



HAL
open science

A CP-Based Approach to Popular Matching

Danuta Sorina Chisca, Mohamed Siala, Gilles Simonin, Barry O ' Sullivan

► **To cite this version:**

Danuta Sorina Chisca, Mohamed Siala, Gilles Simonin, Barry O ' Sullivan. A CP-Based Approach to Popular Matching. AAI 2016, Feb 2016, Phoenix, United States. hal-01689926

HAL Id: hal-01689926

<https://hal.science/hal-01689926>

Submitted on 25 Jan 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A CP-Based Approach for Popular Matching

Danuta Sorina Chisca, Mohamed Siala, Gilles Simonin, Barry O’Sullivan

Insight Centre for Data Analytics, University College Cork, Ireland

{sorina.chisca, mohamed.siala, gilles.simonin, barry.osullivan}@insight-centre.org

Abstract

We propose a Constraint Programming model for solving the popular matching problem. In particular, we show that one can use the *Global Cardinality Constraint* to encode the problem involving possibly ties.

Introduction & Brief Background

The notion of *popular matching* was introduced by (Gardenfors 1975), but this notion takes its roots in 1785 with the notion of Condorcet winner. Popular matching and its extensions have been an exciting area of research in the last decade. An instance of the popular matching problem is a bipartite graph $G = (\mathcal{A} \cup \mathcal{P}, E)$, where \mathcal{A} are the set of applicants, \mathcal{P} the set of posts, and a set of edges E . If $(a, p) \in E_i$ and $(a, p') \in E_j$ with $i < j$, then we say that a prefers p to p' . If $i = j$, we say that a is indifferent between p and p' . This ordering of posts adjacent to a is called a 's preference list. If applicants can be indifferent between posts, we say that preference lists contain ties. Let M be a matching of G , a vertex $u \in \mathcal{A} \cup \mathcal{P}$ is either unmatched in M , or matched to some vertex, denoted by $M(u)$ (i.e. $(u, M(u)) \in M$). We say that an applicant a prefers a matching M' to M if (i) a is matched in M' and unmatched in M , or (ii) a prefers $M'(a)$ to $M(a)$. M' is said more popular than M if the number of applicants that prefer M' to M exceeds the number of applicants that prefer M to M' .

Definition 1 A matching M is popular if and only if there is no matching M' that is more popular than M .

Example 1 $\mathcal{A} = \{a_1, a_2, a_3\}$, $\mathcal{P} = \{p_1, p_2, p_3\}$ and each applicant prefers p_1 to p_2 and p_2 to p_3 . We can notice that this instance does not admit a popular matching.

Motivation

There exist in the literature a number of efficient algorithms for solving popular matching, e.g. (Abraham et al. 2007). However, in real world situations, additional requirements are needed. In many cases, the new problem is intractable and the original algorithms become useless. *Constraint Programming (CP)* is a rich declarative paradigm to tackle these situations. In this approach, the problem is stated as a set

Copyright © 2015, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

constraints (or sub-problems) operating on some variables. One often rely on propagating these constraints to reduce the search space as efficiently as possible.

Stable Matching Problems, firstly introduced by (Gale and Shapley 1962), have been exhaustively studied in the last decades. Different formulations of these problems are proposed, distinguishing between *one-sided* matching (Garg et al. 2010) and *two-sided* matching, like Stable Marriage (SM), (Abraham, Irving, and Manlove 2003). Some attention was put on encoding SM and its variants in a CP context (Gent et al. 2001; Manlove 2005; Prosser 2014). Surprisingly, popular matching has never been studied in the context of constraint programming. We tackle precisely this issue and propose, to the best of our knowledge, the first CP formulations of this problem.

More precisely, we consider two cases of the problem of popular matching: with or without ties, and show that one can encode these problems using one particular *Global Cardinality Constraint*.

Modelling Popular Matching in CP

Similarly to (Abraham et al. 2007), we assume that every applicant $a_i \in \mathcal{A}$, has in its preference list an extra unique post l_i , called last resort, that is worst than any post in \mathcal{P} . In this way every applicant is guaranteed to be matched.

We use one integer variable x_i per applicant a_i . The domain of each x_i represents all posts that are neighbours of a_i (denoted by $N(a_i)$) plus the unique last post l_i . We decide to represent the assignment of a_i by the index of the post (each l_i is indexed by a unique value $|\mathcal{P}| + i$). Therefore, the domain is initialised by $\mathcal{D}(x_i) = \{j | p_j \in N(a_i)\} \cup \{|\mathcal{P}| + i\}$. Assigning a value k to x_i is interpreted as a_i is matched to post p_k if $k \leq |\mathcal{P}|$, and to l_i otherwise.

Our CP model is based on the Global Cardinality Constraint *gcc* (Régin 1996), which restricts the number of occurrences of any value $j \in \bigcup_{i=1}^n \mathcal{D}(x_i)$ in the sequence $[x_1, \dots, x_n]$ to be in an interval $[lb(j), ub(j)]$ where $lb(j) \leq ub(j) \in \mathbb{Z}$.

Preferences Without Ties

For each applicant a_i , we denote by $f(a_i)$ the best post in its preference list. A post $p_j \in \mathcal{P}$ is called *f-post* if $\exists a_i \in \mathcal{A}$ s.t. $f(a_i) = p_j$. We denote by $s(a_i)$ the best post for a_i that is not *f-post*. Our CP model is based on the following Lemma:

Lemma 1 (Abraham et al. 2007) *A matching M is popular iff the following conditions hold: a) Every f -post is matched, b) For each applicant a_i , $M(a_i) \in \{f(a_i), s(a_i)\}$.*

Using Lemma 1, we can model the popular matching problem using one *gcc* constraint. First we reduce the domain of every variable x_i to be exactly $\{f(a_i), s(a_i)\}$. Next, we define $lb(j)$ and $ub(j)$ as follows: a) $lb(j) = 1$ if p_j is an f -post, $lb(j) = 0$ otherwise; and b) $ub(j) = 0$ if $\forall a_i \in \mathcal{A}, f(a_i) \neq p_j$ and $s(a_i) \neq p_j$, $ub(j) = 1$ otherwise.

Theorem 1 *$gcc(lb, ub, [x_1, \dots, x_{|\mathcal{A}|}])$ is satisfiable iff M is a popular matching.*

Preferences With Ties

The definition of $f(a_i)$ becomes the set of top choices for applicant a_i . However the definition of $s(a_i)$ is no longer the same. Indeed it may now contain any number of surplus f -posts. In (Abraham et al. 2007) the authors propose to characterise which f -posts cannot be included in $s(a_i)$, and exploit the Gallai-Edmonds decomposition. Let $G_1 = (\mathcal{A} \cup \mathcal{P}, E_1)$ where $E_1 \subseteq E$ is the subset of edges corresponding to top choices. Let M be a maximum cardinality matching in G_1 . The three set of vertices: *even* (resp. *odd*) is the set of vertices having an even (resp. odd) alternating path (w.r.t. M) in G_1 from an unmatched vertex; and *unreachable* is the set of vertices that are not in $even \cup odd$. Denote by $\mathcal{E}, \mathcal{O}, \mathcal{U}$ the sets of even, odd, and unreachable vertices resp. Our CP model is based on the following Lemma:

Lemma 2 (Edmonds 1965) *Let \mathcal{E}, \mathcal{O} , and \mathcal{U} be the vertices sets defined by G_1 and M above. Then:*

1. $\{\mathcal{E}, \mathcal{O}, \mathcal{U}\}$ is a partition of $\mathcal{A} \cup \mathcal{P}$ and any maximum cardinality matching in G_1 leads to exactly the same sets \mathcal{E}, \mathcal{O} , and \mathcal{U} .
2. Let M be a maximum cardinality matching of G_1 .
 - Every vertex in \mathcal{O} is matched to a vertex in \mathcal{E}
 - Every vertex in \mathcal{U} is matched to another vertex in \mathcal{U}
 - The size of M is $|\mathcal{O}| + |\mathcal{U}|/2$
3. No maximum cardinality matching of G_1 contains an edge between two vertices in \mathcal{O} or a vertex in \mathcal{O} and a vertex in \mathcal{U} . Moreover, there is no edge in G_1 between a vertex in \mathcal{E} with a vertex in \mathcal{U} .

So we define $s(a)$ the set of top-ranked posts in a 's preference list that are *even* in G_1 . We use the following Lemma to model the popular matching problem with ties.

Lemma 3 *A matching M is popular iff the following conditions hold:*

- $M \cap E_1$ is a maximum matching of G_1
- For each applicant a_i , $M(a_i) \in f(a_i) \cup s(a_i)$

We can model the popular matching problem with ties using one *gcc* constraint. First, the domain is pruned with $\mathcal{D}(x_i) \leftarrow f(a_i) \cup s(a_i) \forall i \in [1, |\mathcal{A}|]$. Next, from Lemma 2 we apply the following preprocessing steps:

- Let $\Omega = \{i | a_i \in \mathcal{U}\}$, and $\Psi = \{j | p_j \in \mathcal{U}\}$, then
 - $\forall i \in \Omega, \mathcal{D}(x_i) \leftarrow \mathcal{D}(x_i) \cap \Psi$

- $\forall i \in [1, |\mathcal{A}|] \setminus \Omega, \mathcal{D}(x_i) \leftarrow \mathcal{D}(x_i) \setminus \Psi$
- Let $\Upsilon = \{i | a_i \in \mathcal{E}\}$, $\Theta = \{j | p_j \in \mathcal{E}\}$, $\Phi = \{k | a_k \in \mathcal{O}\}$, and $\Lambda = \{l | p_l \in \mathcal{O}\}$ then
 - $\forall i \in \Upsilon, \mathcal{D}(x_i) \leftarrow \mathcal{D}(x_i) \cap \Lambda$
 - $\forall k \in \Phi, \mathcal{D}(x_k) \leftarrow \mathcal{D}(x_k) \cap \Theta$

The values of $lb(j)$, and $ub(j)$ are defined as follows: a) $lb(j) = 1$, for all j s.t. $p_j \in \mathcal{O} \cup \mathcal{U}$, otherwise $lb(j) = 0$; b) $ub(j) = 0$ for all j s.t. $\forall a_i \in \mathcal{A}, f(a_i) \neq p_j$ and $s(a_i) \neq p_j$, otherwise $ub(j) = 1$.

Theorem 2 *$gcc(lb, ub, [x_1, \dots, x_{|\mathcal{A}|}])$ is satisfiable iff M is a popular matching with ties.*

Proof. [Sketch] The preprocessing steps that we introduced enforce the solution of *gcc* to be a maximum matching of G_1 since every vertex in \mathcal{O} is matched to a vertex in \mathcal{E} , and every vertex in \mathcal{U} is matched to another vertex in \mathcal{U} .

Conclusion and Future Research

We proposed a CP formulation for the problem of popular matching involving possibly ties. As part of the future work, it would be interesting to apply these propositions to solve more general problems embedding popular matching. An example is the popular matching problem with copies where one can add some copies of the posts, possibly with an additional cost.

Acknowledgement. The Insight Centre for Data Analytics is supported by Science Foundation Ireland under Grant Number SFI/12/RC/2289.

References

- Abraham, D. J.; Irving, R. W.; Kavitha, T.; and Mehlhorn, K. 2007. Popular matchings. *SIAM J. Comput.* 37(4):1030–1045.
- Abraham, D. J.; Irving, R. W.; and Manlove, D. 2003. The student-project allocation problem. In *ISAAC Proceedings*, 474–484.
- Edmonds, J. 1965. Paths, trees, and flowers. *Canad. J. Math.* 17:449–467.
- Gale, D., and Shapley, L. S. 1962. College admissions and the stability of marriage. *The American Mathematical Monthly* 69:9–15.
- Gardenfors, P. 1975. Match making: Assignments based on bilateral preferences. *SRBS* 20:166–173.
- Garg, N.; Kavitha, T.; Kumar, A.; Mehlhorn, K.; and Mestre, J. 2010. Assigning papers to referees. *Algorithmica* 58(1):119–136.
- Gent, I. P.; Irving, R. W.; Manlove, D.; Prosser, P.; and Smith, B. M. 2001. A constraint programming approach to the stable marriage problem. In *CP 2001*, 225–239.
- Manlove, D. F. 2005. Modelling and solving the stable marriage problem using constraint programming. In *IJCAI '05*, 10–17.
- Prosser, P. 2014. Stable roommates and constraint programming. In *CPAIOR, 2014. Proceedings*, 15–28.
- Régin, J. 1996. Generalized Arc Consistency for Global Cardinality Constraint. In *AAAI'96*, 209–215.