TPLP **21** (6): 852–869, 2021. © The Author(s), 2021. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited. doi:10.1017/S1471068421000302 First published online 27 September 2021

Knowledge-Based Stable Roommates Problem: A Real-World Application

MÜGE FIDAN and ESRA ERDEM

Faculty of Engineering and Natural Sciences, Sabanci University, Istanbul, Turkey (e-mails: mugefidan@sabanciuniv.edu, esra.erdem@sabanciuniv.edu)

submitted 11 August 2021; accepted 22 August 2021

Abstract

The Stable Roommates problem with Ties and Incomplete lists (SRTI) is a matching problem characterized by the preferences of agents over other agents as roommates, where the preferences may have ties or be incomplete. SRTI asks for a matching that is stable and, sometimes, optimizes a domain-independent fairness criterion (e.g. Egalitarian). However, in real-world applications (e.g. assigning students as roommates at a dormitory), we usually consider a variety of domain-specific criteria depending on preferences over the habits and desires of the agents. With this motivation, we introduce a knowledge-based method to SRTI considering domain-specific knowledge and investigate its real-world application for assigning students as roommates at a university dormitory.

KEYWORDS: stable roommates problem, answer set programming, declarative problem solving

1 Introduction

The Stable Roommates (SR) problem (Gale and Shapley 1962) is a matching problem characterized by the preferences of agents over other agents as roommates: each agent ranks all others in strict order of preference. A solution to SR is then a partition of the agents into pairs that are *acceptable* to each other (i.e. they are in the preference lists of each other), such that the matching is *stable* (i.e. there exist no two agents who prefer each other to their roommates, and thus *block* the matching).

SR is studied with incomplete preference lists (SRI) (Gusfield and Irving 1989), with preference lists including ties (SRT) (Ronn 1990), and with incomplete preference lists including ties (SRTI) (Irving and Manlove 2002). SRT and SRTI are intractable under weak stability (Ronn 1990; Irving *et al.* 2009).

Optimization variants of SRTI are also studied to find more fair stable solutions. For instance, Egalitarian SRTI aims to maximize the total satisfaction of preferences of all agents. Rank Maximal SRTI aims to maximize the number of agents matched with their first preference, and then, subject to this condition, to maximize the number of agents matched with their second preference, and so on. Almost SRTI aims to minimize the total number of blocking pairs (i.e. pairs of agents who prefer each other to their roommates), if a stable matching cannot be found. These optimization variants are NP-hard (Feder 1992; Cooper 2020; Abraham *et al.* 2005).

These optimization variants of SRTI are based on domain-independent measures. However, in real-world applications (e.g. in dormitory applications), there are also domaindependent criteria that necessitates further knowledge: consider, for instance, dormitory applications that request information about the personal habits of the students, as well as their preferences of the living environment.

In our earlier work (Erdem *et al.* 2020), we have developed a formal framework, called SRTI-ASP, that is flexible enough to provide solutions to all variations of SR mentioned above, including the intractable decision/optimization versions: SRT, SRTI, Egalitarian SRTI, Rank Maximal SRTI, Almost SRTI. SRTI-ASP utilizes the expressive languages and efficient solvers of Answer Set Programming (ASP) (Niemelä 1999; Marek and Truszczyński 1999; Lifschitz 2002; Brewka *et al.* 2016) based on answer set semantics (Gelfond and Lifschitz 1988; 1991).

In this study, we extend SRTI-ASP to accommodate additional domain-specific criteria in two ways: Personalized-SRTI and Most-SRTI. In addition, we extend SRTI-ASP to accommodate diversity preferences and constraints.

- For Personalized-SRTI, we introduce a new type of preference ordering considering (i) the importance of each criterion for each agent (e.g. one student may give more importance to sleeping habits whereas another student may give more importance to smoking habits), and (ii) the agents' preferred choices for each domain-specific criterion (e.g. whether a student prefers a roommate who does not smoke). We define an extended preference list for each agent, that combines two types of preference lists: a preference list of the agent over other agents (as in SRTI) and this new type of criteria-based personalized preference list of the agent. Personalized-SRTI considers these extended preference lists to compute personalized stable matchings.
- For Most-SRTI, we introduce a new incremental definition of a stable matching considering (i) the ordering of the most preferred criteria (e.g. identified by large surveys) and (ii) the agents' preferred choices for each domain-specific criterion, with the motivation that the agents with close choices are matched. Most-SRTI aims to compute such most preferred criteria based stable matchings, by utilizing the weak constraints of ASP.
- In addition to the students' preferences over a set of domain-specific criteria, the schools may prefer matchings (or put constraints over matchings) to increase diversity. For example, they may want to match students from different departments, classes, or countries. With this motivation, we extend SRTI-ASP by representing such diversity preferences/constraints using weak/hard constraints of ASP.

We illustrate a real-world application of SRTI-ASP by interacting with at least 200 students at Sabanci University over four surveys: (i) to decide which domain-specific criteria to consider, (ii) to collect the students preferences for domain-specific criteria, and (iii) to evaluate the usefulness of our method.

We also present results of our experiments with objective and subjective measures, to understand the scalability of the proposed two methods, Personalized-SRTI and Most-SRTI.

2 SRTI: Stable roommates problem with ties and incomplete lists

We define SRTI as in our earlier work (Erdem *et al.* 2020). Let A be a finite set of agents. For every agent $x \in A$, let A_x of $A \setminus \{x\}$ be a set of agents that are *acceptable* to x as roommates. For every y in A_x , we assume that x prefers y as a roommate compared to being single.

Let \prec_x be a partial ordering of x's preferences over A_x where incomparability is transitive. We refer to \prec_x as agent x's preference list. For two agents y and z in A_x , we denote by $y \prec_x z$ that x prefers y to z. In this context, ties correspond to indifference in the preference lists: an agent x is *indifferent* between the agents y and z, denoted by $y \sim_x z$, if $y \not\prec_x z$ and $z \not\prec_x y$. We denote by \prec the collection of all preference lists.

A matching for a given SRI instance is a function $M : A \mapsto A$ such that, for all $\{x, y\} \subseteq A$ such that $x \in A_y$ and $y \in A_x$, M(x) = y if and only if M(y) = x. If agent x is mapped to itself, we then say he/she is *single*.

A matching M is blocked by a pair $\{x, y\} \subseteq A \ (x \neq y)$ if

B1 both agents x and y are acceptable to each other,

- B2 x is single with respect to M, or $y \prec_x M(x)$, and
- B3 y is single with respect to M, or $x \prec_y M(y)$.

A matching for SRTI is called *stable* if it is not blocked by any pair of agents.

We can declaratively solve SRTI using ASP as described in our earlier work (Erdem *et al.* 2020). For that, the input $I = (A, \prec)$ of an SRTI instance is formalized by a set F_I of facts using atoms of the forms agent(x) ("x is an agent in A") and prefer2(x, y, z) ("agent x prefers agent y to agent z, i.e. $y \prec_x z$ "). For every agent x, for every $y \in A_x$, we also add facts of the form prefer2(x, y, x) to express that x prefers y as a roommate instead of being single.

Based on the preferences of agents, for each agent, the concept of acceptability is defined:

$$accept(x, y) \leftarrow prefer(x, y, _).$$
$$accept(x, y) \leftarrow prefer(x, _, y).$$

as well as the concept of mutual acceptability:

$$accept2(x, y) \leftarrow accept(x, y), accept(y, x).$$

The output $M : A \mapsto A$ of an SRTI instance is characterized by atoms of the form room(x, y) ("agents x and y are roommates"). The ASP formulation P of SRTI first generates pairs of roommates. For every agent x, exactly one mutual acceptable agent y is nondeterministically chosen as M(x) by the choice rules:

$$1\{room(x, y): agent(y), accept2(x, y)\}1 \leftarrow agent(x). \\ \leftarrow room(x, y), not \ room(y, x).$$

Then, the stability of the generated matching is ensured by the hard constraints:

$$\leftarrow block(x, y) \qquad (x \neq y).$$

where atoms of the form block(x, y) describe the blocking pairs (i.e. conditions B1–B3).

3 Personalized-SRTI: SRTI with personalized criteria

Some universities and colleges send questionnaires to students before making roommate matches, and they match students as roommates taking into account the additional information included in these questionnaires. For instance, the University of Texas has a questionnaire used for applying to dormitories.¹ It contains questions about the sleep preferences, music preferences, and sharing preferences of applicants. In some other surveys, we can see a question about the smoking habits,² preferences regarding the room temperature,³ and willingness to share a room with an international student.⁴ In addition to questions about such different criteria, these questionnaires usually request for the applicants to indicate the most important criteria.¹

With these motivations, we extend SRTI-ASP to include such personal preferences of applicants. We call this extension Personalized-SRTI-ASP.

Personalized-SRTI-ASP considers an aggregate preference list \prec''_x defined over two types of preference lists: \prec_x as defined in the previous section, and \prec'_x to capture the additional preferences as discussed above.

Defining the criteria-based personalized preference lists \prec'_x . Let us first introduce some definitions and notations as follows.

Let B be a finite list $\langle b_1, b_2, \ldots, b_k \rangle$ of criteria. For each criterion $b_i \in B$, let C_i be a finite list $\langle c_{i1}, c_{i2}, \ldots, c_{im} \rangle$ of choices for b_i , that is ordered with respect to a "closeness" measure (i.e. for every choice c_{ij} , the choice $c_{ij'}$ (j < j') is "closer" than the choice $c_{ij''}$ (j' < j'')). The closeness measure is useful for matching agents with closer choices, as roommates. For instance, consider the criteria list $B = \langle$ "cleanliness", "sleep habits" \rangle . For each criterion, the choice lists can be defined as follows: $C_1 = \langle$ "Clean", "Messy" \rangle is the list of choices for "cleanliness", and $C_2 = \langle$ "Goes to bed early", "Goes to bed before midnight", "Goes to bed after midnight" \rangle is the list of choices for "sleep habits".

Let f be a function that maps an agent $x \in A$ and a criterion $b_i \in B$ to a positive integer j $(1 \leq j \leq |C_i|)$, describing the choice c_{ij} of the agent x. Consider the example above, and assume that Ayse is an agent in A. If Ayse's preference for the "cleanliness" criterion is "Clean", then f(Ayse, "cleanliness") = 1. If Ayse's preference for "sleep habits" criterion is "Goes to bed after midnight", then f(Ayse, "sleep habits") = 3.

For every agent $x \in A$, let us denote by $P_x = \langle f(x, b_1), f(x, b_2), \ldots, f(x, b_k) \rangle$ the choices of x for each criterion in B respectively. We refer to P_x as the agent x's (preference) profile. Consider the example shown in Table 1. The preference profile P_{Buse} for agent Buse is $\langle 1, 2, 3, 3, 3 \rangle$ where $B = \langle$ "smoking", "cleanliness", "environment", "sleep habits", "study habits" \rangle . According to P_{Buse} , Buse prefers a roommate that is a "Smoker", 'Messy", "Social and quiet", "Goes to bed after midnight", "Studies in and out of the room."

Every criterion in B may have a different importance for each agent. For instance, agent Ayse may give more importance to "study habits" while agent Buse gives more importance to "cleanliness." To take into account the importance of these criteria, we introduce a weight function w that maps an agent $x \in A$ and a criterion $b_i \in B$ to a non-negative integer such that $w(x, b_i)$ denotes the importance of the criterion b_i for $x \in A$. For every

¹ https://tams.unt.edu/studentlife/roommate-preferences-questionnaire

² https://my.clevelandclinic.org/-/scassets/files/org/professionals/student-housing/

roommate-questionnaire-worksheet

³ https://www.wells.edu/files/public/forms/Housing_Roommate_Questionnaire-fillable.pdf

⁴ https://college.lclark.edu/live/files/27111-2019-20-returning-student-questionnaire

Table 1. A personalized-SRTI instance defined over an agent set $A = \{Ayse, Buse, Cem, Duru\}$, a criteria list $B = \langle$ "smoking", "cleanliness", "room environment", "sleep habits", "study habits" \rangle , and the following choice lists for each criterion, $C_1 = \langle$ "Smoker", "Non-smoker" \rangle , $C_2 = \langle$ "Clean", "Messy" \rangle , $C_3 = \langle$ "Quiet", "Social", "Social and quiet" \rangle , $C_4 = \langle$ "Goes to bed early", "Goes to bed before midnight", "Goes to bed after midnight \rangle , $C_5 = \langle$ "Studies in the room", "Studies out of the room" \rangle .

Agent x	Preference list \prec_x	Profile P_x	Weight list W_x	Extended preference list \prec'_a
Ayse (non-smoker)	$\langle Duru \rangle$	$\langle 2, 1, 1, 1, 1 \rangle$	$\langle 5,4,3,2,1 \rangle$	$\langle {\rm Duru, \ Cem} \rangle$
Buse (smoker)	$\langle \rangle$	$\langle 1, 2, 3, 3, 3 \rangle$	$\langle 1,0,3,4,5 \rangle$	$\langle Duru, Cem \rangle$
Cem (non-smoker)	$\langle Ayse, Buse \rangle$	$\langle 2, 1, 3, 2, 3 \rangle$	(5,5,4,3,2)	$\langle Ayse, Buse, Duru \rangle$
Duru (non-smoker)	$\langle \mathrm{Cem} \rangle$	$\langle 2, 1, 3, 3, 3 \rangle$	$\langle 3,3,3,3,3\rangle$	$\langle \mathrm{Cem},\mathrm{Buse},\mathrm{Ayse}\rangle$

agent $x \in A$, let us denote by the weight list $W_x = \langle w(x, b_1), w(x, b_2), \dots, w(x, b_k) \rangle$ the respective weights of criteria in B for x. Note that $w(x, b_i) > w(x, b_j)$ implies that the criterion b_i is more important than the criterion b_j for agent x. We say that $w(x, b_i) = 0$ to indicate that the criterion b_i is not important for agent x. For the example shown in Table 1, $W_{Buse} = \langle 1, 0, 3, 4, 5 \rangle$: the most important criterion for Buse is "study habits", and the "cleanliness" criterion is not important.

For every agent $x \in A$, with a profile P_x and a weight list W_x , let us denote the criteria of the same weight u > 0 and the agent x's choices for them, by a nonempty set E_u of tuples as follows:

$$E_u = \{ (f(x, \pi_i), \pi_i) \mid u = w(x, \pi_i) > 0, \ \pi_i \in \{b_1, b_2, \dots, b_k\} \}.$$

Then, for every agent $x \in A$, we define a sorted profile P'_x for x, with respect to P_x and W_x , as follows:

$$P'_{x} = \langle E_{u_{1}}, E_{u_{2}}, \dots, E_{u_{m}} \rangle$$
 where $m \leq k$, and, for each $i \ (1 \leq i < m), u_{i} > u_{i+1}$.

In Table 1, the sorted profile for Cem is $P'_{Cem} = \langle \{(2, \text{"smoking"}), (1, \text{"cleanliness"})\}, \{(3, \text{"room environment"})\}, \{(2, \text{"sleep habits"})\}, \{(3, \text{"study habits"})\}\rangle$ considering the importance of each criterion for him: w(Cem, "smoking") = w(Cem, "cleanliness") = 5, w(Cem, "room environment") = 4, w(Cem, "sleep habits") = 3, w(Cem, "study habits") = 2.

For every agent $y \in A \setminus A_x$ (i.e. y is not acceptable to x), if there exists some criterion $b_i \in B$ where $w(x, b_i) > 0$ such that $f(y, b_i) = f(x, b_i)$, then we say that y is choiceacceptable to x. We denote by A'_x the set of all agents in $A \setminus A_x$ that are choice-acceptable for x. In Table 1, since Ayse has no common choice with Buse, Ayse is not choiceacceptable for Buse. On the other hand, Duru has a common choice with Buse: f(Duru,"study habits") = f(Buse, "study habits") = 3; and thus Duru is choice-acceptable for Buse. We assume that x prefers every choice-acceptable y as a roommate compared to being single.

For every agent x with a sorted profile $P'_x = \langle E_{u_1}, E_{u_2}, \ldots, E_{u_m} \rangle$ $(m \leq k)$, for every two agents y and z that are choice-acceptable to x, the agents y and z are *choice-equal* for x relative to the first j sets $E_{u_1}, E_{u_2}, \ldots, E_{u_j}$ in P'_x (denoted $y =_x z|_j$) if the following holds:

- j = 0, or
- $j > 0, y =_x z|_{j-1}$, and, for every $(f(x, \pi_i), \pi_i) \in E_{u_j}, f(x, \pi_i) = f(y, \pi_i) = f(z, \pi_i).$

We say that x prefers y to z with respect to a sorted profile P'_x (denoted $y \prec'_x z$) if the following holds for some j > 0:

- $y =_x z|_{j-1}$, and
- $|\{\pi_i|(f(x,\pi_i),\pi_i)\in E_{u_j}, f(x,\pi_i)=f(y,\pi_i)\}| > |\{\pi_i|(f(x,\pi_i),\pi_i)\in E_{u_j}, f(x,\pi_i)=f(z,\pi_i)\}|.$

For agent x, we say that $y \sim'_x z$ if $y \not\prec'_x z$ and $z \not\prec'_x y$.

A criteria-based personalized preference list \prec'_x is a partial ordering of x's preferences over A'_x with respect to a sorted profile P'_x , where such incomparability is transitive.

For example, in Table 1, for Ayse, $P'_{Ayse} = \langle E_5, E_4, E_3, E_2, E_1 \rangle$ where $E_5 = \{(2, \text{"smoking"})\}, E_4 = \{(1, \text{"cleanliness"})\}, E_3 = \{(1, \text{"room environment"})\}, E_2 = \{(1, \text{"sleep habits"})\}$, and $E_1 = \{(1, \text{"study habits"})\}$. Cem is choice-acceptable for Ayse: f(Cem, "smoking") = f(Ayse, "smoking") = 2. Then, the criteria-based personalized preference list \prec'_{Ayse} is $\langle Cem \rangle$: Ayse prefers Cem as a roommate compared to being single.

For Buse, $P'_{Buse} = \langle E_5, E_4, E_3, E_1 \rangle$ where $E_5 = \{(3, \text{"study habits"})\}$, $E_4 = \{(3, \text{"sleep habits"})\}$, $E_3 = \{(3, \text{"room environment"})\}$, $E_1 = \{(1, \text{"smoking"})\}$. Since Ayse has no common choice with Buse, Ayse is not choice-acceptable for Buse. On the other hand, Cem and Duru are choice-acceptable for Buse. Then, Duru \prec'_{Buse} Cem since

- for the criterion π_1 = "study habits" in E_5 , $f(\text{Buse},\pi_1)=f(\text{Duru},\pi_1)=f(\text{Cem},\pi_1)=3$, and thus Duru =_{Buse} Cem |₁; and
- for the criterion π_2 ="sleep habits" in E_4 , $f(\text{Duru},\pi_2)=f(\text{Buse},\pi_2)=3$ while $f(\text{Buse},\pi_2)\neq f(\text{Cem},\pi_2)=2$. Therefore, $|\{\pi_i| \ (f(\text{Buse},\pi_i),\pi_i)\in E_4, f(\text{Duru},\pi_i)=f(\text{Buse},\pi_i)\}|=1$ is larger than $|\{\pi_i| \ (f(x,\pi_i),\pi_i)\in E_4, \ f(\text{Buse},\pi_i)=f(\text{Cem},\pi_i)\}|=0$.

Then, the criteria-based personalized preference list \prec'_{Buse} is (Duru, Cem).

For Duru, $P'_{Duru} = \langle E_3 \rangle$ where $E_3 = \{(3, \text{"study habits"}), (3, \text{"sleep habits"}), (3, \text{"room environment"}), (2, \text{"smoking"}), (1, \text{"cleanliness"})\}$. Ayse and Buse are choice-acceptable for Duru. Then, Buse \prec'_{Duru} Ayse since $|\{\pi_i| \ (f(\text{Duru}, \pi_i), \pi_i) \in E_3, \ f(\text{Duru}, \pi_i) = f(\text{Buse}, \pi_i)\}|=3$ is larger than $|\{\pi_i| \ (f(x, \pi_i), \pi_i) \in E_{u_1}, \ f(\text{Duru}, \pi_i) = f(\text{Ayse}, \pi_i)\}|=2$. Then the criteria-based personalized preference list \prec'_{Duru} is $\langle \text{Buse}, \text{Ayse} \rangle$.

Defining the extended preference lists \prec''_x . We define \prec''_x as an extended preference list by concatenating \prec_x and \prec'_x depending on the importance given to these two types of lists. For the instance in Table 1, suppose that the preference lists \prec_x are more important. Then the preference list \prec'_x is appended to end of the preference list \prec_x . Then the extended preference list of Buse is $\prec''_{Buse} = \langle \text{Duru, Cem} \rangle$. The extended preference lists for other agents are as shown in Table 1.

Personalized-SRTI is then characterized by (A, \prec'') where A is a finite set of agent, and \prec'' is collection of the extended preference list of each agent $x \in A$. To solve Personalized-SRTI, we utilize SRTI-ASP as described in Section 2.

4 Most-SRTI: SRTI with most preferred criteria

Instead of considering individual importance of the criteria for each agent, we can consider the most preferred criteria (e.g. identified by large surveys) and try to find stable roommate matchings accordingly. For such applications, we introduce a new definition for stable matchings.

Most-SRTI. Let W be a criteria list $\langle b_1, b_2, \ldots, b_k \rangle$ sorted with respect to their overall importance for all agents. For each criterion $b_i \in W$, let C_i be a finite list of choices ordered with respect to a closeness measure, as discussed in the previous section. Let f be a function that maps an agent $x \in A$ and a criterion $b_i \in W$ to a positive integer j $(a \leq j \leq |C_i|)$.

We start with the set \mathcal{M} of all stable matchings of a given SRTI instance (A, \prec) , and define a series of subsets $\mathcal{M}_{max}(i)$ of these matchings to maximize the overall satisfaction of the roommates with respect to the closeness of their choices for the criterion b_1, b_2, \ldots, b_k :

$$\mathcal{M}_{max}(0) = \mathcal{M}$$
$$\mathcal{M}_{max}(i) = \{ M \in \mathcal{M}_{max}(i-1) | 1 \leq i \leq |W|, \forall M' \in \mathcal{M}_{max}(i-1) \text{ s.t. } M' \neq M,$$
$$\sum_{x \in A} |f(x,b_i) - f(M'(x),b_i)| \geq \sum_{x \in A} |f(x,b_i) - f(M(x),b_i)| \}.$$

Then, a stable matching $M \in \mathcal{M}_{max}(|W|)$ is called a most preferred criteria based stable matching with respect to the criteria list W. We call the problem of finding such a stable matching, Most-SRTI.

For example, consider the instance in Table 1. Instead of considering the individual importance of the criteria for each applicant, let us take $W = \langle$ "smoking", "cleanliness", "room environment", "sleep habits", "study habits" \rangle . Hence, we try to find a matching that maximizes first the number of roommates which are close to each others in terms of their smoking criteria, and then, subject to this condition, maximizes the number of roommates which are close to each other in terms of their cleanliness criteria, and then, subject to this condition, maximizes the number of roommates which are close to each other in terms of their subject to this condition, maximizes the number of roommates which are close to each other in terms of their subject to this condition, maximizes the number of roommates which are close to each others in terms of their some environment criteria, and then, subject to this condition, maximizes the number of roommates which are close to each others in terms of their sleep habits criteria, and then, subject to this condition, to maximize the number of roommates which are close to each others in terms of their sleep habits criteria, and then, subject to this condition, to maximize the number of roommates which are close to each others in terms of their sleep habits criteria. A stable matching at the end is called a most-preferred stable matching.

Solving Most-SRTI using ASP. We can solve Most-SRTI in ASP utilizing weighted weak constraints of different priorities. The idea is to introduce weighted weak constraints to express preferences for each criterion, where the higher priorities are given for the most preferred criteria.

For each agent x, for each criterion $b_i \in W$, we describe the choice $r \in C_i$ of x for b_i (i.e. $f(x, b_i) = r$) by atoms. For instance, we introduce atoms of the form bedTime(x, r) to describe that f(x, "sleep habits") = r. Then the preferences of agents can be represented as follows:

- bedTime(x, 1): the agent x prefers a roommate who goes to bed before 11 pm,
- bedTime(x, 2): the agent x prefers a roommate who goes to bed before midnight,
- bedTime(x, 3): the agent x prefers a roommate who goes to bed after midnight.

Using these atoms, the following weak constraint tries to maximize the number of roommates who are close to each other in terms of their sleep habits:

$$\stackrel{\sim}{\leftarrow} \{room(x,y), bedTime(x,r1), bedTime(y,r2)\}.[|r1-r2|@p,x,y]$$
(1)

Here, the priority p is assigned a high value if "sleep habits" is one of the most preferred criteria.

For the "cleanliness" criterion, the preferences of agents can be represented by the following atoms of the forms:

- cleanliness(x, 1): the agent x tends to keep his/her room clean,
- cleanliness(x, 2): the agent x tends to keep his/her room messy.

Using these atoms, the following weak constraints try to maximize the number of roommates who are close to each other in terms of their cleanliness degrees:

$$\widetilde{\leftarrow} \{room(x,y), clean liness(x,r1), clean liness(y,r2)\}.[|r1 - r2|@p,x,y].$$
(2)

Consider, for instance, "smoking" habits. This is an important criterion to match roommates even if they live on a smoke-free campus. According to the following questions:²

- Are you smoker? \circ Yes \circ No
- Are you comfortable with a roommate that is a smoker? \circ Yes \circ No

we can describe the smoking habits of the agents with atoms of the forms smoker(x), nonsmoker(x), and their preferences with the following atoms of the forms:

- comfortableSmoker(x, 1): the agent x is comfortable with a smoker roommate,
- comfortableSmoker(x, 2): the agent x is not comfortable with a smoker roommate.

We can define non-smoker agents who are comfortable with a smoker roommate:

 $smokeComfor(x, y) \leftarrow nonsmoker(x), comfortableSmoker(x, 1), smoker(y).$

We can define agents who is not comfortable with a smoker roommate:

 \neg smokeComfor $(x, y) \leftarrow$ comfortableSmoker(x, 2), smoker(y).

Then the following weak constraints can be added to our ASP formulation to maximize the number of roommates who are comfortable with each others in terms of their smoking habits with the given priority p:

$$\widetilde{\leftarrow} \{room(x,y), \text{not } smokeComfor(x,y), nonsmoker(x), smoker(y)\}.[1@p, x, y] \\ \widetilde{\leftarrow} \{room(x,y), \neg smokeComfor(x,y), smoker(x), smoker(y)\}.[1@p, x, y].$$

$$(3)$$

According to the "room Environment" criterion,⁴ the preferences of agents can be represented by atoms of the form:

- roomEnvironment(x, 1): the agent x wants his/her room to be quiet and study oriented,
- roomEnvironment(x, 2): the agent x wants his/her room to social gathering place for friends to hang out,
- roomEnvironment(x, 3): the agent x wants his/her room to be a combination of social and quiet.

Using these atoms, the following weak constraint tries to maximize the number of roommates who are close to each other in terms of their room description:

 $\underset{\leftarrow}{\leftarrow} \{room(x,y), roomEnvironment(x,r1), roomEnvironment(y,r2)\}.[|r1-r2|@p,x,y].$

(4)

Another important criterion is "study Habits." For this criterion,⁴ the preferences of applicants can be represented by the following atoms of the form studyHabit(x, r):

- studyHabit(x, 1): the agent x expects to study in his/her room,
- studyHabit(x, 2): the agent x expects to study outside of his/her room,
- studyHabit(x, 3): the agent x expects to study both inside and outside of his/her room.

Using these atoms, the following weak constraint tries to maximize the number of roommates which are close to each other in terms of their study environment:

$$\overset{\sim}{\leftarrow} \{room(x,y), studyHabit(x,r1), studyHabit(y,r2)\}.[|r1-r2|@p',x,y].$$
(5)

Here, the priority p' is assigned a lower value if *Study Habits* is not one of the most preferred criteria.

Note that we can combine different domain-independent measures of SRTI with domain-specific measures, by assigning different priorities to them.

5 Diversity preferences

In addition to the student's preferences, the schools may prefer matchings to increase diversity. For example, they may want to match student from various departments, different classes, countries. Also, some students may be forbidden to match with each other (like in the hedonic diversity games (Boehmer and Elkind 2020)) where the school partition the students into two groups for diversity preferences.

Consider, for instance, maximizing the number of roommates from different departments at a university. A student's department can be defined by atoms of the form department(x, d) ("the student x's department is d"). Then, the following weak constraints can be added to our ASP formulation, to maximize the number of roommates from different departments:

 $\stackrel{\sim}{\leftarrow} \{ room(x,y), department(x,d1), department(y,d2), \ d1 \neq d2 \}. [1@p, x, y].$

The school may not want to allow some students to be roommates. Then, such students can be defined by atoms of the form forbidden(x, y) ("students x and y are forbidden to be roommates"), and the following hard constraints can be added to our ASP formulation:

 $\leftarrow forbidden(x, y), \ room(x, y) \qquad (x \neq y).$

Therefore, the diversity-related constraints and preferences can be easily added to SRTI-ASP.



Fig. 1. Scalability of Personalized-SRTI-ASP as the number of criteria and the completeness degrees increase, for instances where the initial completeness degree is 25% (the left figure) and the initial completeness degree is 50% (the right figure).

6 Experimental evaluations

We have experimentally evaluated Personalized-SRTI to understand its scalability over SRTI instances with additional knowledge, and compared Personalized-SRTI with Most-SRTI.

Scalability of Personalized-SRTI. For benchmarks, as a basis, we have used the SRTI instances randomly generated for our earlier experiments (Erdem *et al.* 2020). It is based on the following idea (Mertens 2005): (1) generate a random graph ensemble G(n, p) according to the Erdos-Renyi model (Erdös and Rényi 1960), where *n* is the required number of agents and *p* is the edge probability (i.e. each pair of vertices is connected independently with probability *p*); (2) since the edges characterize the acceptability relations, generate a random permutation of each agents acceptable partners to provide the preference lists. We define the *completeness degree* for an instance as the percentage p * 100.

We have considered instances of different sizes, where the number of agents are 40, 60, 80,100, 150 and 200, and the completeness degrees are 25%, 50%. For each number of agents and for each completeness degree, there are 20 instances. Then, for each instance, for each agent in that instance, we have randomly generated the agent's choices for each criterion, and the importance of each given criteria according to the agent. For each instance, we have considered 2–5 criteria.

In our experiments, we have used CLINGO (Version 5.2.2) on a machine with Intel Xeon(R) W-2155 3.30 GHz CPU and 32GB RAM. The results are shown in Figure 1.

We make the following observations from this figure, similar to our observations (Erdem *et al.* 2020) over SRTI experiments: As the number of agents and the completeness degree increase, the computation times increase. In addition, as the number of criteria increase, the computation times increase.

Note that the initial completeness degree changes as additional knowledge is included about preferences of agents over different criteria. For Personalized-SRTI instance, the completeness degree is around

$$\frac{n \times \frac{d}{100} + n \times (1 - \frac{d}{100}) \times \frac{m}{m+1}}{n} \times 100,$$

Initial completeness degree	A	#Instances with a solution	Personalized-SRT Completeness degree with additional knowledge	I Avg. time (s)	Most-SRTI Avg. time (s)
25%	40	11	83%	0.109	0.015
	60	10		0.379	0.056
	80	13		1.072	0.142
	100	14		2.469	0.353
	150	8		14.556	2.331
	200	10		51.640	11.116
50%	40	11	89%	0.149	0.059
	60	16		0.636	0.257
	80	13		1.903	0.822
	100	12		4.610	2.497
	150	14		42.273	25.563
	200	9		140.378	104.921

Table 2. Personalized-SRTI vs. Most-SRTI

where n is the number of agents, d is the initial completeness degree, and m is the number of criteria. Therefore, the completeness degree of a Personalized-SRTI depends on the initial completeness degree and the number of criteria but not on the number of agents. Consider an instance where d = 25 and m = 3. We expect that the completeness degree be around 81% depending on the preferences of the agents. In fact, the completeness degree in our experiments is 83% (Figure 1).

Personalized-SRTI vs. Most-SRTI. For benchmarks, as a basis, we have used the SRTI instances randomly generated for our earlier experiments (Erdem *et al.* 2020) as described above.

For each instance, for each agent, we have randomly generated the agent's choices for the most popular three criteria, which are cleanliness, sleep habit and study habit. The importance of each given criteria is fixed as 1, 2, 3 respectively.

The results of our experiments for 40–200 agents are shown in Table 2. We can observe that Most-SRTI performs better than Personalized-SRTI. For both approaches, the computation times for finding a stable matching (if one exists) and finding out that there exists no stable matching are comparable to each other. We can make further observation: As the completeness degree increase, the computation times of Most-SRTI more increase than Personalized-SRTI.

7 A real-world application

In collaboration with more than 200 students at Sabanci University, we have investigated the applicability of our methods for Personalized-SRTI.

First, we have conducted a survey to select the most important 5 criteria that should be included in a dormitory application. Next, we have conducted a survey to get the preferences of each student for each criterion. Next, we have conducted a survey to evaluate the usefulness of Personalized-SRTI from the perspective of students. The surveys are given several months apart from each other, considering the availabilities of students.

Part 1: About Yourself

- 1. Do you live in dormitory?
 - Yes
 - \circ No
- 2. Which one do you prefer?
 - \circ Random roommate
 - Roommate with similar expectations

Part 2: About Dormitory Questionnaires

- 3. Which multiple choice questions should be included in the roommate questionnaire?
- o "I like living in a ..." (a) Clean Space (b) Messy Space (c) Indifferent
- "My ideal room temperature is ..." (a) Cold (below 18 C) (b) Fairly cold (18 C-21 C)
 - (c) Fairly warm (21 C-24 C) (d) Warm (above 24 C)
- o "I go to bed ..." (a) Before 11pm (b) Before Midnight (c) After Midnight
- o "I get up ..." (a) Before 8am (b) 8am 10am (c) 10am 12pm (d) 12pm or later
- "Are you a smoker?" & "Are you comfortable with a roommate that is a smoker?" (a) Yes
 - (b) No

• "I would describe myself as ..." & "I would like a roommate who is ..." (a) Shy

- (b) Fairly Shy
- (c) Fairly Outgoing
- (d) Outgoing

• "I want my room to be ..." (a) Quiet and study oriented

(b) A social gathering place for friends to hang out

- (c) A combination of social and quiet
- "I expect to study ..." (a) In my room (b) Outside of my room

(c) Both inside and outside of my room

7.1 First survey: Which criteria should be considered in a dormitory questionnaire?

Students prefer short application forms. With this motivation, first we have conducted a survey to find out which multiple choice questions presented in the second part of Table 3 should be included in a dormitory questionnaire.

We have conducted this survey online (due to pandemic), at Sabanci University: 156 students have participated, 120 of them live in the dormitories, and 36 of them do not.

Figure 2 shows the most popular five questions, chosen by more than 100 students.

7.2 Second survey: What are your preferences?

As a result of the first survey in Table 3, a roommate questionnaire (Table 4) is prepared with respect to the most preferred five criteria. The purpose of this survey is to generate real data for roommate matching: for each student, we gather the importance of criteria as well as their preferences for each criterion.

We have conducted this survey online (due to pandemic), at Sabanci University: 81 students have filled this survey.

According to the survey results, the following order of the given criteria describes the overall importance: smoking habits, cleanliness, room environment, sleep habits, and study habits.

M. Fidan and E. Erdem



Fig. 2. Results of the first survey.

This suggests solving a Most-SRTI problem instance, where the goal is to find a most preferred criteria based stable matching that tries to maximize first the number of roommates which are comfortable with each others in terms of their smoking habits, and then, subject to this condition, the number of roommates which are close to each others in terms of their cleanliness degree, and then, subject to this condition, the number of roommates which are close to each others in terms of their room description, and then, subject to this condition, the number of roommates which are prefer the same bedtime as close as possible, and then, subject to this condition, the number of roommates which are close to each others in terms of their study environment.

As described in Section 4, we add weighted weak constraints to our ASP formulation of SRTI (as described in Section 2) to express preferences for each one of these five criteria, where the higher priorities are given for the most preferred criteria. Since the most important criteria is smoking habits, we add a weak constraint (3) where the priority p is 5. Then, the next important criteria is cleanliness, we add a weak constraint (2) where the priority p is 4. Then, we add a weak constraint (4) where the priority p is 3. Then, we add a weak constraint (1) where the priority p is 2. Finally, we add a weak constraint (5) where the priority p is 1.

Using our ASP program augmented with all these weak constraints, we have experimented over the real data collected in this survey (i.e. preferences of agents for each criterion). A most preferred criteria based stable matching is computed in 4403.51 s, where roommates are comfortable to each others in terms of the first three optimization criteria (smoking habits, cleanliness, room environment) but the importances of the sleep and study habits are 20 and 6 respectively. With anytime search, Figure 3 shows that the most preferred criteria based stable matching is actually computed in 250 s; so the rest of the time is spent for optimality check.

7.3 Third survey: How good are the results of Personalized-SRTI-ASP compared to unstable matchings?

In this survey, we have presented to the participants 3 Personalized-SRTI instances with 3 agents (like in Figure 4). Each instance is presented with 3 matchings, including a personalized stable matching computed by Personalized-SRTI-ASP and 2 unstable matchings.

First Name : Last Name : Email : Gender : I am requesting a specific roommate:	∘ Female ∘ Male
Sleep Habits:	D (11
• I go to bed	 Before 11pm Before Midnight After Midnight
Cleanliness:	
• I like living in a	 Clean Place Messy Place Indifferent
Smoking Habits:	
• Are you a smoker?	∘ Yes ∘ No
• Are you comfortable with a roommate that is a smoker?	∘ Yes ∘ No
Room Environment:	
• I want my room to be	 Quiet and study oriented A social gathering place for friends to hang out A combination of social and quiet
Study Habits:	
• I expect to study	 In my room Outside of my room Both inside and outside of my room
Indicate the importance on the scale 1-5 of each of the following with (1) being very important to you and (5) being of little importance to you: Sleep Habits: Cleanliness: Smoking Habits: Room Environment: Study Habits:	

Table 4.	Roommate	auestionnaire	(Second	survey)
100010 10	10001101100000	900000000000000000000000000000000000000	1 ~ ~ ~ ~ ~ ~ ~ ~ ~ ~	00,009/

We have requested the participants to choose the matching that makes sense the most. If they choose a matching different from the one computed by our method, we have asked for an explanation.



Fig. 3. Computation of an optimal solution with anytime search.



Fig. 4. An example question of the third survey: Given the preferences shown on the left hand side, choose the most reasonable roommates on the right hand side.

We have conducted this survey online (due to pandemic), at Sabanci University: 59 students have participated in this survey. The survey is conducted in three groups (Red, Blue, Green) with different orderings of instances.

The percentages of choosing the stable matching computed by our method is shown in Table 5. According to the results, for 3 questions, many participants have chosen the personalized stable matchings computed by Personalized-SRTI-ASP. This shows that extending the preferences of agents with additional information about their habits and room environments is useful for the SR problems.

Participants $(\#)$	Question 1	Question 2	Question 3	
Red Group (13)	77%	15%	15%	
Blue Group (24)	17%	96%	79%	
Green Group (22)	100%	73%	27%	

Table 5. Results of the Third Survey: Percentages of the participants who chose thepersonalized stable matching computed by Personalized-SRTI-ASP.

We have also made interesting observations from the feedback and explanations provided by the participants, when they choose a matching different from the one computed by our method (over the remaining 2 instances). For instance, for the question shown in Figure 4 (Red Group, Question 2), although both Ayşe and Duru stated that they want Buse as their roommate, 85% of the survey respondents chose Ayşe and Duru as the best roommate pair based on the given preferences. Eight of these participants stated that the reason why they chose Ayşe and Duru is that "They give more importance to both smoking habits and cleanliness habits", one of them stated that "They both prefer non-smoker roommates", and one of them specified the reason as "only cleaning matters". This feedback shows that the participants focus more on the additional information about habits and room environments, rather than specific preferences of roommates. In that sense, extending the preferences of agents with such additional information is useful. Furthermore, these results show that the participants have also considered their own preferences and priorities while choosing the best roommates.

7.4 Fourth survey: How good are the results of Personalized-SRTI-ASP compared to the results of SRTI-ASP?

In this survey, we have presented to the participants 3 Personalized-SRTI instances (like in Section 7.3). These instances consider 4 agents. Each instance comes with 3 matchings to choose from: a personalized stable matching computed by Personalized-SRTI-ASP, a stable matching computed by SRTI-ASP, and an unstable matching. We have requested the participants to choose the matching that makes sense the most.

We have conducted this survey online (due to pandemic), at Sabanci University: 42 students have participated in this survey. The survey is conducted in two groups (Blue, Green) with different orderings of instances.

According to the results (Table 6), while the overall percentage of choosing the stable matchings computed by SRTI-ASP is 29%, the overall percentage of choosing the personalized stable matchings computed by Personalized-SRTI-ASP is 49%. These results illustrate that extending the preferences of agents with additional information about their habits and room environments is useful.

8 Conclusion

We have extended SRTI-ASP to consider domain-specific knowledge about each individual's preferences about a set of criteria (e.g. about the habits of their roommates and the

Participants $(\#)$	Personalized stable matching (Personalized-SRTI-ASP)	Stable matching (SRTI-ASP)	Unstable matching
Green Group (13) Blue Group (29)	$46\% \\ 50\%$	$39\%\ 25\%$	$15\% \\ 25\%$
Overall (42)	49%	29%	22%

Table 6. Results of the fourth survey: Percentages of the participants who chose the personalized stable matching computed by Personalized-SRTI-ASP.

room environments) and about the diversity preferences of dormitories and schools (e.g. for assigning roommates from different departments). We have in particular introduced two methods taking into account these additional preferences. Personalized-SRTI considers personal preferences for each criterion, and the importance of the criteria for each agent, while Most-SRTI considers personal preferences for the most preferred criteria (e.g. obtained by a survey as in our application).

We have also evaluated Personalized-SRTI-ASP over different sizes of randomly generated Personalized-SRTI instances and compared it with Most-SRTI-ASP. We have observed that, although Personalized-SRTI-ASP pays more attention to individuals' preferences, Most-SRTI-ASP performs better in computation time.

We have illustrated a real-world application of Personalized-SRTI-ASP by interacting with at least 200 students at Sabanci University. First, we have conducted a survey to select the most important five criteria that should be included in a dormitory application. Next, we have conducted a survey to get the preferences of each student for each criterion; in this way we have also collected real data for our experiments. Next, we have conducted two surveys to evaluate the usefulness of Personalized-SRTI from the perspective of students. We have observed that many participants have chosen the solutions computed by Personalized-SRTI-ASP and have given more importance to the additional information about habits and room environments. In that sense, extending SRTI to include additional domain-specific knowledge is useful.

Acknowledgments. We would like to thank Mustafa Oguz Afacan and Selin Eyupoglu for useful discussions, the anonymous reviewers for valuable comments, and the participants of the surveys for their useful feedback.

References

- ABRAHAM, D. J., BIRÓ, P. AND MANLOVE, D. F. 2005. "almost stable" matchings in the roommates problem. In *International Workshop on Approximation and Online Algorithms*. Springer, 1–14.
- BOEHMER, N. AND ELKIND, E. 2020. Stable Roommate Problem With Diversity Preferences. In Proceeding of AAMAS, 1780-1782.
- BREWKA, G., EITER, T. AND TRUSZCZYNSKI, M. 2016. Answer set programming: An introduction to the special issue. AI Magazine 37, 3, 5–6.

COOPER, F. 2020. Fair and large stable matchings in the stable marriage and student-project allocation problems. Ph.D. thesis, University of Glasgow.

- ERDEM, E., FIDAN, M., MANLOVE, D. AND PROSSER, P. 2020. A general framework for stable roommates problems using answer set programming. *Theory and Practice of Logic Programming 20*, 6, 911–925.
- ERDÖS, P. AND RÉNYI, A. 1960. On the evolution of random graphs. In Publication of the Mathematical Institute of the Hungarian Academy of Sciences. 17–61.
- FEDER, T. 1992. A new fixed point approach for stable networks and stable marriages. *Journal of Computer and System Sciences* 45, 2, 233–284.
- GALE, D. AND SHAPLEY, L. S. 1962. College admissions and the stability of marriage. *The American Mathematical Monthly* 69, 1, 9–15.
- GELFOND, M. AND LIFSCHITZ, V. 1988. The stable model semantics for logic programming. In Proceedings of ICLP. MIT Press, 1070–1080.
- GELFOND, M. AND LIFSCHITZ, V. 1991. Classical negation in logic programs and disjunctive databases. New Generation Computing 9, 365–385.
- GUSFIELD, D. AND IRVING, R. W. 1989. The Stable Marriage Problem: Structure and Algorithms. MIT Press, Cambridge, MA, USA.
- IRVING, R. W. AND MANLOVE, D. F. 2002. The stable roommates problem with ties. *Journal of Algorithms* 43, 1, 85–105.
- IRVING, R. W., MANLOVE, D. F. AND O'MALLEY, G. 2009. Stable marriage with ties and bounded length preference lists. Journal of Discrete Algorithms 7, 2, 213 – 219.
- LIFSCHITZ, V. 2002. Answer set programming and plan generation. Artificial Intelligence 138, 39–54.
- MAREK, V. AND TRUSZCZYŃSKI, M. 1999. Stable models and an alternative logic programming paradigm. In *The Logic Programming Paradigm: A 25-Year Perspective*. Springer Verlag, 375–398.
- MERTENS, S. 2005. Random stable matchings. Journal of Statistical Mechanics: Theory and Experiment 2005, 10, P10008.
- NIEMELÄ, I. 1999. Logic programs with stable model semantics as a constraint programming paradigm. Annals of Mathematics and Artificial Intelligence 25, 241–273.
- RONN, E. 1990. NP-complete stable matching problems. Journal of Algorithms 11, 2, 285 304.