A Complexity Approach for Core-Selecting Exchange with Multiple Indivisible Goods under Lexicographic Preferences

Etsushi Fujita
Kyushu University
Motooka 744, Fukuoka, Japan
fujita@agent.inf.kyushu-u.ac.jp

Julien Lesca
Paris Dauphine University
Place du Mal de Lattre de Tassigny
75775 Paris Cedex 16, France
julien.lesca@dauphine.fr

Akihisa Sonoda
Kyushu University
Motooka 744, Fukuoka, Japan
sonoda@agent.inf.kyushu-u.ac.jp

Taiki Todo
Kyushu University
Motooka 744, Fukuoka, Japan
todo@inf.kyushu-u.ac.jp

Makoto Yokoo
Kyushu University
Motooka 744, Fukuoka, Japan
yokoo@inf.kyushu-u.ac.jp

Abstract

Core-selection is a crucial property of social choice functions, or rules, in social choice literature. It is also desirable to address the incentive of agents to cheat by misreporting their preferences. This paper investigates an exchange problem where each agent may have multiple indivisible goods, agents' preferences over sets of goods are assumed to be lexicographic, and side payments are not allowed. We propose an exchange rule called augmented top-trading-cycles (ATTC) procedure based on the original TTC procedure. We first show that the ATTC procedure is core-selecting. We then show that finding a beneficial misreport under the ATTC procedure is NP-hard. Under the ATTC procedure, we finally clarify the relationship between preference misreport and splitting, which is a different type of manipulation.

Introduction

Designing rules/mechanisms that achieve good properties is a central research topic in mechanism design and social choice literature. In this paper we study exchange problems with following properties: (i) each agent is initially endowed with a set of indivisible goods, (ii) each agent has a strict preference relation over the set of possible bundles of the goods, (iii) compensation using monetary transfers is prohibited. Exchange rules must be designed so that it prescribe the socially desirable trade of goods. Such exchange problems have many real applications, such as on-campus university housing markets (Chen and Sönmez 2002), nationwide kidney exchanges (Roth, Sönmez, and Ünver 2004), and barter exchanges in disaster areas.

Core-selection is one of the most well-studied properties that rules are expected to achieve. A rule is said to be core-selecting if no group of agents has an incentive to make a cartel and trade their goods among themselves. By definition, core-selecting rules encourage agents to participate in the rules and result in a Pareto efficient trade of goods, which in a sense is a socially optimal outcome. When each agent is assumed to have a single good, Gale's Top-Trading-Cycles (TTC) procedure is core-selecting (Ma 1994).

Another common requirement is strategy-proofness, i.e., each agent has no incentive to misreport her preference. To be more precise, for each agent, submitting her true preference is a dominant strategy regardless of the submitted preferences of other agents. Unfortunately, Sönmez (1999) showed that when there is at least one agent who initially owns more than one good and the agents' preferences are strict, as in our exchange problems, there exists no rule that is strategy-proof and core-selecting.

In the wake of this impossibility, we tackle the incentive issue from the theory of computational complexity. Even if an agent is selfish and hopes to benefit by misreporting, assuming finding a beneficial misreport is hard, e.g., it requires to solve an NP-hard problem, then, as long as its computational power is limited, it will refrain from doing such a manipulation. Under this assumption, in a sense, showing that finding a beneficial preference misreport under a rule is NP-hard can substitute strategy-proofness. Such a complexity approach for incentive issues, especially in social choice, has attracted much attention from computer scientists, and is considered one of the main stream approaches in computational social choice literature (Pini et al. 2011).

In this paper for our exchange problems, we propose a rule called augmented top-trading-cycles (ATTC) procedure. We show that the ATTC procedure is core-selecting under the lexicographic preference domain. We then show that finding a beneficial misreport under the ATTC procedure is NP-hard by a reduction from MONOTONE-3SAT. Finally, we consider a different type of manipulation called *splitting* manipulation, and clarify its relation with preference misreporting. We show that for any splitting manipulation, there exists a corresponding preference misreport that gives the same utility as the splitting to the manipulator.

Related Works

The exchange model we deal with in this paper is also known as housing market (Shapley and Scarf 1974), where each agent initially owns a single house, his/her preference is strict, and monetary transfers are prohibited. In housing markets literature, the TTC procedure is characterized by three properties: individual rationality, Pareto efficiency, and strategy-proofness (Ma 1994). Also it always chooses the unique core assignment (Roth and Postlewaite 1977). Other types of rules' properties have been studied in AI, e.g., fairness (de Keijzer et al. 2009; Endriss et al. 2006), envy-freeness (Chevaleyre, Endriss, and Maudet 2007). On the other hand, when at least one agent initially owns more than one house, as well as the impossibility result presented in the previous section advents, we can no longer guarantee the uniqueness of the core assignment (Sönmez 1999).

One common approach for going beyond such an "impossibility" result is to weaken one of the requirements. For instance, several papers have designed strategy-proof rules by weakening the core-selecting property (Pápai 2003; 2007; Todo, Sun, and Yokoo 2014). Others restrict the agents' preference domain (Sonoda et al. 2014).

Our approach keeps the core-selecting property, but weakens strategy-proofness by focusing on the computational hardness of beneficial manipulation. In various mechanism design/social choice problems, many works consider the computational hardness of beneficial manipulation, such as voting (Bartholdi, Tovey, and Trick 1989) and two-sided matching (Teo, Sethuraman, and Tan 2001; Pini et al. 2011). Although it has been pointed out that such a computational complexity approach is not always sufficient as a barrier for agents' incentives (Faliszewski, Hemaspaandra, and Hemaspaandra 2010; Faliszewski and Procaccia 2010), we believe that discussing complexity in exchange problems is an important first step to develop useful exchange rules for self-interested agents in practice.

In this paper we focus on the lexicographic preference domain. This restriction leads to compact representations of the agents' preferences, and makes sense whenever those preferences are non compensatory (Gigerenzer and Goldstein 1996). Thus such preferences have been well studied in the AI (Booth et al. 2010; Yaman et al. 2011; Conitzer and Xia 2012).

The effect of splitting manipulations has been studied in several algorithmic/economic environments, such as scheduling (Moulin 2008), voting (Conitzer 2008; Todo, Iwasaki, and Yokoo 2011), combinatorial auctions (Yokoo, Sakurai, and Matsubara 2004), two-sided matching (Todo and Conitzer 2013; Afacan 2014), and coalitional games (Aziz et al. 2011; Yokoo et al. 2005; Ohta et al. 2008), some of which are also known as *false-name* manipulations. Especially in research on exchanges, a class of exchange rules resistant to splitting manipulations, as well as another class of manipulations called *hiding* (Atlamaz and Klaus 2007), has been proposed (Todo, Sun, and Yokoo 2014). However, this class of rules does not satisfy the core-selecting property.

Preliminaries

In this section we introduce the exchange problem with multiple indivisible goods studied in this paper as well as several properties of exchange rules that have been discussed in the literature.

Model

We have a set of agents $N = \{1, \dots, n\}$ and a finite set of heterogeneous indivisible goods K. An assignment $\boldsymbol{x} = (x_1, \dots, x_n)$ is a partition of K into n subsets, where x_i is the bundle assigned to agent i. We denote by $\mathcal{X}$ the set of all possible assignments.

The assignment of the goods to the agents is made based on the agent preferences, which are orders/rankings on the subsets of the goods. Since the set of possible bundles grows exponentially with the number of available goods, we need to represent a preference compactly in order to handle a large number of goods. In this paper we focus on lexicographic preferences domain, in which the preference over all subsets of goods can be obtained base on the linear order of the goods.

Definition 1 (Lexicographic Preference Domain)**.** *Let $\mathscr{O}$ be the set of all possible linear orderings over K. The* lexicographic preference P *associated with a linear ordering $\succ \in \mathscr{O}$ is defined as follows: for any $A, B \subseteq K$, we have APB iff there exists $k^* \in A \setminus B$ such that $\{k \in B \mid k \succ k^*\} \subseteq A$ holds. We denote by $\mathscr{L}$ the lexicographic preference domain, i.e., the set of preferences over the bundles of goods that can be represented by lexicographic preferences.*

In the rest of this paper, we assume that agent preferences are lexicographic. Since there exists a bijection between $\mathscr{O}$ and $\mathscr{L}$, whenever there is no ambiguity, for any $i \in N$, we assume P_i denotes the lexicographic preference associated with $\succ_i$ of the considered exchange problem, and R_i denotes the preorder associated with P_i (which includes reflexivity).

An exchange problem $(\boldsymbol{e}, \succ)$ is defined by an initial endowment $\boldsymbol{e} = (e_i)_{i \in N} \in \mathcal{X}$ and by a preference profile $\succ = (\succ_i)_{i \in N} \in \mathscr{O}^n$ of linear orders, where $\succ_i$ denotes the linear order representing the preferences of agent i. For any $i \in N$ and any $\succ'_i \in \mathscr{O}$, we denote by $(\succ'_i, \succ_{-i})$ the preference profile $(\succ_1, \dots, \succ_{i-1}, \succ'_i, \succ_{i+1}, \dots, \succ_n)$.

Properties of Exchange Rules and Known Results

An exchange rule can be described as a function $f : \mathcal{X} \times \mathscr{O}^n \to \mathcal{X}$ that maps any exchange problem to a possible assignment. Let $f_i(\boldsymbol{e}, \succ)$ denote the bundle assigned to agent i under assignment $f(\boldsymbol{e}, \succ)$.

Definition 2 (Individual Rationality)**.** *For an exchange problem $(\boldsymbol{e}, \succ)$, an assignment $\boldsymbol{x} \in \mathcal{X}$ is said to be individually rational if $x_i R_i e_i$ holds for any agent i. An exchange rule f is said to be* individually rational (IR) *if for any exchange problem $(\boldsymbol{e}, \succ)$, $f(\boldsymbol{e}, \succ)$ is individually rational.*

In other words, for every agent, as long as she truthfully reports her preference, an agent never worse off by participating in an IR exchange rule. Under such an exchange rule, every agent is incentivized to participate.

Definition 3 (Pareto Efficiency)**.** *For an exchange problem $(\boldsymbol{e}, \succ)$, an assignment $\boldsymbol{x} \in \mathcal{X}$ is said to be Pareto dominated by another $\boldsymbol{y} \in \mathcal{X}$ if (i) for any agent $i \in N$, $y_i R_i x_i$ holds, (ii) for at least one agent $j \in N$, $y_j P_j x_j$ holds. An exchange rule f is said to be* Pareto efficient (PE) *if for any exchange problem $(\boldsymbol{e}, \succ)$, $f(\boldsymbol{e}, \succ)$ is not Pareto dominated by any assignment.*

When an assignment $\boldsymbol{x}$ is Pareto dominated by another assignment $\boldsymbol{y}$, all the agents weakly prefer $\boldsymbol{y}$ to $\boldsymbol{x}$. Thus, choosing $\boldsymbol{x}$ is suboptimal. In this sense, using a PE exchange rule is "socially optimal", i.e., it never chooses a suboptimal assignment.

Definition 4 (Core Selection)**.** *For an exchange problem* $(\boldsymbol{e}, \succ)$*, a coalition* $T \subseteq N$ *blocks an assignment* $\boldsymbol{x} \in \mathcal{X}$ *if there exists an assignment* $\boldsymbol{y} \in \mathcal{X}$ *such that (i) for any agent* $i \in T$, $y_i \subseteq \bigcup_{j \in T} e_j$*, (ii) for any agent* $i \in T$, $y_i R_i x_i$*, (iii) for at least one agent* $j \in T$, $y_j P_j x_j$*. The* core $\mathcal{C}(\boldsymbol{e}, \succ)$ *is the set of all assignments that are not blocked by any coalition. An exchange rule* f *is said to be* core-selecting (CS) *if for any exchange problem* $(\boldsymbol{e}, \succ)$, $f(\boldsymbol{e}, \succ) \in \mathcal{C}(\boldsymbol{e}, \succ)$ *holds.*

Intuitively, the existence of a set of agents T that blocks an assignment means that they jointly have incentives to form a cartel and get higher utility by leaving behind all the other agents $N \setminus T$. When an exchange rule is CS, no set of agents has such an incentive, and thus they are expected to participate without forming any such cartel. In this sense, CS can be regarded as a refinement of IR to any coalition of agents. By setting $T = N$, the definition coincides with that of PE. Thus, if an exchange rule is CS, it is also PE and IR.

Strategy-proofness, which is a well-known property of exchange rules, is a quite strong incentive constraint. Under a strategy-proof exchange rule, for every agent, reporting her true preference is a dominant strategy. Sönmez (1999) studied a broad class of economic environments that contains exchange problems as a special case. One of his findings reveals that under any strict preference domain, no exchange rule exists that is IR, PE, and strategy-proof when at least one agent is endowed with more than one good. From this result, even under the lexicographic preference domain, no exchange rule exists that is IR, PE, and strategy-proof.

Our focus in this paper is to design an exchange rule that is IR, PE, and "hard to manipulate," while it is inevitably not strategy-proof.

Augmented Top-Trading-Cycles Rule

Let $\delta : K \to N$ be the function that maps for any good $k \in K$ to its owner $\delta(k)$ in $\boldsymbol{e}$. Next we describe our exchange rule that was inspired by the well-known Gale's Top-Trading-Cycle (TTC) procedure:

Definition 5 (Augmented TTC procedure)**.** *For any exchange problem* $(\boldsymbol{e}, \succ)$*, the* Augmented TTC *(ATTC) procedure* φ *runs as follows:*

1. *For the initial step, create a directed graph* $G_1 = (V_1, E_1)$ *such that* $V_1 = K$ *is the set of vertices and* E_1 *is the set of arrows, where each* $(k, k') \in E_1$ *is directed from each* $k \in K$ *to* k'*, which is the most preferred good in* K *according to* $\succ_{\delta(k)}$*. Set* $t \leftarrow 1$*.*
2. *Let* $\mathscr{C}_t$ *be the set of the vertices of* V_t *included in a cycle in* $G_t = (V_t, E_t)$*. For any good* $k \in \mathscr{C}_t$ *and any* $(k, k') \in E_t$ *assign* k' *to* $\delta(k)$*. Create a new set of vertices* $V_{t+1} = V_t \setminus \mathscr{C}_t$ *and a new set of arrows* E_{t+1}*, where each* $(k, k') \in E_{t+1}$ *is directed from any* $k \in V_{t+1}$ *to* k'*, which is the most preferred good in* V_{t+1} *according to* $\succ_{\delta(k)}$*.*

(a) G_1 (b) G_2

Figure 1: Example of the ATTC procedure

3. *If* $V_{t+1} = \emptyset$ *then quit the procedure. Otherwise set* $t \leftarrow t + 1$ *and go back to (2).*

During the ATTC procedure, each agent is divided into several atomic agents, each of which is assigned exactly one good from the original agents' endowments. Then, the standard TTC procedure is applied to these atomic agents. The ATTC procedure generalizes the TTC procedure, since if each agent is initially endowed with a single good, it becomes identical to the TTC procedure. The TTC procedure is well-known because whenever each agent is initially endowed with a single good, this procedure is CS and strategy-proof.

Since at any step t of the ATTC procedure at least one cycle is contained in the ATTC graph G_t, and such a cycle contains at least one good, it is clear that the ATTC procedure runs in polynomial time.

Example 1. *Consider* $N = \{1, 2\}$, $K = \{\alpha, \beta, \gamma\}$*, and an exchange problem* $(\boldsymbol{e}, \succ) = ((\{\alpha, \beta\}, \{\gamma\}), (\succ_1, \succ_2))$ *where* $\gamma \succ_1 \beta \succ_1 \alpha$ *and* $\beta \succ_2 \alpha \succ_2 \gamma$ *hold.*

Figure 1 (a) shows the ATTC graph G_1*. There exists a cycle that contains* β *and* γ*, and the ATTC procedure assigns* γ *to* $\delta(\beta) = 1$ *and* β *to* $\delta(\gamma) = 2$*. Figure 1 (b) shows the ATTC graph* G_2*. There exists a cycle that only contains* α*, and the ATTC procedure assigns* α *to* $\delta(\alpha) = 1$*. Thus,* $\varphi_1(\boldsymbol{e}, \succ) = \{\alpha, \gamma\}$ *and* $\varphi_2(\boldsymbol{e}, \succ) = \{\beta\}$ *hold.*

In Example 1, agent 1 obtains β and γ by misreporting $\succ'_1$ where $\beta \succ'_1 \gamma \succ'_1 \alpha$ holds. Since $\{\beta, \gamma\} P_1 \varphi_1(\boldsymbol{e}, \succ)$ holds, the ATTC procedure is not strategy-proof.

Core Selection

In this section we show that the ATTC procedure inherits the CS property of the TTC procedure for the lexicographic preference domain.

Theorem 1. *The ATTC procedure* φ *is CS.*

Proof. For the sake of contradiction we assume that there exists an exchange problem $(\boldsymbol{e}, (\succ_i)_{i \in N})$ such that the assignment $\boldsymbol{x}$, returned by the ATTC procedure for such problems, is blocked by coalition T. Let $\boldsymbol{y} \in \mathcal{X}$ be an assignment satisfying the conditions (i), (ii), and (iii) in Definition 4.

Let $\mathscr{S}_t$ be the subset of agents of N that obtain a good during round t of the ATTC procedure, let $\mathscr{T}_t = T \cap \mathscr{S}_t$, and let x_i^t be the good obtained by agent $i \in \mathscr{S}_t$ during this round (this x_i^t is unique since the arrows from all the goods belonging to agent i point to the same good during the step t of the ATTC procedure). For any $i \in N$, let L_i represents the list of the goods of K ranked by $\succ_i$, such that $\forall r \in \{1, \ldots, |K|\}$, $L_i(r)$ is the r^{th} best good for agent i. Let $L_i^{-1}(k)$ denote the index of good $k \in K$ in L_i.

For any round t of the ATTC procedure, let H_1^t be the property where $\forall i \in \mathscr{T}_t, x_i^t \in y_i$. Let H_2^t be the property where $\forall i \in \mathscr{T}_t, \forall j \in N, x_i^t \in e_j \Rightarrow j \in T$. Finally let H_3^t be the property where $\forall i \in \mathscr{T}_t, \forall r \in \{1, \dots, L_i^{-1}(x_i^t) - 1\}, L_i(r) \in x_i \Leftrightarrow L_i(r) \in y_i$. We prove H_1^t, H_2^t and H_3^t by induction for any round t.

Base case: For H_1^1, $\forall i \in \mathscr{T}_1$, we have $x_i^1 = L_i(1)$. $y_i R_i x_i$ implies that $x_i^1 \in y_i$. For H_2^1, $\forall i \in \mathscr{T}_1$, let $j \in N$ be such that $x_i^1 \in e_j$. By contradiction, if $j \notin T$, then $x_i^1 \notin \bigcup_{i' \in T} e_{i'}$. But by H_1^1 we know that $x_i^1 \in y_i$. Therefore there is a contradiction, and $j \in T$. Finally, H_3^1 is trivially true since $\forall i \in \mathscr{T}_1, L_i^{-1}(x_i^1) = 1$.

Induction step: First, we prove H_3^t. By contradiction we assume that $\exists i \in \mathscr{T}_t$ and $\exists r \in \{1, \dots, L_i^{-1}(x_i^t) - 1\}$ such that $L_i(r) \in y_i$ and $L_i(r) \notin x_i$ By hypothesis $y_i \subseteq \bigcup_{i' \in T} e_{i'}$, so $\delta(L_i(r)) \in T$. Furthermore $L_i(r)$ cannot belong to G_t because $L_i(r) \succ_i x_i^t$ and x_i^t is the best remaining good for agent i in G_t. So $\exists l \in \{1, \dots, t-1\}$ and $\exists j \in \mathscr{S}_l \setminus \{i\}$ such that $x_j^l = L_i(r)$. Since $\delta(L_i(r)) \in \mathscr{T}_l$ and by repeatedly applying H_2^l along the cycle of G_l containing x_j^l, we know that $j \in T$. By H_1^l we have $L_i(r) \in y_j$ since $j \in \mathscr{T}_l$ and $x_j^l = L_i(r)$. There is a contradiction because we have $L_i(r) \in y_j$ and $L_i(r) \in y_i$ with $i \neq j$. So we proved that $\forall i \in \mathscr{T}_t, \forall r \in \{1, \dots, L_i^{-1}(x_i^t) - 1\}, L_i(r) \in y_i \Rightarrow L_i(r) \in x_i$. This result also implies that $L_i(r) \in x_i \Rightarrow L_i(r) \in y_i$, since based on our hypothesis, we know that $y_i R_i x_i$.

Second, we prove H_1^t. Let $i \in \mathscr{T}_t$. H_3^t implies that $x_i^t \in y_i$ since based on our hypothesis, we know that $y_i R_i x_i$,

Finally, we prove H_2^t. Let $i \in \mathscr{T}_t$ and $j \in N$ such that $x_i^t \in e_j$. By contradiction, if $j \notin T$, then $g_i^t \notin \bigcup_{i' \in T} e_{i'}$. But by H_1^t we know that $x_i^t \in y_i$. Therefore there is a contradiction, and $j \in T$.

Because H_1^t and H_3^t are true for all t, for any agent $i \in T$, $y_i = x_i$ must hold. Thus, $\boldsymbol{x}$ is not blocked by T. This contradicts the assumption and concludes the proof. □

Incentives

In practice, even if an agent is selfish and hopes to benefit by misreporting, her computation power is limited (Bartholdi, Tovey, and Trick 1989). Under this "bounded rationality" assumption, we expect that an agent will refrain from misreporting, if she needs to solve a NP-hard problem to find a beneficial manipulation. We show in this section that finding a beneficial preference misreport in the ATTC procedure is NP-hard, which gives agents reasonable incentives to report their true preferences.

We formalize the manipulation problem on the ATTC procedure as follows:

Definition 6 (BENEFICIAL-MISREPORT)**.**

Instance: *an exchange problem $(\boldsymbol{e}, \succ)$, an agent $i \in N$.*

Objective: *find a misreport $\succ_i' \in \mathscr{O}$ such that $\varphi_i(\boldsymbol{e}, (\succ_i', \succ_{-i})) P_i \varphi_i(\boldsymbol{e}, \succ)$.*

To show the NP-hardness of BENEFICIAL-MISREPORT, we introduce the following problem:

Definition 7 (ALMOST-CONSISTENT-MONOTONE (ACM) 3SAT)**.**

Instance: *a collection $C = \{c_1, \dots, c_m\}$ of m clauses on a set U of variables such that all clauses have exactly three literals. The set of possible clauses is restricted to clauses with only positive literals and to clauses with only negative literals (w.l.o.g., we assume that the indices of the clauses with only positive literals are always lower than the indices of the clauses with only negative literals). Furthermore the first literals of the $m-1$ first clauses are consistent.*

Objective: *find a truth assignment of the variables satisfying all the clauses.*

ACM 3SAT corresponds to the MONOTONE 3SAT problem restricted to instances where the first literals of their $m-1$ first clauses are consistent. MONOTONE 3SAT is NP-hard (Gold 1978).

Lemma 1. ACM 3SAT *is NP-hard.*

We omit the proof due to space limitations.

Before introducing our complexity result we introduce some notations for the ACM 3SAT problem. Let $\mathcal{M} = \{1, \dots, m\}$ and $\mathcal{L} = \{1, 2, 3\}$. Let $\mathcal{I} = \mathcal{M} \times \mathcal{L}$ be the set of literal indices. For any $l \in \mathcal{L}$, let c_r^l denote the l^{th} literal of c_r. Let $\mu \in \mathcal{M}$ be the index of the first clause with only negative literals, i.e., for all $r \in \mathcal{M}$, if $r < \mu$ then c_r only contains positive literals, and if $r \geq \mu$ then c_r only contains negative literals. Finally let $\mathcal{I}(c) = \{(r, l) \in \mathcal{I} \mid c_r^l = c\}$ be the set of all literal indices that corresponds to literal c. To simplify the notations we also consider function $first(c)$ which returns for any literal c the pair of indices $(r', l') \in \mathcal{I}(c)$ with the smallest value for the first component. We also consider for any literal c and any $r \in \mathcal{M}$ such that $\exists l \in \mathcal{L}$ with $(r, l) \in \mathcal{I}(c)$, function $next(c, r, l)$, which returns the pair of indices $(r', l') \in \mathcal{I}(c)$ that follows (r, l) in $\mathcal{I}(c)$ when we order the pair of indices in increasing order by the first component's value. (when such an element does not exist $next$ returns $(0, 0)$). Now we are able to show the main result:

Theorem 2. BENEFICIAL-MISREPORT *is NP-hard.*

Proof. To prove this result we show that, from an instance of ACM 3SAT, we can build in polynomial time an instance of BENEFICIAL-MISREPORT, and from this instance's optimal solution we can construct a truth assignment of U that is consistent with C, if such an assignment exists.

To any literal c_r^l, we associate a good χ_r^l. $\{\chi_r^l\}_{(r,l) \in \mathcal{I}}$ constitutes the endowment of agent i. Except for agent i, we assume that the agents own only one good. To simplify notations we identify a good with its owner. We first construct the set of goods that agent i will try to acquire. We associate to any clause c_r of C good g_r. These goods are the main set of goods used to obtain a solution to the ACM 3SAT problem. For any clause c_r, agent i will be able to obtain good g_r by trading one of her goods from $\{\chi_r^1, \chi_r^2, \chi_r^3\}$. The good used for this exchange by agent i must correspond to a true literal based on the assignment of U we will construct. Whenever agent i can acquire all the goods of $\{g_r\}_{r \in \mathcal{M}}$, we should be able to construct an assignment of U that is consistent with C.

We need to establish some constraints on the set of goods used by agent i to acquire the goods of $\{g_r\}_{r \in \mathcal{M}}$ to obtain consistent literals. No pair of them should correspond to a variable and its negation. To avoid that, for any literal c_r^l we introduce good k_r^l. Also for any $(r,l) \in \mathcal{I}$ such that $r \geq \mu$ and for any $(s,t) \in \mathcal{I}(\neg c_r^l)$ we create 3 goods $\alpha_{r,l}^{s,t}$, $\beta_{r,l}^{s,t}$, and $\gamma_{r,l}^{s,t}$. Figure 2 illustrates our construction. In this figure, (i) a solid-line denotes the most preferred good, (ii) a dash-line the second most preferred, (iii) a dots-line the third most preferred, and (iv) a dashed-dotted line the fourth most preferred. Let c_r^l be a positive literal and $(s,t) \in \mathcal{I}(\neg c_r^l)$. The preferences of the agents are set up so that any path from g_r to χ_r^l (resp. from g_s to χ_s^t), in the ATTC graph, contains the goods of $\{\alpha_{u,v}^{r,l}\}_{(u,v)\in\mathcal{I}(\neg c_r^l)}$ (resp. $\{\gamma_{s,t}^{v,w}\}_{(v,w)\in\mathcal{I}(\neg c_s^t)}$). The agent preferences are also set up to force the exchange between $\beta_{s,t}^{r,l}$ and $\gamma_{s,t}^{r,l}$ whenever $\delta(\beta_{s,t}^{r,l})$ does not acquire $\alpha_{s,t}^{r,l}$. Since $\beta_{s,t}^{r,l}$ never belongs to a path from g_r to χ_r^l in the ATTC graph, whenever χ_r^l is used by agent i to acquire the good g_r (so $\alpha_{s,t}^{r,l}$ is traded but not $\beta_{s,t}^{r,l}$), goods $\gamma_{s,t}^{r,l}$ and $\beta_{s,t}^{r,l}$ are exchanged by their owners and no path is possible anymore from g_s to χ_s^t in the ATTC graph. Consequently in such cases, agent i cannot use χ_s^t anymore to acquire g_s. So we preserved almost the consistency we were looking for but not completely since agent i may still exchange first the goods corresponding to negative literals. Below we explain how we make this situation impossible.

We also need to force agent i to use only one of the goods of $\{\chi_r^1, \chi_r^2, \chi_r^3\}$ to acquire the good g_r. For any $(r,l) \in \mathcal{I}$ we introduce good p_r^l, and for any $s \in \mathcal{M}$ we introduce goods $q_s^{1,2}$ and $q_s^{2,3}$. The preferences of the agents are such that if a good χ_r^l is used by agent i to obtain a good of $\{g_s\}_{s\in\mathcal{M}}$ then the exchange between $q_r^{t,t+1}$ and p_r^{t+1} is forced for any $t \geq l$. In that case, no path from any good of $\{g_s\}_{s\in\mathcal{M}}$ to such χ_r^t is available in the ATTC graph anymore. Furthermore the preferences are such that there is no path from a good of $\{g_s\}_{s\in\mathcal{M}}$ to good χ_r^l before all the goods χ_r^t with $t < l$ are traded by agent i to obtain a good which cannot belong to $\{g_s\}_{s\in\mathcal{M}}$. Therefore we fulfill the condition where no two goods of $\{\chi_r^1, \chi_r^2, \chi_r^3\}$ are used by agent i to obtain a good of $\{g_s\}_{s\in\mathcal{M}}$. The preferences are also set up so that good g_r, with $r < \mu$, can only be obtained by agent i by trading a good of $\{\chi_r^1, \chi_r^2, \chi_r^3\}$. Therefore agent i cannot use one of these three goods to obtain g_s, with $r \neq s$, if she wants to obtain the whole set of goods $\{g_u\}_{u\in\mathcal{M}}$. The goods of $\{\chi_r^1, \chi_r^2, \chi_r^3\}$, with $r \geq \mu$, can only be used by agent i to obtain g_r. So if agent i wants to obtain the whole set of goods $\{g_u\}_{u\in\mathcal{M}}$ then it is clear that, for any $r \in \mathcal{M}$, she has to use an object of $\{\chi_r^1, \chi_r^2, \chi_r^3\}$ to obtain g_r.

Finally, for any $(r,l) \in \mathcal{I}$, we introduce good π_r^l to force the goods of $\{g_s\}_{s\in\mathcal{M}}$ to be traded in increasing order of their indices during the ATTC procedure. By doing so we know that all the goods corresponding to the clauses with only positive literals are traded before the goods corresponding to the clauses with only negative literals start to be traded. This is also why we focused on the monotonic version of 3SAT for this proof. As a consequence we insure the consistency of the truth assignment we are constructing.

We now define the agent preferences on the different goods to obtain the above properties. For this proof we only need to define the linear order on the best goods, and an arbitrary order of the other goods will suffice. In this proof we describe a preference as a tuple representing the preference order on the best goods. For any $r > 1$, the preference of $\delta(g_r)$ is $(g_{r-1}, k_r^1, k_r^2, k_r^3, g_r)$. The preference of $\delta(g_1)$ is $(k_1^1, k_1^2, k_1^3, g_1)$. For any k_r^l, let $first(\neg c_r^l) = (s,t)$. If $r < \mu$ then the preference of $\delta(k_r^l)$ is $(\alpha_{s,t}^{r,l}, k_r^l)$. Otherwise the preference of $\delta(k_r^l)$ is $(\gamma_{r,l}^{s,t}, k_r^l)$. For any $\alpha_{r,l}^{s,t}$, if $next(c_r^l, r, l) = (u,v) \neq (0,0)$ then the preference of $\delta(\alpha_{r,l}^{s,t})$ is $(\alpha_{u,v}^{s,t}, \beta_{r,l}^{s,t}, \alpha_{r,l}^{s,t})$. Otherwise the preference of $\delta(\alpha_{r,l}^{s,t})$ is $(p_r^l, \beta_{r,l}^{s,t}, \alpha_{r,l}^{s,t})$. The preference of $\delta(\beta_{r,l}^{s,t})$ is $(\alpha_{r,l}^{s,t}, \gamma_{r,l}^{s,t}, \beta_{r,l}^{s,t})$. For any $\gamma_{r,l}^{s,t}$, if $next(c_s^t, s, t) = (u,v) \neq (0,0)$ then the preference of $\delta(\gamma_{r,l}^{s,t})$ is $(\beta_{r,l}^{s,t}, \gamma_{r,l}^{u,v}, \gamma_{r,l}^{s,t})$. Otherwise the preference of $\delta(\gamma_{r,l}^{s,t})$ is $(\beta_{r,l}^{s,t}, p_r^l, \gamma_{r,l}^{s,t})$. The preference of $\delta(p_r^1)$ is $(\pi_r^1, q_r^{1,2}, p_r^1)$, of $\delta(p_r^2)$ is $(q_r^{1,2}, \pi_r^2, q_r^{2,3}, p_r^2)$, and the preference of $\delta(p_r^3)$ is $(q_r^{2,3}, \pi_r^3, p_r^3)$. The preference of $\delta(q_r^{1,2})$ is $(p_r^1, p_r^2, q_r^{1,2})$, and the preference of $\delta(q_r^{2,3})$ is $(p_r^2, p_r^3, q_r^{2,3})$. For any $(r,l) \in \mathcal{I}$ with $r > 1$, the preference of $\delta(\pi_r^l)$ is $(g_{r-1}, \chi_r^l, \pi_r^l)$. For any $l \in \mathcal{L}$, the preference of $\delta(\pi_1^l)$ is (χ_1^l, π_1^l).

The last thing we need to establish is the true preferences of agent i. They need to be such that agent i acquires all the goods of $\{g_r\}_{r\in\mathcal{M}}$ for any beneficial misreport, but not if he reveals her true preferences. These true preferences are defined as $(g_1, \ldots, g_m, \chi_m^1, \chi_1^2, \chi_1^3, \chi_2^2, \chi_2^3, \ldots, \chi_m^2, \chi_m^3)$. By revealing them, we can verify that agent i obtains, at the end of the ATTC procedure, all her $3m+1$ best goods except for good g_m. The only way for her to improve this result is to acquire all the goods of $\{g_r\}_{r\in\mathcal{M}}$. We can easily show that (i) we can build in polynomial time this instance because there are only $O(m^2)$ good, and (ii) whenever agent i can beneficially manipulate the ATTC procedure, we can construct in polynomial time from this manipulation a truth assignment of the variables of U which is consistent with C. On the other hand when such manipulation is not possible, we can also easily show that no assignment of the variables of U can be consistent with C. This contradicts the assumption that P $\neq$ NP since the NP-hard problem ACM 3SAT is solvable in polynomial time. $\square$

Extended Model with Private Endowments

In this section, we consider the situation where each agent can use multiple accounts, and the set of her accounts, as well as her initial endowment, are private information. Each agent can cheat exchange rule's outcomes by pretending to be multiple agents under different accounts (splitting accounts). We assume that an agent can declare any preference on each account. Thus, splitting accounts is clearly more general than misreporting the preference of a single account. However, to our surprise, it turns out that the spaces of possible outcomes by misreporting and splitting coincide in the ATTC procedure.

Let us formally define splitting accounts. W.l.o.g., we can assume manipulator i uses $|e_i|$ accounts, each of which is

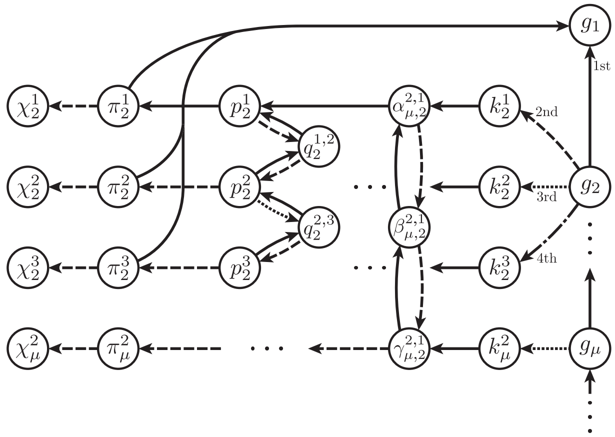

Figure 2: Transformation of instance of ACM 3SAT into instance of BENEFICIAL-MISREPORT

endowed with a single good. This is because in the ATTC procedure, an agent is divided into atomic agents. Thus, the outcome obtained by using less than $|e_i|$ accounts can be also obtained by using $|e_i|$ accounts. Therefore, for an agent $i \in N$, with initial endowment e_i, a splitting manipulation is described as a linear ordering profile $(\succ_k)_{k \in e_i} \in \mathscr{O}^{|e_i|}$. Here, each $\succ_k$ indicates the linear order reported under the account corresponding to the good k. Let $\mathcal{S}(e_i)$ be the set of all possible splitting manipulations for an initial endowment e_i. For an exchange problem $(\boldsymbol{e}, \succ)$, an agent $i \in N$, and a splitting manipulation $\boldsymbol{s}_i \in \mathcal{S}(e_i)$, let $\varphi_i(\boldsymbol{e}, (\boldsymbol{s}_i, \succ_{-i}))$ denote the bundle assigned to agent i when she uses the splitting manipulation $\boldsymbol{s}_i$.

We clarify the relationship between misreporting and splitting in the ATTC procedure. We show that for any splitting manipulation, there exists a preference misreport that returns the same assignment to the manipulator.

Proposition 1 (Splitting $\rightarrow$ Misreport)**.** *Given an exchange problem $(\boldsymbol{e}, \succ)$, a manipulator $i \in N$, and a splitting manipulation $\boldsymbol{s}_i \in \mathcal{S}(e_i)$, Algorithm 1 returns a misreport $\succ'_i \in \mathscr{O}$ such that $\varphi_i(\boldsymbol{e}, (\succ'_i, \succ_{-i})) = \varphi_i(\boldsymbol{e}, (\boldsymbol{s}_i, \succ_{-i}))$.*

Proposition 1 is proven by showing that Algorithm 1 is valid. The first loop of Algorithm 1 provides a splitting manipulation, where agent i obtains exactly the same set of goods, but such that no two goods of agents i appear in the same cycle during the ATTC procedure. This part is convenient to ease the proof. The second loop orders the best goods of the manipulation according to their appearance during the ATTC procedure. This second loop relies on the procedure $insert$, which inserts an element at the tail of a list, and the procedure $follow$, which provides the good following another one in a cycle. The two main ideas behind the algorithm are the following: (i) if a good o is traded during the ATTC procedure to obtain a good o' then putting o' in top of $\succ_{\delta(o)}$ does not change the outcome of the ATTC procedure because $\delta(o)$ owns only o, and (ii) if a good o' cannot be acquired by $\delta(o)$ (even by misreporting) then putting o in top of $\succ_{\delta(o)}$ does not change the outcome of the ATTC procedure. We omit the proof due to space limitations.

Algorithm 1 Splitting-to-Misreport

Input: $\boldsymbol{e} \in \mathcal{X}, \succ \in \mathscr{O}^n, i \in N, \boldsymbol{s}_i \in \mathcal{S}(e_i)$
Output: $\succ'_i \in \mathscr{O}$

while A cycle C containing more than one good of e_i appears during the ATTC procedure with parameters $(\boldsymbol{e}, (\boldsymbol{s}_i, \succ_{-i}))$. **do**
  Let o and o' be two distinct goods of e_i in C.
  $l \leftarrow follows(C, o), l' \leftarrow follows(C, o')$
  Let $\succ'_o$ and $\succ'_{o'}$ be any linear orders where the best goods are respectively l and l'.
  $\boldsymbol{s}_i \leftarrow (\succ'_o, \succ'_{o'}, (\succ_g)_{g \in e_i \setminus \{o,o'\}})$.
end while
Let L'_i be an empty list, and O an empty set of goods.
while $(K \setminus O) \cap e_i \neq \emptyset$ **do**
  Run the ATTC procedure to $(\boldsymbol{e}, (\boldsymbol{s}_i, \succ_{-i}))$ restricted to the goods of $K \setminus O$. Let C be the first cycle encountered which contains a good $\tilde{o}$ of $e_i \setminus O$.
  $L'_i \leftarrow insert(L'_i, follows(C, \tilde{o})), O \leftarrow O \cup C$.
end while
Let $\succ'_i$ be any linear order where the best $|L'_i|$ goods are ranked as in L'_i.

Proposition 1 implies that, while splitting accounts provides much richer manipulations than misreporting preferences (and actually any misreport can be obviously represented as a splitting), the spaces of the possible outcomes by splitting and misreport coincide.

We then consider the complexity of the problem of finding a beneficial splitting manipulation in the ATTC procedure: BENEFICIAL-SPLITTING. From the relationship between misreporting and splitting, and since Algorithm 1 runs in polynomial time, we have the following corollary:

Corollary 1. BENEFICIAL-SPLITTING *is NP-hard.*

Conclusions

In this paper we investigated exchange problems where each agent initially has a set of indivisible goods and a lexicographic preference. We proposed the augmented TTC procedure, which is core-selecting and runs in polynomial time. Concerning agents' incentives, we showed that finding a beneficial misreport in the augmented TTC procedure is NP-hard, while it is inevitably not strategy-proof due to Sönmez's finding. Most of our results are also valid in cases where the ownerships of endowments are private and unobservable so that each agent can use splitting manipulations.

Our future work will characterize the augmented TTC procedure. Since the core assignment is not always unique, there might be a different exchange rule that is also core-selecting and runs in polynomial time. For such a characterization, we must discover unique properties of the ATTC procedure.

Acknowledgments

This work was partially supported by JSPS KAKENHI Grant Number 24220003 and 26730005.

References

Afacan, M. 2014. Fictitious students creation incentives in school choice problems. *Economic Theory* 56(3):493–514.

Atlamaz, M., and Klaus, B. 2007. Manipulation via endowments in exchange markets with indivisible goods. *Social Choice and Welfare* 28(1):1–18.

Aziz, H.; Bachrach, Y.; Elkind, E.; and Paterson, M. 2011. False-name manipulations in weighted voting games. *Journal of Artificial Intelligence Research* 40(1):57–93.

Bartholdi, J.J., I.; Tovey, C.; and Trick, M. 1989. The computational difficulty of manipulating an election. *Social Choice and Welfare* 6(3):227–241.

Booth, R.; Chevaleyre, Y.; Lang, J.; Mengin, J.; and Sombattheera, C. 2010. Learning conditionally lexicographic preference relations. In *Proceedings of the 19th European Conference on Artificial Intelligence*, 269–274.

Chen, Y., and Sönmez, T. 2002. Improving efficiency of on-campus housing: An experimental study. *American Economic Review* 92(5):1669–1686.

Chevaleyre, Y.; Endriss, U.; and Maudet, N. 2007. Allocating goods on a graph to eliminate envy. In *Proceedings of the 22th International Joint Conference on Artificial Intelligence*, 700–705.

Conitzer, V., and Xia, L. 2012. Paradoxes of multiple elections: An approximation approach. In *Proceedings of the 13th International Conference on Principles of Knowledge Representation and Reasoning (KR-12)*.

Conitzer, V. 2008. Anonymity-proof voting rules. In *Proceedings of the 4th Workshop on Internet and Network Economics*, 295–306.

de Keijzer, B.; Bouveret, S.; Klos, T.; and Zhang, Y. 2009. On the complexity of efficiency and envy-freeness in fair division of indivisible goods with additive preferences. In *Algorithmic Decision Theory*. 98–110.

Endriss, U.; Maudet, N.; Sadri, F.; and Toni, F. 2006. Negotiating socially optimal allocations of resources. *Journal of Artificial Intelligence Research* 25:315–348.

Faliszewski, P., and Procaccia, A. D. 2010. Ai's war on manipulation: Are we winning? *AI Magazine* 31(4):53–64.

Faliszewski, P.; Hemaspaandra, E.; and Hemaspaandra, L. A. 2010. Using complexity to protect elections. *Communications of the ACM* 53(11):74–82.

Gigerenzer, G., and Goldstein, D. 1996. Reasoning the fast and frugal way: Models of bounded rationality. *Psychological review* 103(4):650.

Gold, E. M. 1978. Complexity of automaton identification from given data. *Information and Control* 37(3):302 – 320.

Ma, J. 1994. Strategy-proofness and the strict core in a market with indivisibilities. *International Journal of Game Theory* 23(1):75–83.

Moulin, H. 2008. Proportional scheduling, split-proofness, and merge-proofness. *Games and Economic Behavior* 63(2):567–587.

Ohta, N.; Conitzer, V.; Satoh, Y.; Iwasaki, A.; and Yokoo, M. 2008. Anonymity-proof shapley value: extending shapley value for coalitional games in open environments. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems - Volume 2*, 927–934.

Pápai, S. 2003. Strategyproof exchange of indivisible goods. *Journal of Mathematical Economics* 39(8):931 – 959.

Pápai, S. 2007. Exchange in a general market with indivisible goods. *Journal of Economic Theory* 132(1):208 – 235.

Pini, M. S.; Rossi, F.; Venable, K. B.; and Walsh, T. 2011. Manipulation complexity and gender neutrality in stable marriage procedures. *Autonomous Agents and Multi-Agent Systems* 22(1):183–199.

Roth, A. E., and Postlewaite, A. 1977. Weak versus strong domination in a market with indivisible goods. *Journal of Mathematical Economics* 4(2):131–137.

Roth, A. E.; Sönmez, T.; and Ünver, M. U. 2004. Kidney exchange. *The Quarterly Journal of Economics* 119(2):457–488.

Shapley, L., and Scarf, H. 1974. On cores and indivisibility. *Journal of Mathematical Economics* 1(1):23–37.

Sönmez, T. 1999. Strategy-proofness and essentially single-valued cores. *Econometrica* 67(3):677–689.

Sonoda, A.; Fujita, E.; Todo, T.; and Yokoo, M. 2014. Two case studies for trading multiple indivisible goods with indifferences. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence*, 791–797.

Teo, C.-P.; Sethuraman, J.; and Tan, W.-P. 2001. Gale-shapley stable marriage problem revisited: Strategic issues and applications. *Management Science* 47(9):1252–1267.

Todo, T., and Conitzer, V. 2013. False-name-proof matching. In *Proceedings of the 12th International Joint Conference on Autonomous Agents and Multiagent Systems*, 311–318.

Todo, T.; Iwasaki, A.; and Yokoo, M. 2011. False-name-proof mechanism design without money. In *Proceedings of the 10th International Joint Conference on Autonomous Agents and Multiagent Systems*, 651–658.

Todo, T.; Sun, H.; and Yokoo, M. 2014. Strategyproof exchange with multiple private endowments. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence*, 805–811.

Yaman, F.; Walsh, T. J.; Littman, M. L.; and desJardins, M. 2011. Democratic approximation of lexicographic preference models. *Artificial Intelligence* 175(7 - 8):1290 – 1307.

Yokoo, M.; Conitzer, V.; Sandholm, T.; Ohta, N.; and Iwasaki, A. 2005. Coalitional games in open anonymous environments. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence*, 1668–1669.

Yokoo, M.; Sakurai, Y.; and Matsubara, S. 2004. The effect of false-name bids in combinatorial auctions: New fraud in internet auctions. *Games and Economic Behavior* 46(1):174–188.