社会的ジレンマとHCI

2015年
認知科学と人工知能

The Tragedy of the Commons by Garrett Hardin in Science, 1968

ある共有の牧草地があり、5人の村人がそれぞれ20頭ずつ羊を飼っている。ここには羊100頭分の牧草しかなく、それが守られていれば、村人みんなが牧草地の恩恵に預かれる。また、この羊は1頭100万円で取引され、羊が1頭この牧草地に増えることにより餌となる牧草が減り、栄養不足のため99万円で取引される。以後、1頭増える度に、1万円ずつ取引価格は下がっていく。(初期の村人一人の取引高は2000万円)

ある一人の村人が、羊を増やしたほうが自分にメリットがあると考えた。そこでその村人はもう1頭羊を放牧した。結果、
その村人は，2079万円の利益を得た(羊21頭×99万円)
他4人の村人は、1980万円と取引高が減った(羊20頭×99万円)。

それを見た周りの村人も自分の効用を最大限に高めたいという考えた。
5人全員が1頭ずつ羊を増やした場合、一人当たり1995万円(羊21頭×95万円)

それぞれが自由に羊を共有地に放し始めた結果として、牧草地が荒れ果て誰にとっても使えないものになってしまった。

Social Dilemma

- 個人が，Cooperate, Defect (Non-cooperate) のいずれかを選択できる。
- 個人に関して言えば，Defectにより多くの利益がもたらされる。
- ただし，全てのメンバーがDefectを選択すると，個々人の利益は少なくなる。

集団のジレンマ

- 環境破壊
- 大気汚染
- オーバーハンティング
- 戦争
- 保護貿易
- 入試競争
- 違法駐車
- 安売り競争

		他人の環境配慮行動	
		配慮する	配慮しない
自分の環境配慮行動	配慮する（協力）	手間はかかるが環境は良い（A）	手間がかかり環境は悪い（B）
	配慮しない（非協力）	手間がかからず環境は良い（C）	手間はかからないが環境は悪い（D）

		すき家＆松屋	
		安売り	定価
吉野家	安売り	すき家＆松屋：売り上げ低下 吉野家：売り上げ低下	すき家＆松屋：大打撃 吉野家：売り上げ倍増
	定価	すき家＆松屋：売り上げ倍増 吉野家：大打撃	すき家＆松屋：通常通り 吉野家：通常通り

ジレンマではないもの

現在の達成金額	54,695,000 円
目標金額	100,000,000 円
スポンサー募集終了まで	終了しました
支援した人数	1,384 人

Social Dilemma Game

- One Shot Game
- Iterative Game

	B: cooperate	B: defect
A: cooperate	A: 120 Yen/B: 120 Yen	A: 0 Yen/B: 180 Yen
A: defect	A: 180 Yen/B: 0 Yen	A: 60 Yen/B:60 Yen

利他的利己主義

- Iterative Dilemma Game
- 2つの条件
  - 短期的な利益の追求を止めて，相互信頼に基づく長期的利益の重要さに気づくこと
  - 協力関係の構築失敗（報復合戦）に陥らないことの確認に元づき，相互信頼が現れること

ケータイを持ったサル –「人間らしさ」の崩壊

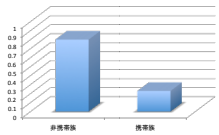
正高信男著

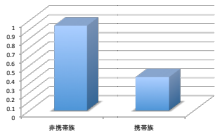
- 2人のプレーヤ
  - 第1プレーヤ，第2プレーヤ
- 所持金5000円
- 相手に5000円を投資すると，相手には倍の1万円が入金される。
- 投資するか否か

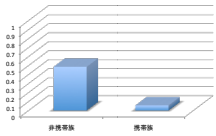
Results

第1 Player

第2 Player







投資群　　非投資群

全米科学振興財団の実験（1984）

- 読者が100ドル，もしくは20ドルのどちらかをハガキに書いて送る
- もし，100ドルの希望者が20パーセント以下であれば，参加者全員に書いた通りの金額を支払う。
- 33,511人の参加者

全米科学振興財団の実験（1984）

- 読者が100ドル，もしくは20ドルのどちらかをハガキに書いて送る
- もし，100ドルの希望者が20パーセント以下であれば，参加者全員に書いた通りの金額を支払う。
- 33,511人の参加者
- 100ドル: 35%(11,758人)

TFT (Tit for Tat)

- 相手の手をそのまま返す。
  - 相手がCooperate → Cooperate
  - 相手がDefect → Defect

トーナメント

- Axelrod, 1980a
  - 15種類の方略
  - ゲーム理論等の専門家
  - TFTが優勝
- Axelrod, 1980b
  - 62種類の方略，TFTが最強であることを通知
  - 再びTFTが優勝

Human Computer Interaction

Research Question

- 人間は，コンピュータエージェントに対しても，利他的利己主義的行動を生み出そうとするだろうか?

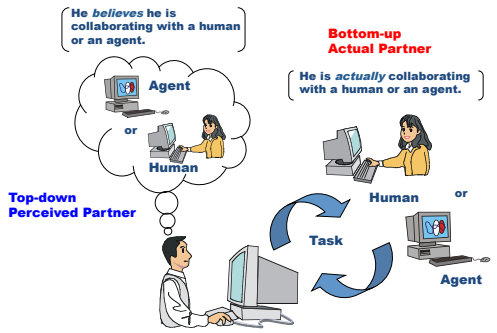
  cf. Media Equation

- 利他的利己主義
  - 短期的な利益の追求を止めて，相互信頼に基づく長期的利益の重要さに気づくこと
  - 協力関係の構築失敗（報復合戦）に陥らないことの確認に元づき，相互信頼が現れること

対人認知

- Bottom-up処理
  - Evidence based
  - 実際の相手の行動
- Top-down処理
  - Schema based
  - プロトタイプ（類型），ステレオタイプ（典型）

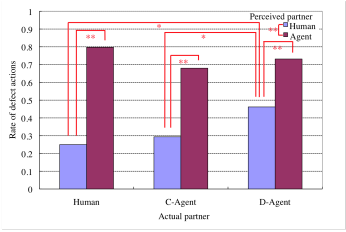
実験セッティング



実験1

- 利得表に基づき実際に獲得資金を配当
- Perceived
  - 人間※，コンピュータエージェント
    ※最初に自己紹介
- Actual
  - 人間，C-agent（1回裏切り），D-agent（5回裏切り）

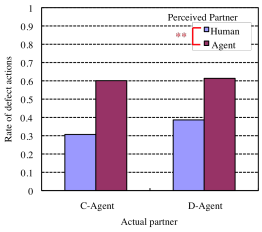
Result



実験2

- 得点を争うゲーム
- Perceived
  - 人間※，コンピュータエージェント
    ※最初に自己紹介
- Actual
  - C-agent（1回裏切り），D-agent（5回裏切り）

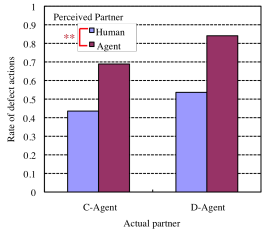
Result



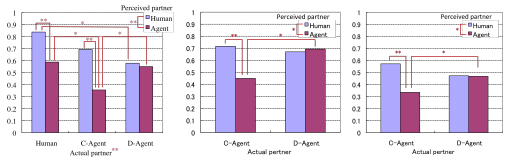
実験3

- 得点を争うゲーム
- Perceived
  - 人間（anonymous），コンピュータエージェント
- Actual
  - C-agent（1回裏切り），D-agent（5回裏切り）

Result



TFTの割合



Research Question

- 人間は，コンピュータエージェントに対しても，利他的利己主義的行動を生み出そうとするだろうか?
  cf. Media Equation
- The answer is,

Research Question

- 人間は，コンピュータエージェントに対しても，利他的利己主義的行動を生み出そうとするだろうか?
  cf. Media Equation
- The answer is,
  No！