Centre for Data Analytics



Revisiting Two-Sided Stability Constraints

Mohamed Siala Barry O'Sullivan

May 18, 2016













• Matching under preferences & Constraint Programming?

- Matching under preferences & Constraint Programming?
- Few CP formulations exist in the literature for stable matching

- Matching under preferences & Constraint Programming?
- Few CP formulations exist in the literature for stable matching
- Not much about local consistency levels

- Matching under preferences & Constraint Programming?
- Few CP formulations exist in the literature for stable matching
- Not much about local consistency levels
- Global constraints for stable matching problems?

Matching Under Preferences

- They are everywhere!
- For instance, assigning students to universities/residents to hospitals/ workers to firms . . .

Matching Under Preferences

- They are everywhere!
- For instance, assigning students to universities/residents to hospitals/ workers to firms . . .
- Bipartie structure with two sided preferences
- Bipartie structure with one sided preferences
- Non-bipartie structure

Matching Under Preferences

- They are everywhere!
- For instance, assigning students to universities/residents to hospitals/ workers to firms . . .
- Bipartie structure with two sided preferences
- Bipartie structure with one sided preferences
- Non-bipartie structure

Stable Marriage [Gale and Shapley, 1962]



Stable Marriage [Gale and Shapley, 1962]



- ੋ, , , , , . . .
- ♀, <mark>♀</mark>, ♀, · · ·
- Two sided preferences
- A matching *M* is stable when no blocking pair exists
- A pair (♂,♀) is blocking a matching M if ♂/♀ prefer each other to their situation in M





Male proposals



Male proposals
 (♂,♀)



Male proposals
 (♂,♀), (♂,♀)



Male proposals
 (♂,♀), (♂,♀), (♂,♀)



Male proposals
 (♂,♀), (♂,♀), (♂,♀), (♂,♀)



Male proposals
 (♂,♀), (♂,♀), (♂,♀), (♂,♀), (♂,♀)



Male proposals
 (♂,♀), (♂,♀), (♂,♀), (♂,♀), (♂,♀), (♂,♀)



- Male proposals
 (♂,♀), (♂,♀), (♂,♀), (♂,♀), (♂,♀), (♂,♀)
- Female proposals



- Male proposals
 (♂,♀), (♂,♀), (♂,♀), (♂,♀), (♂,♀), (♂,♀)
- Female proposals
 (♀,♂)



- Male proposals
 (♂,♀), (♂,♀), (♂,♀), (♂,♀), (♂,♀), (♂,♀)
- Female proposals
 (♀,♂), (♀,♂)



- Male proposals
 (♂,♀), (♂,♀), (♂,♀), (♂,♀), (♂,♀), (♂,♀)
- Female proposals
 (♀,♂), (♀,♂), (♀,♂)



- Male proposals
 (♂,♀), (♂,♀), (♂,♀), (♂,♀), (♂,♀), (♂,♀)
- Female proposals
 (♀,♂), (♀,♂), (♀,♂), (♀,♂)



- Male proposals
 (♂,♀), (♂,♀), (♂,♀), (♂,♀), (♂,♀), (♂,♀)
- Female proposals
 (♀,♂), (♀,♂), (♀,♂), (♀,♂), (♀,♂)

Many to one extension of SM

- Many to one extension of SM
- Two sets of agents: residents r_1, r_2, \ldots and hospitals h_1, h_2, \ldots

- Many to one extension of SM
- Two sets of agents: residents r_1, r_2, \ldots and hospitals h_1, h_2, \ldots
- Preferences without ties

- Many to one extension of SM
- Two sets of agents: residents r_1, r_2, \ldots and hospitals h_1, h_2, \ldots
- Preferences without ties
- Each hospital h_j has a capacity c_j
- Notation: *i* better than *j* w.r.t. a list *L*: $i \prec j$ or $j \succ i$

- Many to one extension of SM
- Two sets of agents: residents r_1, r_2, \ldots and hospitals h_1, h_2, \ldots
- Preferences without ties
- Each hospital h_j has a capacity c_j
- Notation: *i* better than *j* w.r.t. a list *L*: $i \prec j$ or $j \succ i$
- Find a stable matching?

 Some global constraints exist in the literature [Gent et al., 2001, Unsworth and Prosser, 2005, Manlove et al., 2007, Unsworth and Prosser, 2013]

- Some global constraints exist in the literature [Gent et al., 2001, Unsworth and Prosser, 2005, Manlove et al., 2007, Unsworth and Prosser, 2013]
- They are all equivalent in terms filtering (even for SM)

- Some global constraints exist in the literature [Gent et al., 2001, Unsworth and Prosser, 2005, Manlove et al., 2007, Unsworth and Prosser, 2013]
- They are all equivalent in terms filtering (even for SM)
- They all are equivalent to the GL-Lists

- Some global constraints exist in the literature [Gent et al., 2001, Unsworth and Prosser, 2005, Manlove et al., 2007, Unsworth and Prosser, 2013]
- They are all equivalent in terms filtering (even for SM)
- They all are equivalent to the GL-Lists
- What is the level of consistency?
CP Model for Hospital/Resident Problem (Г)

Variables

- x_i: index of the hospital assigned to r_i
- y_{j,k}: index of the resident assigned to the kth position in h_j

Constraints

$$y_{j,k} < y_{j,k+1} \ (\forall j \in [1, n_H], \forall k \in [1, c_j - 1])$$
 (1)

$$y_{j,k} \ge q_{i,j} \implies x_i \le p_{i,j} (\forall j \in [1, n_H], \forall k \in [1, c_j], \forall i \in \mathcal{H}_j)$$
(2)

$$x_i \neq p_{i,j} \implies y_{j,k} \neq q_{i,j} \ (\forall i \in [1, n_R], \forall j \in \mathcal{R}_i, \forall k \in [1, c_j])$$
(3)

$$(x_i \ge p_{i,j} \land y_{j,k-1} < q_{i,j}) \implies y_{j,k} \le q_{i,j} (\forall i \in [1, n_R], \forall j \in \mathcal{R}_i, \forall k \in [1, c_j])$$
(4)

$$y_{j,c_j} < q_{i,j} \implies x_i \neq p_{i,j} \ (\forall j \in [1, n_H], \forall i \in \mathcal{H}_j)$$
(5)

Example

$$\begin{array}{l} \mathcal{R}_1 = [3,2,1] \\ \mathcal{R}_2 = [4,1,3,2] \\ \mathcal{R}_3 = [2,4,3] \\ \mathcal{R}_4 = [1,3,4] \end{array} \middle| \begin{array}{l} \mathcal{H}_1 = [1,2,4] \\ \mathcal{H}_2 = [2,1,3] \\ \mathcal{H}_3 = [3,2,4,1] \\ \mathcal{H}_4 = [4,3,2] \end{array}$$

Initial Domain

• $\mathcal{D}(x_1) = \mathcal{D}(x_3) = \mathcal{D}(x_4) = \{1, 2, 3, 5\}$

•
$$\mathcal{D}(x_2) = \{1, 2, 3, 4, 5\}$$

- $\mathcal{D}(y_{1,0}) = \mathcal{D}(y_{2,0}) = \mathcal{D}(y_{3,0}) = \mathcal{D}(y_{4,0}) = \{0\}$
- $\mathcal{D}(y_{1,1}) = \mathcal{D}(y_{2,1}) = \mathcal{D}(y_{4,1}) = \{1, 2, 3, 5\}$

•
$$\mathcal{D}(y_{3,1}) = \{1, 2, 3, 4, 5\}$$

2-SidedStability($\mathcal{X}, \mathcal{A}, \mathcal{B}, \mathcal{C}$)

- X is the set of variables x₁,..., x_{n_R} defined the same way in Γ,
- $\mathcal{A} = \{\mathcal{R}_1, \ldots, \mathcal{R}_{n_R}\}$
- $\mathcal{B} = \{\mathcal{H}_1, \dots, \mathcal{H}_{n_H}\}$
- $\mathcal{C} = \{c_1, \ldots, c_{n_H}\}$

2-SidedStability($\mathcal{X}, \mathcal{A}, \mathcal{B}, \mathcal{C}$)

- X is the set of variables x₁,..., x_{n_R} defined the same way in Γ,
- $\mathcal{A} = \{\mathcal{R}_1, \ldots, \mathcal{R}_{n_R}\}$
- $\mathcal{B} = \{\mathcal{H}_1, \ldots, \mathcal{H}_{n_H}\}$
- $\mathcal{C} = \{c_1, \ldots, c_{n_H}\}$

We show that AC on Γ enforces BC(D) on any domain ${\cal D}$

Insight Centre for Data Analytics

Definitions

• A support of a constraint *C* in a domain *D* is an assignment of the variables in *D* that satisfies *C*

Definitions

- A support of a constraint *C* in a domain *D* is an assignment of the variables in *D* that satisfies *C*
- A constraint in arc consistent in D iff for every variable x in the scope of C, every value in D(x) has a support in D

Definitions

- A support of a constraint *C* in a domain *D* is an assignment of the variables in *D* that satisfies *C*
- A constraint in arc consistent in D iff for every variable x in the scope of C, every value in D(x) has a support in D
- A constraint in bound(D) consistent in D iff for every variable x in the scope of C, min(D(x)) and max(D(x)) have a support in D

Definitions

- A support of a constraint *C* in a domain *D* is an assignment of the variables in *D* that satisfies *C*
- A constraint in arc consistent in D iff for every variable x in the scope of C, every value in D(x) has a support in D
- A constraint in bound(D) consistent in D iff for every variable x in the scope of C, min(D(x)) and max(D(x)) have a support in D

Bound(D) consistency is stronger than the classic bound consistency property

Arc Consistency using Γ?

Insight Centre for Data Analytics

Arc Consistency using Γ?

- AC removes 5 from $\mathcal{D}(x_1), \mathcal{D}(x_2), \mathcal{D}(x_3), \mathcal{D}(x_4)$
- AC removes 5 from $D(y_{1,1}), D(y_{2,1}), D(y_{3,1}), D(y_{4,1})$
- No more propagation

Arc Consistency using Γ?

- AC removes 5 from $\mathcal{D}(x_1), \mathcal{D}(x_2), \mathcal{D}(x_3), \mathcal{D}(x_4)$
- AC removes 5 from $D(y_{1,1}), D(y_{2,1}), D(y_{3,1}), D(y_{4,1})$
- No more propagation
- Assigning 3 to x₂ has no solution
- **F** hinders propagation!

Theorem [Gusfield and Irving, 1989]

Theorem [Gusfield and Irving, 1989]

• The number of assigned residents per hospital is the same in all stable matchings

Theorem

[Gusfield and Irving, 1989]

- The number of assigned residents per hospital is the same in all stable matchings
- If a resident r_i is unassigned in one stable matching then it is unassigned in all stable matchings.

Theorem

[Gusfield and Irving, 1989]

- The number of assigned residents per hospital is the same in all stable matchings
- If a resident r_i is unassigned in one stable matching then it is unassigned in all stable matchings.
- If a hospital *h_j* is under-subscribed in one stable matching then it is assigned exactly the same residents in all stable matching

Preprocessing

 Compute the GS_lists and prune the domain accordingly (O(L))

Preprocessing

- Compute the GS_lists and prune the domain accordingly (O(*L*))
- Notation: *HFull* the set of hospitals fully subscribed in any stable matching

Necessary and Sufficient Condition for Stability

Theorem

2-SidedStability($\mathcal{X}, \mathcal{A}, \mathcal{B}, \mathcal{C}$) is satisfiable iff

$$\forall 1 \leq j \leq n_H, \sum_{i=1}^{n_R} (\mathcal{R}_i[x_i] == j) \leq c_j \land$$

$$\forall 1 \leq i \leq n_R, \forall 1 \leq j \leq n_H + 1, x_i = j \implies$$

$$\forall k \in [1, j[$$

$$\mathbf{if} h = \mathcal{R}_i[k] \mathbf{then}$$

$$\sum_{m=1}^{n_R} (\mathcal{R}_m[x_m] == h) = c_h$$

$$\land$$

$$\forall I \succ \succ i, \mathcal{R}_I[x_I] \neq h$$

Insight Centre for Data Analytics

Lemma

If Γ is AC then assigning all variables to their minimum value is solution.

Lemma

If Γ is AC then assigning all variables to their minimum value is solution.

Lemma

If Γ is AC then assigning all variables to their maximum value is solution.

Lemma

If Γ is AC then assigning all variables to their minimum value is solution.

Lemma

If Γ is AC then assigning all variables to their maximum value is solution.

Theorem

Enforcing AC on Γ makes 2-SidedStability($\mathcal{X}, \mathcal{A}, \mathcal{B}, \mathcal{C}$) Bound(\mathcal{D}) consistent.

Revisiting Bound(D) Consistency

- Best BC(D) algorithm runs in O(c × L) [Manlove et al., 2007]
- We propose an optimal algorithm running in O(L)

Revisiting Bound(D) Consistency

- Best BC(D) algorithm runs in O(c × L) [Manlove et al., 2007]
- We propose an optimal algorithm running in O(L)

Theorem

2-SidedStability($\mathcal{X}, \mathcal{A}, \mathcal{B}, \mathcal{C}$) is BC(D) iff assigning every variable to its maximum is a solution and assigning every variable to its minimum is a solution.

Lower bound changes

- Assume we have a domain that is BC(D)
- Let *r_i* be a resident whose lower bound has changed
- Let *h_j* be the hospital corresponding to the new lower bound

- Let *h* be any hospital 'between' the old and the new lower bound
- Make sure that any resident worse (in H_h) than r_i cannot be assigned to h

- Let h be any hospital 'between' the old and the new lower bound
- Make sure that any resident worse (in *H_h*) than *r_i* cannot be assigned to *h*

- Let h be any hospital 'between' the old and the new lower bound
- Make sure that any resident worse (in *H_h*) than *r_i* cannot be assigned to *h*

- Let h be any hospital 'between' the old and the new lower bound
- Make sure that any resident worse (in *H_h*) than *r_i* cannot be assigned to *h*



- Let h be any hospital 'between' the old and the new lower bound
- Make sure that any resident worse (in *H_h*) than *r_i* cannot be assigned to *h*

- Let h be any hospital 'between' the old and the new lower bound
- Make sure that any resident worse (in *H_h*) than *r_i* cannot be assigned to *h*

$$\mathcal{H}_h$$
 r₁₀ r₃ r₁ r₇ r₁₄ \varkappa \varkappa \varkappa \varkappa \varkappa κ κ

- Let h be any hospital 'between' the old and the new lower bound
- Make sure that any resident worse (in *H_h*) than *r_i* cannot be assigned to *h*

Make sure that the new hospital h_j has no more than c_j residents whose lower bound corresponds to h_j .

- *MIN_{h_i*: The variables whose minimum corresponds to *h_j*}
- If |MIN_{hj}| = c_{hj} + 1: Remove the worst resident r from MIN_{hi} and prune h_j from the domain of r

Make sure that the new hospital h_j has no more than c_j residents whose lower bound corresponds to h_j .

- *MIN_{hi}*: The variables whose minimum corresponds to *h_j*
- If |MIN_{hj}| = c_{hj} + 1: Remove the worst resident r from MIN_{hi} and prune h_j from the domain of r

Example

Make sure that the new hospital h_j has no more than c_j residents whose lower bound corresponds to h_j .

- *MIN_{hi}*: The variables whose minimum corresponds to *h_j*
- If |MIN_{hj}| = c_{hj} + 1: Remove the worst resident r from MIN_{hi} and prune h_j from the domain of r

Example

Make sure that the new hospital h_j has no more than c_j residents whose lower bound corresponds to h_j .

- *MIN_{hi}*: The variables whose minimum corresponds to *h_j*
- If |MIN_{hj}| = c_{hj} + 1: Remove the worst resident r from MIN_{hi} and prune h_j from the domain of r

Example

$$\mathcal{H}_{h_j} \bigcirc r_8 \bigcirc r_{11} r_9 \bigcirc r_6 \oslash r_1 r_8 \not \approx \not m \leq f_1$$
Step 2

Make sure that the new hospital h_j has no more than c_j residents whose lower bound corresponds to h_j .

- *MIN_{hi}*: The variables whose minimum corresponds to *h_j*
- If |MIN_{hj}| = c_{hj} + 1: Remove the worst resident r from MIN_{hi} and prune h_j from the domain of r

Example

$$\mathcal{H}_{h_j} \bigcirc r_8 \bigtriangledown r_{11} r_9 \backsim r_5 \lor r_6 \qquad \underbrace{r_1 r_8}_{h_j} \checkmark f_8 \qquad \underbrace{r_1 r_8}_{h_j} \checkmark f_1 \qquad f_1 \qquad f_2 \qquad f_2 \qquad f_1 \qquad f_1 \qquad f_2 \qquad f_2 \qquad f_1 \qquad f_2 \qquad f_2 \qquad f_2 \qquad f_2 \qquad f_2 \qquad f_2 \qquad f_3 \qquad f_3 \qquad f_4 \qquad f_6 \quad f_6 \qquad f_6 \qquad$$

Step 2

Make sure that the new hospital h_j has no more than c_j residents whose lower bound corresponds to h_j .

- *MIN_{h_i*: The variables whose minimum corresponds to *h_j*}
- If |MIN_{hj}| = c_{hj} + 1: Remove the worst resident r from MIN_{hi} and prune h_j from the domain of r

Example

- MAX_h: The indexes of residents where the maximum corresponds to h
- $maxofMAX_h = max(MAX_h)$

Lemma

Assigning all variables to their maximum is a solution iff $\forall h \in HFull, |MAX_h| = c_h$, and $\forall i \leq maxofMAX_h$, let $r = \mathcal{H}_h[i]$, and $l = \mathcal{R}_r^{-1}[h]$, then $i \notin MAX_h \implies max(x_r) < l$.

- Assume we have a domain that is BC(D)
- Let *r_i* be a resident whose upper bound has changed
- Let *h* be the hospital corresponding to the previous upper bound
- Find a 'replacement' for r

- Assume we have a domain that is BC(D)
- Let *r_i* be a resident whose upper bound has changed
- Let *h* be the hospital corresponding to the previous upper bound
- Find a 'replacement' for r

Example

 \mathcal{H}_{h_j} . . (b) . . (c) . . (g) (c) r_{11} r_3 . .

- Assume we have a domain that is BC(D)
- Let *r_i* be a resident whose upper bound has changed
- Let *h* be the hospital corresponding to the previous upper bound
- Find a 'replacement' for r

Example

$$\mathcal{H}_{h_j}$$
 . . @ . . @ . . & @ r_{11} r_3
 \uparrow
 $maxofMAX_h$

- Assume we have a domain that is BC(D)
- Let *r_i* be a resident whose upper bound has changed
- Let *h* be the hospital corresponding to the previous upper bound
- Find a 'replacement' for r

Example

 \mathcal{H}_{h_j} . . (2) . . (2) . . (3) . . (5) (6) r_{11} r_3 . . \uparrow $maxofMAX_h$

- Assume we have a domain that is BC(D)
- Let r_i be a resident whose upper bound has changed
- Let *h* be the hospital corresponding to the previous upper bound
- Find a 'replacement' for r

Example

maxofMAX_h

Insight Centre for Data Analytics

• Enforce BC(D)

- Enforce BC(D)
- For any variable *x_i*, we suppose that the upper bound is removed

- Enforce BC(D)
- For any variable *x_i*, we suppose that the upper bound is removed
- Enforce BC(D) on the new domain

- Enforce BC(D)
- For any variable *x_i*, we suppose that the upper bound is removed
- Enforce BC(D) on the new domain
- Any value between the old and the new upper bound does not have a support

- Enforce BC(D)
- For any variable *x_i*, we suppose that the upper bound is removed
- Enforce BC(D) on the new domain
- Any value between the old and the new upper bound does not have a support
- Repeat Until the lower bound of x_i

- Enforce BC(D)
- For any variable *x_i*, we suppose that the upper bound is removed
- Enforce BC(D) on the new domain
- Any value between the old and the new upper bound does not have a support
- Repeat Until the lower bound of x_i
- Complexity: $O(n_R \times L)$

- Enforce BC(D)
- For any variable *x_i*, we suppose that the upper bound is removed
- Enforce BC(D) on the new domain
- Any value between the old and the new upper bound does not have a support
- Repeat Until the lower bound of x_i
- Complexity: $O(n_R \times L)$
- Not incremental

Beyond Constraint Propagation

Insight Centre for Data Analytics

Beyond Constraint Propagation

Hospital/Resident Problem with Forced and Forbidden Pairs

- Unknown complexity
- Straightforward to solve! Just enforce BC(D) on the domain
- O(L) to solve with our approach

Beyond Constraint Propagation

Hospital/Resident Problem with Forced and Forbidden Pairs

- Unknown complexity
- Straightforward to solve! Just enforce BC(D) on the domain
- *O*(*L*) to solve with our approach

Stable Marriage Problem with Forced and Forbidden Pairs

- Best complexity $O(n^4)$ [Dias et al., 2003]
- Particular case of the above approach
- $O(n^2)$ to solve

Experiments

Problem description

- Some couples prefer to be matched together
- Same preferences for these couples
- Find a stable matching maximizing the number of such couples who are matched together

Experiments

Protocol

- Random instances: 2k, 4k, ..., 8k residents; 100, 200, ... 500 hospitals; and different capacities c ∈ {100 + 50 + k|k ∈ [0, 8]}
- BC(D) and AC Implemented in Mistral-2.0
- LEX variable branching + (min/max/random min max) value branching
- Geometric restarts
- 5 different seeds for "random min max"
- 20 minutes time cutoff

Summary of the Results

Set	BC(D)-min			AC-min			BC(D)-max		
	Time	Nodes	Opt	Time	Nodes	Opt	Time	Nodes	Opt
2k	2	16.56	100	19	13.33	100	8	256.38	100
4k	5	18.14	100	151	14.69	100	37	394.86	100
6k	9	18.08	100	393	14.89	93	86	648.82	100
8k	19	18.16	100	332	15.80	79	131	491.50	100
	AC-max			BC(D)-rand			AC-rand		
	Time	Nodes	Opt	Time	Nodes	Opt	Time	Nodes	Opt
2k	28	204.53	100	4	67.36	100	12	81.91	100
4k	202	320.33	100	12	81.91	100	160	65.61	100
6k	564	535.65	85	25	120.58	100	502	99.20	92
8k	530	432.87	69	42	98.88	100	475	88.11	77

Runtime



Speed of Exploration



Conclusion and Future Research

Contributions

- Previous CP propositions maintain only BC(D)
- A better implementation of BC(D) in O(L) time
- An adaptation of BC(D) to achieve AC
- First polynomial algorithm to solve the Hospital/Resident problem with forced and forbidden pairs
- Improving the worst case complexity to solve SM with forced and forbidden pairs by a factor of $O(n^2)$

Conclusion and Future Research

Contributions

- Previous CP propositions maintain only BC(D)
- A better implementation of BC(D) in O(L) time
- An adaptation of BC(D) to achieve AC
- First polynomial algorithm to solve the Hospital/Resident problem with forced and forbidden pairs
- Improving the worst case complexity to solve SM with forced and forbidden pairs by a factor of $O(n^2)$

Future Research

- Better implementation for AC?
- New global constraints are coming for different stable matching problems..



Thank you.

Picture taken from The New York Times

Insight Centre for Data Analytics

May 18, 2016

References I

Dias, V. M. F., da Fonseca, G. D., de Figueiredo, C. M. H., and Szwarcfiter, J. L. (2003).
The stable marriage problem with restricted pairs. <i>Theor. Comput. Sci.</i> , 306(1-3):391–405.
Gale, D. and Shapley, L. S. (1962). College admissions and the stability of marriage. <i>American mathematical monthly</i> , pages 9–15.
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 <i>Proceedings of CP</i> , pages 225–239.
Gusfield, D. and Irving, R. W. (1989). The Stable marriage problem - structure and algorithms. Foundations of computing series. MIT Press.
Manlove, D., O'Malley, G., Prosser, P., and Unsworth, C. (2007). A constraint programming approach to the hospitals / residents problem. In <i>Proceedings of CPAIOR</i> , pages 155–170.

References II



Unsworth, C. and Prosser, P. (2005).

A specialised binary constraint for the stable marriage problem.

In Abstraction, Reformulation and Approximation, 6th International Symposium, SARA 2005, Airth Castle, Scotland, UK, July 26-29, 2005, Proceedings, pages 218–233.



Unsworth, C. and Prosser, P. (2013). An n-ary constraint for the stable marriage problem. *CoRR*, abs/1308.0183.