

提携構造の安定性分析に基づく Q&A コミュニティの インセンティブ設計

藤井 勇基 松原 繁夫

本論文では、Q&A コミュニティにおいて、回答が付かない質問を減らすという目的を実現するインセンティブ設計法を提案する。本稿では、ユーザ駆動の評判システムを特徴とする、プログラミングに関する Q&A コミュニティ Stack Overflow を考察対象とする。従来、この種のインセンティブ設計はコンテスト設計という観点からモデル化されることが多かった。しかし、実際の Q&A コミュニティでは、複数の質問が同時に存在し、また、一つの質問に着目しても複数人が褒章を受け取る場合があり、これをコンテストとして分析することは容易でない。この問題を解決するため、本稿では提携構造形成問題としてモデル化し、分析を行う。本研究の貢献は、回答の付かない質問を減らすという課題に対して、(1) 回答者の質問への割当てを提携構造形成問題としてモデル化し、(2) 実データを表現するのに適した解概念を同定し、(3) シミュレーションを併用して、新たなポイント付与法の性能評価が可能であることを示す点にある。

This paper proposes an incentive design method to reduce questions not answered in the Q&A community. We focus on Stack Overflow, which is a Q&A community on programming, featuring a user-driven reputation system. So far, this kind of incentive design was often studied by modeling the problem as contest design. In the actual Q&A community, however, there are multiple questions simultaneously, and more than one persons, who give answers to the same question, may receive a reward. Such a complicated situation is difficult to analyze as a contest. To overcome this drawback, we try to model and analyze the problem as a coalition formation problem studied well in the cooperative game theory. The contribution of this paper can be summarized as follows. First, we model the assignment problem of answerers to questions as a coalition formation problem. Second, we clarify a solution concept suitable for expressing actual situation. Third, we show that it is possible to evaluate the performance of the new incentive design by using simulation.

1 はじめに

Q&A コミュニティにおいて、回答の付かない質問を減らすには、どのようなインセンティブ設計が望ましいか？インセンティブ設計を行う上で、ゲーム理論と実データをどのように統合すればよいか？本稿では、これらの問題を議論することを目的とする。

本研究では、Q&A コミュニティとして Stack Over-

flow^{†1}を対象にする。Stack Overflow は 2008 年に開設されたプログラミングに関する Q&A コミュニティである。2017 年 8 月時点で登録ユーザ数は 760 万人、質問数と回答数は各々 1,400 万件、2,300 万件に上る。Stack Overflow の特徴はユーザ駆動の評判システムにある。ユーザは他者の質問や回答に対して賛成票、反対票を投じることができる。賛成票を受けた質問者、回答者は各々 1 票当たり 5 点、10 点を獲得し、反対票を受けなければいずれも 2 点を失う。この獲得ポイントに応じて、ユーザには様々な権限が開放される。

Stack Overflow は 1 日当たり訪問数が 700 万を超える、成功している Q&A コミュニティの一つといえる。しかし、回答が与えられる質問は全体の 71%であり、

* Incentive Design of a Q&A community by Analyzing Coalition Stability

This is an unrefereed paper. Copyrights belong to the Authors.

Yuki Fujii and Shigeo Matsubara, 京都大学大学院情報学研究科, Dept. of Social Informatics, Kyoto University.

^{†1} <https://stackoverflow.com/>

質問者が回答を承認した質問は全体の 55%に留まる。回答が付かない理由として、質問の品質が悪いことが疑われるかもしれない。ここで、未回答の質問の獲得ポイントの平均値を調べると 4.8 点であった。つまり、1 票弱の賛成票を得ていることになり、回答が不要な質問と結論づけることは無理がある。よって、回答者を各質問に適切に割り当てることで、Q&A コミュニティの問題解決効率の改善が期待できる。

回答者の質問への割当てに関しては、質問がなされたときに、適した回答者を推薦するという研究が存在するが [5]、本稿では、間接的な制御、すなわち、回答者への評判ポイントの与え方を工夫することで、システム設計者が望む行動を回答者に取らせることを考える。ここで、Q&A コミュニティ参加者の行動動機に関しては様々な議論がある。ポイント獲得を目的とせず、利他的動機で参加するユーザも存在すると考えられる。ただし、行動動機という点で、ポイント獲得が主要な要因の一つであることは間違いない [1]。

さて、ユーザの貢献を引き出すためのインセンティブ設計を考えるとき、既存研究として、褒章を得るためのコンテストとしてモデル化し、均衡を分析する研究が存在する [4][6]。しかし、実際のコミュニティの状況はより複雑であり、コンテストとして分析することは容易でない。質問には、多くの者が閲覧して多数の賛成票が期待できるが、その賛成票の獲得が他の回答者と競合的であるもの、多くの賛成票は期待できないが、他の回答者と競合するリスクは小さいものなど様々なものが存在する。また、1つの質問に複数の回答が付いた場合に、賛成票が得られる回答は1つとは限らず、複数の回答が賛成票を得ることもある。これらすべてをコンテストのモデルに含めることは困難である。

この問題を克服するため、本研究では協力ゲーム理論の枠組みを用いてインセンティブ設計を行う。協力ゲーム理論とは、集団を対象にどのようなグループ分けが実現するか、すなわち、どのような提携構造が安定となるかを分析する方法論である。本稿では、複数の回答者が同じ質問に回答したとき、それを提携とみなし、与えられたポイント付与法のもとでどのような提携構造が安定となるかを分析する。ただし、

これは複数の回答者が協力することに事前合意して回答を構成することを仮定するものではなく、回答者の行動選択が合理的であると仮定し、結果として安定な提携構造が出現すると考えるものである。

本研究では、実データから回答者の嗜好関係を推測し、そのもとでどの解概念が実際の状況を説明するのに適切かを検証する。これは、解概念、すなわち、安定性概念に関するモデル選択といえる。これまでに、協力ゲーム理論と実証分析をより密接に関係付けようとする研究や [7]、交渉問題において、適切な解概念を選択しようとする研究はあるが [8]、実データをもとに提携構造形成問題での解概念選択を議論する点が本研究の新たな点である。

さらには、実データから推測した嗜好関係と、選択された解概念に基づいて、シミュレーションによりポイント付与法を変更した場合に回答者分布がどのように変化するかを調べる。このように、実データをもとに理論とシミュレーションを組み合わせるインセンティブ設計を行う点が本研究の特長である。

本研究の貢献は、回答の付かない質問を減らすという課題に対して、(1) 回答者の質問への割当てを提携構造形成問題としてモデル化し、(2) 実データを表現するのに適した解概念を同定し、(3) シミュレーションを併用して、新たなポイント付与法の性能評価が可能であることを示した点にある。

2 提携構造形成問題

2.1 協力ゲーム

エージェント集合の存在を仮定する。提携 (coalition) とは、エージェントの部分集合であり、どのエージェントとどのエージェントがチームを組むかを表す。また、集団全体がどのような提携に分かれるかを提携構造 (coalition structure) と呼ぶ。協力ゲームは、提携に対して利得が定義されているときに、どのような提携構造が現れるかを説明する理論である。

例えば、3 エージェント a_1, a_2, a_3 がいるとき、 a_1 にとっては、二人提携として $\{a_1, a_2\}$ と $\{a_1, a_3\}$ が考えられ、三人提携として $\{a_1, a_2, a_3\}$ が考えられる。複数名で提携を組むとは限らず、単独で行動する場合を $\{a_1\}$ と表し、これを単独提携と呼

ぶ。提携構造としては、 $\{a_1, a_2, a_3\}$, $\{\{a_1, a_2\}, \{a_3\}\}$, $\{\{a_1, a_3\}, \{a_2\}\}$, $\{\{a_1\}, \{a_2\}, \{a_3\}\}$ の4つが考えられる。 $\{a_1, a_2, a_3\}$ は3エージェントで提携を組む場合を表し、 $\{\{a_1\}, \{a_2\}, \{a_3\}\}$ は3エージェントが各々単独で行動する場合を表す。

協力ゲームという呼称であるが、参加者は自己の利得を犠牲にして他者を助けるといったことは意味しない。定義された利得の元で、各参加者は自己の利得が最大になるようにどの提携に属するかを決定する。このとき、どのような提携構造が実現するか、提携の中でどのように利得が分配されるかを議論するのが協力ゲーム理論である。前者は提携構造形成問題 (coalition formation problem) と呼ばれる。

協力ゲームは、譲渡可能効用を持つゲーム (Game with transferable utility, TU ゲーム) と譲渡可能効用を持たないゲーム (Game with non-transferable utility, NTU ゲーム) に分類される。前者は、金銭などを用いて参加者間で効用をやり取りできる場面を想定しており、後者は、やり取りできない場面を想定している。本研究では、Q&A コミュニティにおけるインセンティブ設計を議論するが、参加者間でポイントの譲渡といったことは認められていないため、NTU ゲームの一つであるヘドニックゲーム (hedonic game) の枠組みを用いることにする。

2.2 ヘドニックゲーム

ヘドニックゲームの特徴は、各エージェントが自己が属する提携にのみ関心を持ち、他の提携に含まれるエージェントには一切関心を持たないという仮定を置く点である [3][2]。ヘドニックゲームは安定結婚問題などのマッチング問題の一般化となっており、計算機科学分野でも興味を集めている。

Q&A コミュニティでは、ユーザは他者と評判ポイントを競うといったことが皆無ではないが、多くの場合は自己の獲得ポイントにのみ興味があって、他者の獲得ポイントには興味を持たないと考えられる。また、以下で述べるように、ヘドニックゲームでは、選好は基数的関係ではなく、序数的関係で表される。実データから選好関係を獲得するとした場合、提携の利得が何ポイントであるか正確に表現するのは難しい

ため、順序関係として表現するのが簡単である。よって、本研究ではQ&A コミュニティにおける回答者の行動をヘドニックゲームとして表現することにする。

ここで、ヘドニックゲームを定義する。 N でエージェントの有限集合を表す。提携 (coalition) とは N の非空の部分集合である。 $\mathcal{N}_{a_i} = S \subseteq N : a_i \in S$ はエージェント a_i を含むすべての提携の集合である。提携構造 (coalition structure) はエージェント集合 N の互いに素な提携への分割 π である。 \succsim で、各エージェント a_i に対する、 \mathcal{N}_{a_i} 上での反射的、完備、かつ、推移的な二項関係 \succsim を規定する選好プロファイル (preference profile) を表すとき、ヘドニック提携形成ゲームは (N, \succsim) で表される。 \succsim_{a_i} を選好関係 (preference relation) と呼ぶ。

提携構造が与えられたとき、単一エージェントの逸脱とは、当該エージェントが現行の提携を離脱し、別の提携 (空の提携を含む) へ移動することを意味する。現行の提携よりも新しい提携を選好するエージェントが存在する場合、エージェントは逸脱への誘因 (incentive to deviate) を持つ。これを、利益のある逸脱 (profitable deviation) と呼ぶ。どのエージェントも利益のある逸脱を持たないとき、許容される単一エージェントの逸脱のクラスに関して、提携構造は安定 (stable) であるという。

つぎに、解概念 (solution concept) について説明する。エージェント a_i を含む2つの提携 $S, T \in \mathcal{N}_{a_i}$ に対して、 a_i が T よりも S を厳密に選好するとき、 $S \succ_{a_i} T$ と表し、 S と T が無差別であるとき、 $S \sim_{a_i} T$ と表し、 $S \succ_{a_i} T$ か $S \sim_{a_i} T$ が成立する場合、 $S \succeq_{a_i} T$ と表す。提携構造 π が与えられたとき、エージェント a_i が含まれる提携を $\pi(a_i)$ で表す。

3章で、解概念として、コア、個人安定性、パレート最適性を比較する。これらに加えて、解概念間の関係の理解に役立つ個人合理性、完全分割、Nash 安定性についても以下で説明する。

2.2.1 個人合理性

定義 1 (個人合理性 (individual rationality)). どのエージェントも少なくとも単独でいるのと同様に提携 $\pi(a_i)$ に属することを望むとき、すなわち、すべての $a_i \in N$ にとって、 $\pi(a_i) \succeq_{a_i} \{a_i\}$ であるとき、分割

π は個人合理的であるという。

個人合理性は提携構造が安定となるための最低限の要求と考えられる。

つぎに、完全分割 (perfect partition) の概念とは、どのエージェントも最も選好する提携に属する場合を指し、そのとき、分割は完全である (perfect) という。どのような逸脱が許容されるとしても、完全分割はつねに安定である。

2.2.2 パレート最適性

定義 2 (パレート最適 (Pareto optimal)). すべてのエージェント a_j に対して、 $\pi'(a_j) \succeq_{a_j} \pi(a_j)$ となり、かつ、少なくとも一人のエージェント a_i に対して、 $\pi'(a_i) \succ_{a_i} \pi(a_i)$ となるような分割 π' がいないとき、分割 π はパレート最適 (Pareto optimal) である。

パレート最適性は、エージェント集合全体としてより選好される分割がないという点で、集団視点での安定性の概念と考えることができる。

2.2.3 コア

定義 3 (コア (core)). コアは、協力ゲーム理論における最も基本的な解概念である。提携 S に含まれるエージェント a_i が分割 π における現行の提携 $\pi(a_i)$ より S を厳密に選好するとき、提携 $S \subseteq N$ は分割 π を強く阻止 (strongly block) するという。強く阻止する提携が存在しない分割がコアであり、コア安定であるという。

阻止に関するより弱い定義が考えられる。提携 S に含まれるエージェント $a_i \in S$ が提携 $\pi(a_i)$ よりも S を弱く選好し、かつ、少なくとも一人のエージェント $a_j \in S$ が現行の提携 $\pi(a_j)$ よりも S を厳密に選好するとき、提携 $S \subseteq N$ は分割 π を弱く阻止 (weakly block) するという。

2.2.4 個人安定性

定義 4 (Nash 安定 (Nash stable)). どのエージェントもある提携から別の提携 (空の提携を含む) に移動することで利益を得ることができないならば、すなわち、すべての $a_i \in N$ にとって、すべての $S \in \pi \cup \{\emptyset\}$ に対して、 $\pi(a_i) \succeq_{a_i} S \cup \{a_i\}$ となるならば、分割 π は Nash 安定である。

$\{\emptyset\}$ は空の提携を表し、エージェント a_i が空の提携に移動するということは、現在所属する提携を離れ

て、単独提携 $\{a_i\}$ を形成することを意味する。

定義 5 (個人安定 (individually stable)). どのエージェントも、提携に含まれる他のメンバの効用を低下させることなく、ある提携から別の提携 (空の提携を含む) に移ることで利益を得ることがないとき、すなわち、すべての $a_i \in N$ に対して、 $S \cup \{a_i\} \succ_{a_i} \pi(a_i)$ となるような、 $S \neq \pi(a_i)$ である提携 $S \in \pi \cup \{\emptyset\}$ が存在するならば、 $S \succ_{a_i} S \cup \{a_i\}$ となる $j \in S$ が存在するとき、これを個人安定 (individually stable) であるという。

2.2.5 解概念間の関係性

コアが成立しているとき、コアに属する場合の利得は単独提携によって改善されることはない。すなわち、分割 π がコアであるなら、 $\pi(a_i) \succeq_{a_i} \{a_i\}$ が成立する。したがって、ある分割がコアであるならば、その分割は個人合理性を満たす。同様に、ある分割 π が Nash 安定であるとき、その分割に対して、 $S \cup \{a_i\} \succ_{a_i} \pi(a_i)$ となるような、提携 $S \subseteq \pi$ が存在しないため π は個人安定性を満たす。また個人安定性を満たす提携構造 π が個人合理性を満たさないと仮定すると $\{a_i\} \succ_{a_i} \pi(a_i)$ が成立するが、このとき $a_j \in \{a_i\}$ 、かつ、 $\{a_i\} \succ_{a_j} \{a_i\} \cup \{a_j\}$ を満たすような a_j は存在しない。したがって、個人安定性を満たす分割は個人合理的である。

一方、パレート最適性、コア、個人安定性の間に包含関係はない。パレート最適であれば必ずコア安定である、コア安定であれば必ず個人安定であるといった関係は成立しない。そのため、各安定性概念がどのような分割を与えるかは個別に議論する必要がある。

モデルの予測力という点では、完全分割など強い安定性概念のもとで議論することが望ましいが、強い安定性を満たす分割は存在しない場合が多い。そのため、以下本稿では、比較的弱い安定性概念である、パレート最適性、コア、個人安定性について議論することとする。

3 提携構造形成問題としての Stack Overflow

本章では、Stack Overflow における回答者の行動をヘドニックゲームとして考え、その行動を特徴付

けるのにどの解概念が適切であるかを調べる。まず、回答間の関係をどのようにユーザの選好関係に反映させるかを述べ、つぎに、Stack Overflow の実データにおけるポイント分布を分析し、提携構造形成問題としてのシミュレーション実施法を述べ、最後に、どの安定性概念が実データを説明するのに適切かを評価する。

3.1 Stack Overflow における回答行動と選好関係の対応づけ

Stack Overflow において、1つの質問に複数の回答があった場合、その回答間の関係を観察すると、補完的（別解）、補完的（補足・修正）、代替的、誤り付加に分類できる。以下、順に説明する。

補完的（別解）

複数の回答が互いに補完し合う関係にあり、1つの回答が別の回答と異なる内容の回答になっている。複数の回答が合わさることで、質問者の問題をより効率的に解決できる。このタイプの回答については、実データ観察の結果、どの回答者も多くのポイントを獲得していた。Stack Overflow では、1つの質問に対する回答数が多くなるほど回答者の獲得ポイントの平均値が高くなる傾向にあり、回答者は一人で回答するより、提携に参加する方がポイントを多く獲得できると考えられる。この関係は、二人の回答者 a_1, a_2 がいるとき、その選好関係を以下のように表現できる。

$$\begin{aligned} \{a_1, a_2\} &\succeq_{a_1} \{a_1\} \\ \{a_1, a_2\} &\succeq_{a_2} \{a_2\} \end{aligned}$$

補完的（補足・修正）

複数の回答が互いに補完し合う関係にあり、一方の回答がもう一方の回答を補足、修正する関係にある。補足や修正はそれらの回答単体では、質問者の問題を解決することはできず、ポイントを獲得できない。しかし、補足・修正が提示されることで、質問者の問題をより効率的に解決できることになり、結果としてより多くのポイント獲得に結びつくと考えられる。回答者 a_1, a_2 がいて、 a_1 の回答を a_2 の回答が補足・修正しているとする。このとき、空集合 \emptyset で参加しないこと、つまり、ポイント獲得できない状況を表す

と、その選好関係は以下のように表現できる。

$$\begin{aligned} \{a_1, a_2\} &\succeq_{a_1} \{a_1\} \\ \{a_2\} &\sim_{a_2} \emptyset \end{aligned}$$

代替的

複数の回答が代替的な関係にある、つまり、一方があれば十分という状況である。2つの回答が類似している場合、2つあることで付加価値が生じるのではなく、質問者の問題解決には一方の回答があれば十分である。ただし、回答者の回答能力により、回答の詳細さや理解しやすさに差が生じる場合はある。このため、獲得ポイントに差が生じると考えられる。つまり、相対的に回答能力の低い回答者は、回答が競合しそうな質問に回答しても、競合しなそうな質問に単独で回答するときと比べて、より多くのポイントを獲得することはできない。この関係を、回答者 a_1 の方が a_2 よりも回答能力が高いと仮定すると、選好関係として以下のように表現できる。

$$\begin{aligned} \{a_1, a_2\} &\succeq_{a_1} \{a_1\} \\ \{a_1, a_2\} &\sim_{a_2} \{a_2\} \end{aligned}$$

誤り混入

複数の回答のうち1つが誤りを含んでいる場合がある。このタイプの場合、回答数が増えることで質問の注目度が大きくなり、また、他の回答と比較されることで、誤りに対する認識が深まり、回答評価が大きく下がると考えられる。回答者として a_1, a_2 がいて a_2 の回答が誤りを含んでいる場合、選好関係は以下のように表現できる。

$$\{a_2\} \succ_{a_2} \{a_1, a_2\}$$

上記のように考えることで、回答間の関係が与えられれば、それをヘドニックゲームの選好関係として表現することが可能である。また、ヘドニックゲームの選好関係が与えられれば、それを回答間の関係として解釈することが可能である。

3.2 回答者数の分布と賛成票数の分布

1質問当たりの回答数の分布を表1に示す。これは、1つ以上の回答が付いた質問に関する分布である。Stack Overflow では、多くの回答が付く質問は少なく、1回答しか付かない質問が全体の57.0%を占め、2回答付くもの、3回答付くものが、各々26.4%

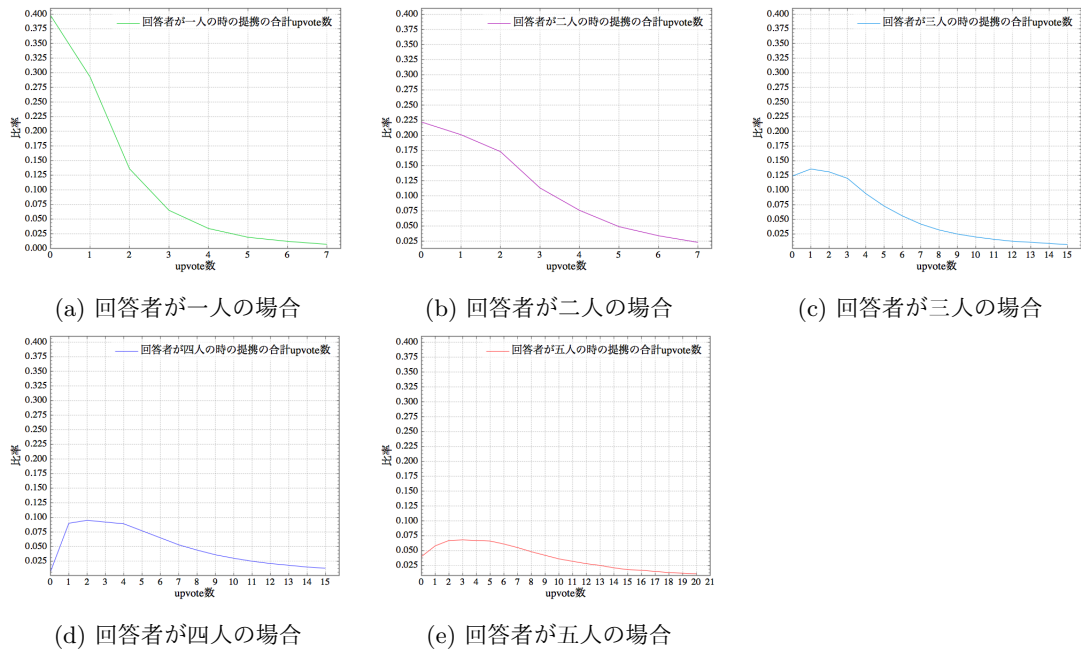


図 1 Stack Overflow における賛成票数の分布

表 1 1 質問当たりの回答数の分布

| 回答数 | 比率 |
|--------|-------|
| 1 回答 | 0.570 |
| 2 回答 | 0.264 |
| 3 回答 | 0.108 |
| 4 回答以上 | 0.058 |

10.8%を占める。

つぎに、賛成票の獲得状況について調べた結果を図 1 に示す。図 1(a) から (e) は、順に回答者数が 1 人の場合から 5 人の場合に対応する。上述のように、賛成票を 1 票得れば、評判ポイントを 10 点獲得できる。各図は、横軸が合計賛成票数を表し、縦軸がその比率を表す。図 1 を見ると、多くの回答を持つ質問の方が、獲得した賛成票数の平均が高くなるのがわかる。これは、閲覧者が高い関心を持つと予想される質問を回答者が見つけて回答していると考えられる。また、多くの場合、賛成票は少数の回答に集中する傾向にある。例えば、5 つ回答を持つ質問では、60%程度の賛成票が一つの回答に投じられている。一方で、まったく賛成票を獲得していない回答も存在するなど不均一な賛成票の獲得結果となっており、競争的な状

況にあることがわかる。

3.3 解概念比較のための実験法

本節では、安定性に関するどの解概念が Stack Overflow での回答状況を説明するのに適切かを調べる方法について述べる。以下の手順を考える。(Step 1) 回答状況に関する実データをもとに、回答者の選好関係を獲得する。(Step 2) 獲得された選好関係のもとで、各解概念を用いて均衡となる提携構造を求める。(Step 3) そこでの回答者数分布が実データにおける回答者分布とどの程度一致するかという観点で、適した解概念を選択する。

Step 1 としての、回答者の選好関係の具体的な獲得法としては、図 1 に示す Stack Overflow 上の各回答者数ごとの賛成票分布を、賛成票数を引数とする確率密度関数とみなし、それに従うように各提携に対する利得の値を決定する。ただし、上述のように、Stack Overflow では同じ質問に答えている回答者間でも賛成票の獲得数に大きな差が見られる。これは、回答者の知識や回答スキルに依存するものと考えられる。そこで、回答者に回答能力を仮定し、回答能力の高い者が多くの賛成票数を割り当て、低い者には少

ない賛成票しか割り当てないように設定する。

より詳細には、単独提携でない場合、実データからサンプリングして、 n 番目に高い能力を持つ回答者が得る割合の平均値を求める。この割合の平均値と、図 1 の分布に従うように生成した提携の合計賛成票数の積を計算し、その整数部分を n 番目に高い能力を持つ回答者が得る賛成票数と設定する。小数部分は、誤差項と考えて、乱数を用いて残りの提携メンバに割り当てる。こうすることで、平均として、Stack Overflow の提携の分配率と整合し、かつ、Stack Overflow における賛成票が整数単位であることを表現できる。その後、割り振られた賛成票の大小関係から、各参加者に選好関係を決定する。

つぎに、Step 2 として、生成された選好関係から、ヘドニックゲームの 3 つの解概念を用いて安定な提携構造を求める。ただし、コア、あるいは、個人安定性を満たす提携構造に関しては、つねに存在するとは限らない。存在しない場合には、最低限満たすべき条件としての個人合理性を満たす提携で代替する。安定な提携構造は複数存在する場合があるが、探索過程で最初に見つかったものを選択する。

上記の手順に従い、1,000 個の選好プロファイルを生成し、安定な提携構造を求める。その提携構造における回答者数の分布と、実データ上の回答者分布の類似度を Kullback-Leibler 情報量を用いて比較する。Kullback-Leibler 情報量とは、2 つの確率分布の差異を計る尺度であり、その値が 0 となると、2 つの確率分布は同一である。この Kullback-Leibler 情報量の観点から、Stack Overflow の回答状況を説明するのに適した解概念を明らかにする。

3.4 解概念比較に関する実験結果

表 2 3 つの解概念と Kullback-Leibler 情報量

| 解概念 | Kullback-Leibler 情報量 |
|---------|----------------------|
| コア | 0.0213 |
| 個人安定性 | 0.0472 |
| パレート最適性 | 0.0510 |

図 2 に、3 つの解概念を用いた安定な提携構造を

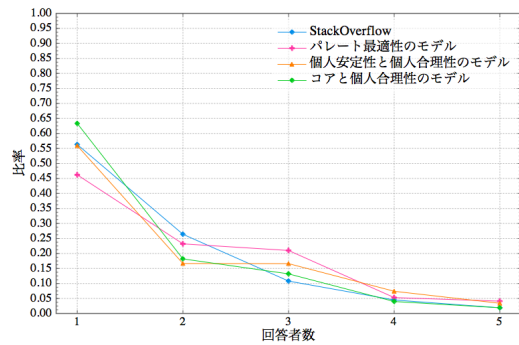


図 2 3 つの解概念における回答数の分布

求めた場合の結果を示す。図の横軸は 1 つの質問に付く回答数であり、縦軸がその比率を表す。青線が実データでの回答者数分布を表す。緑線、橙線、赤線がそれぞれコア、個人安定性、パレート最適性の解概念に対応する。また、表 2 に、各解概念に対する Kullback-Leibler 情報量を示す。Kullback-Leibler 情報量から、解概念の実モデルへの当てはまりのよさを比較すると、コア（コアが存在しない場合は個人合理性）を用いる場合が最良であった。個人安定性やパレート最適性を解概念として用いると、実際よりも大きな提携が形成される様子が観察された。Stack Overflow では、提携規模が大きくなる時、投げられる賛成票数が増えるが、それら賛成票は少数の回答に分配されることが多く、大規模な提携は形成されにくいといえる。パレート最適性を用いる場合、個人視点での逸脱は考慮されないため、実データへの当てはまりが悪くなったと考えられる。4 章では、評判ポイント付与法の改善について議論するが、本章で得られた結果に基づいて、解概念としてコアを用いて分析することにする。

4 ポイント付与に関する改善手法の提案と評価

本章では、未回答質問を減らすという目的を実現するために、評判ポイント付与法をどのように変更すればよいかを調べる。未回答質問が減るということは、多人数提携が分裂して複数の少人数提携になって分割数が増えることに対応する。ここでは、3 章で生成した選好プロファイルに変更を加え、そのもとの、コア

となる提携構造を求めることで、単独提携が増えるかどうかを調べる。

さて、単に単独提携が増えればよいというものではない。補完型の関係にある提携は維持しつつ、代替型の提携を単独提携に変換することが望ましく、ポイント付与法の変更により影響を受ける範囲を限定する必要がある。現行の Stack Overflow では、回答者に対し賛成票数に応じて、10点ごとのポイントが与えられる。そこで、10点未満のポイントを与えることで、影響を受ける範囲を限定し、補完型に関する選好に影響を与えることなく、代替型の提携を分割することを考える。

例えば、 a_1, a_2, a_3 の三人の回答者がいて、 $\{a_1, a_2\} \sim_{a_1} \{a_1\}$ 、かつ、 $\{a_1, a_3\} \succ_{a_1} \{a_1\}$ である場合を考える。このとき、単独提携 $\{a_1\}$ に10点未満のポイントを与えれば、前者の選好関係を $\{a_1\} \sim_{a_1} \{a_1, a_2\}$ に変化させつつ、後者の選好関係 $\{a_1, a_3\} \succ_{a_1} \{a_1\}$ を維持することができる。こうして、代替型の選好関係のみに影響を与えることで、目的の実現を図る。

本稿では、以下の3つの評判ポイント付与法について調べる。

方法 1: 単独回答となった場合に、回答者に10点未満のポイントボーナスとして付与する。

方法 2: 単独回答となった場合で、かつ、回答者の評判スコアが低い場合に、10点未満のポイントボーナスとして付与する。

方法 3: 補完・修正回答者および評判スコアの高い単独回答者に10点未満のポイントボーナスとして付与する。

4.1 方法 1: 単独回答者全員にボーナスを付与する場合

単独回答者全員に、10点未満のポイントボーナスとして付与する場合に、回答者分布がどう変化するかを調べる。これは、複数名の提携に属することと単独提携が無差別 \sim である回答者に対して、単独提携を厳密に選好するように選好関係を変化させることをねらいとするものである。このとき、代替型選好については、少なくとも一人のメンバに関して（複数名

提携） \sim （個人提携）の関係が成立しており、この関係を変化させるため、複数名提携が安定でなくなる。一方、補完型選好であっても、（複数名提携） \sim （個人提携）という関係が含まれる場合は、影響を受ける可能性がある。

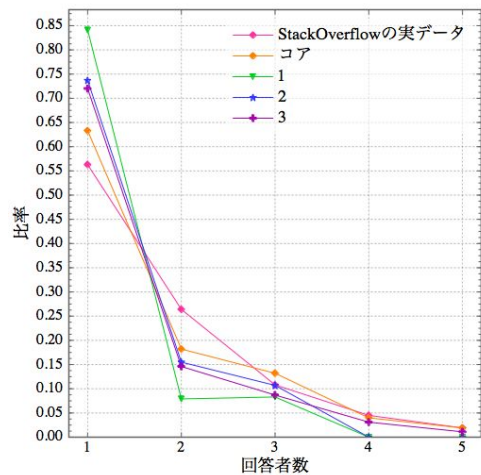


図 3 各ポイント付与法による回答者分布の変化

この方法 1 に関して、コアを解概念とした安定提携構造を求め、その回答数分布を調べた結果を図 3 に示す。図の横軸は回答者数を示す。1 の場合は単独提携に対応する。図の縦軸はその比率を表す。図中の赤線が実データにおける分布を表し、黄線がポイント付与法変更前の分布を表す。緑線、青線、紫線が各々方法 1、方法 2、方法 3 の結果を示す。方法 2、方法 3 の結果については後で述べる。方法 1 の結果を見ると、90%以上の回答が単独回答となり、代替型の提携だけでなく、補完型の提携も破壊されていると考えられる。つまり、副次的効果が大きく、未回答質問の削減には成功するものの、コミュニティで関心の高い質問に対して質の高い回答が得られなくなる恐れがある。

4.2 方法 2: 評判スコアの小さい単独回答者にボーナスを付与する場合

つぎに、評判スコアの小さい単独回答者にのみボーナスを付与する場合に、回答者分布がどう変化するかを調べる。Stack Overflow では、質問や回答などに

よって蓄積されたポイントが評判スコアとして、誰からでも見えるようになっている。この評判スコアの情報を用いる。方法1は単独提携を増やす、すなわち、未回答質問を減らす、過度に単独提携が増えたように見える。この影響範囲を限定するため、単独回答となった場合に全員にボーナスを与えるのではなく、評判スコアの低い回答者のみにボーナスを与えるように変更する。

さて、評価実験では、選好プロファイル生成時に回答能力に関して順序関係を導入するが、評判スコア自体のデータは入手できないため、評判スコアの小さい回答者と判断する何らかの基準が必要となる。ここでは、単独提携と不参加が無差別であるような選好を持つ回答者を評判スコアの小さい回答者として扱い、その者が単独提携となった場合にボーナスを付与する。補完型選好を持つ回答者は、複数名提携に参加してもポイントを獲得できるため、ボーナスの付与ポイントが小さい場合は、依然として複数名提携を選好する。そのため、方法1と比べて、単独提携への選好が抑制される。評判スコアの低い回答者が単独で回答しても、実場面では質問者が承認するような品質の回答は得られないかもしれないが、問題解決の手がかりにはなると考える。

この方法2に関して、コアを解概念とした安定提携構造を求め、その回答数分布を調べた結果を図3に示す。上述のように横軸は回答者数を表し、縦軸はその比率を表す。図中の青線が方法2の結果を表す。方法1と比較すると、単独提携が減り、2人提携や3人提携の比率が増えている。

4.3 方法3: 補完・修正回答者および評判スコアの高い単独回答者にボーナスを付与する場合

補完型（補足・修正）回答では、補足・修正を行う回答者はポイントが得られないと定義した。つまり、補足・修正を行う回答者にとっては、この提携に参加することと不参加であることが無差別となっている。このような回答者に10点未満のポイントを付与することにして、補完型（補足・修正）の提携形成を促進することを考える。例えば、 a_1, a_2 の二人の回答者がいて、選好関係が $\{a_1, a_2\} \succ_{a_1} \{a_1\}$,

かつ、 $\{a_1, a_2\} \sim_{a_2} \{a_2\} \sim_{a_2} \emptyset$ で与えられると仮定する。ここで、方法1で行ったような単独提携全員に対してボーナスを付与すると、後者の関係が $\{a_2\} \succ_{a_2} \{a_1, a_2\} \sim_{a_2} \emptyset$ に変換され、補完型の提携から a_2 は逸脱する。方法3では、 a_2 が a_1 と提携を形成した場合に10点未満のポイントをボーナスとして与えるため、選好関係が $\{a_1, a_2\} \succ_{a_1} \{a_1\}$, かつ、 $\{a_1, a_2\} \succ_{a_2} \{a_2\} \sim_{a_2} \emptyset$ と変更され、補完型（補足・修正）の提携が維持される。

さらに代替型の提携からの逸脱を促進するため、上記のポイント付与と同時に、単独で回答して10点以上得られた場合に、さらに10点未満のボーナスを付与する。例えば、 a_3, a_4 の二人の回答者がいて、選好関係が $\{a_3, a_4\} \succeq_{a_3} \{a_3\}$, かつ、 $\{a_3, a_4\} \sim_{a_4} \{a_4\} \succ_{a_4} \emptyset$ で与えられると仮定する。このとき、方法3では、前者の選好関係が $\{a_3\} \succeq_{a_3} \{a_3, a_4\}$ と変更され、代替型の提携から a_3 は逸脱する。

この方法3に関して、コアを解概念とした安定提携構造を求め、その回答数分布を調べた結果を図3に示す。上述のように横軸は回答者数を表し、縦軸はその比率を表す。図中の紫線が方法3の結果を表す。方法2と比較すると、単独提携、2人提携、3人提携が微減して、4人提携の比率が増えている。方法1に比べて単独提携の増加が緩和されているのは、代替型と補完型（別解）に関係する提携が破壊されたが、補完型（修正・捕捉）に関係する提携が維持されやすくなったためと考えられる。

方法3の欠点は、選好関係が $\{a_1, a_2\} \succ_{a_1} \{a_1\}$ かつ $\{a_1, a_2\} \sim_{a_2} \{a_2\} \succ_{a_2} \emptyset$ で表される補完型（補足・修正）と区別できない代替型の提携を破壊することはできないことと、補完型（別解）の一部の提携も破壊してしまうことである。

本章では、3つの評判ポイント付与法が回答数分布に与える影響を調べた。方法1では単独提携が過度に増えるように見えるといった表現をしたが、どの方法が良いか知るためには、より詳細にコミュニティの状況を調べる必要がある。個々の質問に対する回答の品質向上と未回答質問の削減はトレードオフの関係にあると考えられ、回答数分布だけを見てどれが望ましいかを決定することはできない。本章の貢献として

は、理論と実データを組み合わせることで、ポイント付与法の性質に関する議論が可能となった点にあると考えられたい。

5 おわりに

本論文では、Q&A コミュニティにおいて、回答が付かない質問を減らすという目的を実現するインセンティブ設計法を提案した。まず、1 質問に対する複数回答間の関係を、補完的（別解）、補完的（補足・修正）、代替的、誤り混入の 4 つに分類し、それらが提携構造形成問題における選好関係としてどのように表現できるかを示した。つぎに、実データをもとに選好関係を推定し、そのもとで、コア、個人安定性、パレート最適性の 3 つの安定性概念のどれが実際の状況を表現するのに適切であるかを調べた。その結果、実データとの間の Kullback-Leibler 情報量は、各々 0.0213, 0.0472, 0.0510 となり、3 つの中ではコアが適切であることを示した。最後に、ポイント付与に関する 3 つの改善法をシミュレーションを用いて比較し、インセンティブ設計評価法として、理論と実データを組み合わせることの可能性を示した。

本稿では、ポイント付与法を設計することで、ユーザの行動を誘導しようとした。しかし、コミュニティでは、利己的動機だけではなく利他的動機に基づくユーザも存在する。これは、インセンティブ設計に関して新たな挑戦課題が存在すると思われる。この点に関して詳細な議論を行うことは今後の課題である。

謝辞 本研究は、日本学術振興会科学研究費基盤研

究 (A) (17H00759, 平成 29 年度～31 年度) の補助を受けた。

参考文献

- [1] Anderson, A., Huttenlocher, D., Kleinberg, J., and Leskovec, J.: Discovering Value from Community Activity on Focused Question Answering Sites: A Case Study of Stack Overflow, *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '12, New York, NY, USA, ACM, 2012, pp. 850–858.
- [2] Aziz, H. and Savani, R.: Hedonic Games, *Handbook of Computational Social Choice*, Brandt, F., Conitzer, V., Endriss, U., Lang, J., and Procaccia, A. D.(eds.), Cambridge University Press, 2016, pp. 356–376.
- [3] Bogomolnaia, A. and Jackson, M. O.: The Stability of Hedonic Coalition Structures, *Games and Economic Behavior*, Vol. 38, No. 2(2002), pp. 201 – 230.
- [4] DiPalantino, D. and Vojnovic, M.: Crowdsourcing and all-pay auctions, *Proceedings of the 10th ACM conference on Electronic commerce (EC'09)*, 2009, pp. 119–128.
- [5] Guo, J., Xu, S., Bao, S., and Yu, Y.: Tapping on the Potential of Q&A Community by Recommending Answer Providers, *Proceedings of the 17th ACM Conference on Information and Knowledge Management*, CIKM '08, New York, NY, USA, ACM, 2008, pp. 921–930.
- [6] Immorlica, N., Stoddard, G., and Syrgkanis, V.: Social Status and Badge Design, *Proceedings of the 24th International Conference on World Wide Web*, WWW '15, 2015, pp. 473–483.
- [7] Laver, M. and Shepsle, K. A.: Coalitions and Cabinet Government, *The American Political Science Review*, Vol. 84, No. 3(1990), pp. 873–890.
- [8] Mitsutsune, M. and Adachi, T.: Estimating non-cooperative and cooperative models of bargaining: an empirical comparison, *Empirical Economics*, Vol. 47, No. 2(2014), pp. 669–693.