7F−6

分散AIにおける協調への ゲーム理論的アプローチ

中村豪一　出口光一郎
東京大学 工学部 計数工学科

1、はじめに

お互いに interaction する自律的 self-interested なシステム構成要素（agent）からなる分散 AI などの分散型システムにおいて、agent 間の協調 (total-benefit) が達成されるかということは大きな問題である。これを調べるために、システムの動きをゲームとしてモデル化してその解析をゲーム理論を用いて行うことが考えられる。

Rosenshein ら [1],[2] はゲーム理論に通信の概念を導入し、それの有無それぞれの場合について特殊なゲームにおける合理性を定義している。本論文ではこれを基にして，一般的なゲームにおける合理性と合理的戦略決定手続きを提示し、それの 2x2 利得行列における協調達成への影響を解析し、協調の達成が困難ではないことを示す。

2、利得行列

o agent 間の interaction を単純にモデル化した利得行列がゲーム理論でよく使われる。

表1

		囚人B	
		自白するc	自白しないd
囚人A	自白するc	3\3	0\5
	自白しないd	5\0	1\1

これは囚人のジレンマの利得行列である。例えば、row-c、column-d という行動の組合せを c \ d と表記し、行動列と呼ぶ. 当然, 各 agent（以後 player とも呼ぶ）は相手への利得は知らない。

3、PLAYERの振舞い

Rosenshein らによる player の振舞いを簡潔にまとめて述べる。

ゲームに際して player がまずやることは行動を仮決定することと、それとは全く独立に次に述べる offer group を offer することである。この二つが player の戦略である。

戦略決定を全ての agent が同時かつ独立に実行し、相手の戦略を待ってそれに対応することはしないとする。これを合理的仮定と呼ぶ。

各 agent は「皆さん、こういった行動列のうちのどれかを（私を含めた）皆さんにとって欲しい」という offer group を他の agent に offer できる。offer group は行動列の集合である。ゲームにおいては各 player が offer group を offer しあう。この offer しあうことが通信である。通信無とは agent が offer group を offer しないということになる。

ある agent の offer group の内で、他の全ての agents も offer しているような行動列がある時その集合を deal と呼ぶ。

つまり deal とは全ての agents の offer group の交わりである。

agent は deal が成立した時はそれ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。

4、合理的戦略決定手続き

一般的な合理的戦略決定手続きを提示する。agent-i が行動 m_i を仮決定し、offer group o_i を offer する時の 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。

agent-i の合理的行動 M_i を、$minpay(i, m_i, \phi)$ を最大にする m_i とする。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 は利得行列からすぐに求まる。

agent-i の合理的 offer groupO_i は次の二つの制限を満たす。一つは、O_i は $minpay(i, M_i, o_i)$ を最大にする o_i であることである。重要なことは、offer は協調のためではなくあくまでも自分の可能性のある利得の範囲を高めるという自己目的のためにやるということである。もう一つの制限は、ある行動列が合理的 offer group に属しているのならそれより高い利得を与える行動列も属していなくてはならないという制限である。この二つの条件から、利得の低い行動列から順に、それを offer group に属させないことで利得の最小値が高まるか同じうちはそうしていく、という手続きで合理的 offer group が求まることが分かる。

このようにして求めた合理的行動と合理的 offer group が合理的戦略である。

5、協調の定義

協調を次のように定義する。ある定められた整数 n があり、player-i に与える利得の大きいほうから順に n 番目までの行動列の集合を $C_i(n)$ として協調となる行動列 $m_c(n)$ は $m_c(n) \in C_i(n)$ for all i。つまり、ある行動列の両 player に対する利得がそれぞれ高い方から n 番目までの利得に入っている場合、その行動列は協調となる行動列である。

6、囚人のジレンマ

4章の手続きをとった時、通信の有無によって結果のことなる例として表1の囚人のジレンマを考える。通信無のとき、囚人 A,B 共に利得行列を見て自白しない行動 d をとり、結果ジレンマに陥る。通信有のとき、合理的 offer group の決め方に従って囚人 A は $\{d\backslash c, c\backslash c\}$ を、囚人 B は $\{c\backslash d, c\backslash c\}$ を offer し、deal$\{c\backslash c\}$ が成立して、結果両者自白し協調が達成される。

7、2X2利得行列における協調達成への影響の解析

4章の戦略決定手続き及び5章の協調定義を基に解析を行なう。

協調定義は利得の大小のみによっているので利得行列中の片方に対する利得の大きい方から順に 4,3,2,1 とコード化してもよい。さらに行、列それぞれの置換を行なって row の 4 が左下に来るようにする。

Game Thoric Approach to Cooperations in DAI
Gouichi NAKAMURA, Kouichirou DEGUCHI
Univercity of Tokyo

こうすると 2x2 利得行列は全部で 144 とうりにまで落せる。また、rowplayer への利得に注目すると次の五つのタイプ A,B,C,D,E に分類できる。

A	c	d
a	2\x	1\x
b-	4\o	3\x

B	c	d
a	3\o	1\x
b-	4\o	2\x

D	c	d
a-	2\x	3\o
b	4\o	1\x

A	c	d
a	1\x	2\x
b-	4\o	3\x

C	c	d
a	1\x	3\o
b-	4\o	2\x

E	c	d
a-	3\o	2\x
b	4\o	1\x

ただし、-、o、は 4 章の戦略決定手続きをとった時のそれぞれ、合理的行動、合理的 offer group の要素である行動列、を表す 。column 側についても同様の分類が出来るのでタイプどうしの組合せは 15 になる。この組合せを基にしてて $n = 2$ とした時の協調達成の程度を調べた結果が表 2 である。

組合せ	総数	異	有数	有和	無数	無和	不能
A-A	16	0	16	7	16	7	0
A-B,B-A	16	4	8	6.5	8	6.5	8
A-C,C-A	16	4	12	7	8	6.5	0
A-D,D-A	16	4	12	6.75	8	6	0
A-E,E-A	16	4	8	6.5	8	6	8
B-B	4	3	4	7	1	6	0
B-C,C-B	8	6	8	7	2	6	0
B-D,D-B	8	6	8	7	2	5.5	0
B-E,E-B	8	6	8	7	2	5.5	0
C-C	4	1	2	6.75	1	6	2
C-D,D-C	8	2	4	6.5	2	5.5	4
C-E,E-C	8	6	8	7	2	5.5	0
D-D	4	2	2	6.25	1	5	2
D-E,E-D	8	6	8	7	2	5	0
E-E	4	3	4	7	1	5	0
計	144	57	112	—	64	—	24

表 2 − 左から順に組合せ、その組合せでの総数、通信有無で結果（最終的に選ばれる行動列）の異なる数、通信有で協調達成された数、有時での結果の行動列の両 player に対する利得の和の平均値、通信無で協調達成された数、無時での利得の和の平均値、協調となる行動列を持たない協調不能なものの数 、協調不能なものの内にはゼロサムゲームが 6 含まれる。

通信無において協調が達成されるのに通信有にすると達成されない通信逆効果となるものは一つも無かった。ランダムに利得行列が与えられた時協調が達成される確率は通信有で約 0.78、無で約 0.44、合理性を全く考えずに戦略決定した場合には 0.25 となる。

$n = 1$ の場合についても同様に調べると、協調となる行動列を持つものが 36 とうり、そのうち通信有無で結果が異なるのは 20 とうり、通信有において協調が達成されるのは 36 とうり、それが通信無においては 16 とうりになる。また、通信逆効果となるものは一つも無い。通信無から有にした時に結果の変わる全てが協調達成へと変わる。

（n の値に関係のない）結果の行動列の両 player に対する利得の和の平均値より求めた、ランダムに利得行列が与えられた場合の利得の和の期待値は、通信有の場合には約 6.805、無では 6 であった。合理性を考えずに戦略決定をした場合はこの値は 5 である。利得行列内の利得の和の最高値の平均は 6.875 である。つまり、利得の和の期待値は戦略決定手続きにくにより 5 から 6.875 の値をとる。

8、まとめ

一般的なゲームにおける通信を含んだ合理的戦略決定手続きを提示した。

また、2x2 の利得行列においては次のことが言える。

利得の和の期待値を 6.875 にするにはどちらかが自分に対する利得を全て相手に送る、あるいは両者が第三者に送って調停してもらうということをしなければ不可能であり、かなりの通信量となってしまう。またそのようなことをする agent はもはや自律 (self-interested)agent とは言い難いし、利得の最低値の保証もない。

4 章の合理的戦略決定手続きは、offer group の offer という程度の通信を含んだ、あくまでも agent への利得の可能性を高めるための、それほど複雑ではない手続きであるが、agent がそれに従って self-interest に動く結果、（ランダムに利得行列が与えられた時に） 5 から 6.875 の間の 6.805 という利得の和の期待値、確率約 0.78 でもって協調が達成される。また、通信をしない、あるいは出来ない状況においてもある程度の利得の和の期待値、協調達成確率となる。

逆に次のように言うことも出来る。それほど複雑なものではないこのような手続きでもってこれほどの協調が達成できる、つまり協調となる利得行列が全体に対して大部分を占めるということは、協調が達成されない利得行列を持った利得関係においても、ちょっと利得の値をいじる、つまりちょっと利得関係を変えるだけで協調が達成されるようになる、あるいは通信有でしか協調が達成出来ないような利得関係においても、ある程度利得関係を変えれば通信無で協調が達成されるようになる。このことは分散 AI などの分散システムの設計には重要なことである。

また、三人以上の player によるゲームでの合理的戦略決定は基本的には二人の場合と同じである。ある一人の player を row に、他の player をまとめて column にとった利得行列、をその row に持たせれば、row はそれを使って合理的戦略を決定できる（もちろん column への利得は分からないが row が戦略を決めるのにそれはいらない）。これを全ての player に対して行なえば良い。

終りに。合理的戦略決定手続きは相手方への合理的仮定ともからんで 4 章で述べた以上に複雑なものであるが今回は紙面の都合で割愛した。この論文で述べたゲーム理論を実際の分散システムにうまく応用していくこと、communication の内容を変えることや、相手のモデルを自分の内に作ることなどによる communication 量と computation 量の変化関係、などを考慮に入れた総合的な分散 AI へのゲーム理論的アプローチ、が今後の課題である。

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, *Distriduted Artificial Intelligence*, pp.3-28 (1987).