False-name-proof Matching

Taiki Todo
Department of Computer Science
Duke University
Durham, NC 27705, USA
todo@cs.duke.edu

Vincent Conitzer
Department of Computer Science
Duke University
Durham, NC 27705, USA
conitzer@cs.duke.edu

ABSTRACT

Matching a set of agents to a set of objects has many real applications. One well-studied framework is that of priority-based matching, in which each object is assumed to have a priority order over the agents. The Deferred Acceptance (DA) and Top-Trading-Cycle (TTC) mechanisms are the best-known strategy-proof mechanisms. However, in highly anonymous environments, the set of agents is not known *a priori*, and it is more natural for objects to instead have priorities over *characteristics* (e.g., the student's GPA or home address). In this paper, we extend the model so that each agent reports not only its preferences over objects, but also its characteristic. We derive results for various notions of strategy-proofness and false-name-proofness, corresponding to whether agents can only report weaker characteristics or also incomparable or stronger ones, and whether agents can only claim objects allocated to their true accounts or also those allocated to their fake accounts. Among other results, we show that DA and TTC satisfy a weak version of false-name-proofness. Furthermore, DA also satisfies a strong version of false-name-proofness, while TTC fails to satisfy it without an acyclicity assumption on priorities.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Multi-agent systems*; J.4 [**Social and Behavioral Sciences**]: Economics

General Terms

Algorithms, Economics, Theory

Keywords

Efficiency, fairness, matching

1. INTRODUCTION

1.1 Background

Matching a set of agents to a set of objects has many real applications, such as school choice, hospital-resident matching, and kidney exchange. One well-studied framework is that of priority-based matching, in which each object is assumed to have a priority order over the agents, while each agent also has a preference over the objects.

In this paper, we focus on matching problems in which the objects are not strategic agents, such as *student placement* [3] or *house allocation* [6]. For such problems, a mechanism is a function that maps a profile of reported preferences to an assignment of objects to agents, based on the priorities of objects. We focus on mechanisms that are strategy-proof (for the agents). The (agent-proposing) Deferred Acceptance (DA) and Top-Trading-Cycle (TTC) mechanisms are the best-known strategy-proof mechanisms for these problems (and under some assumptions, there are no others).

In this paper, we focus on matching environments with a high degree of anonymity. While our analysis may be of interest even in traditional settings (e.g., traditional school choice), clearer motivating examples come from Internet environments (for example, selective Massive Open Online Courses, or even M.O.O. Schools, to which students must apply[1]). In such environments, the set of agents is not known to the objects *a priori*. Instead, each object just has information about the set of possible *characteristics* that an agent could have, e.g., the student's GPA or home address. In such environments, it is natural to assume that the objects have priorities over these characteristics rather than over the agents themselves. Then, each agent is required to report its characteristic, as well as its preference over the objects, to the mechanism. Even in a mechanism that satisfies strategy-proofness in the traditional sense (misreporting of preferences), an agent may misreport its characteristic in a way that results in a better allocation. Intuitively, if an agent reports a characteristic that is ranked higher than its original one in the priority of each object, then it seems likely that its assignment will become weakly better.

Furthermore, in such highly anonymous environments, an agent may also have an incentive to use fake accounts (i.e., *false-name* accounts) to manipulate the outcome of the mechanism. Even if each agent does not want more than one object—e.g., a student could not enter more than one school—it is not obvious whether strategy-proof mechanisms necessarily satisfy false-name-proofness. For example, a student may represent him/herself differently to one school than to another school, by using fake accounts, based on what these schools consider important. Furthermore, people applying for political asylum under multiple identities is a recognized problem (here the countries take the roles of schools).

[1]One may question in what sense these would still be "Open," but we hope the reader understands our intent.

Appears in: *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2013), Ito, Jonker, Gini, and Shehory (eds.), May, 6–10, 2013, Saint Paul, Minnesota, USA.*

1.2 Our Contributions

First, we formalize the environment with arbitrarily misreportable characteristics. We provide a necessary and sufficient condition for the existence of strategy-proof matching mechanisms that satisfy the *mutual-best* condition (if an agent and an object prefer each other most, they should be matched). This required condition on the relationship between possible misreports and objects' priorities is rather stringent and suggests that, without restricting misreports on characteristics, almost all mechanisms are vulnerable to strategic behavior by the agents.

Therefore, we introduce a natural restriction on agents' misreports based on the idea of *partial verification* [9]. We then define two natural notions of *false-name-proofness* in this domain. We show that the two most prominent strategy-proof mechanisms in the literature, DA and TTC, are both false-name-proof in the weaker sense. Furthermore, DA also satisfies the stronger version of false-name-proofness, while TTC fails to satisfy it without an acyclicity assumption on priorities.

Our matching model with partially-verifiable characteristics can be applied to numerous situations. A natural and important example is that of school choice problems with several independent scores, such as GPA, TOEFL, etc. These generate a lattice structure on the characteristics that allows us to design mechanisms with desirable properties. It is also quite natural for each object (school or college) to have an order that is compatible with the lattice structure—for example, perhaps the school ranks the applicants according to a specific conical combination of their scores. The standard hospital-resident matching problem is another example. Prospective residents would benefit from overreporting their characteristics, such as their United States Medical Licencing Examination scores, if no verification schemes are implemented.

1.3 Related Work

Matching. The DA mechanism was introduced by Gale and Shapley [8], as a strategy-proof (on the agents' side) and fair matching mechanism. Crawford [5] showed that in the agent-proposing DA, an arrival of a new agent never makes other agents better off. Balinski and Sönmez [3] discussed the effect of a change of priorities in DA. They showed that DA respects improvement in priority, i.e., no agent can become worse off as the result of an improvement in that agent's priority for each object. Ergin [7] gives a necessary and sufficient condition on priorities for DA to be Pareto efficient. There is also a significant amount of research on the TTC mechanism [14]. Kesten [11] showed a necessary and sufficient condition on priorities for TTC and DA to coincide. The proposed condition, called acyclicity, seems quite similar to the one proposed in Ergin [7]. Morrill [12] defined the weak mutual-best condition mentioned above and gave an alternative characterization of the TTC mechanism. Finally, Ehlers et al. [6] give a characterization of population monotonic [15] matching mechanisms under the two additional axioms of strategy-proofness and Pareto efficiency.

False-name-proofness. Most research on false-name-proof mechanism design has focused on combinatorial auctions. Yokoo et al. [17] showed that the well-known VCG mechanism is not false-name-proof in general combinatorial auctions, and Iwasaki et al. [10] gave a low upper bound on the worst-case efficiency that can be obtained using a false-name-proof mechanism satisfying some additional assumptions. Conitzer [4] gave a very negative characterization of false-name-proof voting rules. While these results have shown the difficulty of designing false-name-proof mechanisms for resource allocation with transfers (at least in sufficiently general settings) and public decision making without transfers, respectively, to our knowledge there has been no earlier work on designing false-name-proof mechanisms for resource allocation without transfers, including matching problems. Finally, as discussed by Todo et al. [16], population monotonicity seems to have a strong connection to false-name-proofness even in resource allocation problems.

2. PRELIMINARIES

2.1 Matching Model with Characteristics

In this section, we first define our model of matching with characteristics. Let $\mathcal{N}$ be a set of potential agents (identities) and $N \subseteq \mathcal{N}$ be a set of *attending* agents. Let $n := |N|$ denote the number of attending agents and let each index $i, j, k \in \mathcal{N}$ denote an agent. Let $X = \{x, y, z, \ldots\}$ be a set of m objects (items) to be allocated to attending agents. That is, $m := |X|$. Then, let $a \in \mathcal{A}^n$ be an allocation (assignment) of the set of objects X to a set of attending agents N, where $\mathcal{A}^n$ denotes the set of possible allocations to N. For a given allocation $a \in \mathcal{A}^n$, let $a_i \in X \cup \varnothing$ denote the assignment to agent $i \in N$. Here $a_i = \varnothing$ means that i is not allocated any object (equivalently, is allocated the *null object*) under allocation a. That is, for each agent i, an assignment to i is either a single object or an empty set. We assume that for every allocation $a \in \mathcal{A}^n$, the feasibility condition is satisfied: $\bigcup_{i \in N} a_i \subseteq X$ and $a_i \cap a_j = \varnothing$ for all i and j with $i \neq j$.

We next introduce notation for each agent $i \in N$. Let $\theta_i \in \Theta$ denote the type of agent i. Specifically, a type $\theta_i = (R_i, c_i)$ consists of a *preference relation* R_i and a *characteristic* $c_i \in \mathcal{C}$.[2] A preference relation $R_i \in \mathcal{R}$ is a complete order of $X \cup \varnothing$, where $\mathcal{R}$ indicates the set of all possible preference relations. Let P_i denote the strict part of R_i. Because we assume preferences to be strict, given $x, y \in X \cup \varnothing$, xR_iy and $\neg xP_iy$ implies $x = y$.

A characteristic $c_i \in \mathcal{C}$ reflects the features of agent i, where $\mathcal{C}$ indicates a set of all possible characteristics. Then, $\Theta = \mathcal{R} \times \mathcal{C}$ defines the domain of possible types.

Next, we consider a misreport θ'_i by an agent i with true type $\theta_i = (R_i, c_i)$. We assume that each agent can report any preference R'_i in $\mathcal{R}$, regardless of its true type θ_i. On the other hand, for misreports of characteristics, we introduce a directed graph model. Let $\succeq^{\mathcal{C}}$ denote a binary relation over $\mathcal{C}$, where $\mathcal{G} \ni \succeq^{\mathcal{C}}$ is the set of all possible such relations satisfying the following restrictions. We assume this relation $\succeq^{\mathcal{C}}$ is *transitive*; for any $d, e, f \in \mathcal{C}$, $d \succeq^{\mathcal{C}} e$ and $e \succeq^{\mathcal{C}} f$ implies $d \succeq^{\mathcal{C}} f$. Also, we assume the relation $\succeq^{\mathcal{C}}$ is *reflexive*; for any $d \in \mathcal{C}$, $d \succeq^{\mathcal{C}} d$ holds. Thus, any $\succeq^{\mathcal{C}} \in \mathcal{G}$ corresponds to a directed graph, whose vertices are the set of all possible characteristics $\mathcal{C}$, which contains self-loops for every characteristic, and which is transitive. $\mathcal{G}$ corresponds to the set of all possible such directed graphs.

We now introduce two correspondences $L, M : \Theta \to 2^{\Theta}$ that map an agent's true type to the set of types it can mis-

[2]Note that each agent has only a single characteristic, but this characteristic can be a vector including all of (say) GPA, TOEFL score, etc., so that this is without loss of generality.

report, based on a given directed graph $\succeq^{\mathcal{C}}$. The *misreport correspondence* M allows agents to misreport not only their preferences, but also their characteristics in a way that is consistent with $\succeq^{\mathcal{C}}$, i.e., for all $\theta_i = (R_i, c_i) \in \Theta$, $M(\theta_i) \subseteq \Theta$ is the set of all types $\theta'_i = (R'_i, c'_i) \in \Theta$ satisfying $c_i \succeq^{\mathcal{C}} c'_i$ (while R'_i is arbitrary). When $\succeq^{\mathcal{G}}$ contains all ordered pairs from $\mathcal{C}$, all possible types are allowed by M for each type θ_i. By the reflexivity, for any $\succeq^{\mathcal{C}} \in \mathcal{G}$ and any $\theta_i \in \mathcal{C}$, $\theta_i \in M(\theta_i)$, i.e., truth-telling is always possible. The non-existence of a directed edge from characteristic e to characteristic d can be interpreted as the existence of a partial verification scheme that could detect a misreport of characteristic from e to d. If we could detect any misreport of characteristics, then $\succeq^{\mathcal{C}}$ contains no edges, and the problem coincides with traditional matching without characteristics.

On the other hand, the *limited misreport correspondence* L never allows agents to misreport their characteristics, regardless of $\succeq^{\mathcal{C}}$; i.e., for all $\theta_i = (R_i, c_i) \in \Theta$, $L(\theta_i) \subseteq \Theta$ is the set of all types $\theta'_i = (R'_i, c_i) \in \Theta$ (where R'_i is arbitrary). That is, L represents the misreports that are allowed in the traditional matching model without characteristics.

Next, we consider objects' priorities over characteristics. For any object $x \in X$, let $\pi_x \in \Pi$ be a complete and strict order of all possible characteristics $\mathcal{C}$. For two characteristics $d, e \in \mathcal{C}$, $d\pi_x e$ indicates that object x with order π_x *orders* characteristic d prior to characteristic e. Let $\pi \in \Pi^m$ be a profile of orders, where Π^m is the set of all possible profiles of orders. Let $\succ_x \in \mathcal{P}^N$ be the *priority* that object x has over the set of *attending* agents N, according to its own order π_x and the set of (reported) characteristics of attending agents $c = (c_1, \ldots, c_n)$. That is, the priority $\succ_x$ is a function of $\pi_x \in \Pi$ and $c \in \mathcal{C}^N$. Specifically, $\succ_x$ is derived from π_x and c as follows: $\forall i, j \in N$, $i \succ_x j$ if and only if $c_i \pi_x c_j$. In this case, we say agent i with characteristic c_i has a higher priority for object x than agent j with characteristic c_j. To avoid having to deal with tie-breaking, we assume that in any individual run of the mechanism, no two agents can end up reporting the same characteristic.[3]

2.2 Matching Mechanism

In this section, we formally define mechanisms for the model of matching with characteristics. Since the number of participating identities is not known ex ante, a *mechanism* must be defined for any number of attending identities:

Definition 2.1 (Matching Mechanism). *A* matching mechanism φ *is a sequence of functions* $\{\varphi^n\}_{n=0,\ldots,|\mathcal{N}|}$ *such that* $\varphi^n : \Theta^n \times \Pi^m \to \mathcal{A}^n$.

We require a mechanism to be defined for any possible characteristics graph $\succ^{\mathcal{C}} \in \mathcal{G}$. This assumption takes an important role in some of our results. For example, if we can restrict our attention to a restricted family $\mathcal{G}$ that contains only one $\succeq^{\mathcal{C}}$ that has no arrow between any two characteristics, then misreporting characteristics is impossible, and the problem coincides with the traditional formulation of matching (without characteristics). More generally, restricting $\mathcal{G}$ may allow us to design a larger variety of false-name-proof mechanisms, which could be an interesting future direction.

For a type profile θ and a profile of orders π, the mechanism determines an outcome $\varphi(\theta, \pi)$. Let $\varphi_i(\theta, \pi)$ denote the assignment to agent i. Given that we are motivated by highly anonymous settings, we restrict our attention to *anonymous* mechanisms, in which, if two agents swap their types, their assignments must be also swapped.

[3]We believe the assumption is quite natural in real environments. For example, part of the characteristic could be the time at which the agent submitted its application.

2.3 Properties of Mechanisms

First, let us define a traditional notion of strategy-proofness, which considers only misreporting of preferences, not characteristics.

Definition 2.2. *A mechanism φ is said to be* strategy-proof in the traditional sense *(or to satisfy SP-) if* $\forall \succeq^{\mathcal{C}} \in \mathcal{G}$, $\forall \pi \in \Pi^m$, $\forall N \subseteq \mathcal{N}$, $\forall i \in N$, $\forall \theta_{-i} \in \Theta^{n-1}$, $\forall \theta_i \in \Theta$, *and* $\forall \theta'_i \in L(\theta_i)$, *it holds that* $\varphi_i((\theta_i, \theta_{-i}), \pi) R_i \varphi_i((\theta'_i, \theta_{-i}), \pi)$.

However, in our model an agent can misreport both its preference and its characteristic, demanding a stronger notion of strategy-proofness:

Definition 2.3 (Strategy-proofness). *A mechanism φ is said to be* strategy-proof *(or to satisfy SP) if* $\forall \succeq^{\mathcal{C}} \in \mathcal{G}$, $\forall \pi = \Pi^m$, $\forall N \subseteq \mathcal{N}$, $\forall i \in N$, $\forall \theta_{-i} \in \Theta^{n-1}$, $\forall \theta_i \in \Theta$, *and* $\forall \theta'_i \in M(\theta_i)$, *it holds that* $\varphi_i((\theta_i, \theta_{-i}), \pi) R_i \varphi_i((\theta'_i, \theta_{-i}), \pi)$.

We now consider two important and well-known criteria in matching problems, *fairness* and *efficiency*. Fairness (also called stability) requires that there exists no *blocking pair*, i.e., there exists no pair of an agent and an object that both prefer each other to the current assignment.

Definition 2.4 (Fairness). *A mechanism φ is said to satisfy* fairness *(FA) if* $\forall \succeq^{\mathcal{C}} \in \mathcal{G}$, $\forall \pi = \Pi^m$, $\forall N \subseteq \mathcal{N}$, $\forall i \in N$, *and* $\forall \theta \in \Theta^n$, *there are no* $i, j \in N$ *and* $x \in X$ *such that* $x P_i \varphi_i(\theta, \pi)$, $c_i \pi_x c_j$, *and* $\varphi_j(\theta, \pi) = x$.

Pareto efficiency concerns the welfare of the society. It requires that there exists no outcome that Pareto-dominates the current outcome.

Definition 2.5 (Pareto Efficiency). *A mechanism φ is said to satisfy* Pareto efficiency *(PE) if* $\forall \succeq^{\mathcal{C}} \in \mathcal{G}$, $\forall \pi = \Pi^m$, $\forall N \subseteq \mathcal{N}$, *and* $\forall \theta \in \Theta^n$, *there exists no allocation* $a \in \mathcal{A}^n$ *such that* $\forall i \in N$, *we have* $a_i R_i \varphi_i(\theta, \pi)$, *and* $\exists j \in N$ *such that* $a_j P_j \varphi_j(\theta, \pi)$.

The following proposition shows that for a tiny economy, there exists a mechanism that satisfies SP, PE, and FA simultaneously. In fact, the TTC mechanism is one such mechanism. The proof is omitted, but the result can be easily verified from Theorem 6.

Proposition 2.1. *There always exists a matching mechanism that satisfies SP, PE, and FA if (i)* $|N| = 1$ *or (ii)* $|N| = 2$ *and* $|\mathcal{C}| = 2$.

Thus, in the rest of the paper, we focus on cases where $|N| \geq 2$ and $|\mathcal{C}| \geq 3$.

The following weak *mutual best* axiom has been proposed by, e.g., Morrill [12], as a minimal requirement for simultaneously satisfying fairness and efficiency. It requires that any pair of an agent and an object that are each other's first choices must be matched.

Definition 2.6 (Mutual Best). *A mechanism φ is said to satisfy* mutual best *(MB) if* $\forall \succeq^{\mathcal{C}} \in \mathcal{G}$, $\forall \pi = \Pi^m$, $\forall N \subseteq \mathcal{N}$, *and* $\forall \theta \in \Theta^n$, *it holds that, for all* $i \in N$ *and* $x \in X$,

$$[(\forall j \neq i, c_i \pi_x c_j) \wedge (\forall y \neq x, x P_i y)] \Rightarrow [\varphi_i(\theta, \pi) = x].$$

MB is a natural, and quite weak, requirement. It is easy to check that a mechanism that satisfies both FA and PE must satisfy MB.

3. STRATEGY-PROOFNESS

In this section, we study the design of strategy-proof mechanisms in our model.

3.1 Impossibility of Strategy-proof Matching

We first show a rather negative result: if there exists at least one *cyclic conflict* between a characteristics graph $\succeq^{\mathcal{C}}$ and an order π_x held by an object $x \in X$, then we cannot design SP mechanisms satisfying the minor condition MB. More precisely, there exists no mechanism that satisfies SP and MB if $\exists \succeq^{\mathcal{C}}$, $\exists x \in X$, and $\exists d, e, f \in \mathcal{C}$,

$$[d\pi_x e\pi_x f] \wedge [f \succeq^{\mathcal{C}} d]. \qquad \text{(CC)}$$

Interestingly, the opposite is also true. If there is no cyclical conflict for any $\succeq^{\mathcal{C}} \in \mathcal{G}$ and any $\pi \in \Pi^m$, there is at least one mechanism that satisfies SP and MB. Thus, the non-existence of cyclical conflicts is necessary and sufficient for the existence of mechanisms that satisfy both SP and MB.

THEOREM 1. *There exists a mechanism that satisfies both SP and MB if and only if there exists no cyclical conflict between any graph $\succeq^{\mathcal{C}} \in \mathcal{G}$ and any profile of orders $\pi \in \Pi^m$.*

PROOF (ONLY IF PART). Consider a characteristics graph $\succeq^{\mathcal{C}}$ that has a cyclical conflict with an order π_x held by an object x, i.e, $d\pi_x e\pi_x f$ and $f \succeq^{\mathcal{C}} d$. Then, consider the case where there are only two agents i, j, whose true characteristics are e and f, respectively. Moreover, assume they both have the same preference $R_i = R_j$, which ranks x first among all objects, i.e., xR_iy for all $y \neq x$. If the mechanism is to satisfy the mutual best condition, x must be allocated to agent i with characteristic e.

Next, consider the case where their types are given as $\theta_i = (R_i, e)$ and $\theta_j = (R_j, d)$, respectively. Now, to satisfy MB, x must be allocated to j with characteristic d. Now, because $f \succeq^{\mathcal{C}} d$ holds, $(R_j, d) \in M((R_j, f))$, i.e., agent j with characteristic f can pretend to have characteristic d. Therefore, agent j has an incentive to misreport its characteristic, from f to d, and get the object x. This violates SP. □

PROOF (IF PART). Consider the following simple object-proposing mechanism: each object proposes to the agent whose (reported) characteristic has the highest priority for it among those of attending agents, and each agent receives the object it most prefers among all objects proposing to it and the null object. SP- and MB are obviously satisfied.

Now let us assume, for the sake of contradiction, that when i changes its characteristic from c_i to c'_i, its allocation changes from x to some $y \neq x$ such that yP_ix. Since it can misreport c'_i, $c_i \succeq^{\mathcal{C}} c'_i$ holds in the characteristics graph $\succeq^{\mathcal{C}}$. Now, let $j \neq i$ be the agent whose characteristic c_j has the highest priority in π_y among those of all attending agents, i.e., y proposes to j (when i does not misreport). Since y proposes to j rather than to i, π_y satisfies $c_j\pi_y c_i$. Also, since i gets y when it reports c'_i instead, $c'_i\pi_y c_j$ holds. Thus, we have $c'_i\pi_y c_j\pi_y c_i$ and $c_i \succeq^{\mathcal{C}} c'_i$, which violates the assumption that there exists no cyclical conflict. □

We quickly note the independence of the two axioms. The No-Allocation mechanism satisfies SP, but generally does not satisfy MB. The DA (or TTC) mechanism satisfies MB (and SP-), but generally does not satisfy SP, because an agent can benefit from misreporting its characteristic to end up with a higher priority for an object.

3.2 Restricting the Characteristics Structure

We now introduce a natural restriction on $\succeq^{\mathcal{C}}$ and π to prevent such cyclical conflicts. In short, we focus on characteristics graphs $\succeq^{\mathcal{C}}$ that contain no cycles (other than self-loops) and profiles of orders $\pi = (\pi_x)_{x\in X}$ that are compatible with the characteristic graph.

To be more precise, let $\mathcal{D} \subset \mathcal{G}$ be the set of all characteristics graphs that have no cycles, other than the self-loops needed for reflexivity. (We slightly abuse terminology and refer to such graphs as DAGs.) For a given DAG $\succeq^{\mathcal{C}} \in \mathcal{D}$, we say characteristic $e \in \mathcal{C}$ is *weaker* than characteristic $d \in \mathcal{C}$ if it holds that $d \succeq^{\mathcal{C}} e$.

Next, let us discuss the assumption on the structure of the orders that objects have over characteristics. We assume that for the given DAG $\succeq^{\mathcal{C}} \in \mathcal{D}$, for all $x \in X$, order π_x is *compatible* with $\succeq^{\mathcal{C}}$, i.e.,

$$\forall d, e \in \mathcal{C}, [d \succeq^{\mathcal{C}} e] \Rightarrow [d\pi_x e]$$

Let $\Pi_{\mathrm{C}} \subset \Pi$ denote the set of all orders that are compatible with a given DAG $\succeq^{\mathcal{C}} \in \mathcal{D}$, and let $\pi \in \Pi_{\mathrm{C}}^m$ denote a profile of such compatible orders (one for every object). In what follows, we refer to environments where the characteristics graph is a DAG and the objects' priorities are compatible with this DAG as *directed environments*.

Our restriction can be motivated, for example, by a setting where an agent's characteristic corresponds to that agent's test scores, an agent can choose to score lower on any of these tests but not higher, and objects (e.g., schools) prefer, all other things being equal, agents with higher test scores.

Next, we define a monotonicity property, called *respecting improvement in priority*, which requires that for each agent, its assignment never strictly improves when its priority for each object gets (weakly) worse.

DEFINITION 3.1. *A mechanism φ is said to satisfy* respecting-improvement-in-priority *(RIP) if $\forall \succeq^{\mathcal{C}} \in \mathcal{D}$, $\forall \pi \in \Pi_C^m$, $\forall N \subseteq \mathcal{N}$, $\forall i \in N$, $\forall \theta = (R, c) \in \Theta^n$, and $\forall \pi' \in \Pi_C^m$ such that $\forall x \in X$ and $\forall j \neq i$, (i) $[c_j\pi_x c_i] \Rightarrow [c_j\pi'_x c_i]$ and (ii) $[c_j\pi_x c_k] \Leftrightarrow [c_j\pi'_x c_k]$ for any other $k \in N \setminus \{i, j\}$, it holds that $\varphi_i(\theta, \pi)R_i\varphi_i(\theta, \pi')$.*

For our model of directed environments, we define a similar monotonicity condition below, which we refer to as *respecting improvement in characteristic* (RIC).

DEFINITION 3.2. *A mechanism φ is said to satisfy* respecting-improvement-in-characteristic *(RIC) if $\forall \succeq^{\mathcal{C}} \in \mathcal{D}$, $\forall \pi = \Pi_C^m$, $\forall N \subseteq \mathcal{N}$, $\forall i \in N$, $\forall \theta_{-i} \in \Theta^{n-1}$, $\forall \theta_i = (R_i, c_i) \in \Theta$, and $\forall \theta'_i = (R_i, c'_i)$ such that $c_i \succeq^{\mathcal{C}} c'_i$, it holds that*

$$\varphi_i((\theta_i, \theta_{-i}), \pi)R_i\varphi_i((\theta'_i, \theta_{-i}), \pi)$$

In words, RIC requires that when an agent's characteristic gets weaker, its assignment never gets better. This can also be considered an incentive requirement, because if it fails to hold, an agent can benefit from reporting a weaker characteristic. RIP implies RIC (the proof is omitted due to space limitations).

Now we show that achieving SP is equivalent to achieving both SP- and RIC simultaneously. Since objects are not strategic agents in our model, we omit π in the description of mechanism φ's outcome, i.e., $\varphi_i(\theta) = \varphi_i(\theta, \pi)$, when there is no risk of confusion.

Proposition 3.1. *In directed environments, a mechanism φ satisfies SP if and only if it satisfies SP- and RIC.*

Proof. SP clearly implies SP- and RIC, which correspond to misreporting only one's preferences, and only one's characteristic, respectively. Next, we show the opposite direction, that the combination of SP- and RIC implies SP. If φ satisfies RIC, $\varphi_i(\theta_i, \theta_{-i}) R_i \varphi_i(\theta'_i, \theta_{-i})$ holds for any $\theta'_i \in M(\theta_i)$ s.t., $R'_i = R_i$. Furthermore, SP- guarantees that $\varphi_i(\theta'_i, \theta_{-i}) R_i \varphi_i(\theta''_i, \theta_{-i})$ for any $\theta''_i \in L(\theta'_i)$. Moreover, for any θ_i and $\theta''_i \in M(\theta_i)$, we can find a θ'_i such that $\theta'_i \in M(\theta_i)$, $R'_i = R_i$, and $\theta''_i \in L(\theta'_i)$. Hence, we obtain $\varphi_i(\theta_i, \theta_{-i}) R_i \varphi_i(\theta'_i, \theta_{-i}) R_i \varphi_i(\theta''_i, \theta_{-i})$, establishing SP. □

4. FALSE-NAME-PROOFNESS

We now move on to study false-name-proofness.

4.1 Two Varieties of False-name-proofness

In this section, we define two notions of false-name-proofness. Roughly speaking, a mechanism is false-name-proof if no agent ever has an incentive to use fake accounts, but the exact definition depends on what the agent is able to do with such fake accounts.

In the first notion, called *strong false-name-proofness* (S-FNP), we assume that an agent can create fake accounts with any characteristics. That is, in the definition of S-FNP, an agent i with true type θ_i can create fake identities with any possible type, including ones outside $M(\theta_i)$. However, the agent will only be able to claim the object allocated to its true identity. For example, an agent can apply to a school under a fake account, but if the agent attempts to actually enroll in the school under this fake account, the manipulation will be immediately detected. Hence, our previous negative result for strategy-proofness with unrestricted misreporting does not apply. Nevertheless, this type of manipulation can be helpful to an agent if the additional fake accounts result in the agent's true identity getting a better allocation.

Definition 4.1 (Strong False-name-proofness). *A mechanism φ is said to satisfy* strong false-name-proofness *(S-FNP) if $\forall \succeq^{\mathcal{C}} \in \mathcal{D}$, $\forall \pi = \Pi^m_C$, $\forall N \subset \mathcal{N}$, $\forall i \in N$, $\forall \theta_{-i} \in \Theta^{n-1}$, $\forall \theta_i \in \Theta$, $\forall \theta'_i \in M(\theta_i)$, $\forall j \in \mathcal{N} \setminus N$, and $\forall \theta_j \in \Theta$, it holds that $\varphi_i(\theta_i, \theta_{-i}) R_i \varphi_i(\theta'_i, \theta_{-i}, \theta_j)$.*

In the next notion, simply called *false-name-proofness*, we assume that an agent can create fake accounts, but only report characteristics for these accounts that it could also have reported for its true identity. That is, in the definition of FNP, an agent i with true type θ_i can create fake identities but only with (reported) types in $M(\theta_i)$. On the other hand, we do now allow such a manipulating agent to claim the object allocated to one of its fake accounts rather than the one allocated to its true account. (We assume unit-demand agents, i.e., an agent cannot claim more than one object—e.g., the agent can enroll at at most one school.)

Definition 4.2 (False-name-proofness). *A mechanism φ is said to satisfy* false-name-proofness *(FNP) if $\forall \succeq^{\mathcal{C}} \in \mathcal{D}$, $\forall \pi = \Pi^m_C$, $\forall N \subset \mathcal{N}$, $\forall i \in N$, $\forall \theta_{-i} \in \Theta^{n-1}$, $\forall \theta_i \in \Theta$, $\forall \theta'_i \in M(\theta_i)$, $\forall j \in \mathcal{N} \setminus N$, and $\forall \theta_j \in M(\theta_i)$, it holds that*

$$[\varphi_i(\theta_i, \theta_{-i}) R_i \varphi_i(\theta'_i, \theta_{-i}, \theta_j)] \wedge [\varphi_i(\theta_i, \theta_{-i}) R_i \varphi_j(\theta'_i, \theta_{-i}, \theta_j)].$$

In these definitions, we considered only false-name manipulations in which only a single fake identity was used. It is not difficult to see that this is without loss of generality.[4]

It may not be immediately clear that S-FNP is actually stronger than FNP, because only under the latter notion is a manipulating agent free to claim an object allocated to a fake identity. Nevertheless, the following proposition establishes that S-FNP is in fact stronger than FNP.

Proposition 4.1. *In directed environments, a mechanism φ satisfies FNP if it satisfies S-FNP. However, the opposite is not true.*

Proof. Assume, for the sake of contradiction, that we have an instance where a strongly false-name-proof mechanism φ does not satisfy false-name-proofness: $\exists \succeq^{\mathcal{C}} \in \mathcal{D}$, $\exists \pi = \Pi^m_{\mathrm{C}}$, $\exists N \subset \mathcal{N}$, $\exists i \in N$, $\exists \theta_{-i} \in \Theta^{n-1}$, $\exists \theta_i \in \Theta$, $\exists \theta'_i \in M(\theta_i)$, $\exists j \in \mathcal{N} \setminus N$, and $\exists \theta_j \in M(\theta_i)$,

$$[\varphi_i(\theta'_i, \theta_{-i}, \theta_j) P_i \varphi_i(\theta_i, \theta_{-i})] \vee [\varphi_j(\theta'_i, \theta_{-i}, \theta_j) P_i \varphi_i(\theta_i, \theta_{-i})].$$

If the first condition in the final disjunction held, it would clearly violate the assumption that φ satisfies S-FNP. Thus we can assume the second condition holds, and let $x = \varphi_j(\theta'_i, \theta_{-i}, \theta_j)$, with $x P_i \varphi_i(\theta_i, \theta_{-i})$, be the object assigned to j. Then, consider the following alternative manipulation for agent i: report $\theta''_i = \theta_j \in M(\theta_i)$ for i's true identity and $\theta'_j = \theta'_i$ as a fake account. By the anonymity of the mechanism, we have $x = \varphi_i(\theta''_i, \theta_{-i}, \theta'_j)$, so that $\varphi_i(\theta''_i, \theta_{-i}, \theta'_j) P_i \varphi_i(\theta_i, \theta_{-i})$. But this contradicts that φ satisfies S-FNP.

On the other hand, we will show later that the the TTC mechanism is FNP, but, in general, not S-FNP. □

Proposition 4.2. *In directed environments, a mechanism φ satisfies SP if it satisfies FNP. However, the opposite is not true.*

Proof. The implication is straightforward. To prove that the opposite does not hold, consider the following (anonymous) mechanism: when $|N| \leq 5$, run the TTC mechanism, and when $|N| \geq 6$, run the DA mechanism. The mechanism inherits its strategy-proofness from TTC and DA, but does not satisfy FNP, because an agent can benefit from using fake accounts to switch the mechanism from TTC to DA. □

4.2 Connection to Solidarity Conditions

We now propose two solidarity conditions that turn out to have a strong connection to our two respective notions of false-name-proofness. They are closely related to *population monotonicity*, which requires that the arrival of a new agent affects all agents originally present in the same way (see

[4]If there exists a false-name manipulation for i that uses k fake identities, but not one that uses only $k-1$ additional identities, then we can consider a modified profile in which $k-1$ of the additional identities have been added, in which case agent i has an incentive to add the single remaining identity. Note that for this argument to work in the case of FNP, agent i must not have been planning to take the allocation of one of those $k-1$ identities, but this can always be ensured because by unit demand, agent i takes the allocation of at most one of the k additional identities.

e.g., Thomson [15]). Some papers, including Crawford [5], Ehlers et al. [6], and Kesten [11], discussed similar properties in matching.

We first introduce a stronger condition, called *downward population monotonicity*, which requires that a new arrival never improves the allocation of one of the original agents.

Definition 4.3 (DPM). *A mechanism φ is said to satisfy* downward population monotonicity *(DPM) if $\forall \succeq^{\mathcal{C}} \in \mathcal{D}$, $\forall \pi = \Pi_C^m$, $\forall N \subset \mathcal{N}$, $\forall \theta \in \Theta^n$, $\forall j \in \mathcal{N} \setminus N$, and $\forall \theta_j \in \Theta$, it holds that $\forall i \in N, \varphi_i(\theta) R_i \varphi_i(\theta, \theta_j)$.*

The second condition, called *downward population monotonicity for weaker types*, is a slight modification of downward population monotonicity. It requires the downward population monotonicity condition to hold only for those agents originally present whose characteristics are stronger than that of the new arrival.

Definition 4.4 (DPM for Weaker Types). *A mechanism φ is said to satisfy* downward population monotonicity for weaker types *(DPM-W) if $\forall \succeq^{\mathcal{C}} \in \mathcal{D}$, $\forall \pi = \Pi_C^m$, $\forall N \subset \mathcal{N}$, $\forall \theta \in \Theta^n$, $\forall j \in \mathcal{N} \setminus N$, and $\forall \theta_j \in \Theta$, it holds that $\forall i \in N$ such that $M(\theta_i) \ni \theta_j, \varphi_i(\theta) R_i \varphi_i(\theta, \theta_j)$.*

We next show that these two solidarity conditions are necessary and sufficient conditions for a strategy-proof mechanism to satisfy, respectively, S-FNP and FNP.

Proposition 4.3. *In directed environments, a mechanism satisfies S-FNP if and only if it satisfies SP and DPM.*

Proof. It is straightforward that S-FNP implies SP. Also, if φ fails DPM, then there exists a situation where a new arrival of some agent j raises some agent i's utility. Then, agent i can increase its utility by adding j as a fake identity. Hence, S-FNP implies DPM.

Now, we show that the combination of SP and DPM implies S-FNP. From DPM, $\forall \succeq^{\mathcal{C}} \in \mathcal{D}$, $\forall \pi = \Pi_C^m$, $\forall N \subset \mathcal{N}$, $\forall \theta \in \Theta^n$, $\forall j \in \mathcal{N} \setminus N$, and $\forall \theta_j \in \Theta$, it holds that $\forall i \in N, \varphi_i(\theta) R_i \varphi_i(\theta, \theta_j)$. Furthermore, considering $\theta_{-i} \cup \theta_j$ as the type profile of agents (identities) other than i, we have $\varphi_i(\theta_i, \theta_{-i}, \theta_j) R_i \varphi_i(\theta'_i, \theta_{-i}, \theta_j)$ for any $\theta'_i \in M(\theta_i)$, by SP. Combining the two preference relations results in the condition for S-FNP. □

Proposition 4.4. *In directed environments, a mechanism satisfies FNP if and only if it satisfies SP and DPM-W.*

Proof. It is straightforward that FNP implies SP. Also, if φ fails DPM-W, there exists a situation where a new arrival of some agent j raises some agent i's utility, where $\theta_j \in M(\theta_i)$. Then, agent i can increase its utility by adding j as a fake identity. Hence, FNP implies DPM-W.

Now, we show that the combination of SP and DPM-W implies FNP. From DPM-W, $\forall \succeq^{\mathcal{C}} \in \mathcal{D}$, $\forall \pi = \Pi_C^m$, $\forall N \subset \mathcal{N}$, $\forall \theta \in \Theta^n$, $\forall j \in \mathcal{N} \setminus N$, and $\forall \theta_j \in \Theta$, it holds that $\forall i \in N$ such that $M(\theta_i) \ni \theta_j, \varphi_i(\theta) R_i \varphi_i(\theta, \theta_j)$. Furthermore, from SP, we have $\varphi_i(\theta_i, \theta_{-i}, \theta_j) R_i \varphi_i(\theta'_i, \theta_{-i}, \theta_j)$ for any $\theta'_i \in M(\theta_i)$. Combining the two preference relations, we obtain the first condition of FNP.

Moreover, because both $\theta'_i \in M(\theta_i)$ and $\theta_j \in M(\theta_i)$ hold, agent i with true type θ_i can report $\theta''_i = \theta_j$ from its true identity i, and $\theta'_j = \theta'_i$ from the fake identity j. Thus, $\varphi_i(\theta_i, \theta_{-i}) R_i \varphi_i(\theta''_i, \theta_{-i}, \theta'_j)$ still holds. By the mechanism's anonymity, $\varphi_i(\theta''_i, \theta_{-i}, \theta'_j) = \varphi_j(\theta'_i, \theta_{-i}, \theta_j)$ holds, and thus, $\varphi_i(\theta_i, \theta_{-i}) R_i \varphi_j(\theta'_i, \theta_{-i}, \theta_j)$, the second condition of FNP. □

We can now combine these results with those in the previous section to obtain necessary and sufficient conditions for a mechanism to satisfy S-FNP/FNP.

Theorem 2. *In directed environments, a mechanism satisfies S-FNP if and only if it satisfies SP-, RIC, and DPM.*

Theorem 3. *In directed environments, a mechanism satisfies FNP if and only if it satisfies SP-, RIC, and DPM-W.*

5. FALSE-NAME-PROOF MATCHING

In this section, we discuss specific false-name-proof matching mechanisms. First, we introduce two mechanisms well-studied in the literature: the (agent-proposing) Deferred-Acceptance (DA) and Top-Trading-Cycle (TTC) mechanisms.

Definition 5.1 (Deferred Acceptance [8]). *Each agent $i \in N$ proposes to its first-ranked object, and each object $x \in X$ rejects all agents proposing to it, except for the one whose reported characteristic has the highest priority for it among them. Each rejected agent then proposes to its next most favored object, and each object again rejects all agents including the one previously not rejected (if any), except for the one whose reported characteristic has the highest priority for it among them. This procedure is repeated until no agent is rejected. Each object is allocated to the agent who is not rejected when the procedure terminates.*

Definition 5.2 (Top Trading Cycle [14]). *Each agent $i \in N$ points to its first-ranked object, and each object $x \in X$ points to the agent whose reported characteristic has the highest priority for that object. All resulting cycles are removed, and each agent included in such a cycle receives the object to which it is pointing. Each remaining agent then points to its most favored object among the remaining ones, and each remaining object points to the remaining agent whose reported characteristic has the highest priority for that object. This procedure is repeatedly applied until all agents are removed, or all objects are removed.*

5.1 Fair Matching

We first discuss fair mechanisms. It is known that DA satisfies FA, while TTC does not. Furthermore, Balinski and Sönmez [3] proved that in the traditional matching model without characteristics, DA is the only fair mechanism that satisfies both strategy-proofness and RIP. We now prove that DA also satisfies S-FNP.

Theorem 4. *In directed environments, DA satisfies S-FNP.*

Proof. DA is known to be strategy-proof in the traditional model, which easily implies that it satisfies SP- in our model. Also, it was proven by Balinski and Sönmez [3] that DA satisfies RIP. Thereby, it also satisfies RIC. Together, these results imply that DA satisfies SP in our model. Furthermore, Crawford [5] proved that DA satisfies DPM. Thus, by Theorem 2, DA satisfies S-FNP. □

5.2 Efficient Matching

We next focus on Pareto efficient mechanisms. TTC is known to satisfy PE, while DA does not satisfy it. Also, Morrill [12] showed that in the traditional matching model without characteristics, TTC is the only mechanism that satisfies strategy-proofness, PE, MB, and a natural condition called *independence of irrelevant rankings*. We now show that TTC is false-name-proof.

THEOREM 5. *In directed environments, TTC satisfies FNP.*

PROOF. TTC is known to satisfy SP-, i.e., for each agent i, misreporting only its preference R_i is not beneficial. By Theorem 3, all that remains to show is that TTC also satisfies both RIC and DPM-W.

We first show that TTC satisfies RIC, i.e., for any i, misreporting a weaker characteristic c'_i (with $c_i \succeq^{\mathcal{C}} c'_i$) cannot make i strictly better off. [5] Let t be the round in which i is assigned object x when i reports c_i. Also, let N^t and X^t be the sets of agents and objects that have been matched before t when i reports its characteristic truthfully. Then we can observe, from the definition of TTC and the fact that i is assigned x at t, that at round t, there is no remaining object $y \in X \setminus X^t$ such that yP_ix when i reports c_i.

Lemma A.1 in Appendix shows that if i reports the weaker characteristic c'_i instead, N^t and X^t would still be matched before round t, in the same way. Hence, i will not be allocated an object in X^t; but we already know that there is no object $y \in X \setminus X^t$ such that yP_ix, so the manipulation cannot make i strictly better off. This establishes RIC.

Finally, we show that TTC satisfies DPM-W. Assume an agent j with type $\theta_j = (R_j, c_j)$ is added to the original economy with N agents, and consider the allocation to an arbitrary agent i whose type $\theta_i = (R_i, c_i)$ satisfies $M(\theta_i) \ni \theta_j$, i.e., $c_i \succeq^{\mathcal{C}} c_j$. From the definition of TTC, no agent points to a different object from the original one due to j's participation. Furthermore, since $c_i \pi_x c_j$ holds for all $x \in X$ and for all agents i such that $c_i \succeq^{\mathcal{C}} c_j$, no object points to j until all such i are removed. In other words, the allocations made by TTC do not change until all such i are removed. Thus, for any agent i satisfying $c_i \succeq^{\mathcal{C}} c_j$, its allocation does not change due to the joining of agent j. This establishes DPM-W. □

However, the following example, based on a proof by Ehlers et al. [6], shows that TTC in general is not downward population monotonic, and thus, not strongly false-name-proof.

EXAMPLE 5.1. *Suppose there are two objects x, y and three possible agents 1, 2, and 3. Also suppose there are only three characteristics, c_1, c_2, and c_3. Consider a DAG $\succeq^{\mathcal{C}} \in \mathcal{D}$ such that there is only one (non-self-loop) arrow, from c_1 to c_2, i.e., $c_1 \succeq^{\mathcal{C}} c_2$. The agents' true types are given as $\theta_1 = (R_1, c_1)$, $\theta_2 = (R_2, c_2)$, and $\theta_3 = (R_3, c_3)$, respectively. Then, consider the following priority order for each object: $c_1 \pi_x c_2 \pi_x c_3$ and $c_3 \pi_y c_1 \pi_y c_2$. Also, consider the following preferences for the agents: $yP_1 \varnothing P_1 x$, $xP_2 yP_2 \varnothing$, and $xP_3 yP_3 \varnothing$.*

First, consider the case where agents 2 and 3 attend. Because a trading cycle $x \to 2 \to x$ is constructed, the outcome is $(\varnothing, x, y)$. Next, consider the case where all three agents attend. Since a trading cycle $x \to 1 \to y \to 3 \to x$ is constructed, the outcome is $(y, \varnothing, x)$. Now we can see that the arrival of agent 1 makes agent 2 strictly worse off, but agent 3 strictly better off. Thus, TTC violates population monotonicity in this case.

We note that this example does not exhibit a violation of FNP. Agent 3, with true characteristic c_3, would become strictly better off due to agent 1's joining, but cannot create agent 1 as a fake identity, because there is no arrow from c_3 to c_1 in the characteristics DAG $\succeq^{\mathcal{C}}$. In fact, the only possible false-name manipulation in this example is the one by agent 1 adding agent 2 as a fake identity when originally there are only agents 1 and 3. However, agent 1 originally gets y, since a cycle $x \to 1 \to y \to 3 \to x$ is constructed. Thus, this manipulation is not beneficial to agent 1.

We now introduce a condition on a profile of priority orders over characteristics, called *acyclicity* [11], that is necessary and sufficient for TTC to satisfy DPM and S-FNP.

DEFINITION 5.3 (ACYCLICITY). *A profile of orders $\pi = (\pi_x)_{x \in X} \in \Pi_C^m$ is said to be* acyclic *(or to satisfy acyclicity) if for any three characteristics $d, e, f \in \mathcal{C}$ and any two objects $x, y \in X$, it always holds that $[d \pi_x e \pi_x f] \Rightarrow [d \pi_y f]$.*

When there is only one copy of each object (as we assume in this paper), this acyclicity condition coincides with two other properties, *weak acyclicity* as proposed by Ergin [7] and a *separability* condition discussed by Ehlers et al [6]. The proof is omitted due to space limitations.

We now prove that TTC satisfies S-FNP if and only if we restrict our attention to acyclic profiles of priority orders.

THEOREM 6. *In directed environments, TTC satisfies S-FNP if and only if any $\pi \in \Pi_C^m$ satisfies acyclicity.*

PROOF. Kesten [11] showed that TTC satisfies population monotonicity if and only if acyclicity holds. Also, TTC is known to satisfy PE. Furthermore, it is easy to show that under the condition of PE, satisfying population monotonicity is equivalent to satisfying DPM. Thus, TTC satisfies DPM if and only if acyclicity holds. Furthermore, from Theorem 5, TTC satisfies RIC. Thus, if and only if acyclicity holds, TTC satisfies SP-, RIC, and DPM, which is equivalent to saying that TTC satisfies S-FNP by Theorem 2. □

Furthermore, analogously to a result by Kesten [11], we have the following on the relationship between DA and TTC. The proof quickly follows from Kesten's result.

PROPOSITION 5.1. *In directed environments, TTC is equivalent to DA if and only if any $\pi \in \Pi_C^m$ satisfies acyclicity.*

6. CONCLUSIONS

In this paper, motivated by matching in highly anonymous domains, we proposed an, in our opinion realistic, model of matching with characteristics. We showed that no mechanism satisfies strategy-proofness and the mutual-best condition if (and only if) there is a cyclic conflict between the directed characteristics graph and the priority of an object. We then introduced a natural restriction on agents' misreports of characteristics and objects' priorities, allowing agents only to misreport weaker characteristics and requiring objects to always prefer stronger characteristics. Under this restriction, both DA and TTC satisfy a version of false-name-proofness in which an agent can only report weaker types for all of her identities. Furthermore, DA also satisfies a stronger version of false-name-proofness where an agent can report any characteristics for its fake identities, while TTC fails to satisfy this without an acyclicity assumption on priorities.

[5] Others have claimed that TTC satisfies RIP (e.g., [1]). However, we were not able to find a proof of this in the literature.

We are encouraged by these positive results on the design of false-name-proof mechanisms in this domain, especially as false-name-proofness often leads to negative results in other domains. The model could be generalized in several directions. In this paper, we have focused only on one-to-one matching, where there is only one copy of every object, with strict preferences of agents and responsive priorities of objects. A natural direction is to extend to one-to-many matching [11], where there can be more than one copy of an object. Also, allowing indifferences on preferences [13, 2] seems to be an important extension. Finally, we may ask what the essential elements of the model here are that allow us to obtain positive results on false-name-proofness, and see if we can use these to obtain similar positive results on false-name-proofness in other domains.

Acknowledgements

Todo thanks JSPS for support under JSPS KAKENHI Grant Number 10J01756. Conitzer thanks NSF and ARO for support under grants CCF-1101659, IIS-0953756, W911NF-12-1-0550, and W911NF-11-1-0332.

7. REFERENCES

[1] A. Abdulkadiroglu and Y.-K. Che. The role of priorities in assigning indivisible objects: A characterization of top trading cycles. mimeo, 2010.

[2] H. Aziz and B. de Keijzer. Housing markets with indifferences: A tale of two mechanisms. In *AAAI'12*.

[3] M. Balinski and T. Sönmez. A tale of two mechanisms: Student placement. *Journal of Economic Theory*, 84(1):73 – 94, 1999.

[4] V. Conitzer. Anonymity-proof voting rules. In *WINE'08*.

[5] V. P. Crawford. Comparative statics in matching markets. *Journal of Economic Theory*, 54(2):389 – 400, 1991.

[6] L. Ehlers, B. Klaus, and S. Papai. Strategy-proofness and population-monotonicity for house allocation problems. *Journal of Mathematical Economics*, 38(3):329–339, 2002.

[7] H. I. Ergin. Efficient resource allocation on the basis of priorities. *Econometrica*, 70(6):2489–2497, 2002.

[8] D. Gale and L. S. Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15, 1962.

[9] J. R. Green and J.-J. Laffont. Partially verifiable information and mechanism design. *The Review of Economic Studies*, 53(3):447–456, 1986.

[10] A. Iwasaki, V. Conitzer, Y. Omori, Y. Sakurai, T. Todo, M. Guo, and M. Yokoo. Worst-case efficiency ratio in false-name-proof combinatorial auction mechanisms. In *AAMAS'10*.

[11] O. Kesten. On two competing mechanisms for priority-based allocation problems. *Journal of Economic Theory*, 127(1):155 – 171, 2006.

[12] T. Morrill. An alternative characterization of top trading cycles. *Economic Theory*, pages 1–17, 2012.

[13] T. Quint and J. Wako. On houseswapping, the strict core, segmentation, and linear programming. *Mathematics of Operations Research*, 29(4):861–877, 2004.

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

[15] W. Thomson. Population monotonic allocation rules. In W. A. Barnett, H. Moulin, M. Salles, and N. J. Schofield, editors, *Social Choice, Welfare, and Ethics*, chapter 4. Cambridge University Press, 1995.

[16] T. Todo, A. Iwasaki, and M. Yokoo. False-name-proof mechanism design without money. In *AAMAS'11*.

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

APPENDIX

A. PROOF OF THEOREM 5

LEMMA A.1. *Consider directed environments. For any* $\succeq^{\mathcal{C}} \in \mathcal{D}$, *any* $\pi \in \Pi^m$, *any* $N \subseteq \mathcal{N}$, *any* $\theta \in \Theta^n$, *and any* $i \in N$, *let* t *be the round in TTC when* i, *truthfully reporting* c_i *as its characteristic, is allocated object* $x \in X$. *Also, let* N^t *and* X^t *be the sets of agents and objects, respectively, that are matched before round* t *when* i *reports* c_i. *Then, for any* $c_i' \in \mathcal{C}$ *such that* $c_i \succeq^{\mathcal{C}} c_i'$, *if* i *reports* c_i' *instead of* c_i, N^t *and* X^t *are still matched in the exact same way.*

PROOF. We note that since i is assigned x at t, $i \notin N^t$ and $x \notin X^t$. We also note that, when i reports c_i, in the first $t-1$ rounds, no agent in N^t ever points outside X^t, and no object in X^t ever points outside N^t. This is because if an agent/object points to an object/agent, the former cannot be allocated before the latter is allocated.

We now prove by induction on round s (from $s = 1$ to $t-1$) that

1. any cycle constructed in round s when i reports c_i is also constructed in round s when i reports c_i';
2. any agents and objects in N^t and X^t that have not yet been allocated in round s when i reports c_i have also not yet been allocated in round s when i reports c_i'.

Suppose this is true for all $s < k \leq t-1$; we will prove it for $s = k$. At the beginning of this round, by the induction assumption, the set of remaining agents and objects in N^t and X^t is the same, regardless of whether i reports c_i or c_i'. Moreover, as we observed above, if i reports c_i, then at the beginning of round $s = k$, all remaining agents and objects in N^t and X^t point to other objects and agents in X^t and N^t. But then, in the alternative world where i reports c_i', all remaining agents and objects in N^t and X^t must point to the *same* objects and agents at the beginning of this round, because (1) by the induction assumption any agent or object that is still available in this alternative world is also still available at this point in the original world, (2) agents' preferences are the same in both worlds, and (3) object priorities are the same in both worlds, with the possible exception that i may be ranked lower in the alternative world—but no object in X^t points to i in this round even in the original world. It immediately follows that all the cycles formed in the original world in this round (which are all contained in N^t and X^t) are also formed in the alternative world, and no other agents and objects in N^t and X^t are matched in this round in the alternative world. This completes the proof by induction. □