Matchings with Externalities and Attitudes

Simina Brânzei
Aarhus University
simina@cs.au.dk

Tomasz Michalak
Oxford University and Warsaw University
tomasz.michalak@cs.ox.ac.uk

Talal Rahwan
University of Southampton
tr@ecs.soton.ac.uk

Kate Larson
University of Waterloo
klarson@cs.uwaterloo.ca

Nicholas R. Jennings
University of Southampton
nrj@cs.soton.ac.uk

"A pessimist sees the difficulty in every opportunity; an optimist sees the opportunity in every difficulty."

Winston Churchill (1874–1965)

ABSTRACT

Two-sided matchings are an important theoretical tool used to model markets and social interactions. In many real-life problems the utility of an agent is influenced not only by their own choices, but also by the choices that other agents make. Such an influence is called an externality. Whereas fully expressive representations of externalities in matchings require exponential space, in this paper we propose a compact model of externalities, in which the influence of a match on each agent is computed additively. Under this framework, we analyze many-to-many matchings and one-to-one matchings where agents take different *attitudes* when reasoning about the actions of others. In particular, we study optimistic, neutral and pessimistic attitudes and provide both computational hardness results and polynomial-time algorithms for computing stable outcomes.

Keywords

Matchings, Externalities, Coalitional Games

General Terms

Algorithms, Economics, Theory

Categories and Subject Descriptors

I.2.11 [**Distributed Artificial Intelligence**]: Multiagent Systems; J.4 [**Social and Behavioral Sciences**]: Economics

1. INTRODUCTION

Matching games are an important theoretical abstraction which have been extensively studied in several fields, including economics, combinatorial optimization and computer science. Matchings are often used to model markets, and examples include the classic marriage problem, firms and workers, schools and students, hospitals and medical interns [4].

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.

Previous matching literature has focused primarily on one-to-one and one-to-many models [14]. More recently, however, attention is being paid to more complex models of many-to-many matchings due to their relevance to real-world situations [6, 11]. For example, most labour markets involve at least a few many-to-many contracts [6]. More realistic matching models should take into account the fact that in many settings the utility of an agent is influenced not only by their own choices, but also by the choices that other agents make. Such an influence is called an externality. For instance, companies care not only about the employees they hire themselves, but also about the employees hired by other companies. This aspect is crucial to how competitive a company is in the market, and so externalities must be considered in order to completely understand such situations. While researchers have looked at externalities in one-to-one matchings (*e.g.* [9, 15]) and one-to-many matchings (*e.g.* [5]), typically fully expressive respresentations for the externalities were assumed.[1] Modelling matchings with externalities is computationally challenging, as fully expressive representations require exponential space. This motivates the search for compact representations of externalities.

One of the central questions in matching games is stability [14], which informally means that no group of agents can modify the matching and improve the outcome for themselves. In the presence of externalities, stability becomes a highly complex and challenging phenomenon due to the fact that a deviation by some agents can affect the utilities of all other agents in the system. This can invoke a response that can change the worth of the original deviation dramatically, and so any group considering a deviation should consider all possible responses that can be taken by the remaining agents. Evaluating these may be infeasible, particularly for agents with computational limitations or who have bounds on their rationality. Motivated by both the cooperative game theory literature (*e.g.* [12, 13]), and work on models of bounded rationality [8], we argue that agents will use *heuristics*, which are based on their *attitudes* (*i.e.* optimistic, neutral, or pessimistic), to reason about the actions taken by others.

Our contributions in this paper can be summarized as follows. We formulate a compact model of externalities for matchings, in which the influence of matches on agents is computed additively. Second, we consider key stability con-

[1] An interesting exception is work by Bodine-Baron *et al.* which looked at a one-to-many matching problem where externalities were derived from an underlying social network [2].

cepts for matching, under optimistic, neutral and pessimistic attitudes. We study the computational properties of these stability concepts, provide both hardness results and polynomial algorithms where applicable, and show how the stability concepts under different attitudes are related to each other.

2. THE MODEL

Let $N = M \cup W$ be the set of agents, where $M = \{m_1, \ldots, m_{|M|}\}$ and $W = \{w_1, \ldots, w_{|W|}\}$ are disjoint. A *match*, (m, w), is an edge between two agents $m \in M$ and $w \in W$. We let a matching, $\mathcal{A}$, be a set of all matches. If the number of allowable matches any agent can participate in is unrestricted then we say we have a *many-to-many matching problem*, while if each agent can participate in at most one match then we have a *one-to-one matching problem*. We assume the formation of a match requires the consent of both parties, while severing a match can be done unilaterally by any of its endpoints. The *empty matching* contains no matches, while the *complete matching* contains all possible matches. A matching game with externalities is defined as follows:

Definition 1. *A* matching game with externalities *is represented as a tuple $G = (M, W, \Pi)$, where (M, W) is the set of agents and Π is a real valued function such that $\Pi(\mathcal{A}|z)$ is the utility of agent z when matching $\mathcal{A}$ forms.*

We make no assumptions as to whether the utility of the agents is transferrable or not and thus Definition 1 can be viewed as a generalization of assignment games with externalities [15]. We are interested in settings where an agent's utility is formed by additive externalities.

Definition 2. *A* matching game with additive externalities *is represented as a tuple $G = (M, W, \Pi)$, where (M, W) is the set of agents and Π is a real valued function such that $\Pi(m, w|z)$ is the value that agent z receives from the formation of match (m, w). Given a matching $\mathcal{A}$ over N, the* utility *of an agent z in $\mathcal{A}$ is: $u(z, \mathcal{A}) = \sum_{(m,w)\in\mathcal{A}} \Pi(m, w|z)$.*

Thus, in a matching game with additive externalities, an agent's utility is the sum of values it receives from matches it participates in, along with the sum of all externalities that arise due to the matchings of other agents. We study additive externalities since they are a conceptually straightforward compact representation and assumptions about additive utility functions are wide spread throughout the artificial intelligence and algorithmic game theory literature (*e.g.* [1, 2, 3]).

2.1 Stability Concepts

We are interested in whether matchings are *stable* and whether there exist stable matchings given a particular matching game, G. In general, a matching is stable if no subset of agents has any incentive to reorganize and form new matchings amongst themselves. We distinguish between three standard stability concepts which commonly appear in the matching literature. The first, setwise stability, is the most general and encompasses the other two (corewise stability and pairwise stability). Unless otherwise noted, the stability concept used in this paper is setwise stability, which we interchangeably refer to as set stability.

Definition 3. *Given a matching game $G = (M, W, \Pi)$, a matching $\mathcal{A}$ of G is* setwise stable *if there does not exist a set of agents $B \subseteq N$, which can improve the utility of at least one member of B while not degrading the others by:*

- *rearranging the matches among themselves*
- *deleting a (possibly empty) subset of the matches with agents in $N \setminus B$.*

If such a coalition B, exists, it is called a *blocking coalition*.

Definition 4. *Given a matching game $G = (M, W, \Pi)$, a matching $\mathcal{A}$ of G is* corewise stable *if there does not exist a set of agents $B \subseteq N$, which can improve the utility of at least one member of B while not degrading the others by:*

- *rearranging the matches among themselves*
- *deleting all the matches with agents in $N \setminus B$.*

Definition 5. *Given a matching game $G = (M, W, \Pi)$ and a matching $\mathcal{A}$, $\mathcal{A}$ is* pairwise stable *if there does not exist a blocking coalition of size one or two.*

Pairwise stability is most interesting for one-to-one matchings. In the context of one-to-one matchings, a blocking coalition of size one is equivalent to one agent that can improve utility by cutting its matches. A blocking coalition of size two is equivalent to two agents that can form a new match with each other while possibly cutting their previous match (if any), or who can coordinate to cut their existing matches without forming a new match with each other.

Finally, we note that each member of a blocking coalition B is required to perform at least one action, by severing a match with another agent in N, or by forming a new match with another agent in B. Other definitions of stability could incorporate different layers of deviators, such as agents who perform the deviation and agents who agree to it without actively participating. However, such definitions can be problematic, and require specifying which agents are identified as deviators and how they should be treated depending on their role. For this reason, in this paper we only consider one type of deviators, those who are required to perform at least one action.

2.2 Agents' Attitudes

In matching games without externalities, the actions taken by other agents have only a limited effect on an agent – its utility depends solely on who it is matched with, and does not depend on matches involving others. However, if there are externalities then this is no longer true. The utility of agent m, for example, can depend on the matches involving agent w even if $(m, w) \notin \mathcal{A}$. Therefore, we argue, the stability concepts need to account for the actions taken by agents in $N \setminus B$ after a deviation by coalition B. However, it may be hard to compute the possible reactions to a deviation since there are potentially an exponential number (*i.e.* all possible matchings amongst agents in $N \setminus B$). Instead, in this paper we consider several natural heuristics, based on agents' *attitudes*, that members of B use to reason about, and approximate, the reactions to their deviations.

Neutral Attitude: Agents in blocking coalition B have a *neutral* attitude if they assume that agents in $N \setminus B$ will not react to the deviation. All existing matches amongst non-deviating agents will remain and no new matches will form.

Pessimistic Attitude: Agents in blocking coalition B have a *pessimistic* attitude if they assume that the agents in $N \setminus B$ will take actions so as to punish the members of B. That is, the non-deviators will cut matches with a positive influence on the deviators, and will form any new matches with a negative influence on them.

Optimistic Attitude: Agents in blocking coalition B have an *optimistic* attitude if they assume that any response to their deviation will be in their own best interest. That is, when coalition B considers deviation $\mathcal{A}'$ from $\mathcal{A}$, every agent $i \in B$ evaluates the deviation assuming the agents in $N \setminus B$ will organize themselves in the best possible way for i.

Neutrality, optimism and pessimism are *heuristics* used by agents in blocking coalition B to reason about the reactions of others. We make no assumption that the agents in $N \setminus B$ will actually act in the way expected by members of B, and it possible that reacting in a neutral/pessimistic/optimistic way is irrational. Second, it is entirely possible that there are inconsistencies among the reactions that the members of B expect from $N \setminus B$. The optimistic and pessimistic cases are in fact the best and worst possible reactions to the deviating agents, respectively. As such, these cases represent *upper and lower bounds* on the impact a reaction could have on the deviating agents. Based on this, we argue that the analysis of those two particular cases is essential when assessing the expected reward/risk associated with the deviation.

2.3 Computational Complexity

We are interested in understanding the computational complexity of finding stable outcomes when agents have different attitudes. To address the complexity issues, we look at two problems in particular; *Non-emptiness* and *Membership*.

Definition 6 (Non-Emptiness). *Given a matching game $G = (M, W, \Pi)$ and a stability solution concept $\mathcal{C}$, the non-emptiness question asks whether there exists a matching $\mathcal{A}$ of G which is stable according to $\mathcal{C}$.*

Definition 7 (Membership). *Given a matching game $G = (M, W, \Pi)$, a matching $\mathcal{A}$ of G, and a stability solution concept $\mathcal{C}$, the membership question asks whether $\mathcal{A}$ is stable according to $\mathcal{C}$.*

3. MANY-TO-MANY MATCHINGS

In this section we analyze the stability of many-to-many matchings using setwise stability as the solution concept and comparing neutral, optimistic and pessimistic attitudes. In particular, we analyze the complexity of computing stable outcomes under different attitudes, and describe the relationship between the different setwise stable sets.

3.1 Neutral Stability

We start with examples to provide insight into setwise stability under neutral attitudes. Our first example shows that the neutral stable set can be empty (*i.e.* there are matching games which are inherently unstable).

Example 1. *Let $M = \{m\}$, $W = \{w_1, w_2\}$, and Π as follows: $\Pi(m, w_i|m) = 0$, $\Pi(m, w_i|w_i) = \varepsilon$, and $\Pi(m, w_1|w_2) = \Pi(m, w_2|w_1) = -\Delta$, where $\Delta > \varepsilon > 0$. The empty matching is blocked by (m, w_1), $\{(m, w_1), (w_2)\}$ is blocked by (m, w_2), $\{(m, w_2), (w_1)\}$ is blocked by (m, w_1), and $\{(m, w_1), (m, w_2)\}$ is blocked by the empty matching, thus the neutral stable set is empty.*

The next example contrasts the case where the empty matching belongs to the neutral stable set with the case where it does not.

Example 2. *Let $G = (M, W, \Pi)$ be such that $\Pi(m, w|m) = \Pi(m, w|w) = -\varepsilon < 0$ for all m, w, and $\Pi(m, w|z) = \delta > 0$ for all $z \neq m, w$. If $\varepsilon \gg \delta$, the only matching satisfying neutral set stability is the empty matching, since the formation of any match is very expensive for the participating agents and not compensated by the utility obtained from the externalities. Otherwise, if $\delta \gg \varepsilon$, the grand coalition can block the empty matching through the complete matching, and so the neutral stable set is empty.*

For neutral stability we have the following hardness results.

Theorem 1. *Checking nonemptiness of the neutral stable set is NP-hard.*

Proof. We provide a reduction from the *Knapsack* problem. Let $I = \langle U, s, v, B, K \rangle$, where $U = \{u_1, \dots, u_n\}$ is a finite set, $s(u) \in \mathbb{Z}^+$ the size of element $u \in U$, $v(u) \in \mathbb{Z}^+$ the value of element $u \in U$, $B \in \mathbb{Z}^+$ a size constraint, and $K \in \mathbb{Z}^+$ a value goal, such that $U' \subseteq U$ is a solution if $\sum_{u \in U'} s(u) \leq B$ and $\sum_{u \in U'} v(u) \geq K$. We construct a matching game G such that G has a nonempty stable set if and only if I has a solution. Let $M = \{x_1, \dots, x_n, m_1, m_2\}$, $W = \{y_1, \dots, y_n, w\}$, and Π with non-zero entries:

- $\Pi(x_i, y_i|m_1) = -s(u_i)$ and $\Pi(x_i, y_i|m_2) = v(u_i)$, $\forall i \in \{1, \dots, n\}$
- $\Pi(m_1, w|m_1) = -B$
- $\Pi(m_2, w|m_2) = K - \sum_{u_i \in U} v(u_i)$
- $\Pi(x_j, w|x_j) = -1$ and $\Pi(m_i, y_j|y_j) = -1$, $\forall i \in \{1, 2\}$ and $\forall j \in \{1, \dots, n\}$

Since the instances with $\sum_{u_i \in U} v(u_i) < K$ are trivially solvable, we are interested in those cases where $\sum_{u_i \in U} v(u_i) \geq K$, and so $\Pi(m_2, w|m_2) \leq 0$. If I has a solution U', we claim the following matching belongs to the stable set: $\mathcal{A} = \left(\bigcup_{u_i \in U'} \{(x_i, y_i)\}\right) \cup \{(m_1), (m_2), (w)\}$ (1). First note that U' satisfies the knapsack conditions, and so the utilities of the agents in $\mathcal{A}$ are: $u(m_1, \mathcal{A}) = \sum_{u_i \in U'} \Pi(x_i, y_i|m_1) = \sum_{u_i \in U'} -s(u_i) \geq -B$, $u(m_2, \mathcal{A}) = \sum_{u_i \in U'} \Pi(x_i, y_i|m_2) = \sum_{u_i \in U'} v(u_i) \geq K$, and $u(x_i, \mathcal{A}) = u(y_i, \mathcal{A}) = u(w, \mathcal{A}) = 0, \forall i \in \{1, \dots, n\}$.

All the agents, except possibly for m_1 and m_2, obtain their best possible utility in $\mathcal{A}$. Thus any blocking coalition, B, would have to contain at least one of m_1 and m_2. Since blocking requires each member of B to perform an action, it follows that m_1, m_2 can only be involved in blocking $\mathcal{A}$ by forming a new match. Recall that $\Pi(m_i, y_j|y_j) = -1$, and so agents y_j will never accept a match with either m_1 or m_2. Thus the only matches that m_1 and m_2 could form, if deviating, are (m_1, w) and (m_2, w), respectively.

The best utility that m_1 can get when matched with w is attained in $\mathcal{A}_1 = \{(m_1, w), (x_1), \dots, (x_n), (y_1), \dots, (y_n),$

$(m_2)\}$: $u(m_1, \mathcal{A}_1) = \Pi(m_1, w|m_1) = -B \leq u(m_1, \mathcal{A})$ (2), where the candidate for the blocking coalition is $B_1 = \{m_1, w\} \cup \left(\bigcup_{u_i \in U'} \{x_i, y_i\}\right)$.

The best utility that m_2 can get when matched with w is attained in $\mathcal{A}_2 = \{(m_2, w), (x_1, y_1), \ldots, (x_n, y_n), (m_1)\}$: $u(m_2, \mathcal{A}_2) = \Pi(m_2, w|m_2) + \sum_{i=1}^{n} \Pi(x_i, y_i|m_2) = \left(K - \sum_{u_i \in U} v(u_i)\right) + \sum_{u_i \in U} v(u_i) = K \leq u(m_2, \mathcal{A})$ (3), where the candidate for the blocking coalition is $B_2 = \{m_2, w\} \cup \left(\bigcup_{u_i \notin U'} \{x_i, y_i\}\right)$. From inequalities (2) and (3), m_1 and m_2 cannot improve by deviating from $\mathcal{A}$. Since the other agents have no incentive to deviate, it follows that $\mathcal{A}$ belongs to the stable set.

Conversely, if the stable set of G is non-empty, let $\mathcal{A}$ be a stable matching. First note that $\mathcal{A}$ must satisfy neutral individual rationality, and so it cannot contain any match with negative value for one of the endpoints. Thus the only non-zero matches that can be included in $\mathcal{A}$ are a subset of $\left(\bigcup_{u_i \in U} \{(x_i, y_i)\}\right)$. In addition, any matches of the form (x_i, y_j), with $i \neq j$, can be removed from $\mathcal{A}$ without losing stability. Thus without loss of generality, $\mathcal{A}$ can be written as in Equation (1) for some $U' \subseteq U$. From $\mathcal{A}$ stable, coalitions $\{(m_1, w)\}$ and $\{(m_2, w)\}$ are not blocking, thus Inequalities (2) and (3) hold. Equivalently, the knapsack conditions are satisfied, and so U' is a solution. $\square$

THEOREM 2. *Checking neutral stable set membership is coNP-complete.*

PROOF. We show that the complementary problem, of deciding whether a matching does not belong to the stable set of a game, is NP-complete. Given matching $\mathcal{A}$, one can nondeterministically guess pair $\langle B, \mathcal{A}' \rangle$ such that matching $\mathcal{A}$ is blocked by coalition B through matching $\mathcal{A}'$. To verify that $\langle B, A' \rangle$ is blocking for $\mathcal{A}$, it is sufficient to compute, for each $z \in B$, its utility in $\mathcal{A}$ and $\mathcal{A}'$ (assuming that $N \setminus B$ does not react to the deviation).

To prove hardness, we provide a reduction from *Knapsack* [7]. A Knapsack instance has the form $I = \langle U, s, v, B, K \rangle$, where $U = \{u_1, \ldots, u_n\}$ is a finite set, $s(u) \in \mathbb{Z}^+$ the size of element $u \in U$, $v(u) \in \mathbb{Z}^+$ the value of element $u \in U$, $B \in \mathbb{Z}^+$ a size constraint, and $K \in \mathbb{Z}^+$ a value goal. Let $G = (M, W, \Pi)$ be a matching game with $M = \{x_1, \ldots, x_n, m_1, m_2\}$, $W = \{y_1, \ldots, y_n, w_1, w_2\}$, and Π with non-zero entries:

- $\Pi(x_i, y_i|m_1) = -s(u_i)$, $\Pi(x_i, y_i|w_1) = v(u_i)$, $\forall i \leq n$
- $\Pi(m_2, w_2|m_1) = -B - \varepsilon$ and $\Pi(m_2, w_2|w_1) = K - \varepsilon$, for some $0 < \varepsilon < 1$
- $\Pi(m_1, w|w) = \Pi(m, w_1|m) = -1$, for all $w \in W \setminus \{w_1\}$ and $m \in M \setminus \{m_1\}$

Let $\mathcal{A} = \{(m_2, w_2), (m_1), (w_1), (x_1), \ldots, (x_n), (y_1), \ldots, (y_n)\}$, with utilities:

- $u(x_i, \mathcal{A}) = \Pi(m_2, w_2|x_i) = 0$ and $u(y_i, \mathcal{A}) = \Pi(m_2, w_2|y_i) = 0$, $\forall i \in \{1, \ldots, n\}$
- $u(m_1, \mathcal{A}) = \Pi(m_2, w_2|m_1) = -B - \varepsilon$
- $u(w_1, \mathcal{A}) = \Pi(m_2, w_2|w_1) = K - \varepsilon$
- $u(m_2, \mathcal{A}) = \Pi(m_2, w_2|m_2) = 0$
- $u(w_2, \mathcal{A}) = \Pi(m_2, w_2|w_2) = 0$

All the agents except m_1 and w_1 obtain their maximum utility in $\mathcal{A}$. In addition, m_1 or w_1 can only block by forming the match (m_1, w_1), since all other matches with one of these agents as an endpoint are unfeasible. We claim that I has a solution if and only if $\mathcal{A}$ has a blocking coalition. If I has solution $U' \subseteq U$, consider the grand coalition N and matching: $\mathcal{A}' = \left(\bigcup_{u_i \in U'} \{(x_i, y_i)\}\right) \cup \left(\bigcup_{u_i \notin U'} \{(x_i, y_{i+1}), (x_{i+1}, y_i)\}\right) \cup \{(m_1, w_1), (m_2), (w_2)\}$, where $x_{n+1} = x_1$ and $y_{n+1} = y_1$. The utilities of the agents in $\mathcal{A}'$ are:

- $u(x_i, \mathcal{A}') = 0$ and $u(y_i, \mathcal{A}') = 0$, $\forall i \in \{0, \ldots, n\}$
- $u(m_1, \mathcal{A}') = -\sum_{u_i \in U'} s(u_i) \geq -B > -B - \varepsilon = u(m_1, \mathcal{A})$
- $u(w_1, \mathcal{A}') = \sum_{u_i \in U'} v(u_i) \geq K > K - \varepsilon = u(w_1, \mathcal{A})$
- $u(m_2, \mathcal{A}') = u(w_2, \mathcal{A}') = 0$

Thus the grand coalition is blocking under neutrality, since it can form matching $\mathcal{A}'$ and (weakly) improve the utilities of all its members by doing so.

Conversely, assume $\mathcal{A}$ is blocked by a coalition B through a matching $\mathcal{A}'$. Note that all the agents except m_1 and w_1, obtain their maximum possible utility in $\mathcal{A}$. Agent m_1 cannot block without agent w_1, since all other matches (m_1, w), for $w \neq w_1$ are unfeasible; similarly for w_1. Thus any blocking coalition must include agents m_1, w_1 and match (m_1, w_1). In addition, for one of m_1, w_1 to strictly improve after the deviation, the edge (m_2, w_2) must be removed from $\mathcal{A}'$. The conditions for improvement in $\mathcal{A}'$ are: (a) $u(m_1, \mathcal{A}') = \sum_{(x_i, y_i) \in \mathcal{A}'} \Pi(x_i, y_i|m_1) = \sum_{(x_i, y_i) \in \mathcal{A}'} -s(u_i) \geq u(m_1, \mathcal{A}) = -B - \varepsilon$ and (b) $u(w_1, \mathcal{A}') = \sum_{(x_i, y_i) \in \mathcal{A}'} \Pi(x_i, y_i|w_1) = \sum_{(x_i, y_i) \in \mathcal{A}'} v(u_i) \geq u(w_1, \mathcal{A}) = K - \varepsilon$. Since $v(u_i)$ and $s(u_i)$ are integers, Inequalities (a) and (b) are equivalent to $\sum_{(x_i, y_i) \in \mathcal{A}'} s(u_i) \leq B$ and $\sum_{(x_i, y_i) \in \mathcal{A}'} v(u_i) \geq K$. Thus $U' = \{u_i \in U \,|\, (x_i, y_i) \in \mathcal{A}'\}$ is a solution for the I instance. $\square$

3.2 Pessimistic Stability

We now investigate stability when agents have a pessimistic attitude. We first note that neutral and pessimistic attitudes can lead to quite different stable matchings. Example 2 showed that the neutral stable set was empty whenever $\delta \gg \varepsilon$, but it is not hard to show that the pessimistic stable set is non-empty for this example. However, the pessimistic stable set can still be empty as can be seen from the following example.

EXAMPLE 3. *Let* $G = (M, W, \Pi)$ *with* $M = \{x_1, x_2, m_1, m_2\}$, $W = \{y_1, y_2, w_1, w_2\}$, *and* $\Pi(x_1, y_1|m_1) = -3$, $\Pi(x_2, y_2|m_1) = -5$, $\Pi(x_1, y_1|m_2) = 2$, $\Pi(x_2, y_2|m_2) = 10$, $\Pi(m_1, w_1|m_1) = -4$, $\Pi(m_2, w_2|m_2) = -7$, *and* $\Pi(m_i, y_j|y_j) = -1$, $\Pi(x_i, w_j|x_i) = -1$, $\forall i, j \in \{1, 2\}$.

Even though there is a difference between neutral and pessimistic stable sets, the complexity of checking non-emptiness and membership for the two heuristics is the same.

THEOREM 3. *Checking nonemptiness of the pessimistic stable set is NP-hard.*

THEOREM 4. *Checking pessimistic stable set membership is coNP-complete.*

3.3 Optimistic Stability

When agents have optimistic attitudes, they believe that outcomes will always turn out in the best possible way. This positive outlook provides enough structure to characterize optimistic stable sets.

THEOREM 5. *Any matching in the optimistic stable set is a union of two disjoint matchings* $(M^{'}, W^{'}) \cup (M^{''}, W^{''})$*, where* $M^{'} \cup M^{''} = M$*,* $W^{'} \cup W^{''} = W$*, and every agent in* $(M^{'}, W^{'})$ *obtains their highest possible utility, while* $(M^{''}, W^{''})$ *is the empty matching.*

PROOF. Let $\mathcal{A}$ be a matching in the optimistic setwise-stable set and z an agent . Since $\mathcal{A}$ is stable, z is not blocking. Thus it must be that z is either unmatched, case in which he cannot deviate, or z is matched but already obtains the highest possible utility, and so has no incentive to deviate. Thus N can be partitioned in two subsets, $N^{'}$, the agents that obtain in $\mathcal{A}$ their maximal utility, and $N^{''}$, the unmatched agents. □

The next two results immediately follow from Theorem 5.

COROLLARY 1. *Assume that* $\Pi(m, w|i) < 0$ *for all* $m, w, i \in N$*. The unique candidate for the optimistic stable set is the empty matching.*

COROLLARY 2. *Assume that* $\Pi(m, w|i) \geq 0$ *for all* $m, w, i \in N$*. The only candidates for the optimistic stable set are the empty and complete matchings.*

Theorem 5 describes the structure of the optimistic stable set, if it is non-empty. Unfortunately it does not say anything about whether the set is non-empty, and like with neutral and pessimistic stable sets, it remains hard to determine nonemptiness of the optimistic stable set.

THEOREM 6. *Checking nonemptiness of the optimistic stable set is NP-complete.*

PROOF. First note that checking nonemptiness of the optimistic setwise-stable set is in NP. Given a matching $\mathcal{A}$, the utility of each agent $z \in N$ in $\mathcal{A}$ can be computed in $O(n^2)$. In addition, for every pair $(m, w) \in (M, W)$, we can again compute in $O(n^2)$ the best case utilities of m and w when matched with each other. Verifying if $\mathcal{A}$ is stable can be done by iterating over all pairs (m, w) and checking in $O(1)$ if both m and w can (weakly) improve by deviating under optimism, compared to their current utility in $\mathcal{A}$. The reduction is similar to that in Theorem 1. □

Unlike with neutral and pessimistic attitudes, checking membership under optimistic attitudes can be done in polynomial time.

THEOREM 7. *Checking membership to the optimistic stable set is in P.*

PROOF. Let $\mathcal{A}$ be a matching which can be blocked by a coalition B through some deviation $\mathcal{A}^{'}$. Assume there exists agent $z \in B$ which can strictly improve the optimistic estimation of utility only by cutting matches in $\mathcal{A}^{'}$. Then z can deviate alone by cutting the same matches as in $\mathcal{A}^{'}$ while expecting that the rest of the agents in B will perform deviation $\mathcal{A}^{'}$ (including initiating matches with z as an endpoint). Otherwise, any deviator which strictly improves utility forms a new match. Let z be such an agent and $(z, z^{'})$ a new edge in $\mathcal{A}^{'}$. Then coalition $\{z, z^{'}\}$ can block by forming the edge $(z, z^{'})$, since both agents z and $z^{'}$ expect that the edges in $\mathcal{A}^{'}$ will form after the deviation. □

3.4 Relationship Between Attitudes

We saw, through examples that the neutral and pessimistic stable sets were not the same, and the characterization of the optimistic stable set indicated that it also differed from the others. What was unclear was whether there is any direct relationship between the sets. In this section we explicitly answer that question in the affirmative. Given a many-to-many matching game G, denote by $\mathcal{O}$*-set*(G), $\mathcal{N}$*-set*(G), and $\mathcal{P}$*-set*(G) the optimistic, neutral, and pessimistic stable sets.

THEOREM 8. *Given any matching game* G*, the following inclusions hold* $\mathcal{O}\text{-}set(G) \subseteq \mathcal{N}\text{-}set(G) \subseteq \mathcal{P}\text{-}set(G)$*.*

PROOF. Let B denote a potentially blocking coalition. If a matching belongs to the optimistic set, then B cannot hope to improve even when the rest of the agents organize themselves in the best possible way for B. Thus the optimistic set is included in all the other stable sets. If a matching belongs to the pessimistic set, then B cannot improve when the rest of the society will punish them maximally for the deviation. Under the other stability concepts, B assumes a potentially better reaction from $N \setminus B$. Thus if B's deviation is not profitable in the optimistic or neutral scenarios, it is also not profitable in the pessimistic scenario. Hence the pessimistic set contains all the other sets. □

4. ONE-TO-ONE MATCHINGS

We analyze one-to-one matchings from a neutral, optimistic and pessimistic standpoint just as was done with the many-to-many matchings. However, before we proceed we comment on the choice of stability concept used for one-to-one matchings. In particular, in one-to-one matchings, the setwise stable set coincides exactly with the core since agents in a deviating group must necessarily cut all matches with non-members of the deviating group due to the constraint on the number of allowable partners. Thus, while we continue to use the term stable set to mean the setwise stable set, readers more comfortable with the core solution concept can apply it in this section. We also investigate pairwise stability and explicitly state we are looking at pairwise stability when warranted.

4.1 Neutral Stability

The hardness results in one-to-one matchings parallel the ones in the many-to-many setting.

THEOREM 9. *Checking neutral stable set membership is coNP-complete.*

THEOREM 10. *Checking nonemptiness of the neutral stable set is NP-hard.*

We also study pair-wise stability under the neutral assumption. We first note that there is a separation between the neutral pairwise stable set and the neutral setwise stable set.

EXAMPLE 4. *Let* $G = (M, W, \Pi)$*, where* $M = \{m_1, m_2\}$*,* $W = \{w_1, w_2\}$*, and* $\Pi(m_i, w_j|m_i) = \Pi(m_i, w_j|w_j) = -\varepsilon$*,*

$\Pi(m_i, w_j|z) = W \gg \varepsilon > 0$, *where* $z \in N \setminus \{m_i, w_j\}$, $\forall i, j \in \{1, 2\}$. *The neutral stable set of* G *is empty, while the "empty" matching is pairwise stable.*

While, in general, the neutral pairwise stable set may be empty, under certain conditions it is non-empty and we can compute a stable matching in polynomial time.

THEOREM 11. *A neutral pairwise stable matching can be computed in polynomial time when* $\Pi \geq 0$.

PROOF. Let $\mathcal{A}$ be the matching returned by running the Gale-Shapley algorithm by ignoring externalities, and assume by contradiction $\mathcal{A}$ is not stable. Then there exists deviation $(m, w^{'})$, where m, w are matched in $\mathcal{A}$ with $w^{'}, m^{'}$, respectively. Let $ext(m, \mathcal{A})$ denote the value obtained by m in $\mathcal{A}$ from externalities, and $E_m(w) = \Pi(m, w|m)$. From $\{m, w^{'}\}$ blocking, $u(m, \mathcal{A}) = \Pi(m, w|m) + ext(m, \mathcal{A}) < \Pi(m, w^{'}|m) + ext(m, \mathcal{A}) - \Pi(m^{'}, w^{'}|m)$ and $u(w^{'}, \mathcal{A}) = \Pi(m^{'}, w^{'}|w^{'}) + ext(w^{'}, \mathcal{A}) < \Pi(m, w^{'}|w^{'}) + ext(w^{'}, \mathcal{A}) - \Pi(m, w|w^{'})$. Equivalently, (i) $\Pi(m, w|m) < \Pi(m, w^{'}|m) - \Pi(m^{'}, w^{'}|m) \leq \Pi(m, w^{'}|m)$ and (ii) $\Pi(m^{'}, w^{'}|w^{'}) < \Pi(m, w^{'}|w^{'}) - \Pi(m, w|w^{'}) \leq \Pi(m, w^{'}|w^{'})$. From $(i), (ii)$, $E_m(w) < E_m(w^{'})$ and $E_{w^{'}}(m^{'}) < E_{w^{'}}(m)$, and so $(m, w^{'})$ is blocking under the preferences given by E. Contradiction, since $\mathcal{A}$ is stable on (M, W, E). $\square$

The assumption that Π is non-negative is necessary for the above result since the neutral pairwise-stable set can be empty when Π can be negative.

EXAMPLE 5. *Let* $G = (M, W, \Pi)$, *where* $M = \{m_1, m_2\}$, $W = \{w_1, w_2\}$, *and* Π *as follows:* $\Pi(m_i, w_j|m_i) = 1$ *and* $\Pi(m_i, w_j|w_j) = 1$, *for all* $i, j \in \{1, 2\}$, *and* $\Pi(m_i, w_j|z) = -1$, *for all* $i, j \in \{1, 2\}$ *and* $z \neq m_i, w_j$. *All the edges have positive values for their endpoints, and so the only candidates for neutral pairwise stability are:* $\mathcal{A}_1 = \{(m_1, w_1), (m_2, w_2)\}$ *and* $\mathcal{A}_2 = \{(m_1, w_2), (m_2, w_1)\}$. *However, matching* $\mathcal{A}_1$ *is blocked by the pair* (m_1, w_2), *and* $\mathcal{A}_2$ *is blocked by* (m_1, w_1).

4.2 Pessimistic Stability

For pessimistic stability, we obtain similar results to the ones for neutrality when considering setwise stability.

THEOREM 12. *Checking pessimistic stable set membership is coNP-complete.*

THEOREM 13. *Checking nonemptiness of the pessimistic stable set is NP-hard.*

Pessimistic pairwise stable matchings can be computed in polynomial time when $\Pi \geq 0$.

THEOREM 14. *A pessimistic pairwise stable set can be computed in polynomial time when* Π *is non-negative.*

PROOF. A neutral pairwise stable matching can be computed in polynomial time when $\Pi \geq 0$, and any matching satisfying neutral pairwise stability also satisfies pessimistic pairwise stability. $\square$

In one-to-one matchings with non-negative Π, we can also consider a *restricted notion of pessimism*. Namely, any matching containing singletons (unmatched agents) can weakly improve everyone's utility by pairing the singletons with each other. Based on this observation, the pessimistic deviators can have a less extreme attitude, and assume that while the rest of the agents may punish them for the deviation, they will not stay unmatched in order to do so. In other words, a blocking coalition, B, assumes that the agents in $N \setminus B$ form the worst possible matching for B, among all the matchings of size $\min(|M \cap (N \setminus B)|, |W \cap (N \setminus B)|)$.

```
Input: G = (M, W, Π)
Output: Stable matching with ties broken arbitrarily
forall the z ∈ N do
    if z ∈ M then
        Opp(z) = W
    else
        Opp(z) = M
    foreach t ∈ Opp(z) do
        M' ← M \ {z, t}
        W' ← W \ {z, t}
        foreach (m, w) ∈ M' × W' do
            ψ(m, w) ← Π(m, w|z)
        end
        // Compute A⁻, the worst case matching for z
        // when paired with t
        A⁻ ← Min-Matching(M', W', ψ)
        E⁻_z(t) ← Π(z, t|z)
        foreach (m, w) ∈ A⁻ do
            E⁻_z(t) ← E⁻_z(t) + ψ(m, w)
        end
    end
end
return Gale-Shapley(M, W, E⁻)
```

Algorithm 1: (Restricted) Pessimistic Pairwise Stability.

THEOREM 15. *A restricted pessimistic pairwise stable matching can be computed in polynomial time.*

PROOF. Let $\mathcal{A}$ be the matching returned by Algorithm 1 and assume by contradiction $\mathcal{A}$ is unstable under pessimistic pairwise stability. Then there exists deviating pair $(m, w^{'})$, where m, w are matched in $\mathcal{A}$ with $m^{'}, w^{'}$, respectively. Then it must be the case that for any possible matching $\mathcal{A}^{'}(m, w^{'})$ that includes the pair $(m, w^{'})$, both m and $w^{'}$ are better off in $\mathcal{A}^{'}(m, w^{'})$ than in $\mathcal{A}$. Equivalently, $E^{-}_{m}(w^{'}) > u(m, \mathcal{A}) \geq E^{-}_{m}(w)$ and $E^{-}_{w^{'}}(m) > u(w^{'}, \mathcal{A}) \geq E^{-}_{w^{'}}(m^{'})$. However, $\mathcal{A}$ is stable under E^{-}, and so for any $(m, w^{'}) \notin \mathcal{A}$, either $E^{-}_{m}(w^{'}) \leq E^{-}_{m}(w^{'})$ or $E^{-}_{w^{'}}(m^{'}) \leq E^{-}_{w^{'}}(m)$, contradiction. Thus $\mathcal{A}$ is stable. $\square$

4.3 Optimistic Stability

Since checking membership under optimistic attitudes was polynomial for many-to-many matchings, and one-to-one matchings are a special case, it trivially follows that checking membership under optimistic attitudes remains in P when considering one-to-one matchings. This continues to be true when we make the further restriction to pair-wise stability. Thus, we focus our attention to checking non-emptiness of stable sets under optimistic attitudes.

THEOREM 16. *Checking nonemptiness of the optimistic stable set is NP-complete, even if* $\Pi \geq 0$.

PROOF. First note that similarly to the optimistic setwise-stable set, checking nonemptiness of the optimistic stable set is in NP. Given a matching $\mathcal{A}$, the utility of each agent $z \in N$ in $\mathcal{A}$ can be computed in $O(n^2)$. In addition, for every pair $(m, w) \in (M, W)$, we can again compute in $O(n^2)$ the best case utilities of m and w when matched with each other. Verifying if $\mathcal{A}$ is stable can be done by iterating over all pairs (m, w) and checking in $O(1)$ if both m and w can (weakly) improve by deviating under optimism, compared to their current utility in $\mathcal{A}$. The same reduction as in Theorem 1 applies, by noting that the weights are constructed such that any feasible matching (from the perspective of the x_i and y_i agents) is one-to-one. □

For the next theorem, we consider the weak optimistic stable set. Weak stability requires that all members of a blocking coalition strictly improve their utility in order for the deviation to take place. We show that this is equivalent to pair-wise stability with optimistic attitudes.

THEOREM 17. *The optimistic pair-wise stable set is equivalent to the weak optimistic stable set.*

PROOF. Let $\mathcal{A}$ be a one-to-one matching. We show that if $\mathcal{A}$ has a blocking coalition B under optimism, then there exists a blocking singleton or pair. We consider two cases. If there exists $z \in B$ which improves strictly in $\mathcal{A}^{'}$ only by severing a match. Then agent z can deviate alone, by simply cutting the same match as in $\mathcal{A}^{'}$, and expecting (optimistically) that the rest of the agents will react as in $\mathcal{A}^{'}$. Otherwise, there exists $z \in B$ which improves strictly by forming a new match, say $(z, z^{'})$, and possibly severing an existing match. Then coalition $\{z, z^{'}\}$ is blocking. Thus if $\mathcal{A}$ does not belong to the weak optimistic stable set, it is also not pairwise stable under optimism. The reverse direction is clear, and so the optimistic pairwise stable set is equivalent to the weak optimistic stable set. □

COROLLARY 3. *Checking nonemptimess of the optimistic pairwise set is NP-complete, even when* $\Pi \geq 0$.

This is in contrast with neutral and pessimistic attitudes in one-to-one matchings. Recall that for neutrality and pessimism, Algorithm 1 computes a pairwise stable matching. In the case of neutrality, a pairwise stable matching can be computed by effectively ignoring externalities and running the Gale-Shapley algorithm on the preferences given by the values on the direct edges between every two agents. In the case of pessimism, each agent i attaches to every other agent j the worst possible value when paired with j. Computing a pairwise stable outcome based on the preferences given by these worst case values results in a stable matching.

On the other hand, in the case of optimism, while the agents *can* attach the best possible values to each other, a matching computed based on these best case values may be unstable. Some agents may actually not achieve their best case value when paired to their favourite other agent, due to externalities. Thus a stable matching should simultaneously satisfy more precise conditions, and verifying the existence of such a matching is not done in polynomial time.

5. DISCUSSION

In this work we introduced a compact model for matching problems with externalities, and looked at various stability concepts when agents held different attitudes concerning how they expected others to react to their actions. In particular, we analyzed both many-to-many matchings and one-to-one matchings when agents held neutral, pessimistic or optimistic attitudes. We studied the computational complexity of finding stable outcomes and provided both hardness results and polynomial algorithms where possible. Table 1 summarizes our findings.

The attitudes taken by the agents meaningfully change stable outcomes and influence the complexity of finding these outcomes. Using optimistic and pessimistic attitudes can really be viewed as determining what the best and worst possible reactions to a deviation and so can be seen as types of upper and lower bounds on the impact a reaction has on the deviating agents. The complexity results help illustrate how effective these heuristics may or may not be, if they were to be incorporated into algorithms searching for stable matchings.

There are several directions in which this work can be extended. First, alternative attitudes could be explored. However, we note that our notion of optimism is the simplest model under which checking non-emptiness in NP-hard. The non-emptiness question for refinements of optimism, such as having agents not expect others to offer them direct positive connections in response to a deviation, remains NP-hard. Additionally, membership can become hard under refinements where agents use more sophisticated reasoning. One could also explore alternative models of group stability. One obvious model stipulates that group deviations should be Pareto improving. However, checking the existence of a Pareto improving deviation is NP-hard, and so the refinement does not significant additional insight into the matching problem.[2]

We investigated many-to-many matchings and one-to-one matchings as these can be viewed as the two extremes of matching problems. Studying how agents' attitudes influence one-to-many matchings is an obvious next step. We hypothesize that the results would be very similar as to those presented here, but it would be interesting to see whether adding the restriction that $\Pi \geq 0$ would make the underlying matching problem easier as it did for one-to-one matchings under certain circumstances. Second, we assumed that all deviating agents maintained the same attitude, be it neutral, optimistic or pessimistic. It would be interesting, though challenging, to see what happens with hybrid models where, for example, a certain percentage of agents maintained a particular attitude.

Finally, there are other domains where externalities are important, such as in network formation games [10]. The investigation of real social networks could reveal interesting patterns in the externalities (do the values on the edges tend to be clustered in a certain way?). Even on platforms such as Facebook, agents are influenced by the matchings of others (friendships, subscriptions). Such cumulative effects can be expressed with additive externalities and insights from the data may suggest specialized instances of the weighted graph model, that tend to arise in the real world. Since such games can involve very large numbers of agents, investigating appropriate compact representations for externalities, computationally feasible stability concepts and appropriate models for the rationality bounds of the agents are all important and timely.

[2]This can be seen from the proof of Theorem 2. The grand coalition would have to find a Pareto improving matching, which coincides with a Knapsack solution.

		Neutral	Pessimistic	Optimistic
Many-to-Many Setwise Stability	Membership	coNP-complete	coNP-complete	P
	Non-Emptiness	NP-hard	NP-hard	NP-complete
One-to-One Setwise Stability	Membership	coNP-complete	coNP-complete	P
	Non-Emptiness	NP-hard	NP-hard	NP-complete
One-to-One Pairwise Stability ($\Pi \geq 0$)	Membership	P	P	P
	Non-Emptiness	P	P	NP-complete

Table 1: *Summary of complexity results.*

6. ACKNOWLEDGEMENTS

We would like to thank the anonymous AAMAS reviewers for valuable feedback.

Simina Brânzei was supported in part by the Sino-Danish Center for the Theory of Interactive Computation, funded by the Danish National Research Foundation and the National Science Foundation of China (under the grant 61061130540), and by the Center for research in the Foundations of Electronic Markets (CFEM), supported by the Danish Strategic Research Council. In addition, Simina also acknowledges support from the National Sciences and Engineering Research Council of Canada (NSERC). Kate Larson was supported by NSERC. Tomasz Michalak was supported by the European Research Council under Advanced Grant 291528 ("RACE").

7. REFERENCES

[1] H. Aziz, F. Brandt, and H. G. Seedig. Stable partitions in additively separable hedonic games. In *Proceedings of AAMAS*, pages 183–190, 2011.

[2] E. Bodine-Baron, C. Lee, A. Chong, B. Hassibi, and A. Wierman. Peer effects and stability in matching markets. In *Proceedings of SAGT*, 2011.

[3] D. Braziunas and C. Boutilier. Preference elicitation and generalized additive utility. In *Proceedings of AAAI*, pages 1573–1576, 2006.

[4] R. Burkard, M. Dell'Amico, and S. Martello. *Assignment Problems*. SIAM, 2009.

[5] B. Dutta and J. Masso. Stability of Matchings when Individuals have Preferences over Colleagues. *J. Econ. Theory*, 75:464–475, 1997.

[6] F. Echenique and J. Oviedo. A Theory of Stability in Many-to-Many Matching Markets. *Theoretical Economics*, 1(2):233–273, 2006.

[7] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979.

[8] G. Gigerenzer and R. Selten, editors. *Bounded rationality: The adaptive toolbox*. MIT Press, 2001.

[9] I. Hafalir. Stability of marriage with externalities. *Int. J. Game Theory*, 37:353–370, 2008.

[10] M. O. Jackson. *Social and Economic Networks*. Princeton University Press, 2008.

[11] B. Klaus and M. Walzl. Stable Many-to-Many Matchings with Contracts. *J. Math. Econ.*, 45(7-8):422–434, 2009.

[12] L. Koczy. A Recursive Core for Partition Function Form Games. *Theor. Decis.*, 63:41–51, 2007.

[13] L. Koczy. Sequential Coalition Formation and the Core in the Presence of Externalities. *Game. Econ. Behav.*, 66(1):559–565, 2009.

[14] A. E. Roth and M. O. Sotomayor. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Cambridge University Press, 1990.

[15] H. Sasaki and M. Toda. Two-Sided Marriage Problems with Externalities. *J. Econ. Theory*, 70(1):93–108, 1996.

8. APPENDIX

Additional proofs are provided for completeness.

THEOREM 3. *Checking nonemptiness of the pessimistic stable set is NP-hard.*

PROOF. The same reduction as in Theorem 1 applies, by noting that a coalition is blocking for the constructed matching under pessimistic reasoning if and only if it is blocking under neutral reasoning. □

THEOREM 9. *Checking neutral stable set membership is coNP-complete.*

PROOF. The same reduction as in Theorem 2 applies, by noting that the weights are set such that any feasible matching (from the perspective of the x_i and y_i agents) is one-to-one. That is, a coalition can block the matching constructed in Theorem 2 if and only if it is a one-to-one matching. □

THEOREM 10. *Checking nonemptiness of the neutral stable set is NP-hard.*

PROOF. The same reduction as in Theorem 1 applies, by noting that the weights are set up such that any feasible matching (from the perspective of the x_i and y_i agents) is one-to-one, and so a matching of the game in Theorem 1 is stable if and only if it is a one-to-one matching. □

THEOREM 12. *Checking pessimistic stable set membership is coNP-complete.*

PROOF. The reduction is similar to that of Theorem 1, except for the addition of several dummy agents to ensure that the grand coalition can always block in the one-to-one setting when the *Knapsack* instance has a solution. Given $I = \langle U, s, v, B, K \rangle$, let $G = (M, W, \Pi)$ such that $M = \{x_1, \ldots, x_{2n}, m_1, m_2\}$, $W = \{y_1, \ldots, y_{2n}, w\}$, and Π with non-zero entries:

- $\Pi(x_i, y_i|m_1) = \Pi(x_i, y_{n+i}|m_1) = -s(u_i)$ and $\Pi(x_i, y_i|w) = \Pi(x_i, y_{n+i}|w) = v(u_i)$, $\forall i \in \{1, \ldots, n\}$
- $\Pi(m_1, w|m_1) = -B$
- $\Pi(m_2, w|m_2) = K - \sum_{u_i \in U} v(u_i)$
- $\Pi(x_j, w|x_j) = -1$ and $\Pi(m_i, y_j|y_j) = -1$, $\forall i \in \{1, 2\}$ and $\forall j \in \{1, \ldots, 2n\}$

Let $\mathcal{A} = \{(m_1), (m_2), (w), (x_1), \ldots, (x_{2n}), (y_1), \ldots, (y_{2n})\}$. Similarly to Theorem 1, A belongs to the pessimistic stable set if and only if the *Knapsack* instance has a solution. □