

This is the accepted manuscript corresponding to the article:

Solving Stable Matching Problems using Answer Set Programming

Sofie De Clercq, Steven Schockaert, Martine De Cock, Ann Nowé

Theory and Practice of Logic Programming, volume 16, Special issue 03, pp. 247-268.

© Cambridge University Press

Please consult the final publication for the most recent version of this article.

The final publication is available via

http://journals.cambridge.org/abstract_S147106841600003X

Solving Stable Matching Problems using Answer Set Programming

SOFIE DE CLERCQ

Dept. of Applied Mathematics, Computer Science & Statistics, Ghent University, Belgium
(e-mail: SofieR.DeClercq@ugent.be)

STEVEN SCHOCKAERT

School of Computer Science & Informatics, Cardiff University, UK
(e-mail: SchockaertS1@cardiff.ac.uk)

MARTINE DE COCK

Center for Data Science, UW Tacoma, US
Dept. of Applied Mathematics, Computer Science & Statistics, Ghent University, Belgium
(e-mail: MDeCock@u.washington.edu)

ANN NOWE

Computational Modeling Lab, Vrije Universiteit Brussel, Belgium
(e-mail: ANowe@vub.ac.be)

submitted 1 January 2003; revised 1 January 2003; accepted 1 January 2003

Abstract

Since the introduction of the stable marriage problem (SMP) by Gale and Shapley (1962), several variants and extensions have been investigated. While this variety is useful to widen the application potential, each variant requires a new algorithm for finding the stable matchings. To address this issue, we propose an encoding of the SMP using answer set programming (ASP), which can straightforwardly be adapted and extended to suit the needs of specific applications. The use of ASP also means that we can take advantage of highly efficient off-the-shelf solvers. To illustrate the flexibility of our approach, we show how our ASP encoding naturally allows us to select optimal stable matchings, i.e. matchings that are optimal according to some user-specified criterion. To the best of our knowledge, our encoding offers the first exact implementation to find sex-equal, minimum regret, egalitarian or maximum cardinality stable matchings for SMP instances in which individuals may designate unacceptable partners and ties between preferences are allowed.

KEYWORDS: Answer Set Programming, Logic Rules, Stable Marriage Problem, Optimal Stable Matchings

1 Introduction

The stable marriage problem (SMP) is a well-known matching problem introduced by Gale and Shapley (1962). The input of an SMP instance consists of (i) a set of n men and n women, (ii) for each man a ranking of the women as preferred partners,

and (iii) for each woman a ranking of the men as preferred partners. A *blocking pair* of an SMP instance consists of a man and a woman who are in different marriages but both prefer each other to their actual partners. Given the problem, one can compute a *stable matching* or *stable set of marriages*, which is defined as a set of n couples (marriages) such that there are no blocking pairs. Due to its practical relevance, countless variants on the SMP have been investigated, enabling a wider range of applications. Examples of such applications include the kidney-exchange problem (Irving 2007), which matches donors in incompatible donor-recipient pairs to compatible recipients in other incompatible pairs and vice versa, and the hospital-resident problem (Manlove et al. 2002), which matches residents to the free positions in hospitals. In 2012, Roth and Shapley won the Nobel Prize for Economics for their theory of stable allocations and the practice of market design, which directly resulted from an application of the SMP.

In the literature, typically each time a new variant or generalization of the SMP or a different optimality criterion is considered, a new algorithm is developed; see e.g. (Gusfield 1987; Irving et al. 1987; McDermid and Irving 2012), or (Manlove 2013) for an overview. In this paper, we propose to use answer set programming (ASP) as a general vehicle for modeling a large class of extensions and variations of the SMP. We show how ASP encodings can be used to compute stable matchings, and how this encoding can be extended to compute optimal stable matchings. Although the SMP has been widely investigated, and efficient approximation or exact algorithms are available for several of its variants (Iwama et al. 2010; McDermid and Irving 2012), to the best of our knowledge, our encoding offers the first exact implementation to find sex-equal, minimum regret, egalitarian or maximum cardinality stable matchings for SMP instances with unacceptable partners and ties (see below).

In this paper, we will consider two well-known adaptations of the SMP. First, we will consider problem instances in which every person can specify a set of unacceptable partners. The second alteration consists of allowing ties in the preferences, i.e. one can be indifferent between some possible partners. In the literature, the SMP variant with unacceptable partners – or, equivalently, with incomplete preference lists – is abbreviated as SMI. The variant with ties is denoted as SMT and the variant which allows both extensions as SMTI. Note that the original SMP is a special case of the SMTI, i.e. the set of unacceptable partners is empty for each man and woman, and there are no ties. Therefore, our paper focusses on the SMTI variant, as it is the most general one.

Another way to generalize the SMP is by introducing optimality criteria for stable matchings. This is motivated by the fact that, if multiple stable matchings exist, some may be more interesting than others. In this paper, we focus on sex-equality, minimum regret, egalitarianity and maximum cardinality of the stable matchings, as these are commonly investigated optimality criteria in the matching literature. Note, however, that there exist several other optimality criteria in the context of matchings, such as *popularity* (Gärdenfors 1975), *Pareto optimality* (Gale and Sotomayor 1985; Roth and Sotomayor 1990) or profile-based notions such as rank-maximum (Irving 2003), greedy maximum and generous maximum (Irving

2003). As this list is non-exhaustive, we refer the interested reader to (Manlove 2013) for an overview.

The structure of the paper is as follows. First we give some background about the SMP and ASP in Section 2. Then we introduce our encoding of the SMTI with ASP and prove its correctness in Section 3.1. To illustrate the flexibility of the approach, we show how it can be used to tackle three-dimensional stable matching problems in Section 3.2. In Section 4, we first discuss several optimality criteria and then extend the encoding from Section 3.1, enabling us to find optimal stable matchings. We show how optimal stable matchings of an SMTI instance can be found by solving the corresponding induced disjunctive ASP program and prove the soundness of our approach. This paper is an extended version of (De Clercq et al. 2013) and additionally provides detailed examples, complete correctness proofs and an ASP encoding of the three-dimensional stable matching problem. The three-dimensional stable matching problem is very important for practical applications, such as the kidney exchange program (Biró and McDermid 2010).

2 Background

2.1 The Stable Marriage Problem

To solve the original SMP, Gale and Shapley (Gale and Shapley 1962) constructed an iterative algorithm —known as the Gale-Shapley algorithm, G-S algorithm or deferred-acceptance algorithm— to compute a particular stable matching of an SMP instance. The algorithm works as follows: in round 1 every man proposes to his first choice of all women. A woman, when being proposed, then rejects all men but her first choice among the subset of men who proposed to her. That first choice becomes her temporary husband. In the next rounds, all rejected men propose to their first choice among the subset of women by whom they were not rejected yet, regardless of whether this woman already has a temporary husband. Each woman, when being proposed, then rejects all men but her first choice among the subset of men who just proposed to her and her temporary mate. This process continues until all women have a husband. This point, when everyone has a partner, is always reached after a polynomial number of steps and the corresponding set of marriages is stable (Gale and Shapley 1962). It should be noted, however, that only one of the potentially exponentially many stable matchings is found in this way.

The classical SMP can be generalized by (i) allowing men and women to point out unacceptable partners, i.e. exclude them from their preference list and (ii) dropping the restriction that the number of men n equals the number of women p . In this variant, men and women can remain single in a stable matching. Intuitively, one prefers remaining single over being matched with an unacceptable partner. This variant is also referred to as the SMP with incomplete preference lists, abbreviated as SMI. A stable matching for an SMI instance always exists and can be found in polynomial time (Roth and Sotomayor 1990) by a slightly modified G-S algorithm. As we focus on an extension of the SMI, we refer to the electronic Appendix A for the formal definitions of the classical SMP and the SMI.

The SMI variant can further be generalized by additionally allowing ties in the preference lists. For this variant (SMTI) there are several ways to define stability, but we will use the notion of weak stability (Irving 1994). We denote a set of men as $M = \{m_1, \dots, m_n\}$ and a set of women $W = \{w_1, \dots, w_p\}$. A set of marriages or a matching is a collection of man-woman pairs and singles (persons paired to themselves) such that every man and every woman occurs in just one pair.

Definition 1 (SMTI)

An instance of the SMTI is a pair (S_M, S_W) with $S_M = \{\sigma_M^1, \dots, \sigma_M^n\}$ and $S_W = \{\sigma_W^1, \dots, \sigma_W^p\}$. For every $i \in \{1, \dots, n\}$, σ_M^i is a list of disjoint subsets of $\{1, \dots, p\}$. Symmetrically σ_W^i is a list of disjoint subsets of $\{1, \dots, n\}$ for every $i \in \{1, \dots, p\}$. We call σ_M^i and σ_W^i the preferences of man m_i and woman w_i respectively and we denote the length of the list σ_M^i as $|\sigma_M^i|$. If $k \in \sigma_M^i(j)$, woman w_k is in man m_i 's j^{th} most preferred group of women. All the women in that group are equally preferred by m_i . The case $k \in \sigma_W^i(j)$ is similar. If there is no l such that $j \in \sigma_M^i(l)$, woman w_j is an unacceptable partner for man m_i , and similarly when there is no l such that $j \in \sigma_W^i(l)$. For every k in the set $\sigma_M^i(|\sigma_M^i|)$, man m_i equally prefers staying single to being paired to woman w_k , and symmetrically for the preferences of a woman w_i . This is the only set in σ_M^i that might be empty, and similar for σ_W^i . Man m and woman w form a blocking pair in a set of marriages S if m strictly prefers w to his partner in S and w strictly prefers m to her partner in S . A blocking individual in S is a person who strictly prefers being single to being paired to his partner in S . A weakly stable matching is a set of marriages without blocking pairs or individuals.

A weakly stable matching always exists for an instance of the SMTI and it can be found in polynomial time by arbitrarily breaking the ties (Iwama and Miyazaki 2008). However, as opposed to the SMI variant, the number of matched persons is no longer constant for every stable matching in this variant.

We introduce the following notations:

$$\begin{aligned} \text{acceptable}_M^i &= \underbrace{\sigma_M^i(1) \cup \sigma_M^i(2) \cup \dots \cup \sigma_M^i(|\sigma_M^i| - 1)}_{\text{preferred}_M^i} \cup \underbrace{\sigma_M^i(|\sigma_M^i|)}_{\text{neutral}_M^i} \\ &= \text{preferred}_M^i \cup \text{neutral}_M^i \end{aligned}$$

Furthermore $\text{unacceptable}_M^i = \{1, \dots, p\} \setminus \text{acceptable}_M^i$. We define the ordering $\leq_M^{m_i}$ on $\{w_j \mid j \in \text{acceptable}_M^i\} \cup \{m_i\}$ as $x \leq_M^{m_i} y$ iff m_i prefers person x at least as much as person y . Note that m_i is included in its own preference ordering to encode the possibility of staying single. The strict ordering $<_M^{m_i}$ is defined in the obvious way and analogous notations are used for σ_W^j .

Example 1

Suppose $M = \{m_1, m_2\}$, $W = \{w_1, w_2, w_3, w_4\}$ and $S_M = \{\sigma_M^1 = (\{1, 3\}, \{4\}), \sigma_M^2 = (\{2, 3\}, \{\})\}$. Hence man m_1 prefers women w_1 and w_3 to woman w_4 . There is a tie between woman w_1 and w_3 as well as between woman w_4 and staying single. Woman w_2 is unacceptable for man m_1 . Man m_2 prefers woman w_2 and w_3 to staying single, but finds w_1 and w_4 unacceptable. It holds that $w_1 <_M^{m_1} m_1$, i.e. m_1 prefers marrying w_1 over staying single, $\text{acceptable}_M^1 = \{1, 3, 4\}$, $\text{preferred}_M^1 = \{1, 3\}$, $\text{neutral}_M^1 = \{4\}$ and $\text{unacceptable}_M^1 = \{2\}$.

2.2 Answer Set Programming

Answer set programming or ASP is a form of declarative programming (Brewka et al. 2011). Its transparency, elegance and ability to deal with Σ_2^P -complete problems make it an attractive method for solving combinatorial search and optimization problems. An ASP program is a finite collection of first-order rules

$$A_1 \vee \dots \vee A_k \leftarrow B_1, \dots, B_m, \text{not } C_1, \dots, \text{not } C_n$$

where $A_1, \dots, A_k, B_1, \dots, B_m, C_1, \dots, C_n$ are predicates, possibly negated by \neg , and *not* is the negation-as-failure operator, whose meaning is explained below. The semantics are defined by the *ground version* of the program, consisting of all ground instantiations of the rules w.r.t. the constants that appear in it (see e.g. (Brewka et al. 2011) for a good overview). This grounded program is a propositional ASP program. The building blocks of these programs are *atoms*, *literals* and *rules*. The most elementary are *atoms*, which are propositional variables that can be true or false. A *literal* is an atom or a negated atom, denoted with \neg . Beside strong negation, ASP uses a special kind of negation, namely *negation-as-failure* (naf), denoted with ‘*not*’. For a literal a we call ‘*not a*’ the naf-literal associated with a . The *extended literals* consist of all literals and their associated naf-literals. A *disjunctive rule* has the following form

$$a_1 \vee \dots \vee a_k \leftarrow b_1, \dots, b_m, \text{not } c_1, \dots, \text{not } c_n$$

where $a_1, \dots, a_k, b_1, \dots, b_m, c_1, \dots, c_n$ are literals from a fixed set \mathcal{L} , determined by a fixed set \mathcal{A} of atoms. We call $a_1 \vee \dots \vee a_k$ the head of the rule while the set of extended literals $b_1, \dots, b_m, \text{not } c_1, \dots, \text{not } c_n$ is called the *body*. The rule above intuitively encodes that a_1, a_2, \dots or a_k is true when we have evidence that b_1, \dots, b_m are true and we have no evidence that at least one of c_1, \dots, c_n is true. When a rule has an empty body, we call it a *fact*; when the head is empty, we speak of a *constraint*. A rule without occurrences of *not* is called a *simple disjunctive rule*. A *simple disjunctive* ASP program is a finite collection of simple disjunctive rules and similarly a *disjunctive* ASP program \mathcal{P} is a finite collection of disjunctive rules. If each rule head consists of at most one literal, we speak of a *normal* ASP program.

We define an *interpretation* I of a disjunctive ASP program \mathcal{P} as a subset of \mathcal{L} . An interpretation I *satisfies* a simple disjunctive rule $a_1 \vee \dots \vee a_k \leftarrow b_1, \dots, b_m$ when $a_1 \in I$ or $a_2 \in I$ or \dots or $a_k \in I$ or $\{b_1, \dots, b_m\} \not\subseteq I$. An interpretation which satisfies all rules of a simple disjunctive program is called a *model* of that program. An interpretation I is an *answer set* of a simple disjunctive program \mathcal{P} iff it is a minimal model of \mathcal{P} , i.e. no strict subset of I is a model of \mathcal{P} (Gelfond and Lifschitz 1988). The *reduct* \mathcal{P}^I of a disjunctive ASP program \mathcal{P} w.r.t. an interpretation I is defined as the simple disjunctive ASP program $\mathcal{P}^I = \{a_1 \vee \dots \vee a_k \leftarrow b_1, \dots, b_m \mid (a_1 \vee \dots \vee a_k \leftarrow b_1, \dots, b_m, \text{not } c_1, \dots, \text{not } c_n) \in \mathcal{P}, \{c_1, \dots, c_n\} \cap I = \emptyset\}$. An interpretation I of a disjunctive ASP program \mathcal{P} is an answer set of \mathcal{P} iff I is an answer set of \mathcal{P}^I .

Example 2

Let \mathcal{P} be the ASP program with the following 4 rules:

$$\begin{aligned} & man(john) \leftarrow, \quad person(john) \leftarrow, \quad person(fiona) \leftarrow \\ & woman(X) \vee child(X) \leftarrow person(X), not\ man(X) \end{aligned}$$

The first 3 rules are facts; hence their heads will be in any answer set. The fourth rule encodes that any person who is not a man, is a woman or child. The latter rule is grounded to 2 rules in which X is resp. replaced by *john* and *fiona*. We check that the interpretation $I = \{man(john), woman(fiona), person(john), person(fiona)\}$ is an answer set of the ground version of \mathcal{P} by computing the reduct \mathcal{P}^I . As the grounded rule with $X = john$ is deleted since $man(john)$ is in I , \mathcal{P}^I is:

$$\begin{aligned} & man(john) \leftarrow, \quad person(john) \leftarrow, \quad person(fiona) \leftarrow \\ & woman(fiona) \vee child(fiona) \leftarrow person(fiona) \end{aligned}$$

It is clear that I is a minimal model of this simple program, so I is an answer set of \mathcal{P} . By replacing $woman(fiona)$ by $child(fiona)$ in I , another answer set is obtained.

To automatically compute the answer sets of the programs in this paper, we have used the ASP solver DLV (www.dlvsystem.com), due to its ability to handle predicates, disjunction and numeric values, with built-in aggregate functions (Faber et al. 2008). The numeric values are only used for grounding.

3 Modeling the Stable Marriage Problem in ASP

3.1 Modeling the SMTI in ASP

In this section we model the SMTI, using ASP. A few proposals of using non-monotonic reasoning for modeling the SMP have already been described in the literature. For instance, in (Marek et al. 1990) a specific variant of the SMP is mentioned (in which boys each know a subset of a set of girls and want to be matched to a girl they know) and in (Dung 1995) an abductive program is used to find a stable matching of marriages in which two fixed persons are paired, with strict, complete preference lists. To the best of our knowledge, beyond a few specific examples, no comprehensive study has been made of using ASP or related paradigms in this context. In particular, the generality of our ASP framework for weakly stable matchings of SMTI instances allows to easily adjust the encoding to variants of the SMP, such as the stable roommate problem (Gale and Shapley 1962), in which matches need to be found within one group instead of between two groups, or the three-dimensional stable matching problem (Ng and Hirschberg 1991), which matches triples between three groups instead of pairs between two.

The expression $accept(m, w)$ denotes that a man m and a woman w accept each other as partners. The predicate $manpropose(m, w)$ expresses that man m is willing to propose to woman w and analogously $womanpropose(m, w)$ expresses that woman w is willing to propose to man m . Inspired by the Gale-Shapley algorithm, we look for an ASP formalization to find the stable matchings.

Definition 2 (ASP program induced by SMTI)

The ASP program \mathcal{P} induced by an instance $(\{\sigma_M^1, \dots, \sigma_M^n\}, \{\sigma_W^1, \dots, \sigma_W^p\})$ of the SMTI is the program containing for every $i \in \{1, \dots, n\}, j \in \{1, \dots, p\}$ the following rules:

$$\text{accept}(m_i, w_j) \leftarrow \text{manpropose}(m_i, w_j), \text{womanpropose}(m_i, w_j) \quad (1)$$

$$\text{accept}(m_i, m_i) \leftarrow \{\text{not accept}(m_i, w_k) \mid k \in \text{acceptable}_M^i\} \quad (2)$$

$$\text{accept}(w_j, w_j) \leftarrow \{\text{not accept}(m_k, w_j) \mid k \in \text{acceptable}_W^j\} \quad (3)$$

and for every $i \in \{1, \dots, n\}, j \in \text{acceptable}_M^i$:

$$\text{manpropose}(m_i, w_j) \leftarrow \{\text{not accept}(m_i, x) \mid x \leq_M^{m_i} w_j \text{ and } w_j \neq x\} \quad (4)$$

and for every $j \in \{1, \dots, p\}, i \in \text{acceptable}_W^j$:

$$\text{womanpropose}(m_i, w_j) \leftarrow \{\text{not accept}(x, w_j) \mid x \leq_W^{w_j} m_i \text{ and } m_i \neq x\} \quad (5)$$

Intuitively (1) means that a man and woman accept each other as partners if they propose to each other. Due to (2), a man accepts himself as a partner (i.e. stays single) if no woman in his preference list is prepared to propose to him. Rule (4) states that a man proposes to a woman if he is not paired to a more or equally preferred woman. For $j \in \text{neutral}_M^i$ the body of (4) contains $\text{not accept}(m_i, m_i)$. No explicit rules are stated about the number of persons someone can propose to or accept but as we will see below, in Proposition 1, this is unnecessary. Note that, for $k = \max(n, p)$, the number of grounded rules in the induced ASP program is $\mathcal{O}(k^2)$. We now illustrate our approach with an example.

Example 3

Consider the following instance (S_M, S_W) of the SMTI. Let $M = \{m_1, m_2\}$ and $W = \{w_1, w_2, w_3\}$. Furthermore:

$$\begin{aligned} \sigma_M^1 &= (\{1\}, \{2, 3\}, \{\}) & \sigma_W^1 &= (\{1, 2\}, \{\}) \\ \sigma_M^2 &= (\{2\}, \{1\}) & \sigma_W^2 &= (\{1\}, \{\}) \\ & & \sigma_W^3 &= (\{2\}, \{1\}, \{\}) \end{aligned}$$

The ASP program induced by this SMTI instance is:

$$\begin{aligned} \text{accept}(X, Y) &\leftarrow \text{manpropose}(X, Y), \text{womanpropose}(X, Y) \\ \text{manpropose}(m_1, w_1) &\leftarrow \\ \text{manpropose}(m_1, w_2) &\leftarrow \text{not accept}(m_1, w_1), \text{not accept}(m_1, w_3) \\ \text{manpropose}(m_1, w_3) &\leftarrow \text{not accept}(m_1, w_1), \text{not accept}(m_1, w_2) \\ \text{accept}(m_1, m_1) &\leftarrow \text{not accept}(m_1, w_1), \text{not accept}(m_1, w_2), \\ &\quad \text{not accept}(m_1, w_3) \\ \text{manpropose}(m_2, w_2) &\leftarrow \\ \text{manpropose}(m_2, w_1) &\leftarrow \text{not accept}(m_2, w_2), \text{not accept}(m_2, m_2) \\ \text{accept}(m_2, m_2) &\leftarrow \text{not accept}(m_2, w_2), \text{not accept}(m_2, w_1) \\ \text{womanpropose}(m_1, w_1) &\leftarrow \text{not accept}(m_2, w_1) \end{aligned}$$

$$\begin{aligned}
& \text{womanpropose}(m_2, w_1) \leftarrow \text{not } \text{accept}(m_1, w_1) \\
& \quad \text{accept}(w_1, w_1) \leftarrow \text{not } \text{accept}(m_1, w_1), \text{not } \text{accept}(m_2, w_1) \\
& \text{womanpropose}(m_1, w_2) \leftarrow \\
& \quad \text{accept}(w_2, w_2) \leftarrow \text{not } \text{accept}(m_1, w_2) \\
& \text{womanpropose}(m_2, w_3) \leftarrow \\
& \text{womanpropose}(m_1, w_3) \leftarrow \text{not } \text{accept}(m_2, w_3) \\
& \quad \text{accept}(w_3, w_3) \leftarrow \text{not } \text{accept}(m_1, w_3), \text{not } \text{accept}(m_2, w_3)
\end{aligned}$$

If we run the program in DLV, we get three answer sets, containing respectively:

- $\{\text{accept}(m_1, w_3), \text{accept}(m_2, w_1), \text{accept}(w_2, w_2)\}$,
- $\{\text{accept}(m_1, w_2), \text{accept}(m_2, w_1), \text{accept}(w_3, w_3)\}$,
- $\{\text{accept}(m_1, w_1), \text{accept}(m_2, m_2), \text{accept}(w_2, w_2), \text{accept}(w_3, w_3)\}$.

These answer sets correspond to the three weakly stable matching of marriages of this SMTI instance, namely $\{(m_1, w_3), (m_2, w_1), (w_2, w_2)\}$, $\{(m_1, w_2), (m_2, w_1), (w_3, w_3)\}$ and $\{(m_1, w_1), (m_2, m_2), (w_2, w_2), (w_3, w_3)\}$.

The following proposition states that our ASP encoding is sound, i.e. that there is a bijective correspondence between the answer sets of the induced program and the weak stable matchings of the SMTI. The complete proof is provided in the electronic Appendix C.

Proposition 1

Let (S_M, S_W) be an instance of the SMTI and let \mathcal{P} be the corresponding ASP program. If I is an answer set of \mathcal{P} , then a weakly stable matching for (S_M, S_W) is given by $\{(x, y) \mid \text{accept}(x, y) \in I\}$. Conversely, if $\{(x_1, y_1), \dots, (x_k, y_k)\}$ is a weakly stable matching for (S_M, S_W) then \mathcal{P} has the following answer set I :

$$\begin{aligned}
& \{ \text{manpropose}(x_i, y) \mid i \in \{1, \dots, k\}, x_i \in M, y <_M^{x_i} y_i \vee y = y_i \neq x_i \} \\
& \cup \{ \text{womanpropose}(x, y_i) \mid i \in \{1, \dots, k\}, y_i \in W, x <_W^{y_i} x_i \vee x = x_i \neq y_i \} \\
& \cup \{ \text{accept}(x_i, y_i) \mid i \in \{1, \dots, k\} \}
\end{aligned}$$

A pair (m, w) is *stable* if there exists a stable matching that contains (m, w) . In (Manlove et al. 2002) it is shown that the decision problem ‘is the pair (m, w) stable?’ for a given SMTI instance is an NP-complete problem, even in the absence of unacceptability. It is straightforward to see that we can reformulate this decision problem as ‘does there exist an answer set of the induced normal ASP program \mathcal{P} which contains the literal $\text{accept}(m, w)$?’ (i.e. brave reasoning), which is known to be an NP-complete problem (Baral 2003). Thus our model forms a suitable framework for these kind of decision problems concerning the SMTI.

3.2 Modeling the 3-Dimensional SMTI in ASP

To illustrate further the flexibility of our ASP approach, we consider a variant of the SMP and show how small adaptations of the ASP encoding can solve this variant. Extending the SMP by adding another dimension to the problem was first proposed

in (Knuth 1976). We work out the three-dimensional SMTI, where n men are to be matched with p women and r children. Definition 1 can straightforwardly be generalized to a three-dimensional instance (S_M, S_W, S_C) , in which preference lists of the men are defined over the set of woman-child pairs and similarly the women have preferences over man-child pairs and the children have preferences over man-woman pairs. In the three-dimensional case, σ_M^i becomes a list of disjoint subsets of $\{1, \dots, p\} \times \{1, \dots, r\}$, and analogously for σ_W^j and σ_C^k . Similarly as before, we can define the notions of $acceptable_M^i$ etc. Note that the orderings $\leq_M^{m_i}$, $\leq_W^{w_j}$ and $\leq_C^{c_k}$ are resp. defined on pairs in $W \times C$, $M \times C$ and $M \times W$. A stable matching is now defined as a set of man-woman-child triples and singles, with the properties that no man, woman and child can be found such that each of them prefers the pair formed by the others above their current mates in the matching and no person prefers being single to being matched with its current mates. The practical relevance of this problem is pointed out in (Biró and McDermid 2010).

Extending the ASP program from Definition 2, we can write an ASP program induced by an instance of the three-dimensional matching problem.

Definition 3 (ASP program induced by 3D SMTI)

The ASP program \mathcal{P} induced by an instance $(\{\sigma_M^1, \dots, \sigma_M^n\}, \{\sigma_W^1, \dots, \sigma_W^p\}, \{\sigma_C^1, \dots, \sigma_C^r\})$ of the 3-dimensional SMTI is the program containing the following rules for every $i \in \{1, \dots, n\}, j \in \{1, \dots, p\}$ and $k \in \{1, \dots, r\}$:

$$\begin{aligned} accept(m_i, w_j, c_k) &\leftarrow manprop(m_i, w_j, c_k), womprop(m_i, w_j, c_k), \\ &\quad childprop(m_i, w_j, c_k) \\ accept(m_i, m_i, m_i) &\leftarrow \{not\ accept(m_i, w_u, c_v) \mid (u, v) \in acceptable_M^i\} \\ accept(w_j, w_j, w_j) &\leftarrow \{not\ accept(m_u, w_j, c_v) \mid (u, v) \in acceptable_W^j\} \\ accept(c_k, c_k, c_k) &\leftarrow \{not\ accept(m_u, w_v, c_k) \mid (u, v) \in acceptable_C^k\} \end{aligned}$$

and for every $i \in \{1, \dots, n\}$ and $(j, k) \in acceptable_M^i$:

$$manprop(m_i, w_j, c_k) \leftarrow \{not\ accept(m_i, x, y) \mid (x, y) \leq_M^{m_i} (w_j, c_k); (w_j, c_k) \neq (x, y)\}$$

and for every $j \in \{1, \dots, p\}$ and $(i, k) \in acceptable_W^j$:

$$womprop(m_i, w_j, c_k) \leftarrow \{not\ accept(x, w_j, y) \mid (x, y) \leq_W^{w_j} (m_i, c_k); (m_i, c_k) \neq (x, y)\}$$

and for every $k \in \{1, \dots, r\}$ and $(i, j) \in acceptable_C^k$:

$$childprop(m_i, w_j, c_k) \leftarrow \{not\ accept(x, u, c_k) \mid (x, y) \leq_C^{c_k} (m_i, w_j); (m_i, w_j) \neq (x, y)\}$$

Ng and Hirschberg (Ng and Hirschberg 1991) proved that deciding whether a stable matching exists for the three-dimensional problem – in the absence of unacceptability and ties – is an NP-complete problem. Completely analogously as for the two-sided SMP, one can prove that this encoding yields a bijective correspondence between the answer sets of the ASP program and the stable matchings of the three-dimensional matching problem. Note that, for $k = \max(n, p, r)$, the number of grounded rules in the induced ASP program is $\mathcal{O}(k^3)$.

4 Selecting Optimal Stable Matchings

4.1 Notions of Optimality of Stable Matchings

When several stable matchings can be found for an instance of the SMP, some may be more interesting than others. The stable matching found by the G-S algorithm is *M-optimal* (Roth and Sotomayor 1990), i.e. every man likes this set at least as well as any other stable matching. Exchanging the roles of men and women in the G-S algorithm yields a *W-optimal* stable matching (Gale and Shapley 1962), optimal from the women's point of view.

While some applications may require us to favour either the men or the women, in others it makes more sense to treat both parties equally. To formalize some commonly considered notions of fairness and optimality w.r.t. the SMP, we define the cost $c_x(y)$ of being matched with y for an individual x , where $c_x(y) = k$ if y is x 's k^{th} preferred partner. More precisely, for a man m_i , we define $c_{m_i}(y) = |\{z : z <_M^{m_i} y\}| + 1$ for every $y \in \text{acceptable}_M^i$; for a woman w_j , c_{w_j} is defined analogously. So in case of ties we assign the same list position to equally preferred partners, as illustrated in Example 4.

Example 4

Let $x = m_1$ be a man with preference list $\sigma_M^1 = (\{1\}, \{2, 3\}, \{4\})$ then $c_x(w_1) = 1$, $c_x(w_2) = c_x(w_3) = 2$ and w_4 yields $c_x(w_4) = 4$. The cost for being single would be 4, i.e. $c_x(m_1) = 4$, since m_1 prefers women w_1, w_2 and w_3 to being single, but is indifferent between being paired to w_4 or staying single.

Definition 4 (Optimal Stable Matchings)

Let S be a set of marriages and let $S(x)$ denote the partner of x in S .

- The sex-equality cost of S is $c_{\text{sexeq}}(S) = |\sum_{x \in M} c_x(S(x)) - \sum_{x \in W} c_x(S(x))|$,
- the egalitarian cost of S is $c_{\text{weight}}(S) = \sum_{x \in M \cup W} c_x(S(x))$,
- the regret cost of S is $c_{\text{regret}}(S) = \max_{x \in M \cup W} c_x(S(x))$, and
- the cardinality cost of S is $c_{\text{singles}}(S) = |\{z : (z, z) \in S\}|$.

S is a sex-equal stable matching iff S is a stable matching with minimal sex-equality cost. Similarly, S is an egalitarian (resp. minimum regret, maximum cardinality) stable matching iff S is a stable matching with minimal egalitarian (resp. regret or cardinality) cost.

A *sex-equal stable matching* assigns an equal importance to the preferences of the men and women, i.e. the men are as pleased with the matching as the women. An *egalitarian stable matching* is a stable matching in which the preferences of every individual are considered to be equally important, i.e. it minimizes the difference in happiness of all the men and women. In (Xu and Li 2011) the use of an egalitarian stable matching is proposed to optimally match virtual machines (VM) to servers in order to improve cloud computing by equalizing the importance of migration overhead in the data center network and VM migration performance. A *minimum regret stable matching* is optimal for the person who is worst off, i.e. there does not exist a stable matching such that the person who is most displeased with the

matching is happier than the most displeased person in the minimum regret stable matching. A *maximal or minimal cardinality stable matching* is a stable matching with resp. as few or as many singles as possible. Examples of practical applications include an efficient kidney exchange program (Roth et al. 2005), which matches donors of incompatible pairs to recipients of other incompatible pairs and vice versa, and the National Resident Matching Program (www.nrmp.org) (Manlove et al. 2002), which matches healthcare professionals to graduate medical education and advanced training programs. Maximizing cardinality guarantees that as many recipients as possible will get a compatible donor and as many healthcare professionals as possible will get a position.

Remark 1

It might be somewhat confusing that the term *utilitarian* is more frequently used in sociological and economical contexts for an optimization of the overall happiness (here called egalitarian), while *egalitarian* is more used for an optimization which minimizes the unhappiness of the individuals (tending more to minimum regret). However, we will use the cost terms as defined above, since these are standard in the context of the SMP.

For an overview of literature results concerning the computational complexity of finding optimal stable matchings in the SMP, SMI, SMT and SMTI, we refer to the electronic Appendix B.

Note that other notions of preferred matchings have been described in the literature, such as *popularity* (Gärdenfors 1975), *Pareto optimality* (Gale and Sotomayor 1985; Roth and Sotomayor 1990) or profile-based notions such as rank-maximum (Irving 2003), greedy maximum and generous maximum (Irving 2003). For more details on these and other optimality criteria, we refer the interested reader to (Manlove 2013) for an overview.

4.2 Finding Stable Matchings using Disjunctive Naf-free ASP

As we discuss in Section 4.3, we can extend our ASP encoding of the SMTI such that the optimal stable matchings correspond to the answer sets of an associated ASP program. In particular, we use the saturation technique (Eiter and Gottlob 1995) to filter non-optimal answer sets. Intuitively, the idea is to create a program with 3 components: (i) a first part describing the solution candidates, (ii) a second part also describing the solution candidates since comparison of solutions requires multiple solution candidates within the same answer set whereas the first part in itself produces one solution per answer set, (iii) a third part comparing the solutions described in the first two parts and selecting the preferred solutions by saturation. It is, however, known that the presence of negation-as-failure can cause problems when applying saturation. This is due to the fact that rules containing naf-literals can be altered in the reduct. To address this issue, we use saturation in combination with a disjunctive naf-free ASP program instead of the ASP program in Definition 2. To this end, we use a SAT encoding (Janhunen 2004) of the ASP program in

Definition 2 and define a disjunctive naf-free ASP program in Definition 5 which selects particular models of the SAT problem.

Note that our original normal program is absolutely tight, i.e. there is no infinite sequence l_1, l_2, \dots of literals such that for every i there is a program rule for which l_{i+1} is a positive body literal and l_i is in the head (Erdem and Lifschitz 2003). We use the completion to derive an ASP encoding for finding optimal stable matchings. The completion of a normal ASP program is a set of propositional formulas. For every atom a with $a \leftarrow body_i$ ($i \in \{1, \dots, k\}$) all the program rules with head a , the propositional formula $a \equiv body'_1 \vee \dots \vee body'_k$ is in the completion of that program, where $body'_i$ is the conjunction of literals derived from $body_i$ by replacing every occurrence of ‘not’ with ‘ \neg ’. If an atom a of the program does not occur in any rule head, then $a \equiv \perp$ is in the completion of the program. Similarly the completion of the program contains the propositional formula $\perp \equiv body'_1 \vee \dots \vee body'_l$ where the disjunction extends over all the program constraints $\leftarrow body_i$ ($i \in \{1, \dots, l\}$). Because our program is absolutely tight, we know that every propositional model of the completion will correspond to an answer set of the original program and vice versa (Erdem and Lifschitz 2003). When applied to the induced normal ASP program in Definition 2, the completion contains the following formulas, for all $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, p\}$:

$$accept(m_i, w_j) \equiv manpropose(m_i, w_j) \wedge womanpropose(m_i, w_j),$$

$$accept(m_i, m_i) \equiv \bigwedge_{k \in acceptable_M^i} \neg accept(m_i, w_k),$$

$$accept(w_j, w_j) \equiv \bigwedge_{k \in acceptable_W^j} \neg accept(m_k, w_j),$$

and for all $i \in \{1, \dots, n\}$ and $j \in acceptable_M^i$:

$$manpropose(m_i, w_j) \equiv \bigwedge_{x \leq_M^{m_i} w_j, x \neq w_j} \neg accept(m_i, x),$$

and similarly for all $j \in \{1, \dots, p\}$ and $i \in acceptable_W^j$:

$$womanpropose(m_i, w_j) \equiv \bigwedge_{x \leq_W^{w_j} m_i, x \neq m_i} \neg accept(x, w_j),$$

and for all $i \in \{1, \dots, n\}$ and $j \in unacceptable_M^i$:

$$manpropose(m_i, w_j) \equiv \perp,$$

and similarly for all $j \in \{1, \dots, p\}$ and $i \in unacceptable_W^j$:

$$womanpropose(m_i, w_j) \equiv \perp.$$

Using these formulas, which form the completion of the normal ASP program from Definition 2, we can define an equivalent disjunctive ASP program without negation-as-failure.

Definition 5 (Induced disj. naf-free ASP program)

The disjunctive naf-free ASP program \mathcal{P}_{disj} induced by an SMTI instance (S_M, S_W) contains the following rules for every $i \in \{1, \dots, n\}, j \in \{1, \dots, p\}$:

$$\begin{aligned} & \neg accept(m_i, w_j) \vee manpropose(m_i, w_j) \leftarrow \\ & \neg accept(m_i, w_j) \vee womanpropose(m_i, w_j) \leftarrow \\ & accept(m_i, w_j) \vee \neg manpropose(m_i, w_j) \vee \neg womanpropose(m_i, w_j) \leftarrow \end{aligned}$$

For every $i \in \{1, \dots, n\}, l \in unacceptable_M^i, j \in acceptable_M^i, x \leq_M^{m_i} w_j, x \neq w_j$ \mathcal{P}_{disj} contains:

$$\begin{aligned} & \bigvee_{k \in acceptable_M^i} accept(m_i, w_k) \vee accept(m_i, m_i) \leftarrow \\ & \neg accept(m_i, m_i) \vee \neg accept(m_i, w_j) \leftarrow \\ & \neg manpropose(m_i, w_j) \vee \neg accept(m_i, x) \leftarrow \\ & \bigvee_{x \leq_M^{m_i} w_j, x \neq w_j} accept(m_i, x) \vee manpropose(m_i, w_j) \leftarrow \\ & \neg manpropose(m_i, w_l) \leftarrow \end{aligned}$$

and symmetrical for $j \in \{1, \dots, p\}$ and $womanpropose$.

Note that, for $k = \max(n, p)$, the number of grounded rules in the induced naf-free program is $\mathcal{O}(k^3)$. The following lemma follows from the fact that the completion corresponds to the original program (Erdem and Lifschitz 2003).

Lemma 1

Let \mathcal{P} be the normal ASP program from Definition 2 and \mathcal{P}_{disj} the disjunctive ASP program from Definition 5. It holds that for any answer set I of \mathcal{P} there exists an answer set I_{disj} of \mathcal{P}_{disj} such that the atoms of I and I_{disj} coincide. Conversely for any answer set I_{disj} of \mathcal{P}_{disj} there exists an answer set I of \mathcal{P} such that the atoms of I and I_{disj} coincide.

4.3 ASP Program to Select Optimal Solutions

Let (S_M, S_W) be an SMTI instance with $S_M = \{\sigma_M^1, \dots, \sigma_M^n\}$ and $S_W = \{\sigma_W^1, \dots, \sigma_W^p\}$, and let \mathcal{P}_{norm} be the induced normal ASP program from Definition 2. Our technique for extending this program to a program that can respectively optimize for the sex-equality, egalitarian, minimum regret and maximum cardinality criterion is in each case very similar. We start by explaining it for the case of sex-equality. Our first step is to add a set of rules that compute the sex-equality cost of a set of marriages. For every man m_i and every $x \in acceptable_M^i \cup \{m_i\}$ we use the following rule to determine the cost for m_i :

$$mancost(i, c_{m_i}(x)) \leftarrow accept(m_i, x) \quad (6)$$

and similarly for every w_j and every $x \in acceptable_W^j \cup \{w_j\}$:

$$womancost(j, c_{w_j}(x)) \leftarrow accept(x, w_j) \quad (7)$$

We also use the following rules with i ranging from 1 to n and j from 1 to p :

$$\text{manweight}(Z) \leftarrow \#sum\{B, A : \text{mancost}(A, B)\} = Z, \#int(Z) \quad (8)$$

$$\text{womanweight}(Z) \leftarrow \#sum\{B, A : \text{womancost}(A, B)\} = Z, \#int(Z) \quad (9)$$

$$\text{sexeq}(Z) \leftarrow \text{manweight}(X), \text{womanweight}(Y), Z = X - Y$$

$$\text{sexeq}(Z) \leftarrow \text{manweight}(X), \text{womanweight}(Y), Z = Y - X \quad (10)$$

Note that $\#sum$, $\#max$, $\#int$ and $\#count$ are DLV aggregate functions (Faber et al. 2008). The ‘ A ’ mentioned as variable in $\#sum$ indicates that a cost must be included for every person (otherwise the cost is included only once when persons have the same cost). Rule (8) determines the sum of the male costs and similarly (9) determines the sum of the female costs. According to Definition 4 the absolute difference of these values yields the sex-equality cost, as determined by rules (10). Since numeric variables are restricted to positive integers in DLV, we omit conditions as ‘ $X \geq Y$ ’ or ‘ $X < Y$ ’. The program \mathcal{P}_{norm} extended with rules (6) – (10) is denoted $\mathcal{P}_{ext}^{sexeq}$.

We construct a program \mathcal{P}_{sexeq} , composed by subprograms, that selects optimal solutions. Let \mathcal{P}'_{disj} be the disjunctive naf-free ASP program, induced by the same SMTI instance, in which a prime symbol is added to all literal names (e.g. *accept* becomes *accept'*). Define a new program $\mathcal{P}'_{ext}^{sexeq}$ with all the rules of \mathcal{P}'_{disj} in which every occurrence of $\neg atom$ is changed into *natom* for every atom *atom*, i.e. replace all negation symbols by a prefix ‘ n ’. For every occurring atom *atom* in $\mathcal{P}'_{ext}^{sexeq}$, add the following rule to exclude non-consistent solutions:

$$\text{sat} \leftarrow \text{atom}, \text{natom} \quad (11)$$

For instance, the rule $\text{sat} \leftarrow \text{accept}'(m_1, w_1), \text{naccept}'(m_1, w_1)$ is added. Finally add rules (6) – (10) with prime symbols to the literal names to $\mathcal{P}'_{ext}^{sexeq}$ but replace rule (8) and rule (9) by:

$$\text{mansum}'(n, X) \leftarrow \text{mancost}'(n, X)$$

$$\text{mansum}'(J, Z) \leftarrow \text{mansum}'(I, X), \text{mancost}'(J, Y), Z = X + Y, \#succ(J, I)$$

$$\text{manweight}'(Z) \leftarrow \text{mansum}'(1, Z)$$

$$\text{womansum}'(p, X) \leftarrow \text{womancost}'(p, X)$$

$$\text{womansum}'(J, Z) \leftarrow \text{womansum}'(I, X), \text{womancost}'(J, Y), Z = X + Y,$$

$$\#succ(J, I)$$

$$\text{womanweight}'(Z) \leftarrow \text{womansum}'(1, Z) \quad (12)$$

The DLV aggregate function $\#succ(J, I)$ is true whenever $J+1 = I$. We replace the rules with the aggregate function $\#max$ by these rules to make sure the saturation happens correctly. When saturation is used, the DLV aggregate functions $\#max$, $\#sum$ and $\#count$ would not yield the right criterion values. Moreover, DLV does not accept these aggregate functions in saturation because of the cyclic dependency of literals within the aggregate functions created by the rules for saturation. These adjusted rules, however, will not pose any problems because of the successive way they compute the criterion values. This becomes more clear in the proof of Proposi-

tion 2. We define the ASP program \mathcal{P}_{sexeq} as the union of $\mathcal{P}_{ext}^{sexeq}$, $\mathcal{P}'_{ext}^{sexeq}$ and \mathcal{P}_{sat} . The ASP program \mathcal{P}_{sat} contains the following rules to select minimal solutions based on sex-equality:

$$sat \leftarrow sexeq(X), sexeq'(Y), X \leq Y \quad (13)$$

$$\leftarrow not\ sat \quad (14)$$

$$mancost'(X, Y) \leftarrow sat, manargcost'_1(X), manargcost'_2(Y)$$

$$womancost'(X, Y) \leftarrow sat, womanargcost'_1(X), womanargcost'_2(Y) \quad (15)$$

$$manpropose'(X, Y) \leftarrow sat, man(X), woman(Y)$$

$$womanpropose'(X, Y) \leftarrow sat, man(X), woman(Y)$$

$$accept'(X, X) \leftarrow sat, man(X)$$

$$accept'(X, X) \leftarrow sat, woman(X)$$

$$accept'(X, Y) \leftarrow sat, man(X), woman(Y) \quad (16)$$

and analogous to (16) a set of rules with prefix ‘ n ’ for the head predicates. Finally we add the facts $manargcost'_1(1 \cdot n) \leftarrow$, $manargcost'_2(1 \cdot (p+1)) \leftarrow$, $womanargcost'_1(1 \cdot p) \leftarrow$, $womanargcost'_2(1 \cdot (n+1)) \leftarrow$, $man(x) \leftarrow$ for every man x and $woman(x) \leftarrow$ for every woman x to \mathcal{P}_{sat} . The rule $manargcost'_1(1 \cdot n) \leftarrow$ is DLV syntax for the n facts $manargcost'_1(1) \leftarrow, \dots, manargcost'_1(n) \leftarrow$. Intuitively the rules of \mathcal{P}_{sat} express the key idea of saturation. First every answer set is forced to contain the atom sat by rule (14). Then the rules (15) – (16) and the facts make sure that any answer set should contain all possible literals with a prime symbol that occur in \mathcal{P}_{sexeq} . Rule (13) will establish that only optimal solutions will correspond to minimal models and thus lead to answer sets. For any non-optimal solution, the corresponding interpretation containing sat will never be a minimal model of the reduct. It is formally proven in Proposition 2 below that \mathcal{P}_{sexeq} produces exactly the stable matchings with minimal sex-equality cost.

Furthermore, only small adjustments to \mathcal{P}_{sexeq} are needed to create programs \mathcal{P}_{weight} , \mathcal{P}_{regret} , and $\mathcal{P}_{singles}$ that respectively produce egalitarian, minimum regret and maximum cardinality stable matchings. Indeed, the ASP program \mathcal{P}_{weight} can easily be defined as \mathcal{P}_{sexeq} in which the predicates $sexeq$ and $sexeq'$ are respectively replaced by $weight$ and $weight'$ and the rules (10) are replaced by (17), determining the egalitarian cost of Definition 4 as the sum of the male and female costs:

$$weight(Z) \leftarrow manweight(X), womanweight(Y), Z = X + Y \quad (17)$$

Similarly the ASP program \mathcal{P}_{regret} is defined as \mathcal{P}_{sexeq} in which the predicates $sexeq$ and $sexeq'$ are resp. replaced by $regret$ and $regret'$ and rules (8) – (10) are replaced by the following rules:

$$manregret(Z) \leftarrow \#max\{B : mancost(A, B)\} = Z, \#int(Z) \quad (18)$$

$$womanregret(Z) \leftarrow \#max\{B : womancost(A, B)\} = Z, \#int(Z) \quad (19)$$

$$regret(X) \leftarrow manregret(X), womanregret(Y), X > Y$$

$$regret(Y) \leftarrow manregret(X), womanregret(Y), X \leq Y \quad (20)$$

Rule (18) determines the regret cost but only for the men. Similarly (19) determines the regret cost for the women. The regret cost as defined in Definition 4 is the maximum of these two values, determined by the rules in (20). Again we adjust rules (18) and (19) for the program part $\mathcal{P}'_{ext}{}^{regret}$ by replacing them with a variant based on the successor function:

$$\begin{aligned}
manmax'(n, X) &\leftarrow mancost'(n, X) \\
manmax'(J, X) &\leftarrow manmax'(I, X), mancost'(J, Y), X \geq Y, \#succ(J, I) \\
manmax'(J, Y) &\leftarrow manmax'(I, X), mancost'(J, Y), X < Y, \#succ(J, I) \\
manregret'(Z) &\leftarrow manmax'(1, Z) \\
womanmax'(p, X) &\leftarrow womancost'(p, X) \\
womanmax'(J, X) &\leftarrow womanmax'(I, X), womancost'(J, Y), X \geq Y, \#succ(J, I) \\
womanmax'(J, Y) &\leftarrow womanmax'(I, X), womancost'(J, Y), X < Y, \#succ(J, I) \\
womanregret'(Z) &\leftarrow womanmax'(1, Z)
\end{aligned} \tag{21}$$

Finally we define the ASP program $\mathcal{P}_{singles}$ as \mathcal{P}_{sexeq} in which the predicates *sexeq* and *sexeq'* are resp. replaced by *singles* and *singles'*. Furthermore we replace rules (6) – (10) by (22), determining the number of singles:

$$singles(Z) \leftarrow \#count\{B : accept(B, B)\} = Z, \#int(Z) \tag{22}$$

This time we adjust rule (22) for the program part $\mathcal{P}'_{ext}{}^{singles}$ as follows:

$$\begin{aligned}
single'(p + i, 1) &\leftarrow accept'(m_i, m_i), & single'(p + i, 0) &\leftarrow naccept'(m_i, m_i) \\
single'(j, 1) &\leftarrow accept'(w_j, w_j), & single'(j, 0) &\leftarrow naccept'(w_j, w_j) \\
singlesum'(n + p, X) &\leftarrow single'(n + p, X) \\
singlesum'(J, Z) &\leftarrow singlesum'(I, X), single'(J, Y), Z = X + Y, \#succ(J, I) \\
singles'(Z) &\leftarrow singlesum'(1, Z)
\end{aligned} \tag{23}$$

Note that, for $k = \max(n, p)$, the number of grounded rules in the induced ASP program is $\mathcal{O}(k^3)$ for minimum regret and maximum cardinality, but $\mathcal{O}(k^4)$ for sex-equality and egalitarianity. The latter programs have a higher number of grounded rules because of how the weights are counted in the first and second program part. We illustrate our method with an example.

Example 5

We reconsider Example 3. This SMTI instance had 3 stable matchings of marriages:

- $S_1 = \{accept(m_1, w_3), accept(m_2, w_1), accept(w_2, w_2)\}$,
- $S_2 = \{accept(m_1, w_2), accept(m_2, w_1), accept(w_3, w_3)\}$,
- $S_3 = \{accept(m_1, w_1), accept(m_2, w_2), accept(w_2, w_2), accept(w_3, w_3)\}$.

It is easy to compute the respective regret costs as $c_{regret}(S_1) = 2$ and $c_{regret}(S_2) = c_{regret}(S_3) = 3$. The corresponding ASP program selecting this minimum regret stable matching is the program consisting of the rules in Example 3 in addition to:

$$man(m_1) \leftarrow, \quad man(m_2) \leftarrow,$$

$woman(w_1) \leftarrow, \quad woman(w_2) \leftarrow, \quad woman(w_3) \leftarrow$
 $mancost(1, 1) \leftarrow accept(m_1, w_1), \quad womancost(1, 1) \leftarrow accept(m_1, w_1)$
 $mancost(1, 2) \leftarrow accept(m_1, w_2), \quad womancost(1, 1) \leftarrow accept(m_2, w_1)$
 $mancost(1, 2) \leftarrow accept(m_1, w_3), \quad womancost(1, 2) \leftarrow accept(w_1, w_1)$
 $mancost(1, 4) \leftarrow accept(m_1, m_1), \quad womancost(2, 1) \leftarrow accept(m_1, w_2)$
 $mancost(2, 2) \leftarrow accept(m_2, w_1), \quad womancost(2, 2) \leftarrow accept(w_2, w_2)$
 $mancost(2, 1) \leftarrow accept(m_2, w_2), \quad womancost(3, 2) \leftarrow accept(m_1, w_3)$
 $mancost(2, 2) \leftarrow accept(m_2, m_2), \quad womancost(3, 1) \leftarrow accept(m_2, w_3)$
 $womancost(3, 3) \leftarrow accept(w_3, w_3)$

$manregret(Z) \leftarrow \#max\{B : mancost(A, B)\} = Z, \#int(Z)$

$womanregret(Z) \leftarrow \#max\{B : womancost(A, B)\} = Z, \#int(Z)$

$regret(X) \leftarrow manregret(X), womanregret(Y), X > Y$

$regret(Y) \leftarrow manregret(X), womanregret(Y), X \leq Y$

$naccept'(M, W) \vee manpropose'(M, W) \leftarrow man(M),$
 $woman(W)$

$naccept'(M, W) \vee womanpropose'(M, W) \leftarrow man(M),$
 $woman(W)$

$accept'(M, W) \vee nmanpropose'(M, W) \vee nwomanpropose'(M, W) \leftarrow man(M),$
 $woman(W)$

$accept'(m_1, w_1) \vee accept'(m_1, w_2) \vee accept'(m_1, w_3) \vee accept'(m_1, m_1) \leftarrow$
 $accept'(m_2, w_1) \vee accept'(m_2, w_2) \vee accept'(m_2, m_2) \leftarrow$
 $naccept'(m_1, m_1) \vee naccept'(m_1, w_1) \leftarrow$
 $naccept'(m_1, m_1) \vee naccept'(m_1, w_2) \leftarrow$
 $naccept'(m_1, m_1) \vee naccept'(m_1, w_3) \leftarrow$
 $naccept'(m_2, m_2) \vee naccept'(m_2, w_1) \leftarrow$
 $naccept'(m_2, m_2) \vee naccept'(m_2, w_2) \leftarrow$
 $accept'(m_1, w_1) \vee accept'(m_2, w_1) \vee accept'(w_1, w_1) \leftarrow$
 $accept'(m_1, w_2) \vee accept'(w_2, w_2) \leftarrow$
 $accept'(m_1, w_3) \vee accept'(m_2, w_3) \vee accept'(w_3, w_3) \leftarrow$
 $naccept'(w_1, w_1) \vee naccept'(m_1, w_1) \leftarrow$
 $naccept'(w_1, w_1) \vee naccept'(m_2, w_1) \leftarrow$
 $naccept'(w_2, w_2) \vee naccept'(m_1, w_2) \leftarrow$
 $naccept'(w_3, w_3) \vee naccept'(m_1, w_3) \leftarrow$
 $naccept'(w_3, w_3) \vee naccept'(m_2, w_3) \leftarrow$
 $nmanpropose'(m_1, w_2) \vee naccept'(m_1, w_1) \leftarrow$
 $nmanpropose'(m_1, w_2) \vee naccept'(m_1, w_3) \leftarrow$

$$\begin{aligned}
& nmanpropose'(m_1, w_3) \vee naccept'(m_1, w_1) \leftarrow \\
& nmanpropose'(m_1, w_3) \vee naccept'(m_1, w_2) \leftarrow \\
& \quad manpropose'(m_1, w_1) \leftarrow \\
& accept'(m_1, w_1) \vee accept'(m_1, w_3) \vee manpropose'(m_1, w_2) \leftarrow \\
& accept'(m_1, w_1) \vee accept'(m_1, w_2) \vee manpropose'(m_1, w_3) \leftarrow \\
& \quad nmanpropose'(m_2, w_1) \vee naccept'(m_2, w_2) \leftarrow \\
& nmanpropose'(m_2, w_1) \vee naccept'(m_2, m_2) \leftarrow \\
& \quad manpropose'(m_2, w_2) \leftarrow \\
& accept'(m_2, w_2) \vee accept'(m_2, m_2) \vee manpropose'(m_2, w_1) \leftarrow \\
& \quad nwomanpropose'(m_1, w_1) \vee naccept'(m_2, w_1) \leftarrow \\
& nwomanpropose'(m_2, w_1) \vee naccept'(m_1, w_1) \leftarrow \\
& \quad accept'(m_1, w_1) \vee womanpropose'(m_2, w_1) \leftarrow \\
& \quad accept'(m_2, w_1) \vee womanpropose'(m_1, w_1) \leftarrow \\
& \quad \quad womanpropose'(m_1, w_2) \leftarrow \\
& \quad nwomanpropose'(m_1, w_3) \vee naccept'(m_2, w_3) \leftarrow \\
& \quad \quad womanpropose'(m_2, w_3) \leftarrow \\
& \quad \quad accept'(m_2, w_3) \vee womanpropose'(m_1, w_3) \leftarrow \\
& \quad \quad \quad nmanpropose'(m_2, w_3) \leftarrow \\
& \quad \quad \quad nwomanpropose'(m_2, w_2) \leftarrow
\end{aligned}$$

$$\begin{aligned}
sat \leftarrow & manpropose'(X, Y), nmanpropose'(X, Y), man(X), \\
& woman(Y)
\end{aligned}$$

$$\begin{aligned}
sat \leftarrow & womanpropose'(X, Y), nwomanpropose'(X, Y), \\
& man(X), woman(Y)
\end{aligned}$$

$$sat \leftarrow accept'(X, Y), naccept'(X, Y), man(X), woman(Y)$$

$$sat \leftarrow accept'(X, X), naccept'(X, X), man(X)$$

$$sat \leftarrow accept'(X, X), naccept'(X, X), woman(X)$$

$$\begin{aligned}
mancost'(1, 1) \leftarrow & accept'(m_1, w_1), & womancost'(1, 1) \leftarrow & accept'(m_1, w_1) \\
mancost'(1, 2) \leftarrow & accept'(m_1, w_2), & womancost'(1, 1) \leftarrow & accept'(m_2, w_1) \\
mancost'(1, 2) \leftarrow & accept'(m_1, w_3), & womancost'(1, 2) \leftarrow & accept'(w_1, w_1) \\
mancost'(1, 4) \leftarrow & accept'(m_1, m_1), & womancost'(2, 1) \leftarrow & accept'(m_1, w_2) \\
mancost'(2, 2) \leftarrow & accept'(m_2, w_1), & womancost'(2, 2) \leftarrow & accept'(w_2, w_2) \\
mancost'(2, 1) \leftarrow & accept'(m_2, w_2), & womancost'(3, 2) \leftarrow & accept'(m_1, w_3) \\
mancost'(2, 2) \leftarrow & accept'(m_2, m_2), & womancost'(3, 1) \leftarrow & accept'(m_2, w_3) \\
& & womancost'(3, 3) \leftarrow & accept'(w_3, w_3)
\end{aligned}$$

$$manmax'(2, X) \leftarrow mancost'(2, X)$$

$$manmax'(J, X) \leftarrow manmax'(I, X), mancost'(J, Y), X \geq Y, \#succ(J, I)$$

$$\begin{aligned}
& \text{manmax}'(J, X) \leftarrow \text{manmax}'(I, X), \text{mancost}'(J, Y), X \geq Y, \#succ(J, I) \\
& \text{manregret}'(Z) \leftarrow \text{manmax}'(1, Z) \\
& \text{womanmax}'(3, X) \leftarrow \text{womancost}'(3, X) \\
& \text{womanmax}'(J, X) \leftarrow \text{womanmax}'(I, X), \text{womancost}'(J, Y), X \geq Y, \\
& \quad \#succ(J, I) \\
& \text{womanmax}'(J, X) \leftarrow \text{womanmax}'(I, X), \text{womancost}'(J, Y), X \geq Y, \\
& \quad \#succ(J, I) \\
& \text{womanregret}'(Z) \leftarrow \text{womanmax}'(1, Z) \\
& \text{regret}'(X) \leftarrow \text{manregret}'(X), \text{womanregret}'(Y), X > Y \\
& \text{regret}'(Y) \leftarrow \text{manregret}'(X), \text{womanregret}'(Y), X \leq Y \\
& \text{sat} \leftarrow \text{regret}(X), \text{regret}'(Y), X \leq Y \\
& \quad \leftarrow \text{not sat} \\
& \text{manargcost}'_1(1..2) \leftarrow, \quad \text{womanargcost}'_1(1..3) \leftarrow \\
& \text{manargcost}'_2(1..4) \leftarrow, \quad \text{womanargcost}'_2(1..3) \leftarrow \\
& \text{mancost}'(X, Y) \leftarrow \text{sat}, \text{manargcost}'_1(X), \text{manargcost}'_2(Y) \\
& \text{womancost}'(X, Y) \leftarrow \text{sat}, \text{womanargcost}'_1(X), \text{womanargcost}'_2(Y) \\
& \text{manpropose}'(X, Y) \leftarrow \text{sat}, \text{man}(X), \text{woman}(Y) \\
& \text{nmanpropose}'(X, Y) \leftarrow \text{sat}, \text{man}(X), \text{woman}(Y) \\
& \text{womanpropose}'(X, Y) \leftarrow \text{sat}, \text{man}(X), \text{woman}(Y) \\
& \text{nwomanpropose}'(X, Y) \leftarrow \text{sat}, \text{man}(X), \text{woman}(Y) \\
& \text{accept}'(X, Y) \leftarrow \text{sat}, \text{man}(X), \text{woman}(Y) \\
& \text{accept}'(X, X) \leftarrow \text{sat}, \text{man}(X) \\
& \text{accept}'(X, X) \leftarrow \text{sat}, \text{woman}(X) \\
& \text{naccept}'(X, Y) \leftarrow \text{sat}, \text{man}(X), \text{woman}(Y) \\
& \text{naccept}'(X, X) \leftarrow \text{sat}, \text{man}(X) \\
& \text{naccept}'(X, X) \leftarrow \text{sat}, \text{woman}(X)
\end{aligned}$$

Computing the unique answer set of this disjunctive ASP program with DLV and filtering it to the literals *accept* and *regret*, yields $\{\text{accept}(m_2, w_1), \text{accept}(m_1, w_3), \text{accept}(w_2, w_2), \text{regret}(2)\}$, corresponding exactly to the minimum regret stable matching S_1 of the SMTI instance and the corresponding regret cost.

We prove that there exists a bijective correspondence between the answer sets of the induced disjunctive ASP program and the optimal stable matchings of the SMTI (see electronic Appendix D).

Proposition 2

Let the criterion *crit* be an element of $\{\text{sexeq}, \text{weight}, \text{regret}, \text{singles}\}$. For every answer set I of the program $\mathcal{P}_{\text{crit}}$ induced by an SMTI instance the set $S_I = \{(m, w) \mid \text{accept}(m, w) \in I\}$ forms an optimal stable matching of marriages w.r.t. criterion *crit* and the optimal criterion value is given by the unique value v_I for

which $\text{crit}(v_I) \in I$. Conversely for every optimal stable matching $S = \{(x_1, y_1), \dots, (x_k, y_k)\}$ with optimal criterion value v there exists an answer set I of $\mathcal{P}_{\text{crit}}$ such that $\{(x, y) \mid \text{accept}(x, y) \in I\} = \{(x_i, y_i) \mid i \in \{1, \dots, k\}\}$ and v is the unique value for which $\text{crit}(v) \in I$.

Remark 2

If we remove from $\mathcal{P}_{\text{sexeq}}$ the rules (9) – (10) and replace rule (13) by the rule $\text{sat} \leftarrow \text{manweight}(X), \text{manweight}'(Y), X \leq Y$, then we obtain the M-optimal stable matchings. Analogously we can obtain the W-optimal stable matchings.

If a criterion is to be maximized, the symbol \leq in rule (13) is simply replaced by \geq . E.g. for $\text{crit} = \text{singles}$ we will get minimum cardinality stable matchings.

5 Conclusion

We have shown how ASP programs can be used to encode a number of variations and generalizations of the SMP. Apart from the availability of efficient ASP solvers, the main advantage of our approach is its flexibility, allowing us to find solutions for a wide range of stable matching problems. We can, for instance, compute stable matchings of variants such as the three-dimensional stable matching problem, as well as select stable matchings based on optimality criteria, even for problems with unacceptable partners and ties. We have illustrated our method for sex-equality, egalitarianity, minimum regret and maximum cardinality, but the approach can readily be adapted to other optimality criteria (e.g. popular matchings) or to different matching problems (e.g. the roommate problem). To the best of our knowledge, no other exact algorithms exist to find an optimal stable matching for an SMP instance with ties, regardless of the presence of unacceptability and regardless of whether the optimality notion is sex-equality, egalitarianity, minimum regret or maximum cardinality. Therefore, our encoding offers the first exact implementation for solving the aforementioned problems.

References

- ABRAHAM, D., IRVING, R., MEHLHORN, K., TELIKEPALLI, K., AND MEHLHORN, K. 2007. Popular matchings. *SIAM Journal on Computing* 37, 4, 1030–1045.
- BARAL, C. 2003. *Knowledge Representation, Reasoning, and Declarative Problem Solving*. Cambridge University Press, New York, NY, USA.
- BIRÓ, P. AND MCDERMID, E. 2010. Three-sided stable matchings with cyclic preferences. *Algorithmica* 58, 1, 5–18.
- BREWKA, G., EITER, T., AND TRUSZCZYŃSKI, M. 2011. Answer set programming at a glance. *Communications of the ACM* 54, 12, 92–103.
- DE CLERCQ, S., SCHOCKAERT, S., DE COCK, M., AND NOWÉ, A. 2013. Modeling stable matching problems with answer set programming. In *Proceedings RuleML 2013*. Lecture Notes in Computer Science, vol. 8035. 68–83.
- DUNG, P. 1995. An argumentation-theoretic foundation for logic programming. *The Journal of Logic Programming* 22, 2, 151–177.

- EITER, T. AND GOTTLÖB, G. 1995. On the computational cost of disjunctive logic programming: Propositional case. *Annals of Mathematics and Artificial Intelligence* 15, 3-4, 289–323.
- ERDEM, E. AND LIFSCHITZ, V. 2003. Tight logic programs. *Theory and Practice of Logic Programming* 3, 499–518.
- FABER, W., PFEIFER, G., LEONE, N., DELL’ARMI, T., AND IELPA, G. 2008. Design and implementation of aggregate functions in the DLV system. *Theory and Practice of Logic Programming* 8, 5-6, 545–580.
- GALE, D. AND SHAPLEY, L. 1962. College admissions and the stability of marriage. *The American Mathematical Monthly* 69, 1, 9–15.
- GALE, D. AND SOTOMAYOR, M. 1985. Some remarks on the stable matching problem. *Discr. Appl. Math.* 11, 223–232.
- GÄRDENFORS, P. 1975. Match making: Assignments based on bilateral preferences. *Behavioral Science* 20, 3, 166–173.
- GELFOND, M. AND LIFSCHITZ, V. 1988. The stable model semantics for logic programming. In *ICLP/SLP*. Vol. 88. 1070–1080.
- GUSFIELD, D. 1987. Three fast algorithms for four problems in stable marriage. *SIAM J. Comput.* 16, 1, 111–128.
- IRVING, R. 1994. Stable marriage and indifference. *Discr. Appl. Math.* 48, 3, 261–272.
- IRVING, R. 2003. Greedy matchings. Tech. Rep. TR-2003-136, Dept. of Computing Science, University of Glasgow. April.
- IRVING, R. 2007. The cycle roommates problem: a hard case of kidney exchange. *Inf. Process. Lett.* 103, 1, 1–4.
- IRVING, R., LEATHER, P., AND GUSFIELD, D. 1987. An efficient algorithm for the “optimal” stable marriage. *J. ACM* 34, 3, 532–543.
- IWAMA, K. AND MIYAZAKI, S. 2008. A survey of the stable marriage problem and its variants. In *Proc. of the Intern. Conf. on Informatics Educ. and Research for Knowledge-Circulating Society*. ICKS’08. IEEE Computer Society, 131–136.
- IWAMA, K., MIYAZAKI, S., AND YANAGISAWA, H. 2010. Approximation algorithms for the sex-equal stable marriage problem. *ACM Transactions on Algorithms (TALG)* 7, 1, 2.
- JANHUNEN, T. 2004. Representing normal programs with clauses. In *In Proc. of the 16th European Conference on Artificial Intelligence*. IOS Press, 358–362.
- KNUTH, D. 1976. *Mariages Stables*. Les Presses de L’Université de Montréal.
- MANLOVE, D. 2013. *Algorithmics of matching under preferences*. Series on Theoretical Computer Science, vol. 2. World Scientific Publishing.
- MANLOVE, D., IRVING, R., IWAMA, K., MIYAZAKI, S., AND MORITA, Y. 2002. Hard variants of stable marriage. *Theoretical Computer Science* 276, 1-2, 261–279.
- MAREK, V., NERODE, A., AND REMMEL, J. 1990. A theory of nonmonotonic rule systems I. *Ann. Math. Artif. Intell.* 1, 241–273.
- MCDERMID, E. AND IRVING, R. 2012. Sex-equal stable matchings: Complexity and exact algorithms. *Algorithmica*, 1–26.
- NG, C. AND HIRSCHBERG, D. 1991. Three-dimensional stable matching problems. *SIAM Journal on Discrete Mathematics* 4, 245–252.
- ROTH, A., SÖMNEZ, T., AND ÜNVER, M. 2005. Pairwise kidney exchange. *J. Econ. Theory* 125, 2, 151 – 188.
- ROTH, A. AND SOTOMAYOR, M. 1990. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Cambridge University Press.
- XU, H. AND LI, B. 2011. Egalitarian stable matching for VM migration in cloud computing. In *Computer Communications Workshops (INFOCOM WKSHPS), 2011 IEEE Conference on*. IEEE, 631–636.

Solving Stable Matching Problems using Answer Set Programming: Electronic Appendix

SOFIE DE CLERCQ

Dept. of Applied Mathematics, Computer Science & Statistics, Ghent University, Belgium
(e-mail: sofier.declercq@ugent.be)

STEVEN SCHOCKAERT

School of Computer Science & Informatics, Cardiff University, UK
(e-mail: s.schockaert@cs.cardiff.ac.uk)

MARTINE DE COCK

Center for Data Science, UW Tacoma, US
Dept. of Applied Mathematics, Computer Science & Statistics, Ghent University, Belgium
(e-mail: mdecock@u.washington.edu)

ANN NOWE

Computational Modeling Lab, Vrije Universiteit Brussel, Belgium
(e-mail: anowe@vub.ac.be)

submitted 1 January 2003; revised 1 January 2003; accepted 1 January 2003

Appendix A Definition SMP and SMI

Definition 1 (SMP)

An instance of the SMP is a pair (S_M, S_W) with $S_M = \{\sigma_M^1, \dots, \sigma_M^n\}$ and $S_W = \{\sigma_W^1, \dots, \sigma_W^n\}$. For every $i \in \{1, \dots, n\}$, σ_M^i and σ_W^i are permutations of $\{1, \dots, n\}$. We call σ_M^i and σ_W^i the preferences of man m_i and woman w_i respectively. If $k = \sigma_M^i(j)$, woman w_k is man m_i 's j^{th} most preferred woman. The case $k = \sigma_W^i(j)$ is similar. Man m and woman w form a blocking pair in a set of marriages S if m strictly prefers w to his partner in S and w strictly prefers m to her partner in S . A weakly stable matching is a set of marriages without blocking pairs or individuals.

Definition 2 (SMI)

An instance of the SMI is a pair (S_M, S_W) with $S_M = \{\sigma_M^1, \dots, \sigma_M^n\}$ and $S_W = \{\sigma_W^1, \dots, \sigma_W^p\}$. For every $i \in \{1, \dots, n\}$, σ_M^i is a permutation of a subset of $\{1, \dots, p\}$. Symmetrically σ_W^i is a permutation of a subset of $\{1, \dots, n\}$ for every $i \in \{1, \dots, p\}$. We call σ_M^i and σ_W^i the preferences of man m_i and woman w_i respectively. If $k = \sigma_M^i(j)$, woman w_k is man m_i 's j^{th} most preferred woman. The case $k \in \sigma_W^i(j)$ is similar. If there is no l such that $j \in \sigma_M^i(l)$, woman w_j is an unacceptable partner for man m_i , and similarly when there is no l such that

$j \in \sigma_W^i(l)$. Man m and woman w form a blocking pair in a set of marriages S if m strictly prefers w to his partner in S and w strictly prefers m to her partner in S . A blocking individual in S is a person who strictly prefers being single to being paired to his partner in S . A weakly stable matching is a set of marriages without blocking pairs or individuals.

Appendix B Complexity results

Table B 1 presents an overview of known complexity results concerning finding an optimal stable set¹.

Table B 1. *Literature complexity results for finding an optimal stable set*

	sex-equal	egalitarian
SMP	NP-hard (Kato 1993)	P ($O(n^4)$) (Irving et al. 1987))
SMI	NP-hard (McDermid and Irving 2012)	
SMT		
SMTI		
	min. regret	max. card.
SMP	P ($O(n^2)$) (Gusfield 1987))	P ($O(n^2)$) (Gale and Shapley 1962))
SMI		P (Gale and Sotomayor 1985)
SMT	NP-hard (Manlove et al. 2002)	
SMTI		NP-hard (Manlove 1999; Manlove et al. 2002)

Between brackets we mention in Table B 1 the complexity of an algorithm that finds an optimal stable set if one exists, in function of the number of men n . To the best of our knowledge, the only exact algorithm tackling an NP-hard problem from Table B 1 finds a sex-equal stable set for an SMP instance in which the strict preference lists of men and/or women are bounded in length by a constant (McDermid and Irving 2012). To the best of our knowledge, no exact implementations exist to find an optimal stable set for an SMP instance with ties, regardless of the presence of unacceptability and regardless which notion of optimality from Table B 1 is used. Our approach yields an exact implementation of all problems mentioned in Table B 1.

Appendix C Proof of Proposition 1

Proposition 1

Let (S_M, S_W) be an instance of the SMTI and let \mathcal{P} be the corresponding ASP program. If I is an answer set of \mathcal{P} , then a weakly stable matching for (S_M, S_W) is given by $\{(x, y) \mid \text{accept}(x, y) \in I\}$. Conversely, if $\{(x_1, y_1), \dots, (x_k, y_k)\}$ is a weakly stable matching for (S_M, S_W) then \mathcal{P} has the following answer set I :

$$\{\text{manpropose}(x_i, y) \mid i \in \{1, \dots, k\}, x_i \in M, y <_M^{x_i} y_i \vee y = y_i \neq x_i\}$$

¹ We assume that P \neq NP

$$\begin{aligned} & \cup \{ \text{womanpropose}(x, y_i) \mid i \in \{1, \dots, k\}, y_i \in W, x <_W^{y_i} x_i \vee x = x_i \neq y_i \} \\ & \cup \{ \text{accept}(x_i, y_i) \mid i \in \{1, \dots, k\} \} \end{aligned}$$

Proof

Let (S_M, S_W) and \mathcal{P} be as described in the proposition. Because of the symmetry between the men and women we restrict ourselves to the male case when possible.

Answer set \Rightarrow weak stable set We prove this in 4 steps.

1. For every $i \in \{1, \dots, n\}$, every $j \in \{1, \dots, p\}$ and for every answer set I of \mathcal{P} , it holds that $\text{accept}(m_i, w_j) \in I$ implies that $j \in \text{acceptable}_M^i$ and $i \in \text{acceptable}_W^j$.

This can be proved by contradiction. We will prove that for every man m_i and every $j \in \text{unacceptable}_M^i$, $\text{accept}(m_i, w_j)$ is in no answer set I of the induced ASP program \mathcal{P} . For $\text{accept}(m_i, w_j)$ to be in an answer set I , the reduct must contain some rule with this literal in the head and a body which is satisfied. The only rule for which this can be the case is the one of the form (1), implying that $\text{manpropose}(m_i, w_j)$ should be in I . But since j is not in acceptable_M^i there is no rule with $\text{manpropose}(m_i, w_j)$ in the head and so $\text{manpropose}(m_i, w_j)$ can never be in I .

2. For every answer set I of \mathcal{P} and every man m_i , there exists at most one woman w_j such that $\text{accept}(m_i, w_j) \in I$. Similarly, for every woman w_j there exists at most one man m_i such that $\text{accept}(m_i, w_j) \in I$. Moreover, if $\text{accept}(m_i, m_i) \in I$ then $\text{accept}(m_i, w_j) \notin I$ for any w_j , and likewise when $\text{accept}(w_j, w_j) \in I$ then $\text{accept}(m_i, w_j) \notin I$ for any m_i .

This can be proved by contradiction. Suppose first that there is an answer set I of \mathcal{P} that contains $\text{accept}(m_i, w_j)$ and $\text{accept}(m_i, w_{j'})$ for some man m_i and two different women w_j and $w_{j'}$. The first step implies that j and j' are elements of acceptable_M^i . Either man m_i prefers woman w_j to woman $w_{j'}$ ($w_j \leq_M^{m_i} w_{j'}$), or the other way around ($w_{j'} \leq_M^{m_i} w_j$) or man m_i has no preference among them ($w_j \leq_M^{m_i} w_{j'}$ and $w_{j'} \leq_M^{m_i} w_j$). The first two cases are symmetrical and can be handled analogously. The last case follows from the first case because it has stronger assumptions. We prove the first case and assume that man m_i prefers woman w_j to woman $w_{j'}$. The rules (4) imply the presence of a rule $\text{manpropose}(m_i, w_{j'}) \leftarrow \dots, \text{not accept}(m_i, w_j), \dots$ and this is the only rule which can make $\text{manpropose}(m_i, w_{j'})$ true (the only rule with this literal in the head). However, since $\text{accept}(m_i, w_j)$ is also in the answer set, the body of this rule is not satisfied so $\text{manpropose}(m_i, w_{j'})$ can never be in I . Consequently $\text{accept}(m_i, w_{j'})$ can never be in I since the only rule with this literal in the head is of the form (1) and this body can never be satisfied, which leads to a contradiction.

Secondly assume that $\text{accept}(m_i, w_j)$ and $\text{accept}(m_i, m_i)$ are both in an answer set I of \mathcal{P} . Again step 1 implies that $j \in \text{acceptable}_M^i$. Because of the rules (2), \mathcal{P} will contain the rule $\text{accept}(m_i, m_i) \leftarrow \dots, \text{not accept}(m_i, w_j), \dots$. An analogous reasoning as above implies that since $\text{accept}(m_i, w_j)$ is in the answer set I , $\text{accept}(m_i, m_i)$ can never be in I .

3. For every man m_i , in every answer set I of \mathcal{P} exactly one of the following conditions is satisfied:

- (a) there exists a woman w_j such that $\text{accept}(m_i, w_j) \in I$,
- (b) $\text{accept}(m_i, m_i) \in I$,

and similarly for every woman w_i .

Suppose I is an arbitrary answer set of \mathcal{P} and m_i is an arbitrary man. We already know from step 2 that a man cannot be paired to a woman while being single, so both possibilities are disjoint. Therefore, suppose there is no woman w_j such that $\text{accept}(m_i, w_j)$ is in I . \mathcal{P} will contain the rule (2). Because of our assumptions and the definition of the reduct, this rule will be reduced to $\text{accept}(m_i, m_i) \leftarrow$, and so $\text{accept}(m_i, m_i)$ will be in I .

4. For an arbitrary answer set I of \mathcal{P} the previous steps imply that I produces a set of marriages without blocking individuals. Weak stability also demands the absence of blocking pairs. Suppose by contradiction that there is a blocking pair (m_i, w_j) , implying that there exist $i \neq i'$ and $j \neq j'$ such that $\text{accept}(m_i, w_{j'}) \in I$ and $\text{accept}(m_{i'}, w_j) \in I$ while $w_j <_{M}^{m_i} w_{j'}$ and $m_i <_{W}^{w_j} m_{i'}$. The rules of the form (1), which are the only ones with the literals $\text{accept}(m_i, w_{j'})$ and $\text{accept}(m_{i'}, w_j)$ in the head, imply that literals $\text{manpropose}(m_i, w_{j'})$ and $\text{womanpropose}(m_{i'}, w_j)$ should be in I . But since $w_j <_{M}^{m_i} w_{j'}$ and because of the form of the rules (4) there are fewer conditions to be fulfilled for $\text{manpropose}(m_i, w_j)$ to be in I than for $\text{manpropose}(m_i, w_{j'})$ to be in I . Therefore, $\text{manpropose}(m_i, w_j)$ should be in I as well. A similar reasoning implies that $\text{womanpropose}(m_i, w_j)$ should be in I . But now the rules of the form (1) imply that $\text{accept}(m_i, w_j)$ should be in I , contradicting step 2 since $\text{accept}(m_i, w_{j'})$ and $\text{accept}(m_{i'}, w_j)$ are already in I .

Weak stable set \Rightarrow answer set Suppose we have a stable set of marriages $S = \{(x_1, y_1), \dots, (x_k, y_k)\}$, implying that every y_i is an acceptable partner of x_i and the other way around. The rules of the form (1) do not alter when forming the reduct, but the other rules do as those contain naf-literals. Notice first that the stability of S implies that there cannot be an unmarried couple (m, w) , with m a man and w a woman, such that $\text{manpropose}(m, w)$ is in I and $\text{womanpropose}(m, w)$ is in I . By definition of I this would mean that they both strictly prefer each other to their current partner in S . This means they would form a blocking pair, but since S was stable, that is impossible. Therefore, the rules of the form (1) will be applied exactly for married couples (m_i, w_j) , since by definition of I $\text{manpropose}(m_i, w_j)$ and $\text{womanpropose}(m_i, w_j)$ are both in I under these conditions. For other cases the rule will also be fulfilled since the body will be false. This reasoning implies that the unique minimal model of the reduct w.r.t. I should indeed contain $\text{accept}(m_i, w_j)$ for every married couple (m_i, w_j) in S . Since S is a stable set of marriages, every person is either married or single. If a man m_i is single, there will be no other literal of the form $\text{accept}(m_i, \cdot)$ in I , so rule (2) will reduce to a fact $\text{accept}(m_i, m_i) \leftarrow$, which is obviously fulfilled by I . Similarly if a woman w_j is single. Any other rule of the form (2) or (3) is deleted because m_i or w_j is not single in that case and thus there is some literal of the form $\text{accept}(m_i, w)$ for some woman w and some literal

of the form $accept(m, w_j)$ for some man m in I , falsifying the body of the rules. If m_i is single, then $accept(m_i, m_i)$ is in I and this is the only literal of the form $accept(m_i, \cdot)$ in I , so the rules of the form (4) will all be reduced to facts. The heads of these fact rules should be in the minimal model of the reduct and are indeed in I since the women w for which $manpropose(m_i, w)$ is in I are exactly those who are strictly preferred to staying single. The rules of the form (4) for women w_j in $neutral_M^i$ will all be deleted in this case, because $accept(m_i, m_i)$ is in I . If man m_i is married to a certain woman w_j in the stable set S then the rules of the form (4) will reduce to facts of the form $manpropose(m_i, w) \leftarrow$ for every woman w who is strictly preferred to w_j and will be deleted for every other woman appearing in the head, because those rules will contain $not\ accept(m_i, w_j)$ in the body. Again I contains these facts by definition, as the minimal model of the reduct should. We can use an analogous reasoning for the women. The presence of the literals of the form $manpropose(\cdot, \cdot)$, $womanpropose(\cdot, \cdot)$ and $accept(\cdot, \cdot)$ in I is thus required in the unique minimal model of the reduct w.r.t. I . We have proved that every literal in I should be in the minimal model of the reduct and that every rule of the reduct is fulfilled by I , implying that I is an answer set of \mathcal{P} . \square

Appendix D Proof of Proposition 2

Proposition 2

Let the criterion $crit$ be an element of $\{sexeq, weight, regret, singles\}$. For every answer set I of the program \mathcal{P}_{crit} induced by an SMTI instance the set $S_I = \{(m, w) \mid accept(m, w) \in I\}$ forms an optimal stable matching of marriages w.r.t. criterion $crit$ and the optimal criterion value is given by the unique value v_I for which $crit(v_I) \in I$. Conversely for every optimal stable matching $S = \{(x_1, y_1), \dots, (x_k, y_k)\}$ with optimal criterion value v there exists an answer set I of \mathcal{P}_{crit} such that $\{(x, y) \mid accept(x, y) \in I\} = \{(x_i, y_i) \mid i \in \{1, \dots, k\}\}$ and v is the unique value for which $crit(v) \in I$.

Proof

Let (S_M, S_W) be an SMTI instance.

Answer set \Rightarrow optimal stable matching Let I be an arbitrary answer set of \mathcal{P}_{crit} and let S_I be as formulated. It is clear that the only rules in \mathcal{P}_{crit} that influence the literals of the form $manpropose(\cdot, \cdot)$, $womanpropose(\cdot, \cdot)$ and $accept(\cdot, \cdot)$ are the rules in \mathcal{P}_{norm} . Hence I should contain an answer set I_{norm} of \mathcal{P}_{norm} as a subset. Proposition 1 implies that I_{norm} corresponds to a stable matching $S_I = \{(m, w) \mid accept(m, w) \in I_{norm}\}$. Moreover, the only literals of the form $manpropose(\cdot, \cdot)$, $womanpropose(\cdot, \cdot)$ and $accept(\cdot, \cdot)$ in I are those in I_{norm} , so $S_I = \{(m, w) \mid accept(m, w) \in I\}$. If $crit = sexeq$, it is straightforward to see that the literals of the form $accept(\cdot, \cdot)$ in I_{norm} uniquely determine which literals of the form $mancost(\cdot, \cdot)$, $womancost(\cdot, \cdot)$, $manweight(\cdot)$, $womanweight(\cdot)$ and $sexeq(\cdot)$ should be in the answer set I . These literals do not occur in rules of \mathcal{P}_{crit} besides those in $\mathcal{P}_{ext}^{sexeq}$. Note that the rules which do contain these literals imply that there will be just one literal of the form $sexeq(\cdot)$ in I , namely $sexeq(v)$ with v the sex-equality cost of S_I .

Analogous results can be derived for $crit \in \{weight, regret, singles\}$. It remains to be shown that S_I is an optimal stable matching. Suppose by contradiction that S_I is not optimal, so there exists a stable matching S^* such that $v_I > v^*$, with v^* the criterion value of S^* to be minimized. We prove that this implies that I cannot be an answer set of \mathcal{P}_{crit} , contradicting our initial assumption.

Proposition 1 and Lemma 1 imply that there exists an interpretation I_{disj}^* of the ASP program \mathcal{P}_{disj} induced by (S_M, S_W) that corresponds to the stable matching S^* . Moreover this interpretation is consistent, i.e. it will not contain $atom$ and $\neg atom$ for some atom $atom$. This implies that the interpretation I'_{disj} defined as I_{disj}^* in which $\neg atom$ is replaced by $natom$ for every atom $atom$ will falsify the body of the rules of the form (11) of \mathcal{P}_{ext}^{crit} . An analogous reasoning as above yields that the literals of the form $accept'(\cdot, \cdot)$ in I'_{disj} uniquely determine which literals of the form $mancost'(\cdot, \cdot)$, $womancost'(\cdot, \cdot)$, $mansum'(\cdot, \cdot)$, $womansum'(\cdot, \cdot)$, $manweight'(\cdot)$, $womanweight'(\cdot)$ and $sexeq'(\cdot)$ should be in I'_{disj} . With those extra literals added to I'_{disj} , we find that I'_{disj} satisfies all the rules of \mathcal{P}_{ext}^{crit} . Moreover, $crit(v^*)$ is the unique literal of the form $crit(\cdot)$ in I'_{disj} . Note that I'_{disj} does not contain the atom sat .

Define the interpretation $J = I_{norm} \cup I'_{disj}$. From the previous argument it follows that J will satisfy every rule of $\mathcal{P}_{ext}^{crit} \cup \mathcal{P}_{ext}^{crit}$ since the predicates occurring in both programs do not overlap. Moreover J contains $crit(v_I)$ and $crit'(v^*)$ and these are the only literals of the form $crit(\cdot)$ or $crit'(\cdot)$. Since $v_I > v^*$ the rules of the form (13) will be satisfied by J since their body is always false. Call J' the set $J \cup \{a \mid (a \leftarrow) \in \mathcal{P}_{sat}\}$. Since J' does not contain sat , the rules of \mathcal{P}_{sat} will all be satisfied by J' , with exception of the rule $\leftarrow not\ sat$.

The rule of the form (14) implies that I , as an answer set of \mathcal{P}_{crit} , should contain sat . Now the set of rules (15) – (16) imply that I should also contain the literals $mancost'(\cdot, \cdot)$, $womancost'(\cdot, \cdot)$ and $manpropose'(\cdot, \cdot)$, $womanpropose'(\cdot, \cdot)$, $accept'(\cdot, \cdot)$ with the corresponding literals prefixed by n for every possible argument stated by the facts in \mathcal{P}_{sat} . The rules (12), (21) and (23) in \mathcal{P}_{ext}^{crit} , by which we replaced rules (8) – (9), (18) – (19) and (22), guarantee that for every possible set of marriages and its corresponding criterion value c , I will contain $crit(c)$ and all associated intermediate results. For example, for $crit = sexeq$, the rules will guarantee that I also contains $mansum'(\cdot, \cdot)$, $manweight(\cdot)$, $womansum(\cdot, \cdot)$ and $womanweight(\cdot)$ for every argument that could occur in a model of \mathcal{P}_{ext}^{crit} . Note that this would not be the case if we used the original rules with $\#sum$, $\#max$ and $\#count$ in \mathcal{P}_{ext}^{crit} , since these rules would lead to only one value c_M for which e.g. $manweight(c_M)$ should be in I , and similarly only one value c_W for which $womanweight(c_W)$ should be in I . Consequently there would be only one value c such that $crit(c)$ should be in I . This value would not necessarily correspond to v^* and so we would not be able to conclude that $I'_{disj} \subseteq I$. However, with the current formulation of the rules we can conclude that $I'_{disj} \subseteq I$. We already reasoned in the beginning of the proof that $I_{norm} \subseteq I$ holds so it follows that $J \subseteq I$. Since the literals of $J' \setminus J$ are stated as facts of \mathcal{P}_{ext}^{crit} , they should be in I , hence $J' \subseteq I$. Moreover $J' \subset I$ since $sat \in I \setminus J'$.

We use the notation $red(\mathcal{P}, I)$ to denote the reduct of an ASP program \mathcal{P} w.r.t.

an interpretation I . There is no rule in $\mathcal{P}'_{ext}{}^{crit}$ with negation-as-failure in the body, hence $red(\mathcal{P}'_{ext}{}^{crit}, I) = red(\mathcal{P}'_{ext}{}^{crit}, J') = \mathcal{P}'_{ext}{}^{crit}$. We already reasoned that J' satisfies all the rules of the latter. We also reasoned that I does not contain any other literals of the form $accept(\cdot, \cdot)$ than those which are also in I_{norm} , and by construction the same holds for J' . Hence $red(\mathcal{P}_{ext}{}^{crit}, I) = red(\mathcal{P}_{ext}{}^{crit}, J')$ and by construction J' satisfies all the rules of this reduct. It is clear that $red(\mathcal{P}_{sat}, I)$ is \mathcal{P}_{sat} without the rule $\leftarrow not\ sat$, since $sat \in I$. Again we already argued that J' satisfies $red(\mathcal{P}_{sat}, I)$. Hence J' satisfies all the rules of $red(\mathcal{P}_{crit}, I)$, implying that I , which strictly contains J' , cannot be an answer set of \mathcal{P}_{crit} since it is not a minimal model of the negation-free ASP program $red(\mathcal{P}_{crit}, I)$ (Gelfond and Lifschitz 1988).

Optimal stable matching \Rightarrow answer set Let $S = \{(x_1, y_1), \dots, (x_k, y_k)\}$ be an optimal stable matching with optimal criterion value v . To see that the second part of the proposition holds it suffices to verify that the following interpretation I is an answer set of \mathcal{P}_{crit} , with the notation $P_{x_i}(y)$ as the index a for which $y \in \sigma_M^l(a)$ if $x_i = m_l$ and symmetrically $P_{y_i}(x)$ as the index a for which $x \in \sigma_W^l(a)$ if $y_i = w_l$. If $x_i = y_i$ we set $P_{x_i}(y_i) = P_{y_i}(x_i) = |\sigma_M^i|$ if x_i is a man and $P_{x_i}(y_i) = P_{y_i}(x_i) = |\sigma_W^i|$ otherwise. So let I be given by: $I = I_1 \cup I_2$ with

$$\begin{aligned}
I_1 = & \{accept(x_i, y_i) \mid i \in \{1, \dots, k\}\} \cup \{crit(v)\} \cup \{sat\} \\
& \cup \{womanpropose(x_i, y_i) \mid x_i \neq y_i\} \cup \{manpropose(x_i, y_i) \mid x_i \neq y_i\} \\
& \cup \{manpropose(x_i, y) \mid i \in \{1, \dots, k\}, x_i = m_l, \exists a < P_{x_i}(y_i) : y \in \sigma_M^l(a)\} \\
& \cup \{womanpropose(x, y_i) \mid i \in \{1, \dots, k\}, y_i = w_l, \exists a < P_{y_i}(x_i) : x \in \sigma_W^l(a)\} \\
& \cup \{mancost(l, P_{x_i}(y_i)) \mid crit \neq singles, i \in \{1, \dots, k\}, x_i = m_l\} \\
& \cup \{womancost(P_{y_i}(x_i), l') \mid crit \neq singles, i \in \{1, \dots, k\}, y_i = w_l\} \\
& \cup \{manweight(c_M(S)), womanweight(c_W(S)) \mid crit \in \{sexeq, weight\}\} \\
& \cup \{manregret(c_{regret, M}(S)), womanregret(c_{regret, W}(S)) \mid crit = regret\}
\end{aligned}$$

and

$$\begin{aligned}
I_2 = & \{manargcost'_1(z) \mid 1 \leq z \leq n\} \cup \{manargcost'_2(z) \mid 1 \leq z \leq p+1\} \\
& \cup \{womanargcost'_1(z) \mid 1 \leq z \leq p\} \cup \{womanargcost'_2(z) \mid 1 \leq z \leq n+1\} \\
& \cup \{man(x) \mid x \in M\} \cup \{woman(x) \mid x \in W\}
\end{aligned} \tag{D1}$$

$$\begin{aligned}
& \cup \{mancost'(i, j) \mid crit \neq singles, 1 \leq i \leq n, 1 \leq j \leq p+1\} \\
& \cup \{womancost'(j, i) \mid crit \neq singles, 1 \leq i \leq n+1, 1 \leq j \leq p\}
\end{aligned} \tag{D2}$$

$$\begin{aligned}
& \cup \{manpropose'(x, y), womanpropose'(x, y) \mid x \in M, y \in W\} \\
& \cup \{nmanpropose'(x, y), nwomanpropose'(x, y) \mid x \in M, y \in W\} \\
& \cup \{accept'(x, y) \mid x \in M, y \in W\} \cup \{accept'(x, x) \mid x \in M \cup W\} \\
& \cup \{naccept'(x, y) \mid x \in M, y \in W\} \cup \{naccept'(x, x) \mid x \in M \cup W\}
\end{aligned} \tag{D3}$$

$$\cup \{crit'(val) \mid val \in \arg(crit)\}$$

$$\cup \{single'(i, j) \mid crit = singles, 1 \leq i \leq n+p, j \in \{0, 1\}\}$$

$$\cup \{singlesum'(i, j) \mid crit = singles, 1 \leq i \leq n+p, 1 \leq j \leq n+p-i+1\}$$

$$\cup \{mansum'(i, j) \mid crit \in \{sexeq, weight\}, 1 \leq i \leq n,$$

$$\begin{aligned}
& n - i + 1 \leq j \leq (n - i + 1)(p + 1) \\
\cup \{ & \text{womansum}'(j, i) \mid \text{crit} \in \{\text{sexeq}, \text{weight}\}, 1 \leq j \leq p, \\
& p - j + 1 \leq i \leq (p - i + 1)(n + 1) \} \\
\cup \{ & \text{manweight}'(z) \mid \text{crit} \in \{\text{sexeq}, \text{weight}\}, n \leq z \leq n(p + 1) \} \\
\cup \{ & \text{womanweight}'(z) \mid \text{crit} \in \{\text{sexeq}, \text{weight}\}, p \leq z \leq p(n + 1) \} \\
\cup \{ & \text{manmax}'(i, j) \mid \text{crit} = \text{regret}, 1 \leq i \leq n, 1 \leq j \leq p + 1 \} \\
\cup \{ & \text{womanmax}'(j, i) \mid \text{crit} = \text{regret}, 1 \leq j \leq p, 1 \leq i \leq n + 1 \} \\
\cup \{ & \text{manregret}'(z) \mid \text{crit} = \text{regret}, 1 \leq z \leq p + 1 \} \\
\cup \{ & \text{womanregret}'(z) \mid \text{crit} = \text{regret}, 1 \leq z \leq n + 1 \}
\end{aligned} \tag{D4}$$

The notation $\arg(c)$ stands for the possible values the criterion can take within this problem instance:

- if $\text{crit} = \text{sexeq}$ then $\arg(\text{crit}) = \{0, \dots, \max(np + n - p, np + p - n)\}$,
- if $\text{crit} = \text{weight}$ then $\arg(\text{crit}) = \{n + p, \dots, 2np + p + n\}$,
- if $\text{crit} = \text{regret}$ then $\arg(\text{crit}) = \{1, \dots, \max(p, n) + 1\}$,
- if $\text{crit} = \text{singles}$ then $\arg(\text{crit}) = \{0, \dots, n + p\}$.

To verify whether this interpretation is an answer set of $\mathcal{P}_{\text{crit}}$, we should compute the reduct w.r.t. I and check whether I is a minimal model of the reduct (Gelfond and Lifschitz 1988). It can readily be checked that I satisfies all the rules of $\text{red}(\mathcal{P}_{\text{crit}}, I)$. It remains to be shown that there is no strict subset of I which satisfies all the rules. First of all, all the facts of $\mathcal{P}_{\text{crit}}$ must be in the minimal model of the reduct, explaining why the sets of literals (D1) should be in I . The only rules with negation-as-failure are part of $\mathcal{P}_{\text{ext}}^{\text{crit}}$.

As in the previous part of the proof, it is straightforward to see that I_1 is the unique minimal model of the reduct of $\mathcal{P}_{\text{ext}}^{\text{crit}}$ w.r.t. I , considering that the literals in I_2 do not occur in $\mathcal{P}_{\text{ext}}^{\text{crit}}$. So any minimal model of $\text{red}(\mathcal{P}_{\text{crit}}, I)$ must contain I_1 .

The key rule which makes sure that I is a minimal model of the reduct is (13). The rules (11) imply that for each model of $\text{red}(\mathcal{P}_{\text{crit}}, I)$ that does not contain sat , the literals of $\mathcal{P}'_{\text{ext}}^{\text{crit}}$ in that model will correspond to a stable matching of the SMTI instance. In that case rule (13) will have a true body, since S is optimal, implying that sat should have been in the model. And the presence of sat in any minimal model implies the presence of the set of literals (D4) in any minimal model of the reduct. This can be seen with the following reasoning. Due to the presence of the facts (D1) and sat in any minimal model of the reduct, rules (15) imply the presence of the literals (D2) in any minimal model. For the same reason rules (16) imply that the literals (D3) should be in any minimal model of $\text{red}(\mathcal{P}_{\text{crit}}, I)$. For $\text{crit} = \text{sexeq}$ the presence of the literals of the form (D2) in any minimal model of the reduct together with rules (12) imply that $\text{mansum}'(i, j)$ should be in any minimal model for every $i \in \{1, \dots, n\}$ and $j \in \{n - i + 1, \dots, (n - i + 1)(p + 1)\}$: for $i = n$, the first rule of (12) implies that $\text{mansum}'(n, x)$ is in any minimal model for every x such that $\text{manargcost}'_2(x)$ is in it, i.e. any $x \in \{1, \dots, p + 1\}$. Now the second rule of (12) implies that $\text{mansum}'(n - 1, x)$ is in any minimal model for every $x + y$ such that $\text{manargcost}'_2(x)$ and $\text{mansum}'(n, y)$ are in it, i.e. any

$x + y \in \{2, \dots, 2(p + 1)\}$. If we continue like this, it is straightforward to see that every literal of the form $mansum'(\cdot, \cdot)$ of I_2 should be in any minimal model. The third rule of (12) now implies that $manweight'(x)$ should be in any minimal model for every x such that $mansum'(1, x)$ is in it, i.e. $x \in \{n, \dots, n(p + 1)\}$. The same reasoning can be repeated for the literals $womansum'$ and $womanweight'$. At this point rules (10) imply that $sexeq'(|x - y|)$ should be in any minimal model which contains $manweight'(x)$ and $womanweight'(y)$. Note that only one of the two rules in (10) will apply for every x and y since the numerical variables in DLV are positive. Considering the arguments for which $manweight'$ and $womanweight'$ should be in any minimal model, it follows that $sexeq'(x)$ should be in any minimal model for every $x \in \{0, \dots, \max(p(n + 1) - n, n(p + 1) - p)\}$, which is exactly $\arg(crit)$. For the other criteria, an analogous reasoning shows that the presence of all literals of I_2 is required in any minimal model of the reduct.

Considering the fact that we have proved that all literals of I should be in any minimal model of the reduct and I fulfils all the rules of the reduct, we know that I is a minimal model of the reduct and thus an answer set of \mathcal{P}_{crit} . \square

References

- GALE, D. AND SHAPLEY, L. 1962. College admissions and the stability of marriage. *The American Mathematical Monthly* 69, 1, 9–15.
- GALE, D. AND SOTOMAYOR, M. 1985. Some remarks on the stable matching problem. *Discr. Appl. Math.* 11, 223–232.
- GELFOND, M. AND LIFSCHITZ, V. 1988. The stable model semantics for logic programming. In *ICLP/SLP*. Vol. 88. 1070–1080.
- GUSFIELD, D. 1987. Three fast algorithms for four problems in stable marriage. *SIAM J. Comput.* 16, 1, 111–128.
- IRVING, R., LEATHER, P., AND GUSFIELD, D. 1987. An efficient algorithm for the “optimal” stable marriage. *J. ACM* 34, 3, 532–543.
- KATO, A. 1993. Complexity of the sex-equal stable marriage problem. *Japan Journal of Industrial and Applied Mathematics (JJIAM)* 10, 1–19.
- MANLOVE, D. 1999. Stable marriage with ties and unacceptable partners. Tech. rep., University of Glasgow, Department of Computing Science.
- MANLOVE, D., IRVING, R., IWAMA, K., MIYAZAKI, S., AND MORITA, Y. 2002. Hard variants of stable marriage. *Theoretical Computer Science* 276, 1-2, 261–279.
- MCDERMID, E. AND IRVING, R. 2012. Sex-equal stable matchings: Complexity and exact algorithms. *Algorithmica*, 1–26.