Online appendix for the paper

Solving stable matching problems using answer set programming

published in Theory and Practice of Logic 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:* 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 9 December 2014; revised 2 July 2015; accepted 14 December 2015

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 stricly 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 set[1].

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 $\{(x, y) \mid accept(x, y) \in I\}$. Conversely, if $\{(x_1, y_1), \ldots, (x_k, y_k)\}$ is a weakly stable matching for (S_M, S_W) then $\mathcal{P}$ has the following answer set I:

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

[1] We assume that P $\neq$ NP

$\cup\{womanpropose(x, y_i) \,|\, i \in \{1, \ldots, k\}, y_i \in W, x <_W^{y_i} x_i \vee x = x_i \neq y_i\}$
$\cup\{accept(x_i, y_i) \,|\, i \in \{1, \ldots, k\}\}$

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.
$\boxed{\text{Answer set} \Rightarrow \text{weak stable set}}$ We prove this in 4 steps.

1. *For every $i \in \{1, \ldots, n\}$, every $j \in \{1, \ldots, p\}$ and for every answer set I of $\mathcal{P}$, it holds that $accept(m_i, w_j) \in I$ implies that $j \in acceptable_M^i$ and $i \in acceptable_W^j$.*
   This can be proved by contradiction. We will prove that for every man m_i and every $j \in unacceptable_M^i$, $accept(m_i, w_j)$ is in no answer set I of the induced ASP program $\mathcal{P}$. For $accept(m_i, w_j)$ 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 $manpropose(m_i, w_j)$ should be in I. But since j is not in $acceptable_M^i$ there is no rule with $manpropose(m_i, w_j)$ in the head and so $manpropose(m_i, w_j)$ 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 $accept(m_i, w_j) \in I$. Similarly, for every woman w_j there exists at most one man m_i such that $accept(m_i, w_j) \in I$. Moreover, if $accept(m_i, m_i) \in I$ then $accept(m_i, w_j) \notin I$ for any w_j, and likewise when $accept(w_j, w_j) \in I$ then $accept(m_i, w_j) \notin I$ for any m_i.*
   This can be proved by contradiction. Suppose first that there is an answer set I of $\mathcal{P}$ that contains $accept(m_i, w_j)$ and $accept(m_i, w_{j'})$ 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 $manpropose(m_i, w_{j'}) \leftarrow \ldots, not\, accept(m_i, w_j), \ldots$ and this is the only rule which can make $manpropose(m_i, w_{j'})$ true (the only rule with this literal in the head). However, since $accept(m_i, w_j)$ is also in the answer set, the body of this rule is not satisfied so $manpropose(m_i, w_{j'})$ can never be in I. Consequently $accept(m_i, w_{j'})$ 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 $accept(m_i, w_j)$ and $accept(m_i, m_i)$ are both in an answer set I of $\mathcal{P}$. Again step 1 implies that $j \in acceptable_M^i$. Because of the rules (2), $\mathcal{P}$ will contain the rule $accept(m_i, m_i) \leftarrow \ldots, not\, accept(m_i, w_j), \ldots$. An analogous reasoning as above implies that since $accept(m_i, w_j)$ is in the answer set I, $accept(m_i, m_i)$ 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 $accept(m_i, w_j) \in I$,*
   (b) *$accept(m_i, m_i) \in I$,*

   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 $accept(m_i, w_j)$ 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 $accept(m_i, m_i) \leftarrow$, and so $accept(m_i, m_i)$ 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 $accept(m_i, w_{j'}) \in I$ and $accept(m_{i'}, w_j) \in I$ 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 $accept(m_i, w_{j'})$ and $accept(m_{i'}, w_j)$ in the head, imply that literals $manpropose(m_i, w_{j'})$ and $womanpropose(m_{i'}, w_j)$ 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 $manpropose(m_i, w_j)$ to be in I than for $manpropose(m_i, w_{j'})$ to be in I. Therefore, $manpropose(m_i, w_j)$ should be in I as well. A similar reasoning implies that $womanpropose(m_i, w_j)$ should be in I. But now the rules of the form (1) imply that $accept(m_i, w_j)$ should be in I, contradicting step 2 since $accept(m_i, w_{j'})$ and $accept(m_{i'}, w_j)$ are already in I.

$\boxed{\text{Weak stable set} \Rightarrow \text{answer set}}$ Suppose we have a stable set of marriages $S = \{(x_1, y_1), \ldots, (x_k, y_k)\}$, 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 $manpropose(m, w)$ is in I and $womanpropose(m, w)$ 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 $manpropose(m_i, w_j)$ and $womanpropose(m_i, w_j)$ 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 $accept(m_i, w_j)$ 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 $accept(m_i, .)$ in I, so rule (2) will reduce to a fact $accept(m_i, m_i) \leftarrow$, 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 $accept(m_i, w)$ 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 $accept(m_i, .)$ 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 $neutral^i_M$ 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 $manpropose(.,.)$, $womanpropose(.,.)$ and $accept(.,.)$ 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 $\mathcal{P}$. □

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

Proof
Let (S_M, S_W) be an SMTI instance.
$\boxed{\text{Answer set} \Rightarrow \text{optimal stable matching}}$ 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 $manpropose(.,.)$, $womanpropose(.,.)$ and $accept(.,.)$ 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 $manpropose(.,.)$, $womanpropose(.,.)$ and $accept(.,.)$ 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 $accept(.,.)$ in I_{norm} uniquely determine which literals of the form $mancost(.,.)$, $womancost(.,.)$, $manweight(.)$, $womanweight(.)$ and $sexeq(.)$ should be in the answer set I. These literals do not occur in rules of $\mathcal{P}_{crit}$ besides those in $\mathcal{P}^{sexeq}_{ext}$. Note that the rules which do contain these literals imply that there will be just one literal of the form $sexeq(.)$ 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 $\mathcal{P}'^{crit}_{ext}$. An analogous reasoning as above yields that the literals of the form $accept'(.,.)$ in I'_{disj} uniquely determine which literals of the form $mancost'(.,.)$, $womancost'(.,.)$, $mansum'(.,.)$, $womansum'(.,.)$, $manweight'(.)$, $womanweight'(.)$ and $sexeq'(.)$ should be in I'_{disj}. With those extra literals added to I'_{disj}, we find that I'_{disj} satisfies all the rules of $\mathcal{P}'^{crit}_{ext}$. Moreover, $crit(v^*)$ is the unique literal of the form $crit(.)$ 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 $\mathcal{P}^{crit}_{ext} \cup \mathcal{P}'^{crit}_{ext}$ 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 $crit(.)$ or $crit'(.)$. 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 $mancost'(.,.), womancost'(.,.)$ and $manpropose'(.,.)$, $womanpropose'(.,.)$, $accept'(.,.)$ 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 $\mathcal{P}'^{crit}_{ext}$, 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 garantuee that I also contains $mansum'(.,.)$, $manweight(.)$, $womansum(.,.)$ and $womanweight(.)$ for every argument that could occur in a model of $\mathcal{P}'^{crit}_{ext}$. Note that this would not be the case if we used the original rules with $\#sum$, $\#max$ and $\#count$ in $\mathcal{P}'^{crit}_{ext}$, 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 $\mathcal{P}^{crit}_{ext}$, 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 $\mathcal{P}'^{crit}_{ext}$ with negation-as-failure in the body, hence $red(\mathcal{P}'^{crit}_{ext}, I) = red(\mathcal{P}'^{crit}_{ext}, J') = \mathcal{P}'^{crit}_{ext}$. We already reasoned that J' satifies all the rules of the latter. We also reasoned that I does not contain any other literals of the form $accept(.,.)$ than those which are also in I_{norm}, and by construction the same holds for J'. Hence $red(\mathcal{P}^{crit}_{ext}, I) = red(\mathcal{P}^{crit}_{ext}, J')$ 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).

$\boxed{\text{Optimal stable matching} \Rightarrow \text{answer set}}$ Let $S = \{(x_1, y_1), \ldots, (x_k, y_k)\}$ 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 $y \in \sigma^l_M(a)$ if $x_i = m_l$ and symmetrically $P_{y_i}(x)$ as the index a for which $x \in \sigma^{l'}_W(a)$ if $y_i = w_{l'}$. If $x_i = y_i$ we set $P_{x_i}(y_i) = P_{y_i}(x_i) = |\sigma^i_M|$ if x_i is a man and $P_{x_i}(y_i) = P_{y_i}(x_i) = |\sigma^i_W|$ otherwise. So let I be given by: $I = I_1 \cup I_2$ with

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

and

$$
\begin{aligned}
I_2 = &\{manargcost'_1(z) \mid 1 \leq z \leq n\} \cup \{manargcost'_2(z) \mid 1 \leq z \leq p + 1\} \\
&\cup \{womanargcost'_1(z) \mid 1 \leq z \leq p\}\} \cup \{womanargcost'_2(z) \mid 1 \leq z \leq n + 1\} \\
&\cup \{man(x) \mid x \in M\} \cup \{woman(x) \mid x \in W\} && \text{(D1)} \\
&\cup \{mancost'(i, j) \mid crit \neq singles, 1 \leq i \leq n, 1 \leq j \leq p + 1\} \\
&\cup \{womancost'(j, i) \mid crit \neq singles, 1 \leq i \leq n + 1, 1 \leq j \leq p\} && \text{(D2)} \\
&\cup \{manpropose'(x, y), womanpropose'(x, y) \mid x \in M, y \in W\} \\
&\cup \{nmanpropose'(x, y), nwomanpropose'(x, y) \mid x \in M, y \in W\} \\
&\cup \{accept'(x, y) \mid x \in M, y \in W\} \cup \{accept'(x, x) \mid x \in M \cup W\} \\
&\cup \{naccept'(x, y) \mid x \in M, y \in W\} \cup \{naccept'(x, x) \mid x \in M \cup W\} && \text{(D3)} \\
&\cup \{crit'(val) \mid val \in \arg(crit)\} \\
&\cup \{single'(i, j) \mid crit = singles, 1 \leq i \leq n + p, j \in \{0, 1\}\} \\
&\cup \{singlesum'(i, j) \mid crit = singles, 1 \leq i \leq n + p, 1 \leq j \leq n + p - i + 1\}\} \\
&\cup \{mansum'(i, j) \mid crit \in \{sexeq, weight\}, 1 \leq i \leq n,
\end{aligned}
$$

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

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

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

To verify whether this interpretation is an answer set of $\mathcal{P}_{crit}$, 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 $\mathit{red}(\mathcal{P}_{crit}, I)$. 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 $\mathcal{P}_{crit}$ 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 $\mathcal{P}_{ext}^{crit}$.
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 $\mathcal{P}_{ext}^{crit}$ w.r.t. I, considering that the literals in I_2 do not occur in $\mathcal{P}_{ext}^{crit}$. So any minimal model of $\mathit{red}(\mathcal{P}_{crit}, I)$ 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 $\mathit{red}(\mathcal{P}_{crit}, I)$ that does not contain sat, the literals of $\mathcal{P}_{ext}'^{crit}$ 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 $\mathit{red}(\mathcal{P}_{crit}, I)$. For $\mathit{crit} = \mathit{sexeq}$ the presence of the literals of the form (D2) in any minimal model of the reduct together with rules (12) imply that $\mathit{mansum}'(i,j)$ should be in any minimal model for every $i \in \{1, \ldots, n\}$ and $j \in \{n-i+1, \ldots, (n-i+1)(p+1)\}$: for $i = n$, the first rule of (12) implies that $\mathit{mansum}'(n,x)$ is in any minimal model for every x such that $\mathit{manargcost}_2'(x)$ is in it, i.e. any $x \in \{1, \ldots, p+1\}$. Now the second rule of (12) implies that $\mathit{mansum}'(n-1,x)$ is in any minimal model for every $x+y$ such that $\mathit{manargcost}_2'(x)$ and $\mathit{mansum}'(n,y)$ are in it, i.e. any

$x+y \in \{2, \ldots, 2(p+1)\}$. If we continue like this, it is straightforward to see that every literal of the form $mansum'(.,.)$ 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. $x \in \{n, \ldots, n(p+1)\}$. 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 $x \in \{0, \ldots, \max(p(n+1)-n, n(p+1)-p)\}$, 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 ans 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.