行動に着目するシミュレーションを連成するマルチエージェントシミュレーションによって、環境問題を人間、環境、社会のシステムとして表現可能か否かを検討するものである。本稿は、その試行錯誤の一つとして、トレードオフの関係にある快適性と親環境性を評価基準として行動の学習をするエージェント達が、環境や社会と作用しあいながら、それらの状態や自分の行動を変化させていく様子を観察し、環境問題における人間、環境、社会の関わり方を考察する。なお本稿では、エージェントが多数派の行動に影響を受けることとして人間と社会の関係を表現する。

## 2. モデル

本稿ではN戸の住宅群とその住民達が共用する外部空間をモチーフに環境問題における社会的ジレンマモデルを構築する。設定および物性値は問題点を明確にするため恣意的なものを含む。各戸には1人の住民(エージェント)が住み、冷房が設置されている。冷房の屋外機は共用外部空間に設置されている(図1)。エージェントは、昼夜を問わず、窓の開閉や冷房の操作などの熱環境調整行動をなす。行動は室温や窓・冷房の状態や社会の状態とに基づいて選択され、自分の行動評価基準に照らし合わせて計算ステップ毎に算出される報酬をもとに学習される。環境調整行動は室内環境や共用外部空間の熱環境を変化させる。

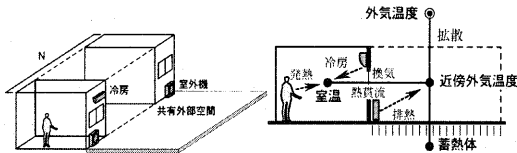


図1 モデル概要

### 2.1. 伝熱モデル

熱移動は以下の状態方程式を用いて逐次計算する。

$$M \frac{dx}{dt} = C \cdot x + C_0 \cdot x_0 + R \cdot g_0 \quad \dots(1)$$

各室及び近傍外気、蓄熱体の各節点はそれぞれ熱容量をもち、 $M$  はそれらの熱容量を表す行列である。 $C$  は各節点間の熱コンダクタンスを表す行列、 $C_0$  は各節点と設定室温や設定外気温などの規定温度間の熱コンダクタンスを表す行列である。 $R$  は熱が各節点に吸収される比率を表す行列である。 $x$  は各節点の温度を示す状態ベクトル、 $x_0$  は設定室温度や設定外気温度を示すベクトル、 $g_0$  は発熱や除熱、排熱を表すベクトルである。計算の基となる物性値は表1のように与える。

### 2.2. エージェント

エージェントは環境と社会の状態を示す9ビットの入力(表2)に基づいて行動を決定する。一回になす行動は3種類の行動の組合せである(表3)。各状態での全行動に対し行動価値関数(Q値)を与える。行動のQ値に基づいて選択される<sup>注3)</sup>。エージェントは行動の結果得られた報酬を基にQ学習を行う<sup>注4)</sup>。報酬はrを快適性、親環境性、社会性の観点から(2)式のように定める。また報酬重みの組 $\langle w_c, w_e, w_s \rangle$ を態度とよぶ。

$$r = w_c \cdot e_c + w_e \cdot e_e + w_s \cdot e_s \quad \dots(2)$$

## 3. シミュレーション

### 3.1. ケース

前節で説明したモデルを用いてシミュレーションを行う。外気温度は22~30℃に変化する周期1日の正弦波として与える。10分間隔で環境の状態とエージェ

表1 物性値

住戸数	54戸
壁面積	4[m]×3[m]=12[m <sup>2</sup> ]
各室体積	4[m]×3[m]×4[m]=48[m <sup>3</sup> ]
近傍体積	4[m]×4[m]×5[m]×54戸=4320[m <sup>3</sup> ]
住宅壁厚	20[cm](コンクリート15[cm],断熱材5[cm])
熱伝導率(コンクリート)	1.51[W/(m·K)]
熱伝導率(断熱材)	0.042[W/(m·K)]
熱伝達率(内表面)	9.3[W/(m <sup>2</sup> ·K)]
熱伝達率(外表面)	23.3[W/(m <sup>2</sup> ·K)]
熱貫流率	0.66[W/(m <sup>2</sup> ·K)]
貫流熱	0.66[W/(m <sup>2</sup> ·K)]×12[m <sup>2</sup> ]×2[面] <sup>注1)</sup> =15.8[W/K]
空気比熱	1.01[kj/(kg·K)]
空気密度	1.30[kg/m <sup>3</sup> ]
換気回数(窓閉)	0.1[回/h]
換気回数(窓開)	3[回/h]
外気の拡散係数	0.44[m <sup>2</sup> /s](但し接地境界層厚100m)
各室内発熱	150[W]
冷房除去熱(弱)	200[W]
冷房除去熱(強)	400[W]
COP値	2.0-0.1×(θ <sub>0</sub> -18) <sup>注2)</sup>

表2 知覚可能な状態

知覚可能な状態(9bit, 512状態)
$S = \langle s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8, s_9 \rangle, s_i \in \{0,1\}$
$\langle s_1, s_2 \rangle$ : 室温の状態(2bit)
00: 室温が22℃以下, 01: 室温が22℃~26℃(快適域)
10: 室温が26℃~30℃, 11: 室温が30℃以上
$\langle s_3, s_4 \rangle$ : 近傍外気温の状態(2bit)
00: 近傍外気温が24℃以下, 01: 近傍外気温が24℃~28℃
10: 近傍外気温が28℃~32℃, 11: 近傍外気温が32℃以上
$\langle s_5 \rangle$ : 窓の状態(1bit)
0: 閉, 1: 開
$\langle s_6 \rangle$ : 冷房の状態(1bit)
0: 運転, 1: 停止
$\langle s_7 \rangle$ : 冷房の設定(1bit)
0: 強, 1: 弱
$\langle s_8 \rangle$ : 1日の平均気温(1bit)
0: 29℃以下, 1: 29℃より高い
$\langle s_9 \rangle$ : 他のエージェントの動向(1bit)
0: 過半数が冷房使用, 1: 過半数が冷房不使用

表3 可能な行動

可能な行動a(3種類, 8通り)
$a \in \{ \langle w, \{Cw\}, \{Cac\}, \{Chl\}, \{Cw, Cac\}, \{Cw, Chl\}, \{Cac, Chl\}, \{Cw, Cac, Chl\} \rangle \}$
Cw: 窓の状態変更(開→閉, 閉→開)
Cac: 冷房の状態変更(運転→停止, 停止→運転)
Chl: 冷房の設定変更(強→弱, 弱→強)

表4 報酬における評価

$e_c$ : 快適性評価
$\langle s_1, s_2 \rangle = \langle 0,0 \rangle$ (室温22℃以下) → $e_c = -1$
$\langle s_1, s_2 \rangle = \langle 0,1 \rangle$ (室温22~26℃, 快適域) → $e_c = 1$
$\langle s_1, s_2 \rangle = \langle 1,0 \rangle$ (室温26℃~30℃) → $e_c = 0$
$\langle s_1, s_2 \rangle = \langle 1,1 \rangle$ (室温30℃以上) → $e_c = -1$
$e_e$ : 親環境性評価
$\langle s_6, s_7, s_8 \rangle = \langle 1,0,0 \rangle$ or $\langle s_6, s_7, s_8 \rangle = \langle 1,1,0 \rangle$ (冷房不使用, 温暖化無) → $e_e = 1$
$\langle s_6, s_7, s_8 \rangle = \langle 1,0,1 \rangle$ or $\langle s_6, s_7, s_8 \rangle = \langle 1,1,1 \rangle$ (冷房不使用, 温暖化有) → $e_e = 0.5$
$\langle s_6, s_7, s_8 \rangle = \langle 0,0,0 \rangle$ (冷房弱使用, 温暖化無) → $e_e = -0.25$
$\langle s_6, s_7, s_8 \rangle = \langle 0,1,0 \rangle$ (冷房強使用, 温暖化無) → $e_e = -0.5$
$\langle s_6, s_7, s_8 \rangle = \langle 0,0,1 \rangle$ (冷房弱使用, 温暖化有) → $e_e = -0.75$
$\langle s_6, s_7, s_8 \rangle = \langle 0,1,1 \rangle$ (冷房強使用, 温暖化有) → $e_e = -1$
$e_s$ : 社会性評価
$\langle s_6, s_9 \rangle = \langle 0,0 \rangle$ (冷房使用かつ冷房使用が多数派) → $e_s = 1$
$\langle s_6, s_9 \rangle = \langle 1,1 \rangle$ (冷房不使用かつ冷房不使用が多数派) → $e_s = 1$
$\langle s_6, s_9 \rangle = \langle 0,1 \rangle$ (冷房使用かつ冷房不使用が多数派) → $e_s = -1$
$\langle s_6, s_9 \rangle = \langle 1,0 \rangle$ (冷房不使用かつ冷房使用が多数派) → $e_s = -1$

ントの行動を400日間分計算する。態度をパラメータとしてその構成を変え、以下の各ケースを計算する。

**CASE1: 態度<1,0,0>のみ**

全エージェントが快適性の評価のみに基づいて行動する。この態度をもつエージェントを快適追求型エージェントとよぶ。

**CASE2: 態度<0,1,0>のみ**

全エージェントが親環境性の評価のみに基づいて行動する。この態度をもつエージェントを親環境型エージェントとよぶ。

**CASE3: 態度<1,0,0>, 態度<0,1,0>がそれぞれ 50%**

快適追求型エージェントと親環境型エージェント

が半数ずついる。

**CASE4: 態度 <0.5,0.5,0>のみ**

全エージェントは快適性と親環境性を等しい重みで評価し行動する。

**CASE5: 態度<1,0,0>, 態度<0,1,0>, 態度<0,0,1>がそれぞれ 33.3%**

快適追求型エージェント、親環境型エージェント、付和雷同型エージェント（態度<0,0,1>）がそれぞれ全人口の三分之一を占める。

**CASE6: 態度<0.33,0.33,0.33>のみ**

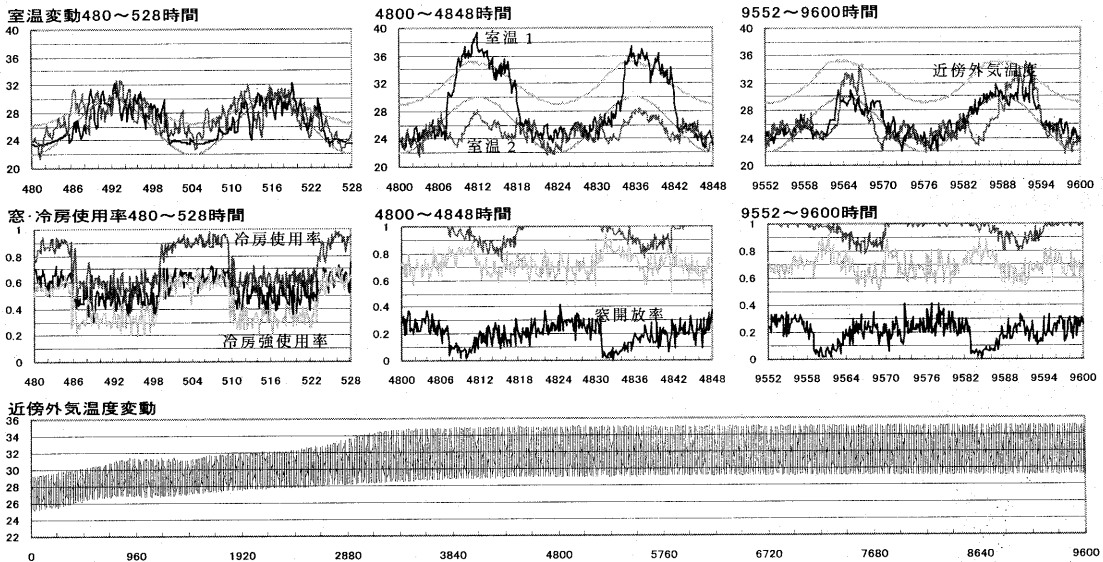


図 2 CASE 1 態度<1,0,0> (快適性追求型エージェント)

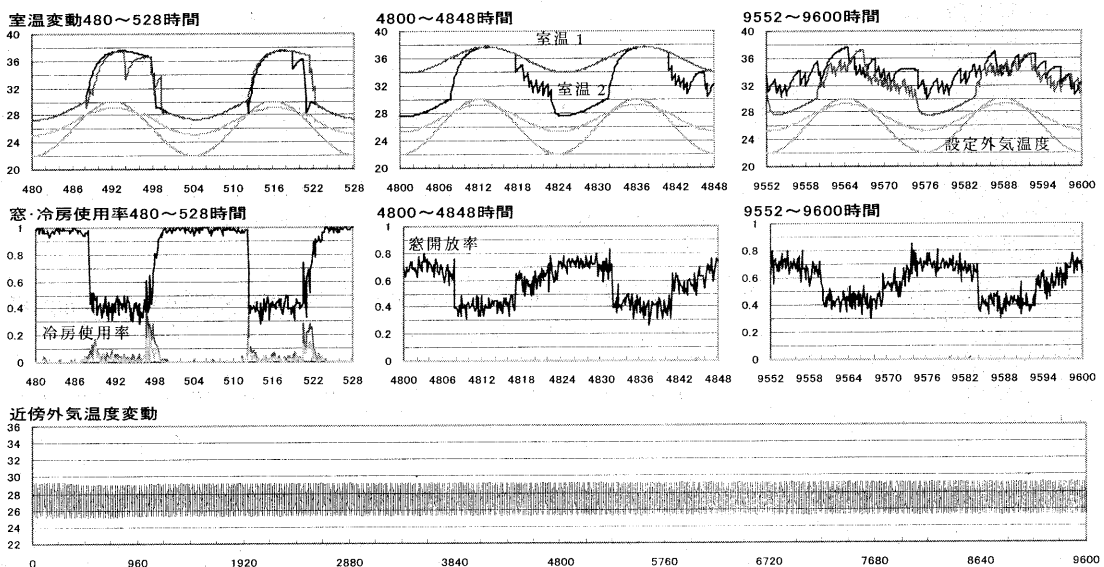


図 3 CASE 2 態度<0,1,0> (親環境型エージェント)

全エージェントは快適性、親環境性、社会性のそれぞれを等しい重みで評価し行動する。

**CASE7:**態度<0.5,0,0.5>態度<0,0.5,0.5>がそれぞれ50% 快適性と社会性のみ等しい重みで評価するエージェント、親環境性と社会性のみ等しい重みで評価するエージェントがそれぞれ半数ずつを占める。

### 3.2. エントロピーの算出

エージェントや系の振舞いに対し、影響する要因を特定するために、エージェントの行動のエントロピーを求め、エージェントが知覚する環境や社会の状態の

情報量を算出する。情報量およびその算出の基となるエントロピーは以下のように求める<sup>15)</sup>。

$$I(X, Y) = H(X) - H(X|Y) \quad \dots(3)$$

$$H(X) = - \sum_{x \in X} p(x) \log p(x) \quad \dots(4)$$

$$H(X|Y) = - \sum_{x \in X, y \in Y} p(x, y) \log p(x|y) \quad \dots(5)$$

快適性に関する入力の情報量:  $I_{com}=I(X, Y_c)$

親環境性に関する入力の情報量:  $I_{env}=I(X, Y_e)$

社会性に関する入力の情報量:  $I_{soc}=I(X, Y_s)$

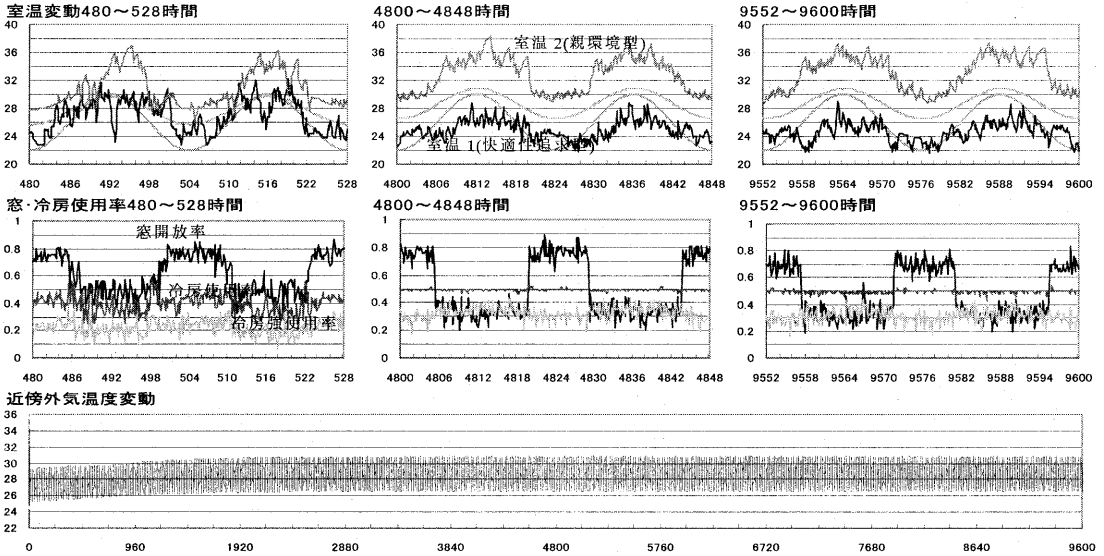


図 4 CASE 3 態度<1,0,0>1/2, 態度<0,1,0>1/2

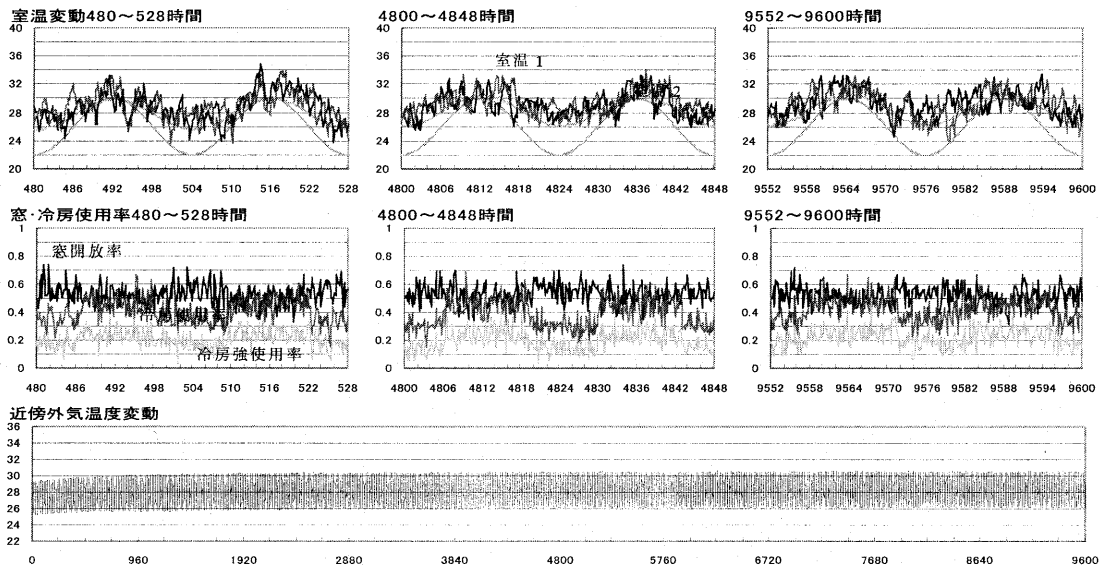


図 5 CASE 4 態度<0.5,0.5,0>

$$X = \{1, 2, 3, 4, 5, 6, 7, 8\}$$

$$Y_c = \{(s_1, s_2); s_1, s_2 \in \{0, 1\}\}$$

$$Y_c = \{(s_5, s_6, s_7); s_5, s_6, s_7 \in \{0, 1\}\}$$

$$Y_s = \{(s_5, s_8); s_5, s_8 \in \{0, 1\}\}$$

### 3.3. 結果

図2～図9に各ケースの室温の変動、窓・冷房の使用率の変動(任意の1試行)、近傍外気温度の変動(5試行の平均)を示す。

**CASE1:** 多くの快適性追求による冷房使用が外気の温暖化を招き、エネルギー消費を増加させる。都市の温

暖化する現象の一端が示されている。

**CASE2:** 温暖化は起こっていない。

**CASE3:** 快適性追求型エージェントは冷房を連続的に使用し室温を快適域に保つが、親環境型エージェントが冷房を使用しない。CASE1ほどは温暖化しない。

**CASE4:** 快適性、親環境性のいずれの観点からも、CASE3より好ましくない結果が見られる。

**CASE5:** 本稿の計算条件では、付和雷同型エージェントは冷房を使用しないことを学習した。

**CASE6:** 温暖化の進行にわずかながらうねりがみられ

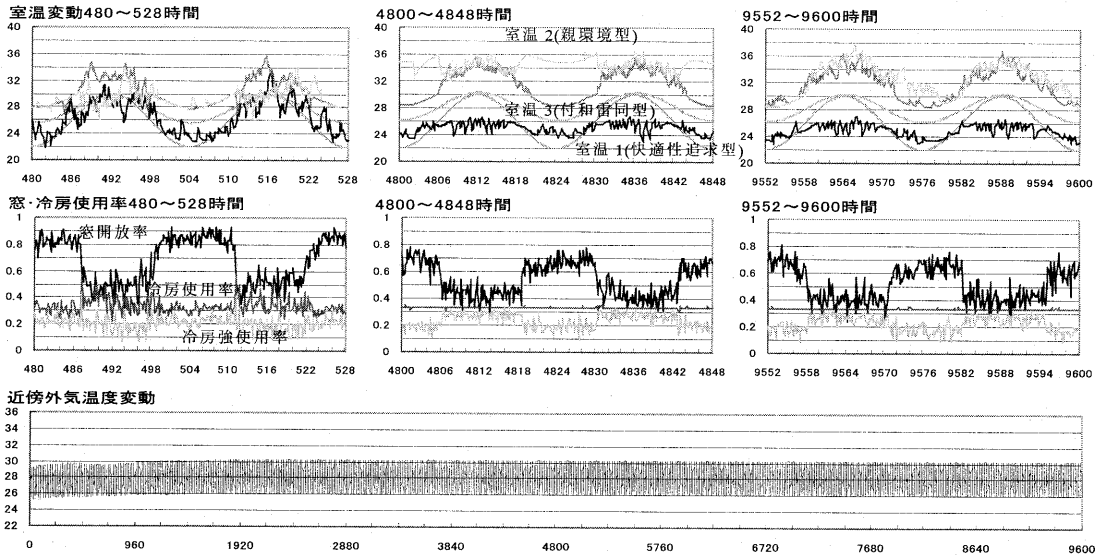


図 6 CASE 5 態度<1,0,0>1/3, 態度<0,1,0>1/3, 態度<0,0,1>1/3

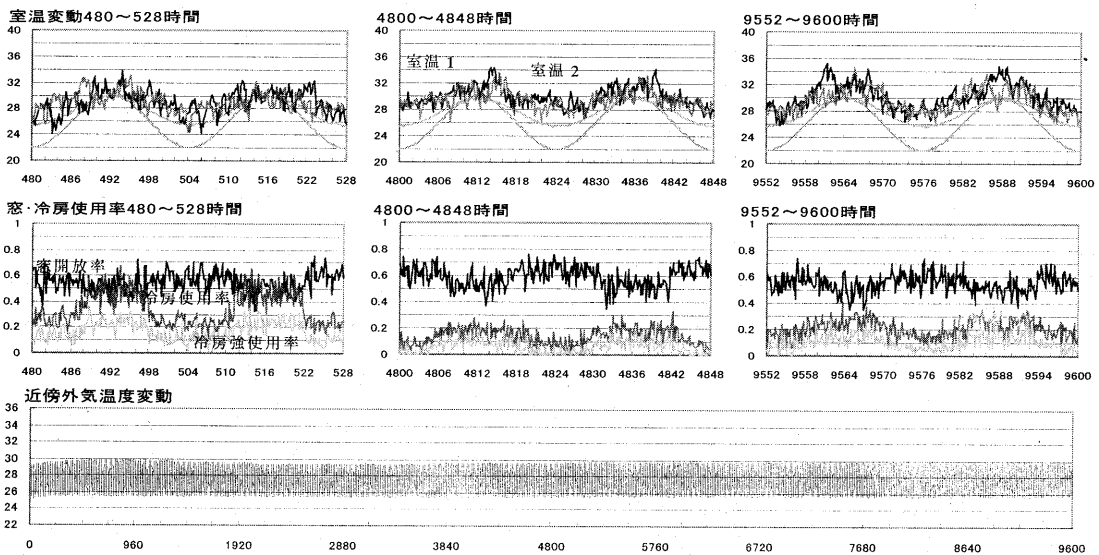


図 7 CASE 6 態度<0.33,0.33,0.33>

る。冷房の使用率は一旦下がるがその後徐々に増加している。

**CASE7:** 本稿の計算条件では、冷房を使用しない行動が過半数となる。

図8は態度<1,0,0>のエージェントのケース毎の入力の情報量を示した図である。同じ態度をもつエージェントであってもケースによりその行動に寄与する情報にわずかながら違いが見られる。

#### 4. まとめ

態度から見たエージェントの構成の違いによる、行動および系のふるまいの違いを観察した。

**謝辞** 本研究の一部は科研費萌芽研究（#14658122）、基盤研究（#14205087）による。関係各位に記して謝辞を表す。

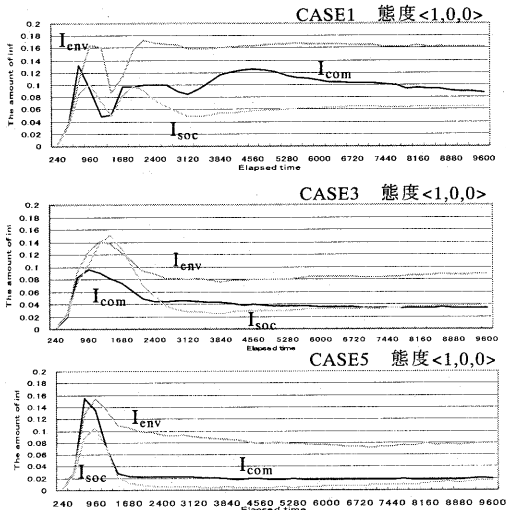


図8 情報量

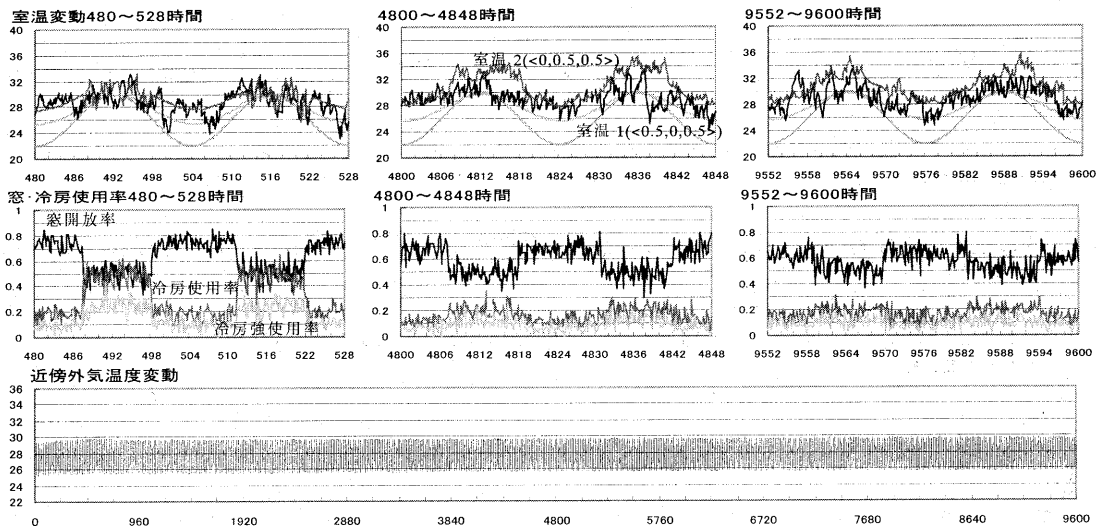


図9 CASE7 態度<0.5,0,0.5> > 1/2, 態度<0,0.5,0.5> > 1/2

#### 【注】

注1) 実際の住宅は床と天井を合わせた6面全てにおいて熱の授受が行われるが、本モデルは近傍外気に接する壁からの熱の授受のみを計算する。そのためこのままではモデルの各家は過剰な断熱状態にあるといえる。その過剰な断熱状態を回避するためにあえて近傍外気との熱のやりとりを壁2面として計算することとする。

注2) 実際の冷房のCOP値と比較し小さく、また外気温のみで決定されるという極端な設定である。

注3) ある状態sにおける行動 $a_i$ はその行動価値関数 $Q(s, a_i)$ の値に基づき以下の確率にて選択される。

$$\frac{\exp(Q(s, a_i)/\tau)}{\sum_{j=1}^8 \exp(Q(s, a_j)/\tau)}$$

ただし温度 $\tau$ は1とする。

注4)  $Q(s, a)$ の初期値はすべてゼロである。この $Q(s, a)$ を更新していくことで学習を行い、エージェントはエージェントのもつ信念のもと環境に適合した行動を身にかけていく。エージェントはステップ事に行動の結果変化した状態を覚知し、その状態を評価し得られた報酬rにより以下のように $Q(s, a)$ を更新する<sup>[7]</sup>。

$$Q(s, a) \leftarrow Q(s, a) + \alpha[r + \gamma \max_{a'} Q(s', a') - Q(s, a)]$$

学習率 $\alpha$ を0.1、割引率 $\gamma$ を0.9とする。

#### 【参考文献】

- [1] 村上周三, "CFDによる建築・都市の環境設計工学," 東京大学出版, 2000.
- [2] 奥山博康, "熱・換気回路網モデル計算プログラム NETS の検証," IBPSA-Japan 講演梗概集, 2002.
- [3] 萩島理, 谷本潤, 片山忠久, 大原健志 "改良・建築・都市・土壌連成系モデル (AUSSSM) による都市高温化の構造解析 第1報 モデルの理論構成及び標準解," 日本建築学会計画系論文集, No.550, pp79-86, 2001.12.
- [4] 出口弘, "複雑系としての経済学," 日科技連, 2002.
- [5] 大内東, 山本雅人, 川村秀憲, "マルチエージェントシステムの基礎と応用-複雑系工学の計算パラダイム-", コロナ社, 2002.
- [6] 広瀬幸雄, "環境と消費の社会心理学-共益と私益のジレンマ-", 名古屋大学出版会, 1995.
- [7] Richard S Sutton, Andrew G. Barto, 三上貞芳・皆川雅章共訳 "強化学習" 森北出版株式会社, 2000.
- [8] 藤井晴行, "屋内環境の変動と居住者の環境調整行動との相互作用の計算モデル," 日本建築学会計画系論文集, No.540, pp15-22, 2001.2.