Using Constraint Programming to Assign Students to First-year Seminars

Xiang Yao (yaox@dickinson.edu)
Advisor: Timothy Wahls (wahlst@dickinson.edu)

Computer Science Major
Dickinson College, 2013
May 4th, 2014
Introduction

Background

• Assign students to first-year seminars at Dickinson College
• More than six hundred students in each incoming class
• Each student can choose six seminars with different rankings

Goal & Objective

• Achieve higher ranking of the students’ seminar choices, better gender balance and balance between international and domestic students
Constraint Satisfaction Problem

- A constraint satisfaction problem is a set of variables and set of constraints on those variables
- Each constraint limits the possible values for each variable
- Objective function: a function that maps a solution to a number that measures the goodness of the solution

Example:

Let X1, X2, X3 and X4 be real numbers

\[ X_1 + X_2 + X_3 + X_4 < 2 \]

\[ X_1 + X_2 - X_3 + X_4 > 0 \]

Objective Function: maximize \( X_1 - X_2 - X_3 + X_4 \)
Related Work

Constraint Programming Approach

- Marte (2002) developed a basic finite domain constraint model to solve school timetabling problems.
- Delgado and Perez (2005) build an application for course assignment with 1600 events.

Operations Research Approach

- Forrester, Hutson, and To (2013) solve the course assignment problem for Dickinson College with balanced gender, student origin, and course size.
- Forrester and Hutson (2014) take the students’ ranking into consideration.
Method Used: Finite Domain Constraint

- Each variable containing a value restricted to a finite set of integers
- Constraints limit the possible values

Example:
- $x < 3 \land x \neq y$
- Finite domain: $x$ is $\{1, 2, 3, 4\}$, $y$ is $\{1, 2, 3\}$
- Possible Solution: $x = 2$, $y = 3$
Method Used: Finite Set Constraint

- Variables containing sets
- Have a lower bound and an upper bound
  - Lower bound: elements definitely in the set
  - Upper bound: elements possibly in the set
- The cardinality of the set: limit the # of elements

Example:
- Let x and y be finite set variable
- {} $\subseteq$ x $\subseteq$ \{1,2,3,4,5\} and {} $\subseteq$ y $\subseteq$ \{1,2,3,4,5\}
- x $\cup$ y = \{1,2,3,4,5\}, #x = 3
- Possible Solution: x = \{1,2,3\} and y = \{3,4,5\}
Finite Domain Model

Variable Definition:
n: the total number of students
m: the total number of seminars
$X_i$: the finite domain variable for each student, $0 \leq i \leq n$
$\text{Sem}_k$: the set of students assigned to seminar $k$, $0 \leq k \leq m$

$\text{StudentRank}_{i,j} = k$, if student $i$'s $j^{th}$ ranked seminar is seminar $k$, where $1 \leq i \leq n$ and $1 \leq j \leq 6$

Constraints for the Finite Domain Model

$\forall$ $i$ such that $0 \leq i \leq n$, $X_i \in \text{StudentRank}_i$
$\text{Student} = \{X_1, X_2, X_3...X_n\}$
$\forall$ $k \in \{1,2...m\}$, $\text{Sem}_k = \{i \mid X_i \in \text{Student} \land X_i = k\}$
For each $k \in \{1,2...m\}$, $|\text{Sem}_k| \leq 16$

Objective Function for the Finite Domain Model

$\sum_{i=1}^{n} j^2$, such that $\text{StudentRank}_{i,j} = X_i$
Finite Set Model

**Variable Definition:**

- $n$: the total number of students
- $m$: the total number of seminars

$\text{StudentRank}_{i,j} = k$, if student $i$’s $j^{\text{th}}$ ranked seminar is seminar $k$, where $1 \leq i \leq n$ and $1 \leq j \leq 6$

$\text{GenderList}_i$: the gender of student $i$, for male $\text{GenderList}_i = 1$ and for female $\text{GenderList}_i = 0$

$\text{InternList}_i$: the nationality of student $i$, for domestic student $\text{InternList}_i = 1$ and for foreign students $\text{InternList}_i = 0$

$\text{Seminar}_k$: finite set variables for the students in seminar $k$

$M_k, F_k, D_k, I_k$: finite set variables for male, female, domestic and international students in seminar $k$

$R_{k,j}$: finite set variable for the students who are assigned to seminar $k$ as their $j^{\text{th}}$ choice
Finite Set Model

Constraints for the Finite Set Model

{} ≺ Seminar_k ≺ \{i \mid 1 \leq i \leq n \land k \in \text{StudentRank}_i\}
{} ≺ M_k ≺ \{i \mid 1 \leq i \leq n \land k \in \text{StudentRank}_i \land \text{GenderList}_i=1\}
{} ≺ F_k ≺ \{i \mid 1 \leq i \leq n \land k \in \text{StudentRank}_i \land \text{GenderList}_i=0\}
{} ≺ D_k ≺ \{i \mid 1 \leq i \leq n \land k \in \text{StudentRank}_i \land \text{InternList}_i=1\}
{} ≺ I_k ≺ \{i \mid 1 \leq i \leq n \land k \in \text{StudentRank}_i \land \text{InternList}_i=0\}

∀ k1 and k2 where 1 ≤ k1 < k2 ≤ m,
Seminar_{k1} \cap Seminar_{k2} = {} 

M_k \cup F_k = D_k \cup I_k = Seminar_k

Objective Function for the Finite Set Model

Objective:
\[ 23 \sum_{k=1}^{m} (|M_k| - |F_k|)^2 + 8 \sum_{k=1}^{m} (|D_k| - |I_k|)^2 + \sum_{k=1}^{m} \left( -43 \times |R_{k,1}| - 25 \times |R_{k,2}| - 15 \times |R_{k,3}| - 9 \times |R_{k,4}| - 5 \times |R_{k,5}| - 3 \times |R_{k,6}| \right) \]

R_{k,1} \cup R_{k,2} \cup R_{k,3} \cup R_{k,4} \cup R_{k,5} \cup R_{k,6} = Seminar_k

8 \leq |Seminar_k| \leq 17
Results

- In Table 1.1, finite set model achieve relative balance for gender, nationality and seminar sizes from 2011 student data.

- In Table 1.2, finite domain model achieve better seminar rankings for students’ top choices.

Table 1.1

<table>
<thead>
<tr>
<th></th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>S5</th>
<th>S6</th>
<th>S7</th>
<th>S8</th>
<th>S9</th>
<th>S10</th>
<th>S11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>8</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Intern</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Size</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>15</td>
<td>16</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>S12</td>
<td>S13</td>
<td>S14</td>
<td>S15</td>
<td>S16</td>
<td>S17</td>
<td>S18</td>
<td>S19</td>
<td>S20</td>
<td>S21</td>
<td>S22</td>
</tr>
<tr>
<td>Male</td>
<td>4</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>4</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Intern</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Size</td>
<td>12</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>15</td>
<td>9</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>S23</td>
<td>S24</td>
<td>S25</td>
<td>S26</td>
<td>S27</td>
<td>S28</td>
<td>S29</td>
<td>S30</td>
<td>S31</td>
<td>S32</td>
<td>S33</td>
</tr>
<tr>
<td>Male</td>
<td>7</td>
<td>5</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>Intern</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Size</td>
<td>16</td>
<td>13</td>
<td>16</td>
<td>15</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>S34</th>
<th>S35</th>
<th>S36</th>
<th>S37</th>
<th>S38</th>
<th>S39</th>
<th>S40</th>
<th>S41</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Intern</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Size</td>
<td>16</td>
<td>14</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
</tbody>
</table>

Table 1.2

<table>
<thead>
<tr>
<th></th>
<th># of 1st</th>
<th># of 2st</th>
<th># of 3st</th>
<th># of 4st</th>
<th># of 5st</th>
<th># of 6st</th>
<th>Runtime</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finite Domain</td>
<td>400(63%)</td>
<td>85(13%)</td>
<td>58(9%)</td>
<td>51(8%)</td>
<td>27(4%)</td>
<td>15(2%)</td>
<td>8 hours</td>
</tr>
<tr>
<td>Finite Set</td>
<td>139(22%)</td>
<td>119(19%)</td>
<td>108(17%)</td>
<td>95(15%)</td>
<td>90(14%)</td>
<td>15(13%)</td>
<td>8 hours</td>
</tr>
</tbody>
</table>
Future Work

- Set hard constraints on the number of students in a seminar
- Further search for more appropriate labeling and searching strategies
Conclusion

• Finite Set Model:
  • Relative balance for gender and nationality
  • Low ranking for seminar choices

• Finite Domain Model:
  • Achieve better seminar rankings
  • Slow running time with only two outputs
  • Not consider gender and nationality
Reference


- Forrester, R and Hutson, K Balancing Student and Faculty Preferences in the Assignment of First-Year Seminars, submitted to International Journal of Information Technology and Decision Making
