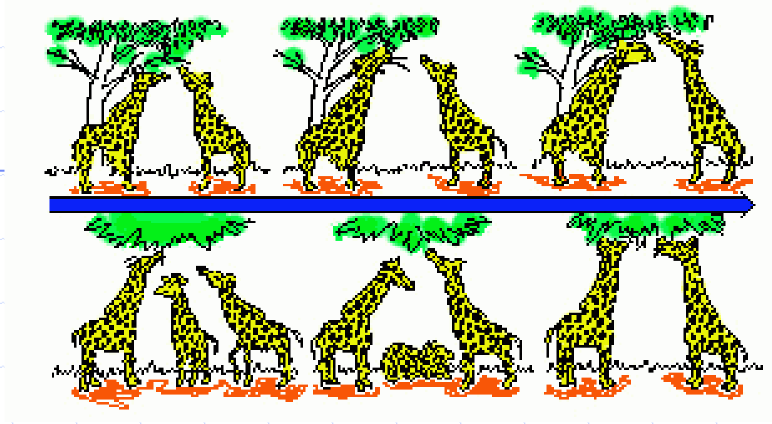
囚人のジレンマ

東京大学大学院
情報理工学系研究科
電子情報学専攻
伊庭斉志



iba lab.

囚人のジレンマの損得

(a) A から見た損得表

	B_1 B が自白しない (C)	B_2 B が自白する (D)
A_1 A が自白しない (C)	-2 懲役 2 年	-5 懲役 5 年
A_2 A が自白する (D)	-1 懲役 1 年	-3 懲役 3 年

(b) B から見た損得表

	B_1 B が自白しない (C)	B_2 B が自白する (D)
A_1 A が自白しない (C)	-2 懲役 2 年	-1 懲役 1 年
A_2 A が自白する (D)	-5 懲役 5 年	-3 懲役 3 年

均衡点

- Def: 相手が均衡点の戦略をとるとき、自分もそれをとらないと必ず損をする

- John　Nash (ナッシュ)
  - 1994年 ノーベル経済学賞
  - A beautiful mind (2001)

Nash均衡点の例：
平均値推定ゲーム

- 多数のプレイヤーがそれぞれ1から100までの実数を1つ選ぶ。選ばれた数字の平均値を0.7倍した数字に最も近い数字を選んだ人がゲームの勝者である。
- どのような数字を答えるのが最適か？

Nash均衡点の例：
平均値推定ゲーム

- 多数のプレイヤーがそれぞれ１から１００までの実数を１つ選ぶ。選ばれた数字の平均値を0.7倍した数字に最も近い数字を選んだ人がゲームの勝者である。
  - 人によって好みの数字はあるかもしれないが、全員の選んだ数値の平均値はおおよそ50だろう。
  - なので50×0.7＝35が、最も近い数値である。
  - But　全員が同じように推論した場合、その0.7倍である35×0.7≒25が候補になる。
  - しかしここでも全員が同じなら、さらに0.7倍の17が候補に。次は12が候補となる。
  - この思考のプロセスを続けていくと、最終的には全員が0を選ぶ。
  - これが**ナッシュ均衡**である。

Nash均衡点

$$\begin{array}{c|cc} & A & B \\ \hline A & a & b \\ B & c & d \end{array}$$

- 相手が均衡点の戦略をとるとき、自分もそれをとらないと必ず損をする

- a>c ならばAは狭義のNash均衡点
- a≧c ならばAはNash均衡点
- d>b ならばBは狭義のNash均衡点
- d≧b ならばBはNash均衡点

例：囚人のジレンマ

$$
\begin{array}{cc}
 & \begin{array}{cc} C & D \end{array} \\
\begin{array}{c} C \\ D \end{array} & \begin{pmatrix} 3 & 0 \\ 5 & 1 \end{pmatrix}
\end{array}
$$

(a) A から見た損得表

	B_1 Bが自白しない (C)	B_2 Bが自白する (D)
A_1 Aが自白しない (C)	-2 懲役 2 年	-5 懲役 5 年
A_2 Aが自白する (D)	-1 懲役 1 年	-3 懲役 3 年

大きな数ほど良い（懲役が短い）ように変換する

例：囚人のジレンマ

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} 3 & 0 \\ 5 & 1 \end{pmatrix} \end{array}$$

- 両プレイヤがを*C*選べば、一方のプレイヤが*D*に変更することで利得を増加できる。
- 両プレイヤが*D*を選べば、どちらのプレイヤも*C*に変更することで利得を改善できない。
- ∴ *D*はNash均衡点
- *D*は*C*よりも優位

囚人のジレンマ

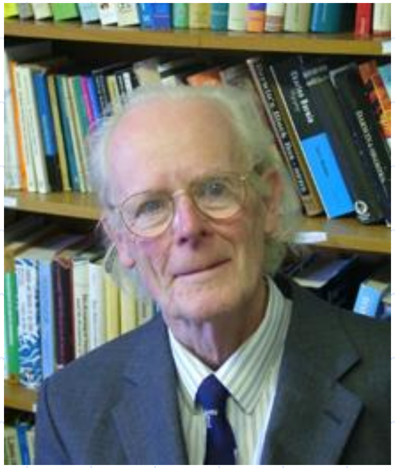
(a) A から見た損得表

	B_1 Bが自白しない(C)	B_2 Bが自白する(D)
A_1 Aが自白しない(C)	−2 懲役2年	−5 懲役5年
A_2 Aが自白する(D)	−1 懲役1年	−3 懲役3年

- (D,D):均衡点
- But 相手が自白しないと信じて自白しなければ、(C,C) ： どちらにとっても均衡点よりよい！！
- しかしお互い隔離されて協力できない！
- 囚人のジレンマ:仲間が自白するなら自白すべきだし、自白しないなら自白しない。→ ジレンマ!!
- 繰り返すとどうなるか？
  - 繰り返し囚人のジレンマ
  - Iterated Prisoners' Dilemma (IPD)

進化的安定な戦略

- ESS：Evolutionary Stable State
- John Maynard Smith
  - 1920-2004
  - British theoretical evolutionary biologist and geneticist.



進化的安定な戦略 (ESS)

$$\begin{array}{cc} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{array}$$

- 戦略Aをプレイする大きな集団を考える
- この集団に戦略Bをプレイする個体が突然変異で生じたとする
- このときAの集団がBの侵入を許さないような条件はなにか？

$f_A = ax_A + bx_B$: Aに対する期待利得

$f_B = cx_A + dx_B$: Bに対する期待利得

x_A : Aの頻度（Aと対戦する確率）

x_B : Bの頻度（Bと対戦する確率）

進化的安定な戦略 (ESS)

$$\begin{array}{c|cc} & A & B \\ \hline A & a & b \\ B & c & d \end{array}$$

- Aの集団　　　　　　大きい
- 侵入者Bの頻度 : ε　　　非常に小さい
- Aの期待利得がBの期待利得を上回るための十分条件:

  $a(1-\varepsilon)+b\varepsilon>c(1-\varepsilon)+d\varepsilon$

- ∴εは小さいのでa>cが得られる
- a=cのときにはb>dとなる

進化的安定な戦略 (ESS)

$$\begin{array}{c} \\ A \\ B \end{array}\begin{array}{c} \begin{array}{cc} A & B \end{array} \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{array}$$

- a>c または (a=cかつb>d)ならば戦略AはESSであり、淘汰によってBはAの集団に侵入できない。

- ESSは集団が非常に大きく、侵入者が少数のときのみに成り立つ考え方である

２つ以上の戦略のとき

$E(S_i, S_j)$：戦略S_jに対する戦略S_iの利得

ESS戦略をとる集団に対しては、淘汰によってどのような戦略も侵入できない。同様な性質は狭義のNash均衡に対しても成り立つが、Nash均衡に対しては成り立たない。

または$E(S_k, S_k) = E(S_i, S_k)$かつ$E(S_k, S_i) > E(S_i, S_i)$

ESS

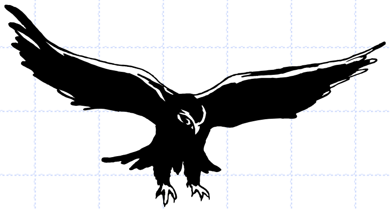
無敵の戦略

$$\forall i \neq k : E(S_k, S_k) > E(S_i, S_k) \text{かつ} E(S_k, S_i) \geq E(S_i, S_i)$$

無敵

- 他のすべての戦略と比べて優位
- 狭義のNash均衡点
- 望みうる最強の戦略
- but高望み、得られるのは稀

例：タカ・ハトゲーム



Hawk

vs

Dove

- 儀式的闘争
- タカ(H):争いを拡大する
- ハト(D):相手が争いを拡大した時には退却する
- 争いに勝って得る利得はb
- けがをして被る損失はc

Cf.「攻撃：悪の自然誌」K.ローレンツ

$$
\begin{array}{cc}
 & \begin{array}{cc} H \quad & \quad D \end{array} \\
\begin{array}{c} H \\ \\ D \end{array} & \begin{pmatrix} \dfrac{b-c}{2} & b \\ 0 & \dfrac{b}{2} \end{pmatrix}
\end{array}
$$

「攻撃：悪の自然誌」

- コンラート・ツァハリアス・ローレンツ（1903年-1989年）　1973年　ノーベル医学・生理学賞
- オーストリアの動物行動学者
- 刷り込みの発見者
- 近代動物行動学を確立
- 「儀式化された闘争」：狼、鹿。。。
- 平和の象徴の鳩？
- 群淘汰？？？

例:タカ・ハトゲーム

$$\begin{array}{ccc} & A & B \\ \begin{array}{c} A \\ B \end{array} & \multicolumn{2}{c}{\begin{pmatrix} 3 & 1 \\ 5 & 0 \end{pmatrix}} \end{array}$$

- 両プレイヤがAを選べば、一方のプレイヤがBに変更することで利得を増加できる。
- 両プレイヤがBを選べば、一方のプレイヤがAに変更することで利得を増加できる。
- $\therefore$ AとBはともにNash均衡点ではない

$$\begin{array}{ccc} & H & D \\ \begin{array}{c} H \\ D \end{array} & \multicolumn{2}{c}{\begin{pmatrix} \frac{b-c}{2} & b \\ 0 & \frac{b}{2} \end{pmatrix}} \end{array}$$

例：タカ・ハトゲーム

$$
\begin{array}{c|cc}
 & H & D \\
\hline
H & \frac{b-c}{2} & b \\
D & 0 & \frac{b}{2}
\end{array}
\qquad
\begin{pmatrix} \frac{b-c}{2} & b \\ 0 & \frac{b}{2} \end{pmatrix}
$$

- b<cならば両戦術ともにNash均衡ではない
- みなが「タカ」⇒「ハト」をプレイすることが最善
- みなが「ハト」⇒「タカ」をプレイすることが最善
- ∴タカとハトは共存できる

- 淘汰ダイナミクス：タカの頻度がb/cである点が安定平衡点

タカ・ハトゲーム：混合戦略

$$\begin{array}{c|cc} & H & D \\ \hline H & \multicolumn{2}{c}{\multirow{2}{*}{}} \end{array}$$

$$\begin{array}{cc} & \begin{array}{cc} H & D \end{array} \\ \begin{array}{c} H \\ D \end{array} & \begin{pmatrix} \frac{b-c}{2} & b \\ 0 & \frac{b}{2} \end{pmatrix} \end{array}$$

- 確率pでタカ、1-pでハトをプレイする混合戦略
- 戦略p2に対する戦略p1の利得

$$E(p_1, p_2) = \frac{b}{2}\left(1 + p_1 - p_2 - \frac{c}{b} p_1 p_2\right)$$

- この関数から、以下の戦略が進化的に安定

$$p^* = \frac{b}{c}$$

定理:Nash均衡は常に存在する

$n \times n$の利得行列A

n個の純粋戦略$S_1, \ldots, S_n$

すべての混合戦略と純粋戦略を合わせた集合を考えると、その中にNash均衡点が必ず存在する

例：チキンゲーム

- ２台の車が互いに向かい合って高速度で走行する
- 初めにハンドルを切った(A)ほうが敗者
- 両者がハンドルを切らなければ(B)双方が重大な損害をこうむる
  - 勝った報酬はb
  - 衝突の損害は-c
  - 両者がハンドルを切れば確率は1/2

$$\begin{array}{c c} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} -c & b \\ 0 & \dfrac{b}{2} \end{pmatrix} \end{array}$$

例：チキンゲーム

- タカ・ハトゲームと同じ結論が導かれる
- 相手と異なる戦略をとることが最適応答である
- AとBの混合戦略がESSとなる

$$\begin{array}{cc} & \begin{matrix} A & B \end{matrix} \\ \begin{matrix} A \\ B \end{matrix} & \begin{pmatrix} -c & b \\ 0 & \dfrac{b}{2} \end{pmatrix} \end{array}$$

例：雪だまりゲーム

- 雪だまりで道路がふさがれて2台の車が立ち往生している
- 協力(C)：車を降りて雪かきをする
- 裏切り(D)：車にとどまり相手を待つ
- 両者が協力すると仕事量は半分
- 両者が裏切ればずっと立ち往生
  - 帰宅できる利益はb
  - 雪かきの労力は-c

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} b-\dfrac{c}{2} & b-c \\ b & 0 \end{pmatrix} \end{array}$$

例：雪だまりゲーム

- $b>c$ならばタカ・ハトゲームと同じ結論が導かれる
- 相手が協力する場合、裏切ることが最適応答となる
- 逆に、相手が裏切る場合は協力が最適応答である
- $b<c$なら相手の戦略に関係なく裏切りが最適応答である。この場合は囚人のジレンマとなる。

$$
\begin{array}{cc}
 & \begin{array}{cc} C & \quad D \end{array} \\
\begin{array}{c} C \\ D \end{array} & \begin{pmatrix} b-\frac{c}{2} & b-c \\ b & 0 \end{pmatrix}
\end{array}
$$

例：第3のゲーム

$$\begin{array}{ccc} & A & B \\ A & \multicolumn{2}{c}{\multirow{2}{*}{}} \end{array}$$

$$\begin{array}{c} \\ A \\ B \end{array}\begin{array}{c} \begin{array}{cc} A & B \end{array} \\ \begin{pmatrix} 5 & 0 \\ 3 & 1 \end{pmatrix} \end{array}$$

- 両プレイヤが*A*を選べば、どちらのプレイヤも*B*に変更することで利得を増加できない。
- 両プレイヤが*B*を選べば、どちらのプレイヤも*A*に変更することで利得を増加できない。
- ∴*A*と*B*はともにNash均衡点である。

囚人のジレンマ

(a) A から見た損得表

	B_1 Bが自白しない(C)	B_2 Bが自白する(D)
A_1 Aが自白しない(C)	−2 懲役2年	−5 懲役5年
A_2 Aが自白する(D)	−1 懲役1年	−3 懲役3年

- (D,D):均衡点
- But 相手が自白しないと信じて自白しなければ、(C,C)　:　どちらにとっても均衡点よりよい！！
- しかしお互い隔離されて協力できない！
- 囚人のジレンマ:仲間が自白するなら自白すべきだし、自白しないなら自白しない。→ ジレンマ!!
- 繰り返すとどうなるか？
  - 繰り返し囚人のジレンマ
  - Iterated Prisoners' Dilemma (IPD)

IPDの利得表（P1に対するもの）

表2・5　IPDの利得表とコード

	相手（P_2）が自白しない $P_2 = \mathrm{C}$	相手（P_2）が自白する $P_2 = \mathrm{D}$
自分（P_1）が自白しない $P_1 = \mathrm{C}$	利得：3 コード：R	利得：0 コード：S
自分（P_1）が自白する $P_1 = \mathrm{D}$	利得：5 コード：T	利得：1 コード：P

ジレンマ性

T>R>P>S　及び

R>(T+S)/2

IPD:繰り返し囚人のジレンマ

- １回だけではない
- 協調的な振る舞いが出現する
- 例：前に裏切ると次には復讐されるから協調を余儀なくされる

IPDのトーナメント（1979）

- 14通の応募（心理学者、数学者などなど。。。）
- ＋「でたらめ戦略　by Axelrod」　：　合計15戦略

- 各戦略は自分自身との試合を含め総当たりする
  - 偶然性の効果を除くため同じ相手と5試合する
- 対戦は1試合につき200回まで
  - つきあいが十分続くため
- 重要なこと
  - 一回の試合の勝ち負けは問題でない
  - 成績や順位は最終的にあげた総得点で決まる

IPDのトーナメント(1979)

- 優勝者:ラポポート(トロント大学、心理・哲学者)
- しっぺ返し
- Basicでわずか4行のプログラム

- 自分から決して相手を裏切らない(礼儀正しさ)
- 相手に裏切られても一度だけ裏切ってあとは根に持たない(肝要さ)

しっぺ返し

- Tit for Tat (TFT)
- Basic で4行
  - 1. はじめは協調
  - 2. 相手の前の手を繰り返す

IPDのトーナメント Part2

- 前回の教訓を公表
- 広くコンピュータ雑誌を通して募集
- 6カ国から幅広い人々が参加
- Tit-for-two-tats: ２発に1発返す（John Maynard Smith）

- 優勝者は？
- またしても、ラポポートの書いたしっぺ返しだった！！

なぜTFTは強いのか？

- 自分から裏切らない肝要さ～気のいいこと
- →上位半分を占めていた
- 下位半分の大半は、自分から裏切るやつ
- 最下位は「でたらめ戦略」

- これはなぜか？

気のいいやつ vs. 気のいいやつ

- どちらからも裏切らない
- 最後の200回まで「協調」する

- 両者の得点はともに3x200＝600点

気のいいやつ vs. 裏切り者

- 「しっぺ返し」vs.「ヨッスJOSS」
- JOSS
  - ほとんどTFTだが、乱数に従って気まぐれに裏切る
  - そのほかはTFTと同じく仲直りする
- 浮気？？

TFT vs. JOSS

- TFT:　CCCCCCDCDCDCD
- JOSS: CCCCCDCDCDCDC

- そのあとは「裏切り」と「協調」の互い違いが続く
- この間の得点は
  - JOSS:5 0 5 0 …
  - TFT：0 5 0 5 …
- もし両者が協調し続けていれば3点だが、
  - 3>(0+5)/2となってしまう！！

TFT の強さは？

- 一度も相手をやっつけていない！！
  - 得点は相手と同点か以下
- お返しは一発だけ！！
  - 「10倍返し」ではない　TFnT – n倍返し
  - 相手の和解のチャンスを無視しない
- カモを必要としない
  - 相手を搾取しない
  - TFTどうしで楽しく反映できる

TFT の重要な仮定

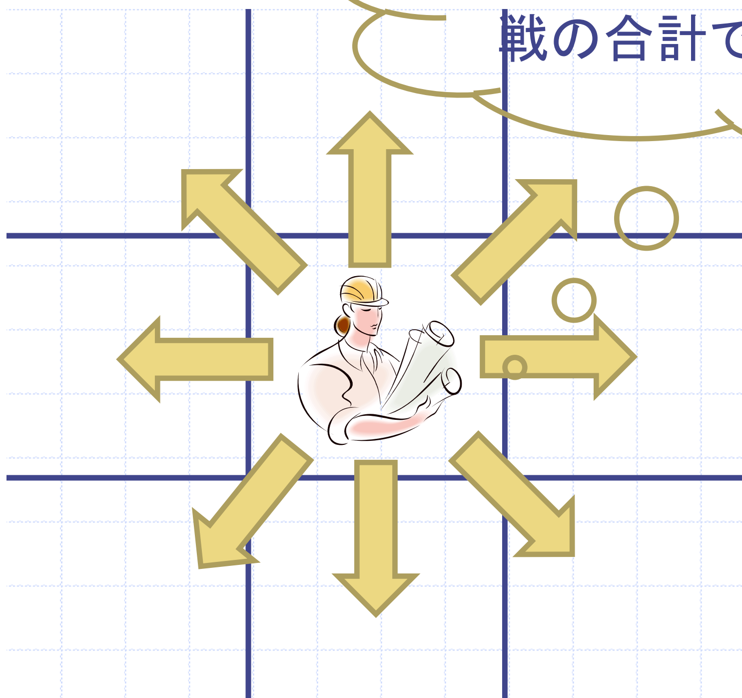
- 終わりが分からない
  - 最後っ屁戦略？
- つきあいが長いこと
- 学生街の定食屋
- 海辺のラーメン屋
- 農耕民族

IPDのいくつかの戦略

- ALLD:いつも裏切る(D)
- ALLC:いつも協力する(C)
- TFT
- GRIM:初回はC,相手が裏切るまでC,裏切ると永遠にDとなり、決して許さない
- GRIM*:GRIMと同じだが最後は必ずD(最後っ屁)
- RANDOM
- WSLS(勝ち残り･負け逃げ)
- ….

Swarmデモ：囚人のジレンマの空間ゲーム

各プレイヤはム近傍のプレイヤと対戦する。利得は全対戦の合計である。



近傍：
- ムーア近傍
- グラフ
- ランダム　etc.

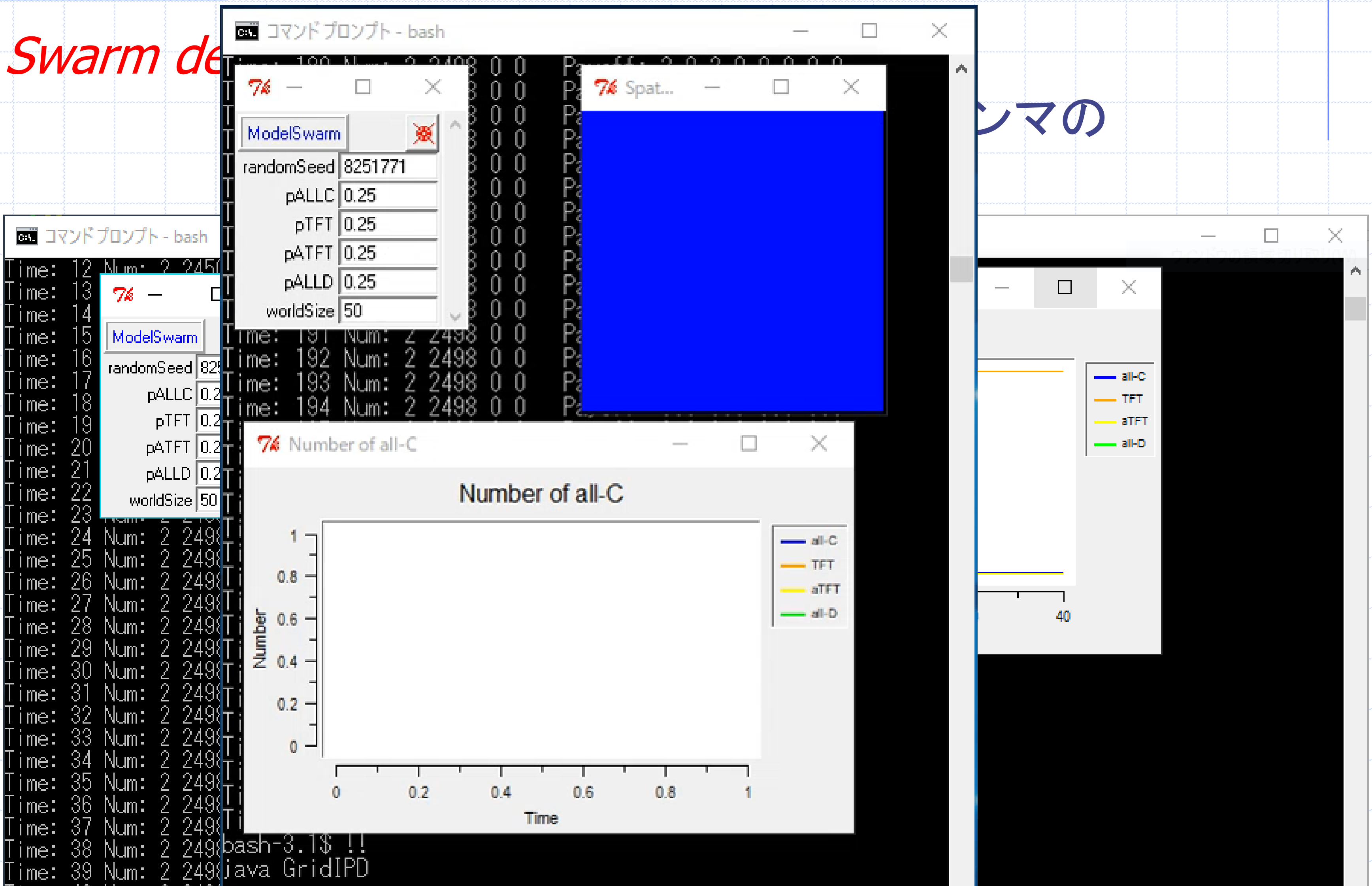
ノイズの影響：
- 認識の間違い
- 戦略の間違い

- 各プレイヤは自分およびとなりのプレイヤの中でもっとも利得の高かった戦略を次に採用する ⇒ 突然変異の可能性あり
- すべてのプレイヤは同期して更新する

デモプログラム：IPDの戦略の進化

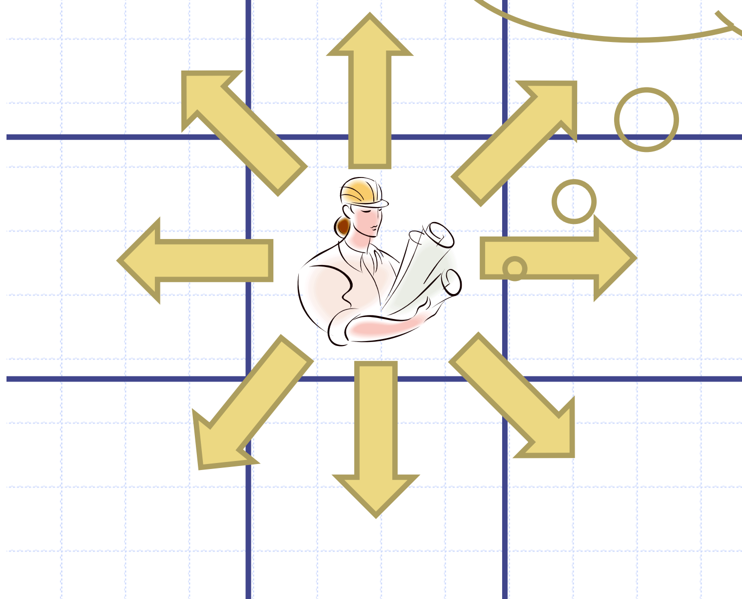
Swarm de

ンマの



Swarmデモ：万華鏡とビックバン

各プレイヤはムーア近傍（隣接する8つのプレイヤ）のプレイヤと対戦する。利得は全対戦の合計である。



利得表

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} 1 & 0 \\ b & 0 \end{pmatrix} \end{array}$$

ただし$3/2 < b < 5/3$

- 各プレイヤは自分およびとなりのプレイヤの中でもっとも利得の高かった戦略を次に採用する
- すべてのプレイヤは同期して更新する

IPDのいくつかの戦略

- ALLD:いつも裏切る(D)
- ALLC:いつも協力する(C)
- TFT
- GRIM:初回はC,相手が裏切るまでC,裏切ると永遠にDとなり、決して許さない
- GRIM*:GRIMと同じだが最後は必ずD(最後っ屁)
- RANDOM
- WSLS(勝ち残り･負け逃げ)
- ....

ALLD vs. GRIM

- m回の対戦の利得行列
- もしmR>T+(m-1)PであればGRIMはALLDに対し狭義のNash均衡点
- 集団全員がGRIMのときALLDは侵入できない

$$
\begin{array}{c c}
 & \begin{array}{cc} GRIM & ALLD \end{array} \\
\begin{array}{c} GRIM \\ ALLD \end{array} & \begin{pmatrix} mR & S+(m-1)P \\ T+(m-1)P & mP \end{pmatrix}
\end{array}
$$

ALLD vs. GRIM

- m回の対戦の利得行列
- もしmが以下の閾値を越えると、GRIMはALLDによる侵入に対して安定である。

$$m > \frac{T-P}{R-P}$$

- しかし mP>S+(m-1)Pであるために、ALLDも狭義のNash均衡点である

$$\begin{array}{cc} & \begin{array}{cc} GRIM & ALLD \end{array} \\ \begin{array}{c} GRIM \\ ALLD \end{array} & \begin{pmatrix} mR & S+(m-1)P \\ T+(m-1)P & mP \end{pmatrix} \end{array}$$

GRIM* vs. GRIM

- m回の対戦の利得行列
- GRIM*のほうがGRIMより有利である
- GRIMばかりの集団はGRIM*に進化的に侵入される

$$
\begin{array}{c|cc}
 & GRIM & GRIM* \\
GRIM & mR & (m-1)R+S \\
GRIM* & (m-1)R+T & (m-1)R+P
\end{array}
$$

GRIM*　vs.　GRIM

- m回の対戦の利得行列
- 1度全員がGRIM*になると、最終回から1回前でも同様に相互裏切りとなる
- もし最終回から1回前も相互裏切りなら、最後から2番目の回も相互裏切りとなる
- つまりGRIMよりもGRIM*が有利となり、GRIM*よりもGRIM**が有利である。
- この議論はどんどん続けられ、結局はALLDが有利となる。
- つまり、ALLDは狭義のNash均衡解＋唯一のESS

TFTの強さのひみつ

- m回の対戦の利得行列
- TFT vs ALLD
- 利得行列はGRIM vs ALLDと同じ

$$
\begin{array}{cc}
 & \begin{array}{cc} GRIM & ALLD \end{array} \\
\begin{array}{c} GRIM \\ ALLD \end{array} & \begin{pmatrix} mR & S+(m-1)P \\ T+(m-1)P & mP \end{pmatrix}
\end{array}
$$

TFT vs ALLD

- m回の対戦の利得行列
- もしmが以下の閾値を越えると、GRIMはALLDによる侵入に対して安定である。

$$m > \frac{T-P}{R-P}$$

$$\begin{array}{cc} & \begin{array}{cc} GRIM & ALLD \end{array} \\ \begin{array}{c} GRIM \\ ALLD \end{array} & \begin{pmatrix} mR & S+(m-1)P \\ T+(m-1)P & mP \end{pmatrix} \end{array}$$

TFT vs ALLD

- TFTは相手が協力すれば再び協力し始めるという点で、GRIMと比べて有利になっている。
- GRIMと違ってTFTは永遠の裏切りの世界から逃れられる

$$\begin{array}{cc} & \begin{array}{cc} GRIM & ALLD \end{array} \\ \begin{array}{c} GRIM \\ ALLD \end{array} & \begin{pmatrix} mR & S+(m-1)P \\ T+(m-1)P & mP \end{pmatrix} \end{array}$$

TFTの弱点

- 間違いを修正できない
  TFT:CCCDCDCDDD…
  TFT:CCCCDCDDDD…
- 間違いが起こると、ゲームは相互協力から協力と裏切りを繰り返すようになる
- さらに間違えると相互裏切りとなる
- もっと間違いが起こると協力に戻る
- しかし長期的に長期的には、ランダムに協力と裏切りを選ぶ戦略2個体の利得と同じになる
- つまり、間違いによってTFTの効率が落ちてしまう

TFTの弱点

- TFT vs ALLC

$$\begin{array}{c|cc} & TFT & ALLC \\ \hline TFT & mR & mR \\ ALLC & mR & mR \end{array}$$

- すべての手番で両方とも協力するので、TFTは狭義のNash均衡でもなくESSでもない。
- 有限集団では、遺伝的浮動によってTFTの数よりALLCの方が多くなる。
- その結果、最後にはALLDに乗っ取られてしまう。

TFTの弱点

- TFTは、ノイズがあるときには、ときどき裏切りを許すような「寛容なしっぺ返し（GTFT）」に取って代わられる。
- しかしTFT, GTFTともにALLCに対して脆弱である。
- その結果ALLDに最後には乗っ取られてしまう。

IPDの研究はまだまだ続く

- Pavlov戦略（１９９３）
- Ｎ人参加によるIPD
  - 協調ゲーム理論
  - 結託の出来る場合

パブロフ戦略

- WSLS（勝ち残り･負け逃げ）
- 前回の2人の手番がCCかDDの時には協力し、CDかDCのときには裏切る
- 高い利得（T, R）なら自分の前回の手番を繰り返すが、低い利得（P, S）なら変更する。
- 現在のところの「世界チャンピョン」

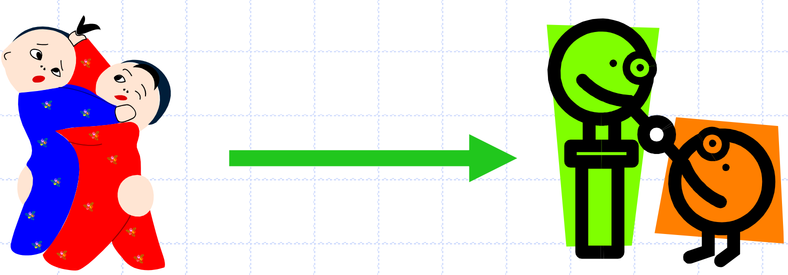
パブロフ戦略

- 間違いを訂正することができる
- WSLS: CCCDDCCC…

  WSLS: CCCCDCCC…

人間の心理と一致？

ちょっとした衝突のあとの友情関係



消える囚人のジレンマ

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} 6 & 7 \\ 2 & 3 \end{pmatrix} \end{array}$$

Aからみた損得表

ジレンマ

$\Longleftrightarrow$

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} 6 & 2 \\ 7 & 3 \end{pmatrix} \end{array}$$

Bからみた損得表

量子ゲームによる囚人のジレンマ

- アリスvs.ボブ
  - アリスはボブの行動を予測して自分の手を決める

- ボブが裏切ると確信している場合にはアリスもおそらく裏切る
- ボブが協力するときにも、アリスはそれなりの確率で裏切る

- 心理実験によると、約80% の確率で裏切る
- ボブがどうするか不明なときどうするか？
  - 裏切り・協力の可能性が五分五分のとき

量子ゲームによる囚人のジレンマ

- 裏切り･協力の可能性が五分五分のとき
- 古典的な論理では90%(= (80+100)/2）の確率で裏切る（当然原理）
- 心理学的実験によると、被験者の裏切りの確率は約40%となった
- この点は従来の論理的推論では説明が困難だった
- 量子ゲームによる説明

なぜ協調が起こるのか？

- 利己的遺伝子
- ミーム
- 進化的安定戦略





幸島、宮崎：
2007-6

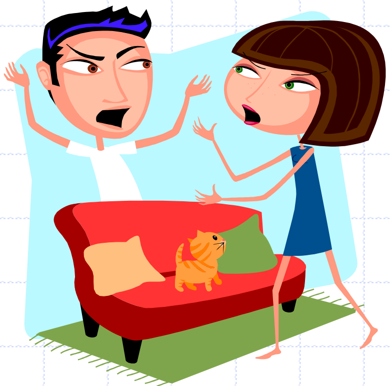
文化の遺伝子

- ミーム
- リチャード・ドーキンス
  - 利己的遺伝子
  - 拡張された表現型

最近の研究から：最後通牒ゲーム

- 2人のプレイヤー
- まだお互い誰か知らされていない
- 2人は一定額のお金を分け合うチャンスを与えられる
- チャンスは1度きり

最後通牒ゲーム



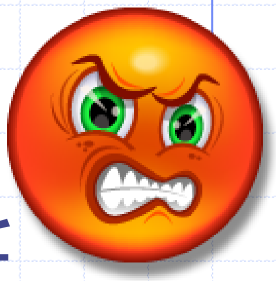
- Aは20＄を渡され、0から20ドルの間の好きな額をBに提示する
- BはAに提示された額を受け入れるか断るかを決める
- Bが受け入れたら二人ともAの提案通りにお金をわけあう
- Bがことわったら二人は何ももらえない

- Aの最適な戦略は何か？

最後通牒ゲーム

- 経済学的な最適戦略
- 1セントでもないよりましだ
- よってBにとっては1セントだって受け入れるべきだ
- したがい、Aとしては1セントだけを提示して、19.99ドルは自分に取っておくのがいい

最後通牒ゲーム



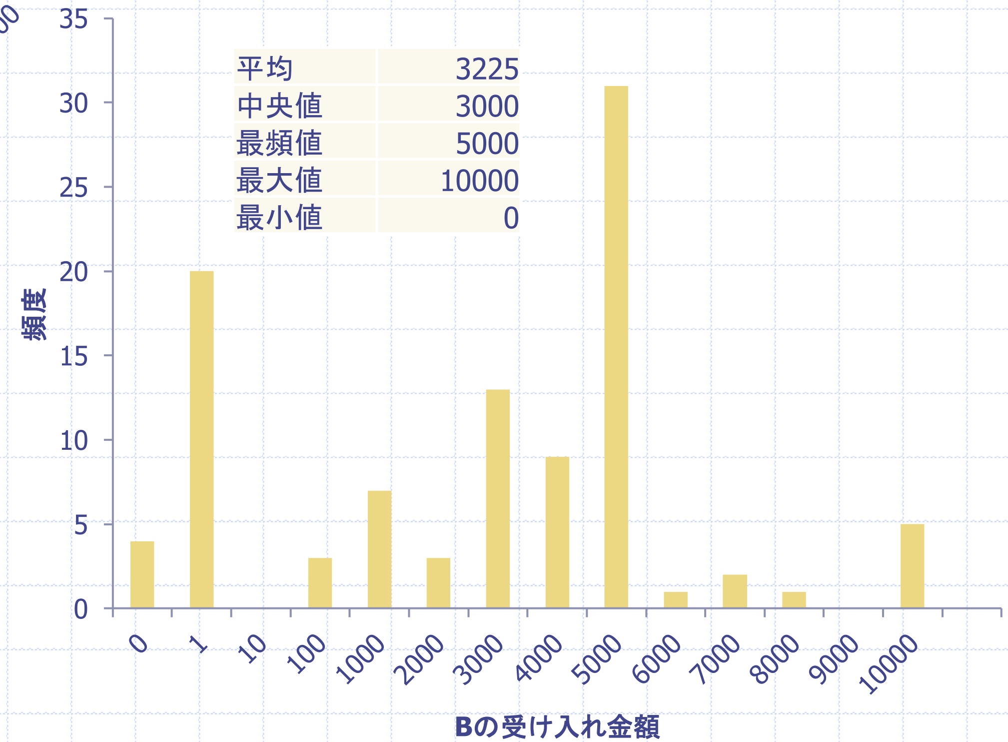
- しかし一般の人はこのようにはしない
- Bは通常3ドル未満を提示されると断る
- 明らかに足元を見られた金額を提示されると腹を立て、お金を取り逃がしても怒りをぶつけてしまう

- そういう足元をみた提示は普通見られない
- AはBに6ドルを上回る額を
  もないよりましだ

思いやり・協調行動が
人間にはもともとある!!

- 相手に断られないためで、気前がいいのが普通である

最後通牒ゲームの結果（@U Tokyo）



独裁者ゲーム

- Aは20ドルを渡され、誰か知らないBという人にそのお金を分け合っていいと言われる。
- Aは0ドルから20ドルまでの好きな金額をBに渡す。
- 結果を決めるのはA（＝独裁者）のみ

独裁者ゲーム

- 多くの国・地域（シカゴ、西モンゴル、タンザニア）で実験してみると、Aは多くの金額を与えた
- 平均で4ドルほど、つまり全体の20%を相手に渡した

- 人類は、実は思いやりと
のか？

超合理的で自分のことしか考えないけだもの

- 経済の教科書にあるような「経済人」は死んだ
- この結果は標準的な経済学の根幹を揺るがした

行動経済学

- ダニエル･カーネマンとエイモス･トヴェルスキー
  - 2002年ノーベル経済学賞受賞
- 損失忌避の実証

- 100ドル得るよりも、100ドルを失う方が心理的に動揺する
- よって、**標準的な人**は、100ドルを得るチャンスよりも、100ドルが失われないことの保証にたいして、多くの金を費やすことを実証した

最近の脳科学研究から

- 電気刺激を用いて脳の右背外側前頭前皮質の働きを阻害した。
- すると、被験者は不公平だと思いながら、最後通牒ゲームにおいて私欲に駆られて少ない提示額を受け入れた。
- 脳のこの領域は自己利益(どんな金額でも了承する)を抑制し、意思決定プロセスに利己的な欲求が入らないようにしている。

最近の脳科学研究から

- オキシトシンを噴霧された被験者はあまり利己的な決定を下さなかった
  - 自分のチームに貢献しようとし、より積極的に他チームのメンバーを妨害した
- オキシトシンは博愛主義をもたらすのか？
  - そうではなく、自分の属する社会集団内での帰属意識を高め、協調行動を促す。
  - 集団外部の者への攻撃性は高くなる。
- オキシトシン分泌という神経生物学的なメカニズムが、身内集団における調和や協調関係を促し、維持するために進化した。