

## 人間の選好モデルに基づくメカニズム設計 の基盤技術に関する研究

産業技術総合研究所  
人工知能研究センター  
櫻井 祐子

## 本研究の成果

- クラウドソーシングとケーキ分割問題を対象に人間の選好モデルを考慮したメカニズム設計を検討
  - 本発表では、クラウドソーシングに関する研究成果を中心に発表

**AIST**

## メカニズム設計 (I)

- ある環境に存在する人間やエージェントの集団に対して、集団として望ましい結果を得るためのメカニズムの設計方法に関する研究

情報環境・人間



+

集団意思決定メカニズム  
(例: 入札メカニズム)





結果 (例: 勝者と落札額)

国立研究開発法人 産業技術総合研究所 3

**AIST**

## メカニズム設計 (II)

- 社会科学(ゲーム理論, ミクロ経済学)と情報科学(マルチエージェントシステム)の学際領域**
  - インターネットオークション, 検索連動型広告オークション, クラウドソーシングなど, 従来の経済学では取り扱われていなかった, 新たな情報サービスを対象に含む
  - 計算コスト, 選好の記述コスト等も考慮する必要があり, 情報科学からの貢献が必須



社会科学 (ゲーム理論, ミクロ経済学) と 情報科学 (マルチエージェントシステム) の学際領域は **メカニズム設計** である。

国立研究開発法人 産業技術総合研究所 4



## ヒューマンコンピューテーション

- 人間と計算機の協調問題解決基盤
  - 人間のみ／計算機のみでは  
解決困難な課題に対して、  
両者の能力を融合して解決することを目指す
- 人間=計算資源の一部
- **クラウドソーシング**
  - ヒューマンコンピューテーションを実現するプラットフォーム
  - 人間の労働力へのアクセス手段

**AIST**

## クラウドソーシング

- 2005年にジェフ ハウ (Jeff Howe)  
(アメリカ「Wired誌」の編集者)が提唱
- Crowd (群集)とSourcing (業務委託) を合成した造語
  - クラウドソーシング: **不特定多数の人々**に業務を委託
  - アウトソーシング: **特定の業者**に外部委託
- クラウドソーシングのメリット
  - 量: 大量の「人間の知」に安価にアクセス可能
  - 質: 「群衆の知」の利用(「三人寄れば文殊の知恵」)

国立研究開発法人 産業技術総合研究所 7

**AIST**

## クラウドソーシング関連の研究数の推移

Year	Crowdsourcing	Human Computation	Total
2007	~200	0	~200
2008	~500	0	~500
2009	~1000	0	~1000
2010	~2000	0	~2000
2011	~4000	~500	~4500
2012	~6000	~1000	~7000
2013	~9000	~2000	~11000
2014	~12000	~3000	~15000
2015	~13500	~3000	~16500
2016	~14500	~3000	~17500

- 人工知能学会全国大会にて、2013～2017年に「ヒューマンコンピューテーションとクラウドソーシング」のオーガナイズドセッションも開催

国立研究開発法人 産業技術総合研究所 8

## なぜ、AIにクラウドソーシングが必要か？

- 教師付き学習には、人間が正解を与えた訓練データが必要
- 完成したシステムの評価が必要



## クラウドソーシングにおける課題

### 作業品質制御

- 不誠実に回答するワーカが存在
- 統計的品質制御アプローチ: 事後の対応策
  - 同じタスクを複数のワーカに依頼し冗長化する
  - 多数決などで正解を推定する
  - 余分な金銭的・人為的コストが必要

## クラウドソーシングにおけるメカニズム設計

- メカニズム設計アプローチ: 事前の対応策
    - ワーカが
      - 自らの能力に応じて作業を実行し,
      - 作業結果を正直に申告すること
- が最適行動となるインセンティブを与える

## メカニズム設計アプローチの課題

- 従来のメカニズムデザインでは, エージェント (優れた計算能力と合理的判断を行う) を仮定
- 人間を対象とした場合, 人間の特性や処理能力を考慮したメカニズム設計が必要
  - 入力情報の簡潔化
  - メカニズム表現方法の容易化

考えてみてください

- 何人の人が写っているでしょうか？



考えてみてください

- 何人の人が写っているでしょうか？



A.100人～150人    \_\_\_%

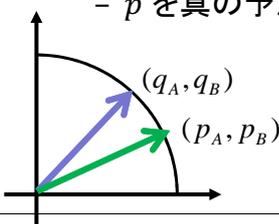
B.150人～200人    \_\_\_%

C.200人～250人    \_\_\_%

**AIST**

## 予測(確率)を正直に申告させる インセンティブ設計

- 予測を正直に申告したとき, 期待報酬(スコア)が最大化されることを保証するルールが存在 [Savage 1971]
- **球状スコアリングルール**: コサイン類似度に基づく
  - 2つの選択肢A, Bに対して, いずれか一方が正解
  - ワーカは予測確率( $q_A, q_B$ )を申告
  - $p$  を真の予測確率,  $q$  を申告する予測確率としたとき



$$p_A \frac{q_A}{\sqrt{q_A^2 + q_B^2}} + p_B \frac{q_B}{\sqrt{q_A^2 + q_B^2}}$$

期待スコアはエージェントの  
真の予測と申告予測の内積

15

**AIST**

## 既存のスコアリングルールの課題



A. 100人~150人    \_\_\_ %  
 B. 150人~200人    \_\_\_ %  
 C. 200人~250人    \_\_\_ %

- ワーカ 1: 「A」 50%, 「B」 10%, 「C」 40%
- ワーカ 2: 「A」 50%, 「B」 40%, 「C」 10%
- 「A」が正解の場合,  
**この二人のワーカが得るスコアは同じ**
- 選択肢間の類似性が存在
  - A(晴れ)に対して, B(曇り)の方がC(雨)より近い
- **正直申告において期待スコアが最大化できることを保証しつつ, リクエスタが設定する選択肢間の類似度に応じてスコアを設定できるようにしたい**

16

**リクエスタ／ワーカのメリット**

- **リクエスタにとって**, 正解が「A(晴れ)」のとき, 「B(曇り)」は「準正解」とみなすことが考えられる
  - ワーカ 1: 「A」 50%, 「B」 10%, 「C」 40%
  - ワーカ 2: 「A」 50%, 「B」 40%, 「C」 10%
- ワーカ 2の方がBの確率を高く予測しているため, ワーカ 2のスコアを高くすることが可能
 

⇒ リクエスタにとって, **スコア(報酬)設定の自由度が上がる**
- ワーカは「正解」だけでなく「準正解」でもスコア(報酬)を受けることが可能
 

⇒ **ワーカにとって, リスクを回避することができる**

17

**一般化球状スコアリングルール**

- 既存の球状スコアリングルールを一般化
- 期待スコアはエージェントの真の予測と申告予測の内積で算出されることに着目

期待スコア:  $\frac{(p, q)_A}{\|q\|_A}$       スコア:  $Q_A(\mathbf{r}, i) = \sum_{j=1}^m q_j \frac{a_{ij}}{\|q\|_A}$

既存メカニズムの報酬行列 A

		結果		
		A	B	C
エージェント	A	0.5	0.1	0.4
	B	0.5	0.4	0.1
	C	0.5	0.4	0.1

提案メカニズムの報酬行列 A

		結果		
		A	B	C
エージェント	A	0.5	0.1	0.4
	B	0.5	0.4	0.1
	C	0.5	0.4	0.1

報酬行列は  
正定値対称行列であることが十分条件

18

**AIST**

### Amazon Mechanical Turkの実験結果

- 各都市の翌日の最高気温を予測
- 3つの範囲を定め, どの範囲に最高気温が含まれるかの確率を申告
- 球状スコアリングルールと一般化球状スコアリングルールどちらで報酬を得たいかを選択

都市	ワーカ数		正解率	
	既存ルール	提案ルール	既存ルール	提案ルール
京都	1	9	50%	38%
パリ	2	8	30%	36%
ニューヨーク	2	8	35%	47.5%
クパティーノ	2	8	35%	30%

国立研究開発法人 産業技術総合研究所 19



**AIST**

## ケーキ分割問題

- 参加者によって選好の異なる連続的な財(ケーキや土地, 公共施設の利用時間等)をどのように分割して参加者に割り当てるかを考える問題



午前中に練習したい!

長い時間練習したい!

3時間連続で使いたい!

参加者の選好

国立研究開発法人 産業技術総合研究所

21

**AIST**

## 研究概要

- 参加者の選好(効用関数)の虚偽申告に関する研究が盛ん
- 参加者の効用関数について, 従来, 加法性を仮定
- 本研究では, 効用関数がAll-or-nothing (0-1)関数を対象
- 正直に効用関数を申告することが最善となるケーキ分割メカニズムを提案
  - 会議室予約など, ある一定の時間の確保が必要といった選好を表明可能

国立研究開発法人 産業技術総合研究所

22

## まとめ

- クラウドソーシングにおいて、何らかの予測(確率)を申告させるタスクを対象に、正解だけに報酬(スコア)を与えるのではなく、正解／準正解／不正解に応じてスコアを与える方法を提案
- ケーキ分割問題において、効用関数がAll-or-nothing (0-1)関数のとき、正直に効用関数を申告することが最善となるメカニズムを提案

## 研究成果

- 櫻井 祐子, 横尾 真: クラウドソーシングのための一般化スコアリングルールの提案, 合同エージェントワークショップ&シンポジウム (JAWS2015), 2015 **最優秀論文賞**
- Yuko Sakurai, et al. : Flexible Reward Plans for Crowdsourced Tasks. The 18th Principles and Practice of Multi-Agent Systems (PRIMA 2015), 400-415, 2015
- Takamasa Ihara, et al.: Strategy-Proof Cake Cutting Mechanisms for All-or-Nothing Utility. The 18th Principles and Practice of Multi-Agent Systems (PRIMA 2015), 118-133, 2015
- 伊原 尚正, 東藤 大樹, 櫻井 祐子, 横尾 真: 二値効用下でのケーキ分割問題, 人工知能学会論文誌, 2017