人　工　知　能 77－8
（１９９１．７．１７）

分散 AI へのゲーム理論による取組みと協調

中村 豪一　出口 光一郎
東京大学工学部

分散 AI などの人工的分散システムにおける各システム構成要素の分散意思決定には、構成要素同志の相互作用や行動決定に必要な情報の散在による困難が伴う。本論文ではシステムをゲーム理論的にモデル化した上で、各システム構成要素が自分の持っている情報のみに基づいて利己的に分散意思決定を行なうが、結果としてシステム全体にとって協調をもたらす分散意思決定方法について述べるとともに、システム構成要素の相互作用関係の協調度を記述する指標を提示する。

Game Theoretical Approach to Distributed Artificial Intelligence and Cooperation

Goichi Nakamura and Koichiro Deguchi
Faculty of Engneering, Univercity of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113 Japan

In a distributed artificial intelligence, it is difficult for each agent in the system to make its decisions because of interactions amang all the agents and dispersion of information needed for the decision making. In this paper, first, we model the system behavior as game playing, and describe how to make decisions distributedly for each agent. Each agent makes decisions self-interestedly and only with its own information, but as the result a kind of cooperation can be obtained amang agents. Also, we propose new indexes which describe the cooperation level of the system.

1 はじめに

この論文では全体として何らかの問題解決を行なう分散AIなどの人工的分散システムにおける各システム構成要素の意思決定法について考える。分散システムでは一般にそのシステム構成要素同士が相互作用する、つまり自分の行動の結果が自分以外のシステム構成要素の行動によって影響を受ける、という状況がある。

この状況をモデル化する手段としてゲーム理論を使う。

ところで，システム構成要素が相互作用するという状況下でシステムがうまく働くには協調が成立しなくてはいけない。すなはち各構成要素の行動が全体としては協調となっていなくてはいけない。

そのような行動決定をするのには相手がどんな状況にあるのか、どういった行動をとろうとしているのか、といった情報がいる。当然そういった情報はシステム全体に散らばっている。システムの内部状態、あるいは外部環境の変化が激しい場合はそれを各システム構成要素が推測しようとしても難しい。

そこで、

1. システム全体の情報を一箇所にあつめてそこで集中意思決定を行なう。あるいは
2. 各システム構成要素が自分の持っている情報のみに基づいて、協調ということは考えずに利己的に分散意思決定を行ない結果として協調になることを狙う、という二つの方法が考えられる。

本論文では第二章でゲームの理論の基本とともに2. の方法について説明する。そして、第三章で 通信を導入して、2. を改良した方法について述べ、それが、結果として協調になるという狙いを達成する方法であることを示す。それとともにシステム構成要素間の相互作用関係の指標となる概念を提示する。

2 通信無の行動決定法

2.1 ゲーム

構成要素が相互作用しあう分散AIなどの分散システムをゲームとしてモデル化する。

o ゲームに参加するplayer(システム構成要素)の集合を P とする。

o P に属するplayerはそれぞれいくつかの行動選択肢を持つ。player-iの行動選択肢を m_i と、全ての m_i から成る集合を M_i と、表記する。各playerは自分以外のplayerがどんな行動選択肢をもつのかを知っている。持ち論、自分の行動選択肢も知っている。それぞれのplayerの行動を組合せたものを行動組合せと呼び、m_P と表記する。m_{P-i} はplayer-i以外のplayerの行動組合せである。全ての m_P, 全ての m_{P-i} から成る集合をそれぞれ M_P,M_{P-i} と表記する。

o 行動組合せ m_P の各platerにとっての望ましさを表したものが利得である。利得の値が大きいほど望ましい。ゲームにおいて、playerはなるべく高い利得を求める。

player間の相互作用関係をゲームとしてモデル化する際の表現の一つに利得行列がある。例えば、次の利得行列は二人の、各playerが二つの行動選択肢を持つ二人二手ゲームの利得行列による表現である。playerP1は行動選択肢M1,M2を、playerP2は行動選択肢M3,M4を持つ。

ゲーム 1

P2

P1		M3	M4
	M1	2.5	3.0
	M2	2.0	1.0

P1

P2		M1	M2
	M3	1.0	2.0
	M4	3.0	2.5

利得行列中の数字は行側のplayerへの利得である。二つの利得行列をまとめて次のように書くこともある。

ゲーム 1

P2

P1		M3	M4
	M1	2.5 \ 1.0	3.0 \ 3.0
	M2	2.0 \ 2.0	1.0 \ 2.5

ゲームにおいては各playerは独立、同時にとるべき行動を決める。つまり先手後手はない。

また、各 player は自分の利得は知っているが相手の利得は知らない。このように各 player の持つ情報が異なっているゲームを不完備情報ゲームという。

このような状況のなかで、各 player は利己的に自分にとって望ましい結果(なるべく高い利得)のみを求めてゲームに参加する。

2.2 pay

oplayer-i が行動 m_i、i 以外の player が行動組合せ m_{P-i} をとるときの player-i に対する利得を $p(i, m_i, m_{P-i})$ とする。
$pay(i, m_i) = \{p(i, m_i, m_{P-i}) : m_{P-i} \in M_{P-i}\}$ と pay を定義する。pay は player がある行動をとる時にゲームの結果受けとる可能性のある利得の集合である。

ゲームに先手後手はないので、各 player がとる行動を決める時点では相手の player がどの行動をとってくるのか分からない。また、相手の利得も知らないのでその予測も出来ない。よって、$pay(i, m_i)$ は player-i の利得行列(player-i が行側にある利得行列)において m_i の行にある利得全てからなる集合となる。ゲーム1でみてみると、$pay(P1, M1) = \{2.5, 3.0\}$, $pay(P1, M2) = \{2.0, 1.0\}$ となる。player はなるべく高い利得を求めるのだから P1 は M1 をとると決まる。同様に P2 は M4 をとると決まる。
しかし、次のゲーム2ではこのようにはいかない。

ゲーム2

		P2 M3	P2 M4
P1	M1	2.0 \ 1.0	3.0 \ 3.0
P1	M2	2.5 \ 2.5	1.0 \ 2.0

そこで、合理性を定義する。

2.3 合理性と合理的戦略

player がどういう戦略をとればよいのかの一般的選定基準として合理性を定義する。合理性の定義は各 player は自分にとって高い利得のみを求めるということに合致するものでなくてはいけない。ところで、ゲーム理論では player がとる手を戦略と呼ぶ。この章においては行動が戦略である。min は集合の最小要素を示す、として、

・定義

$$\max_{m_i \in M_i} \min pay(i, m_i)$$

となる m_i を player-i の合理的戦略とし、この戦略をとるとき player-i は合理性を満たすとする。この合理性をマクシミン合理性と呼ぶ。

つまり、マクシミン合理性とは可能性のある利得の最低保証水準をなるべく高めようというかなり慎重な戦略選定基準である。これ以外の合理性もいろいろ考えられるがそれは第四章で述べる。

ゲーム2では $pay(P1, M1) = \{2.0, 3.0\}$, $pay(P1, M2) = \{2.5, 1.0\}$。P1 にとっては M1 が合理的行動となる。$pay(P2, M3) = \{1.0, 2.5\}$,$pay(P2, M4) = \{3.0, 2.0\}$。P2 にとっては M4 が合理的行動となる。そして、行動組合せ $(P1-M1, P2-M4)$ がとられ、P1 は利得 3.0、P2 も利得 3.0 を得る。

各 player はこのマクシミン合理性にのっとって戦略を決める。ある player に対する利得で同じものがあると合理的戦略が一意に決まらないことがあるが、そのような時は、先の定義で合理的戦略となる行動の集合を $\{M1_i\}$ として
$\max_{m_i} \min_2 pay(i, m_i) : m_i \in \{M1_i\}$
(ただし、$\min_2$ は集合の第二最小要素を表すとする) となる m_i、それでも一意に決まらなければ $\{M2_i\}$ を考えて $\min_3$ … として合理的戦略を決めていけば良い。しかし、議論を簡単にするため以後、同一 player に対する利得で同じ利得はないとする。

2.4 利得

この合理性によって戦略を決める場合、行動組合せの選好順序、つまり行動組合せの player からみた望ましさの順序、が分かればよいので、利得は一つの利得行列内での相対的大きさに意味があるのであって絶対的大きさには意味がない。よって、各利得行列ごとにその中の利得を小さい方から順に 1, 2, 3,… と変換していくことが出来る。例えば、

	P2	
P1	M3	M4
M1	2.0	3.0
M2	2.5	1.0

$\Longrightarrow$

	P2	
P1	M3	M4
M1	2	4
M2	3	1

と変換出来る。

絶対的大きさに意味のある利得とそれを使った合理的戦略決定法については第四章で述べる。第三章まで、利得行列という時はこの変換後の、相対的大きさしか意味のない利得からなる、利得行列をいうことにする。

2.5 多 players でのゲーム理論

三人以上の player によるゲームの利得行列表現は自分を行側に、自分以外の player をまとめて列側にとった利得行列になる。戦略決定の方法は二人の場合と同じである。例えば、三人ゲームは次のようになる (利得は適当に入れたもので特に意味はない)。

P2,P3

P1	M3,M5	M3,M6	M4,M5	M4,M6
M1	5	6	7	8
M2	4	2	3	1

P3,P1

P2	M5,M1	M5,M2	M6,M1	M6,M2
M3	2	6	4	7
M4	8	5	3	1

P1,P2

P3	M1,M3	M1,M4	M2,M3	M2,M4
M5	1	3	7	4
M6	8	2	6	5

2.6 全体合理性

ゲームの結果得られる行動組合せのシステム全体にとっての望ましさを決めるために全体合理性を定義する。全体合理性に対して先ほどの合理性は個人合理性ともいえる。全体合理性の定義にはいくつか考えられるが、個人合理性がマクシミン合理性であること、行動組合せの選好順序のみを表す利得を定義に使えること、システム全体の協調を問題にしていること、を考えて次のように定義した。

行動組合せ m_P がとられる時の player-i に対する利得を $p(i,m_P)$ として、

定義

$\min_{i\in P}\ p(i,m_P)$ を行動組合せ m_P の全体合理度とする。また、

$\max_{m_P\in M_P}\ \min_{i\in P}\ p(i,m_P)$ となる m_P を最大全体合理性を満たす行動組合せ、またこの式の値を最大全体合理度、とする。

最大全体合理性をみたす行動組合せはシステム全体にとって最も望ましい行動組合せである。最大全体合理度はその望ましさの度合で、ゲームについて一意に定まる値である。2*2 利得行列でこの値が最も低くなるゲームとしては次のものがある。

ゲーム 3

	P2	
P1	M3	M4
M1	2 \ 3	4 \ 1
M2	3 \ 2	1 \ 4

$m_P = (M1, M3) or (M2, M3)$ にたいして最大全体合理度 2 となっている。

逆に高くなるゲームが次のものである.

ゲーム 4

	P2	
P1	M3	M4
M1	2 \ 1	4 \ 4
M2	3 \ 3	1 \ 2

$m_P = (M1, M4)$ にたいして最大全体合理度 4 となっている。

このようにあるゲームの最大全体合理度はそのゲームにおいて引き出し得る最大限の協調を示す一つのバロメーターといえる。最大全体合理度が低いとそのゲームはゲーム 3 のような競合的ゲームであり、高いとゲーム 4 のような協調可能なゲーム (協調とよべる行動組合せを持つゲーム) である。

問題は (個人) 合理的戦略を各 player がとった時にきまる行動組合せが最大全体合理性を満たすものであるかということである。例えば、先ほどのゲーム 4 では行動組合せ (M1,M4) が最大全体合理性を満たす行動組合せとなっており、各 player が個人合理性に基づいて戦略決定した結果の行動組合せ (M1,M4) を含んでいる (「含んでいる」と表現したのは最大全体合理性を満たす行動組合せが複数の場合もあるからである)。しかし、次のような利得行列ではそうならない。

ゲーム 5

		P2	
		M3	M4
P1	M1	$3 \setminus 3$	$1 \setminus 4$
	M2	$4 \setminus 1$	$2 \setminus 2$

このゲームは囚人のジレンマ型のゲームと呼ばれている。最大全体合理性を満たす行動組合せは $(M1, M3)$、最大全体合理度は 3 だが、各 player が個人合理性にのっとって戦略決定した結果の行動組合せは $(M2, M4)$、全体合理度は 2 となり食い違う。このように個人合理性と全体合理性が合致するゲームとしないゲームがある。

2.7 実験結果

ゲームは 二人二手ゲームでは $4!*4!$ とうり, 三人二手ゲームでは $8!*8!*8!$ とうり... ある。それぞれの規模のゲームにおいて、ランダムに利得行列を選んでいって、最大全体合理度と、各 player が個人合理性に基づいて戦略決定した結果の行動組合せの全体合理度 (以後これを結果全体合理度と呼ぶ)、その両者が合致するゲームの割合を調べた。

両者が合致するゲームの割合をまとめたのが次の表である。二手、三手... は一人の player のもつ行動選択肢の数である。

表 1

	二手	三手	四手	五手
二人	0.58	0.23	0.12	0.08
三人	0.22	0.10	0.08	-
四人	0.09	0.02	-	-
五人	0.04	-	-	-

ゲームの規模が大きくなるにつれて割合が小さくなる、いわば個人合理性と全体合理性が離れていくのが分かる。

三人三手ゲームにおける最大全体合理度とゲームの結果全体合理度の二次元分布が図 1 である。

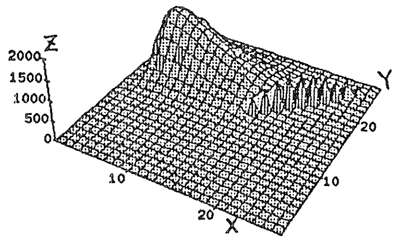

図 1。X 軸が結果全体合理度、Y 軸が最大全体合理度、Z 軸が割合 $\times 10^5$。

この図で $Y = X$ 上にある部分の和が両全体合理度が合致するゲームの割合になる。図から、両全体合理度が合致するものとしないものにゲームが二分されることが分かる。このことは三人三手ゲームに限らず全ての規模のゲームにおいて確かめられた。

3 通信有の行動決定法

前章の方法は通信を考えた方法ではなかった。もちろん通信しないで個人と全体の合理性が合致すればそれにこしたことはないが、実験結果からは前章の方法が個人と全体の合理性が合致する、各 player が個人合理性に基づいて戦略を決定した結果として全体合理性 (協調) が達成されるの狙う、ということを十分達成するものではない。
そこで通信の概念を導入し、通信を組み入れた行動決定方法を考えることにする。通信内容としていろいろなものが考えられるが、本論文ではまず offer という通信を含んだ行動決定方法を述べ、これが先ほどの狙いを満たす方法であることを示す。次に利得行列を通信する行動決定方法について述べる。

3.1 offer を組み入れた戦略と行動決定 [1]

前章では戦略は行動であったが、この章では戦略は行動と次に述べる offer group を組合せたものとなる。また、前章では戦略として決定された行動がそのまま最終的にとられる行動となったが、この章ではそうはならない。行動決定の局面が戦略決定と最終的な行動決

定の二つに分かれる。図2にその流れが図示してある。

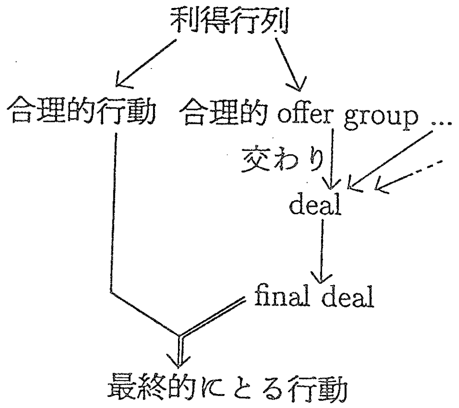

図2。 offer を組み入れた行動決定

player はある行動を戦略の一部として仮決定する。

また、各 player は「皆さん、こういった行動組合せのうちのどれかを（私を含めた）皆さんにとって欲しい」という、戦略の一部である offer group を他の player に offer する。offer group は行動列の集合である。 player-i の offer group を o_i と表記する。 $o_i = \{m_P\}$。ゲームにおいて各 player が offer group を offer することが通信となる。

各 player が offer group を通信することによって、全 player の offer group の交わりが導出される。

$$\{m_P : m_P \in \bigcap_{i \in P} o_i\}$$

これを deal と呼ぶ。
deal が複数の行動組合せからなる場合には、その内で全体合理度が最大となるもののみを残して final deal に絞る。この絞り込みを経てもまだなお final deal が複数の行動組合せからなる場合にはそのなかからランダムに一つを選んで final deal とする。

player は deal が成立した時は final deal に従った行動をとり、成立しなかった時は仮決定してあった行動をとる。これが最終的に選ばれる行動となる。

3.2 合理性と合理的戦略

前章で、戦略決定の基準としてマクシミン合理性を定義したが、ここでも行動と offer group を組み合わせた戦略の決定基準としてマクシミン合理性を定義する。それとともに合理的戦略の決定手続きを示す。

player-i が行動 m_i を仮決定し、 offer group o_i を offer するという戦略を立てた時の i に対する可能性のある利得の集合を $pay(i, m_i, o_i)$ とする。

定義 1)

$$\max_{m_i, o_i} \min pay(i, m_i, o_i)$$

を満たす m_i, o_i を agent-i の合理的戦略とし、この戦略をとるとき player-i は合理性を満たすとする。

重要なことは offer という通信は決してシステム全体のことを考えた譲歩ではなく、あくまでも自分の利得を高めるという自己目的、個人合理性のために行なうものである、ということである。

player-i の合理的戦略の一方を成す offer group (以後合理的 offer group と呼ぶ) について先の定義に次の項目を加える。

2) ある $m_P \in o_i$ について、次のような行動組合せ m'_P が存在するなら o_i は合理的 offer group ではない。
$m'_P \notin o_i$ かつ $p(i, m_P) < p(i, m'_P)$
この項目がないと、各 player が M 個の行動選択肢を持つ L 人ゲームにおいて一人の player が offer する offer group の候補は 2 の M^L 乗とうりになるが、この項目があると M^L とうりに候補は減る。例えば、ゲーム 2 において P1 の 合理的 offer group の候補は
$\{(M1, M4)\}$
$\{(M1, M4), (M2, M3)\}$
$\{(M1, M4), (M2, M3), (M1, M3)\}$
$\{(M1, M4), (M2, M3), (M1, M3), (M2, M4)\}$
の四つになる。
$\{(M1, M4), (M1, M3)\}$ は合理的 offer group とはならない。行動組合せ (M1,M3) を P1 が offer group に含めるのであればそれよりも高い利得をもたらす行動組合せ (M2,M3) も offer group に含めるとするのである。

player-i の仮決定する行動を m_i としたとき、定義 1) により合理的 offer group は $\max_{o_i} \min pay(i, m_i, o_i)$ を与える o_i となる。ゲーム 2 で見てみると、P1 の仮決定する行動を M1 とした時、四つの offer group に対する pay は次のようになる。

$pay(P1, M1, \{(M1, M4)\}) = \{2,4\}$

$pay(P1, M1, \{(M1, M4), (M2, M3)\} = \{2,3,4\}$

$pay(P1, M1, \{(M1, M4), (M2, M3), (M1, M3)\} = \{2,3,4\}$

$pay(P1, M1, \{(M1, M4), (M2, M3), (M1, M3), (M2, M4)\} = \{1,2,3,4\}$。四番目の offer group が合理的 offer group でないことが分かる。

ところで、offer group について次の定理が成り立つ。

定理

o_i, o'_i に対する pay を pay, pay' として

$\min pay < \min pay'$ かつ $\max pay \leq \max pay'$

または

$\min pay \leq \min pay'$ かつ $\max pay < \max pay'$

が成り立つ時、o'_i は o_i を優越するという。

player-i が行動 m_i を仮決定した時、$\min pay(i, m_i) = p(i, m_P)$ となる m_P を含めた、1) と 2) を満たす offer group o_i は、他の 1) と 2) を満たす offer group に対して優越することはないが優越されることはある。

定理を直観的にいうと、offer しなくても最低限保証される利得しか与えない行動組合せをわざわざ offer する必要はないとなる。

略証

$\min pay(i, m_i) = p(i, m_P)$ となる m_P を含み、かつ 1) と 2) を満たす offer group を o_i、そこから m_P を抜いた offer group を o'_i とする。先ほどの四つの offer group のうち三番目と二番目に相当する。$pay(i, m_i, o_i) = pay(i, m_i, o'_i)$ であるが、次の三つの場合に分けて考える。

1.m_P が final deal に含まれないようなゲームであった場合、m_P を offer group に含めることはゲームの結果に何も影響を及ぼさない。

2.m_P が単独で final deal となるゲームであった場合、$pay(i, m_i, o_i) = p(i, m_P)$, $pay(i, m_i, o'_i) = pay(i, m_i)$ となり、$\min pay(i, m_i) = p(i, m_P)$ であることより o'_i が o_i を優越する。

3.m_P が他の行動組合せと共に final deal となるゲームであった場合 $\min pay(i, m_i, o_i) < \min pay(i, m_i, o'_i)$ かつ、$\max pay(i, m_i, o_i) = \max pay(i, m_i, o'_i)$ となり o'_i が o_i を優越する。証終。

定理の中の o_i は「各 player はなるべく高い利得を求める」ことに反するものであるので 3) そのような offer group は合理的 offer group ではない、という項目を定義に加える。

先ほどの例で (M1,M3) という行動組合せは、$\min pay(P1, M1)$ と同じ利得 2 を P1 に与えるから、三番目の offer group は合理的 offer group でなくなる。

1),2),3) を満たす offer group が合理的 offer group である。一般に合理的 offer group は複数ある。ゲームにおいては一つの offer group を選んで offer するのだが、この選択基準は「合理的」とは関係はない。ゲームの人数、行動選択肢の数を考慮して適当な基準を設ける。さしあたりの基準として要素数最大の合理的 offer group、先ほどの例では二番目の offer group を offer するという基準にする。

player-i の合理的戦略の一部をなす行動 (以後合理的行動と呼ぶ) のほうは、$\max_{m_i} \min pay(i, m_i, \phi)$ を与える m_i となる. ϕ は offer group が空集合である、つまり通信しないということを表す. この M_i は前章における合理的戦略と同一となる。

以上合理的戦略決定の手続きを示したが、この手続きによる戦略が定義 1) を満たすものであることは容易に確かめられる。例えば、合理的行動を M2 と P1 がしてしまうと $pay(P1, M2, ...)$ に利得 1 が入ってきて、マクシミン合理的にならない。

前章の個人合理性と全体合理性が一致しない例である囚人のジレンマ型ゲーム (ゲーム 5) について考えてみる。仮決定する合理的は P1 は M2、P2 は M4、offer する合理的 offer group は P1 は $\{(M2, M3), (M1, M3)\}$、P2 は $\{(M1, M4), (M1, M3)\}$ となる。この時、deal $\{(M1, M3)\}$ が成立し P1-M1,P2-M3 という行動組合せが最終的にとられる。この行動組合せは最大全体合理性を与える行動組合せであるので、offer という通信を含む行動

決定法が囚人のジレンマ型ゲームにおいて個人合理性と全体合理性を合致させているといえる。

3.3 実験結果

まず，offer しないのとするのとでは player が結果として得られる利得の分布がどう変わるのかを行動選択肢三の二人ゲームにおいて調べた結果を見る。

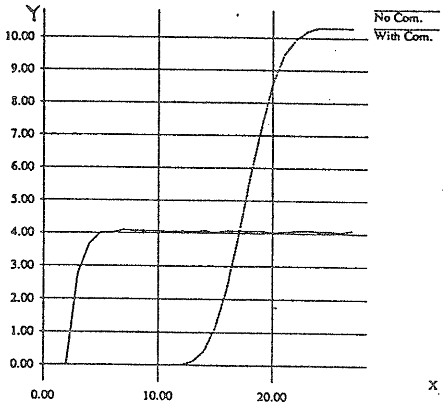

図3。X 軸がゲームの結果得られる利得、Y 軸が割合 $\times 10^2$。No Com. は前章の方法、With Com. はこの章の方法。

With Com. の方が利得が大きい方に偏っている。「offer は全体合理性のために譲歩をしているのではない」ことが実験結果からも確かめられる。

次に，前章と同様にして個人合理性と全体合理性の合致するものの割合を調べた。

表2

	二手	三手	四手	五手
二人	0.95	0.97	0.99	0.99
三人	0.97	1.00	1.00	-
四人	0.99	1.00	-	-
五人	1.00	-	-	-

offer をすればほとんどの場合特に規模の大きなゲームにおいて最大全体合理度と結果全体合理度が一致している。また、三人三手ゲームにおける両全体合理度の二次元分布が図4である。

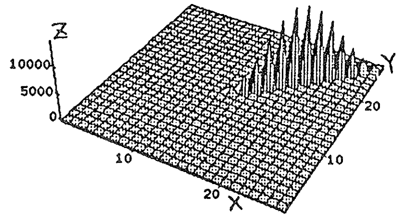

図4。X 軸がゲームの結果の行動組合せの全体合理度、Y 軸が最大全体合理度、Z 軸が割合 $\times 10^5$。

図1に比べて X=Y 上にきれいに並んでいるのが分かる。

まとめると offer という通信を含む戦略はあくまで個人合理性に基づくものであるが、結果としてとられる行動組合せは全体合理性を満たしている、結果として協調となることを狙うということを達成している、といえる。

3.4 集中行動決定

offer するという通信ではなく、各 player が利己的な分散行動決定を完全放棄して、自分の利得行列をそのまま送るという通信による集中行動決定方法も考えられる。全ての player の利得行列を一箇所に集めればゲーム全体の情報がそこに集まることになり全体合理性を満たす行動組合せが確実に決定される。

が、この方法は通信量が多くなる可能性がある。offer group はその交わりつまり deal と、final deal を作るための deal の各要素 (行動組合せ) の全体合理度、が分かれば良いのだから、offer を含む行動決定法では次のような通信を行なえば良い。player の部分集合 Q を経てきた通信を player-i が受けとる時、その通信の中身は Q に属する player の offer group の交わりとその各要素 m_P に対する $\min_{j\in Q} p(j,m_P)$ の値である。player-i は自分の offer group との交わりをとり、その交わりの各 m_P に対する $\min_{j\in Q} p(j,m_P)$ と $p(i,m_P)$ の小さい方の値、を付けて次の player に送る。こうすると (総通信量ではなく) 一回一回の通信の量は通信を経るたびに減っていく。

一方、利得行列の通信はそれとは逆になる。

結局、どちらの通信による行動決定法がよいかは通信網やゲームの規模による。

3.5 offer group について

「さしあたりの基準として要素数最大の合理的 offer group」を offer する、と第二節で述べた。一般にゲームの規模が大きくなると各 player の要素数最大の合理的 offer group も大きくなる。例えば三人三手ゲームにおいては平均で 22.01 の大きさになる。

第二節で述べた合理的 offer group の定義は合理的と思われる offer group の必要条件をのべたもので、十分条件ではない。そこで、合理的 offer group の内で、最大のものの何割の大きさを持つものという基準で合理的 offer group のうちから実際に offer するものを選び、ゲームを行なうとして三人三手ゲームにおいて実験を行なった結果が図 5 である。

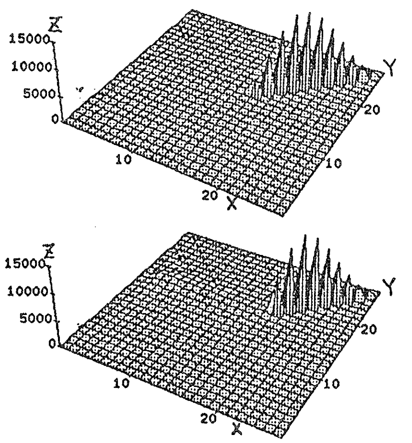

図 5。X 軸が結果全体合理度、Y 軸が最大全体合理度、Z 軸が割合 $\times 10^5$。大きさ最大の合理的 offer group に対して、上の図が 5 割、下の図が 4 割、の大きさの合理的 offer group を offer 。

分かりにくいが、下の図で、X が 3 から 10、Y が 16 から 18、のあたりが少し盛り上がっている。一方上の図は図 4 とほとんど同じである。通信量をけちっていって小さな合理的 offer group を offer するようにしていくと、ある所までは結果に変わりがないが、それを越えるとゲーム 3 のような最大全体合理度の低い競合型ゲームからうまくいかなくなっていくということがわかる。これと同じ傾向が他の規模のゲームにおいても確認された。競合型ゲームが高い確率で現れると予想される時にはあまり通信量をけちらない方が良いと言える。

4 その他の合理性

第二章でマクシミン合理性を定義したが、行動選定基準としての合理性は他にも考えられる。なお、この章では通信は考えない。

4.1 効用による合理性

今までは利得は選好順序を表し、その値の絶対的大きさには意味が無かった。しかし，たとえば一万円、十万円など絶対的大きさが player にとっての望ましさの度合を表すような利得が求まることも考えられる。そのような利得を効用という。

マクシミンという合理性は利得の最低保証水準をなるべく高くするというかなり慎重な考え方であった。この合理性以外に利得が効用として与えられるならば，効用の平均値を最大にする合理性などが考えられる。

この合理性を一般的に定義する。戦略 m_i をとった時の得られる可能性のある効用の平均値を $apay(i,m_i)$、標準偏差を $spay(i,m_i)$、として

<u>定義</u>

$$\max_{m_i}(apay(i,m_i) - \gamma * spay(i,m_i))$$

となる m_i が合理的戦略である。γ はリスク回避度である。

例えば、$\gamma = 0$ として次のゲームを考える。

		P2	
		M3	M4
P1	M1	5.0 \ 5.0	0.0 \ 1.0
	M2	1.0 \ 1.0	1.5 \ 1.5

このゲームでは行動組合せ (M1,M3) が決定される．第二章のマクシミン合理性に基づけば (M2,M4) が決定される。

$\gamma = 0$ として，この合理性とマクシミン、それぞれの合理性に基づいてゲームを行なった時に player の得られる利得の分布が図 6 である。ただし、ゲームは三人二手ゲームである。

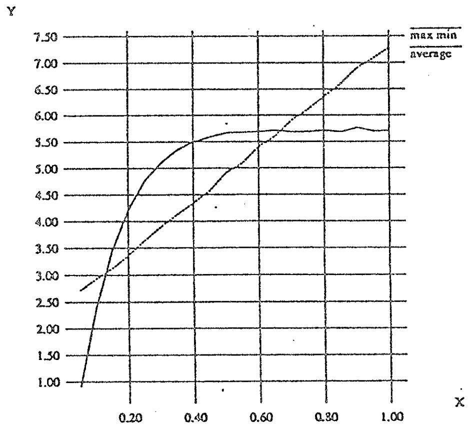

図 6。X 軸がゲームの結果得られる利得、Y 軸が割合 $\times 10^2$。

二箇所で線が交錯しているのでどちらの合理性が良い個人合理性かということはいえない。また、第二章での全体合理性の定義に基づいて全体合理度について調べても、結果全体合理度が低いものの割合がマクシミン合理性による方法よりも多くなったが、高いものの割合も多くなった。よって、全体合理性に関してもどちらが良いということはいえない。どちらの合理性をとるかはシステム設計者の考え方による。

5 終りに

分散行動決定方法とその有効性について述べて，個人と全体の橋渡しをする有効な戦略決定手続きと行動決定方法を提示した。

また、競合的ゲームにおいては通信量がより多くいることも述べた。

第二章で示した最大全体合理性、及びゲームの結果の行動組合せの全体合理性という概念は player 間、つまりシステム構成要素間の相互作用関係の一つの指標になる。図 1 の XY 平面のどこに来るのかによってそのゲームが競合的であるとか、高い協調性を内在しているとか、通信なしの分散行動決定でどこまで出来るゲームであるのか、といったことが分かる。

本論文では各 player の行動選択肢の数は有限であったが、これを無限にしたものは無限ゲームといわれる。無限ゲームでは行動は通常いくつかの連続量で表さる。無限ゲームについても本論文の行動決定方法は使える。が、第二章の表 1 と第三章の表 2 からわかることであるが無限ゲームにおいては通信が必ず必要になる。

また、不完備情報ということについてであるが、本論文では利得について不完備情報であるとした。利得の完備情報性が高い、つまり player が自分以外の player の利得についてもある程度知っている、という状況にあるとすると、人工システムにおいては各 player は戦略決定を完全に利己的には行なわず、分かる範囲で全体合理性のことを考えて戦略決定を行なうようにすれば良い。player が完璧に利己的である場合には player がお互いの利得を知り合うことでかえって、個人とシステム全体両方に悪い結果をもたらす場合もあるが、人工システムではそこまで強い利己性をシステムの各構成要素にもたせる理由はないのでこのような危惧はない。

最後に、今回は戦略決定は独立，同時つまり同期的に行なわれるとしたが，一般的にはそれはゲームのなかで非同期的に行なわれる。その場合の行動決定法が今後の課題であると思っている。

参考文献

[1] J.S.Rosenschein, et al: “Deals Among Rational Agents,”Proceedings of the 1985 International Joint Conference on Artificial Intelligence, pp.91-99, 1985.

[2] M.L.Ginsberg: “Decision Procedures,” M.N.Huhns,editer, Distributed Artificial Intelligence, pp.3-28, 1987.

[3] 中村 他: “分散 AI における協調へのゲーム理論的アプローチ,” 第四十二回情報処理学会全国大会, 1991.