A Constraint Processing Approach to the Assignment of Graduate Teaching Assistants (GTAs) in CSE

Robert L. Glaubius and Berthe Y. Choueiry
Constraint Systems Laboratory
Computer Science & Engineering (CSE)
glaubius@cse.unl.edu

Joint work with M. Augustyn, Ch. Hammack, H. Zou, Ch. Daniel, B.Y. Choueiry

Partially supported by CSE

Problem definition

Given:

- Many classes in CSE require teaching assistants
  - Lecture courses: graders
  - Labs, recitations: instructors & graders
- CSE hires Graduate Teaching Assistants (GTAs)
  - Experienced (material, language, etc.)
  - New ($\pm$ degree of preparedness)

Question:

Assign GTAs to classes in a 'satisfactory' manner.

Background

- Time horizon: 1 semester, 8 classes with half-semester duration
- About 70 classes, 35 GTAs
- Currently solved by hand:
  Chair, Vice-chair, Graduate Program Secretary, GTAs
- Shortcomings:
  - inconsistencies
    (e.g., GTAs assigned to courses that they are taking)
  - unbalanced load across students
  - difficult to quantify a 'satisfactory' solution
  - time-consuming task, about 3 weeks to stabilize

An interactive system

Goals

- Substantial reduction in administrative overhead
- Decreased number of inconsistent assignments
- High expected GTA satisfaction

How well have we done?

- 3 weeks assignment time $\longrightarrow$ 2 days
- Relatively few assignment changes (less than 4, swapping)

Project: a multi-faceted effort

1. Data collection: *Augustyn, Glaubius, Choueiry*
   choice of attributes, preferences, ...

2. Interfaces: *Hammack, Zou, Daniel, Choueiry*
   student-end, administrator-end, security, ...

3. Modeling: *Augustyn, Glaubius, Choueiry*
   constraints, data structures, ...

4. Processing: *Glaubius, Kavan, Zou, Choueiry*
   propagation, search, evaluation criteria, ... (on-going)

Data collection

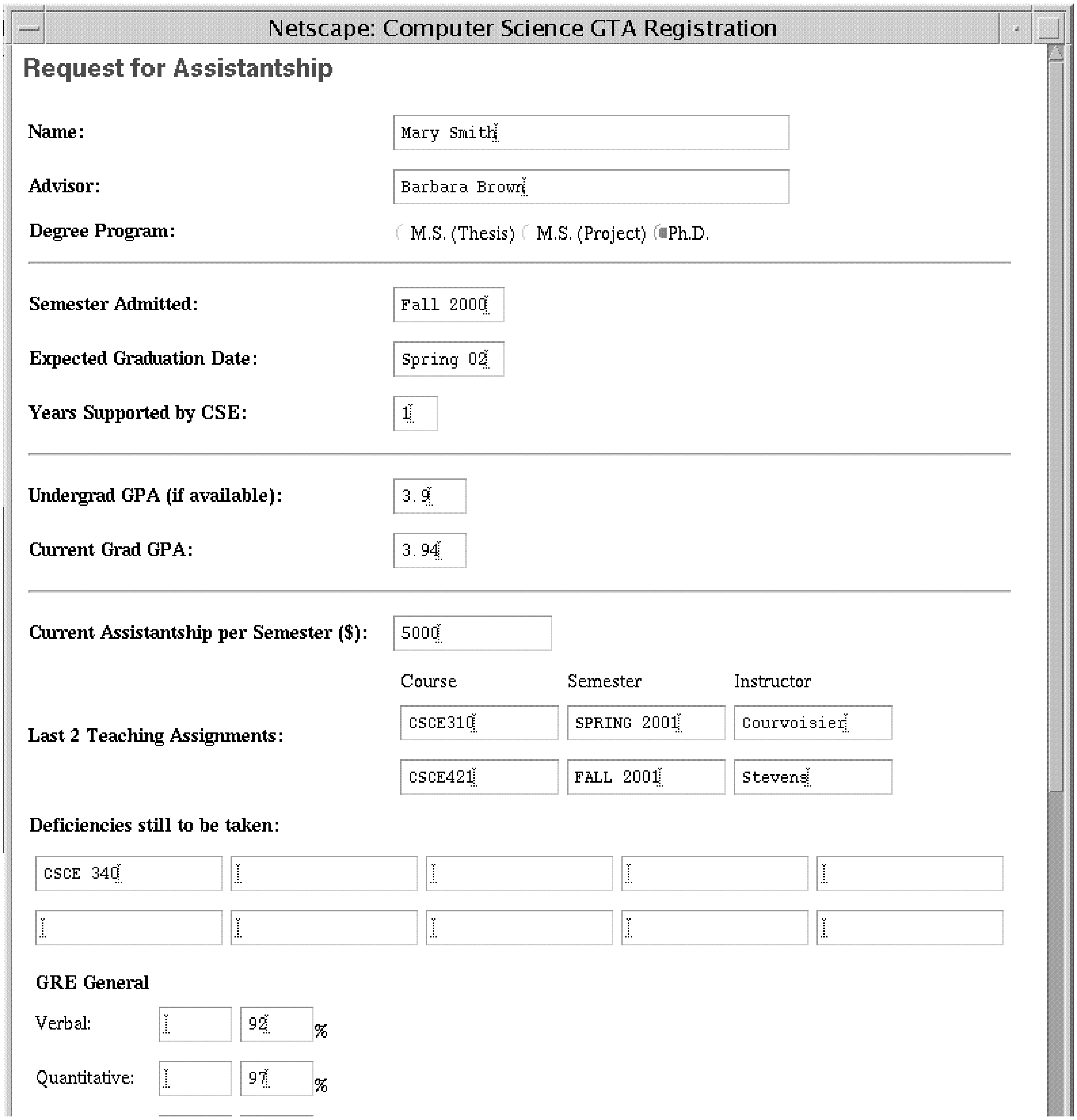
GTAs

- New attribute: course preference $\{0, 1, \ldots, 5\}$
- semester enrollment, deficits, half or full teaching assistantship
- ITA certification, ...

Courses

- type: lab, recitation, lecture
- load: large, medium, small classes
- duration within the semester, ...

Courtesy of Ch. Hammack



Netscape: Computer Science GTA Registration

Request for Assistantship

Name:	Mary Smith
Advisor:	Barbara Brown
Degree Program:	○ M.S. (Thesis) ○ M.S. (Project) ● Ph.D.
Semester Admitted:	Fall 2000
Expected Graduation Date:	Spring 02
Years Supported by CSE:	1
Undergrad GPA (if available):	3.9
Current Grad GPA:	3.94
Current Assistantship per Semester ($):	5000

Last 2 Teaching Assignments:

Course	Semester	Instructor
CSCE310	SPRING 2001	Courvoisier
CSCE421	FALL 2001	Stevens

Deficiencies still to be taken:

CSCE 340				

GRE General

Verbal:		92 %
Quantitative:		97 %

Courtesy of Ch. Hammack

Netscape: Computer Science GTA Registration

Class	Section	Course Name	Preference	Will Enroll in:	Justification for 0
101	001	Computer Science Fundamentals	5	☐	
101L	001	Computer Science Fundamentals Lab	2	☐	
101L	002	Computer Science Fundamentals Lab	2	☐	
101L	003	Computer Science Fundamentals Lab	5	☐	
101L	004	Computer Science Fundamentals Lab	5	☐	
150	150	Intro to Computer Programming	0	☐	too much work
150	151	Intro to Computer Programming Lab	5	☐	
150	152	Intro to Computer Programming Lab	5	☐	
155	150	Intro to Computer Science I	2	☐	
155	151	Intro to Computer Science I Lab	1	☐	
⋮					⋮
897	002	Masters Project–Software Design	3	☐	
923	001	Dev and Analysis of Eff. Algorithms	0	☑	
952	001	Advanced Computer Networks	1	☐	
963	001	Software Process Eng–JDE	3	☐	
979	001	Adv in Neural Networks and Gen. Algorithms	0	☑	
990	003	Seminar–Bioinformatics	1	☐	
990	004	Seminar–Network Systems	0	☐	no experience

Previous Next

Approach: Constraint Satisfaction Problem (CSP)

Motivation

a flexible & expressive paradigm

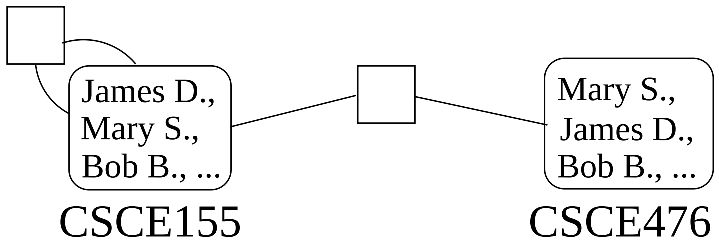
Contributions

- Modeling:
  elicitation and modeling of constraints and preferences
- Processing:
  systematic search & advanced propagation mechanisms

Elements of Constraint Processing

Given:

- *Variables*: CSCE155, CSCE476, . . .
- *Values*: Mary Smith, James Doe, Bob Brown, . . .
- *Constraints*: James can't teach between 10:00-11:00 a.m.,
  Mary can't teach more than 2 courses, . . .



Question:

Assign a value (GTA) to *every* variable (course)

. . . such that all constraints are satisfied *decision problem*

. . . everyone is pleased *optimization problem*

Illustrating example

Variables (courses):

	Course	Type	Size	Time
V_1	Intro to CS	lab	medium	MWF 08:00-09:00
V_2	Data Struct. & Algo.	lecture	large	MWF 11:00-12:00
V_3	Discrete Structures	recitation	medium	MWF 08:00-09:00
V_4	Intro to AI	lecture	medium	MWF 08:00-09:00

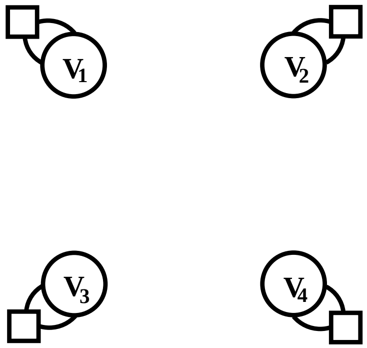
Values (GTAs):

Name	Preferences				ITA?	Unavailable
	V_1	V_2	V_3	V_4		
James Doe	3	5	3	4	No	MWF 10:00-11:00
Mary Smith	4	0	4	0	Yes	MWF 12:00-1:00
Bob Brown	5	3	5	4	Yes	MWF 8:00-9:00

	Course	Type	Size	Time
V_1	Intro to CS	lab	medium	MWF 08:00-09:00
V_2	Data Struct. & Algo.	lecture	large	MWF 11:00-12:00
V_3	Discrete Structures	recitation	medium	MWF 08:00-09:00
V_4	Intro to AI	lecture	medium	MWF 08:00-09:00

Name	Preferences				ITA?	Unavailable
	V_1	V_2	V_3	V_4		
James Doe	3	5	3	4	No	MWF 10:00-11:00
Mary Smith	4	0	4	0	Yes	MWF 12:00-1:00
Bob Brown	5	3	5	4	Yes	MWF 8:00-9:00

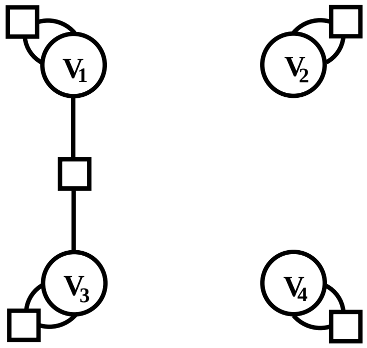
Unary constraint: James can't teach between 10:00 and 11:00 a.m.



	Course	Type	Size	Time
V_1	Intro to CS	lab	medium	MWF 08:00-09:00
V_2	Data Struct. & Algo.	lecture	large	MWF 11:00-12:00
V_3	Discrete Structures	recitation	medium	MWF 08:00-09:00
V_4	Intro to AI	lecture	medium	MWF 08:00-09:00

Name	Preferences				ITA?	Unavailable
	V_1	V_2	V_3	V_4		
James Doe	3	5	3	4	No	MWF 10:00-11:00
Mary Smith	4	0	4	0	Yes	MWF 12:00-1:00
Bob Brown	5	3	5	4	Yes	MWF 8:00-9:00

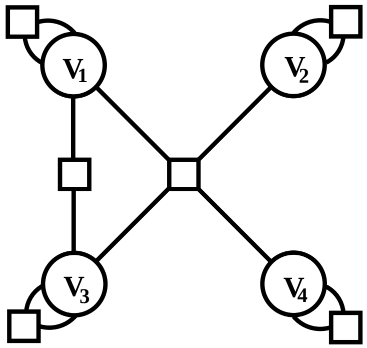
Binary constraint: Bob can't teach two labs that overlap in time



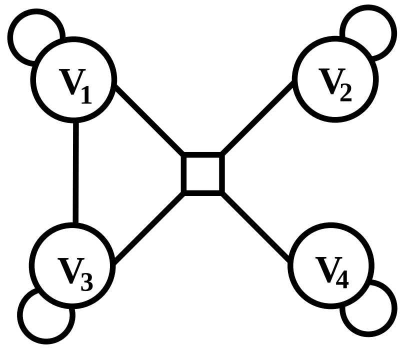
	Course	Type	Size	Time
V_1	Intro to CS	lab	medium	MWF 08:00-09:00
V_2	Data Struct. & Algo.	lecture	large	MWF 11:00-12:00
V_3	Discrete Structures	recitation	medium	MWF 08:00-09:00
V_4	Intro to AI	lecture	medium	MWF 08:00-09:00

Name	Preferences				ITA?	Unavailable
	V_1	V_2	V_3	V_4		
James Doe	3	5	3	4	No	MWF 10:00-11:00
Mary Smith	4	0	4	0	Yes	MWF 12:00-1:00
Bob Brown	5	3	5	4	Yes	MWF 8:00-9:00

Global constraint: Mary can grade at most 2 courses



Modeling the constraints



Unary: James can't teach between 10:00-11:00 a.m. *(5 types)*

Binary: Bob can't teach two labs that overlap in time *(1 type)*

Global (non-binary): Mary can grade at most 2 courses *(3 types)*

Total: about 200 constraint instances

Major, critical endeavor:
⟶ elicitation and innovative modeling of non-binary constraints

Modeling the objectives

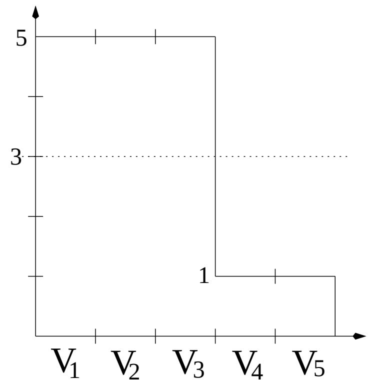
Students specify preference value for each course ($\in \{0, 1, \ldots, 5\}$)
'Satisfaction' measured using preference values

1. Typically, we have a shortage of GTAs, so we prefer the solution that **maximizes the number of courses** covered
   - $\surd$ 5 courses with medium satisfaction
   - $\times$ 3 courses with high satisfaction
2. For an equal number of courses covered, **maximise**
   1. arithmatic mean of satisfaction,
   2. geometric mean of satisfaction, or
   3. lowest satisfaction value

   breaking ties randomly

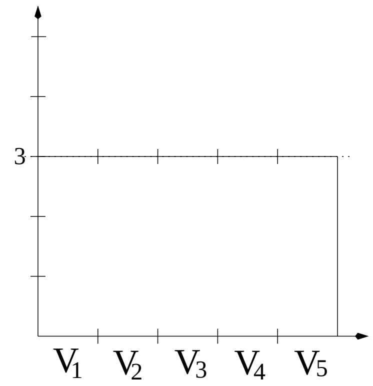
Maximize mean satisfaction: geometric vs. arithmetic

$\langle V_1, 5\rangle, \langle V_2, 5\rangle, \langle V_3, 5\rangle, \langle V_4, 1\rangle, \langle V_5, 1\rangle$



$\sqrt[5]{125} = 2.6$ vs. 3 $\quad\times$

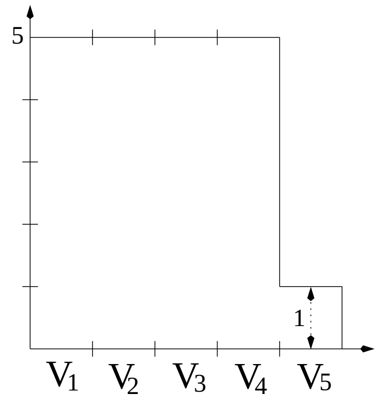
$\langle V_1, 3\rangle, \langle V_2, 3\rangle, \langle V_3, 3\rangle, \langle V_4, 3\rangle, \langle V_5, 3\rangle$



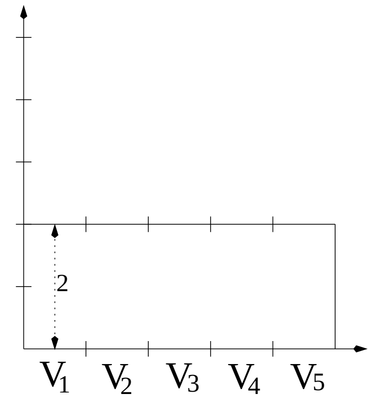
$\sqrt[5]{243} = 3$ vs. 3 $\quad\surd$

Maximize lowest satisfaction

$\langle V_1, 5\rangle, \langle V_2, 5\rangle, \langle V_3, 5\rangle, \langle V_4, 5\rangle, \langle V_5, 1\rangle$



$\langle V_1, 2\rangle, \langle V_2, 2\rangle, \langle V_3, 2\rangle, \langle V_4, 2\rangle, \langle V_5, 2\rangle$



Solution method

- Basic computational mechanisms
  - Backtrack search: examines possibilities systematically
  - Ordering heuristics: avoid dead-ends, find better solutions
  - General propagation algorithms (e.g., `FC`)
  - Branch and bound strategy: find a first solution, exclude relatively ‘poor’ possibilities by comparing to the current retained solution
- Advanced computational mechanisms
  - Least discrepancy search: better coverage of solution space
  - Advanced propagation algorithms (e.g., n-`FC`, `all-diffs`)

Current status

- Modeling: several iterations
- Validating: Fall 2001 (good start), Spring 2002 (on-going)

Future research

- Reformulation: aggregating courses to reduce problem size
- Decomposition and localization of interactions *Kavan*
- Local search *Zou*
- Compact solutions *Beckwith, Buettner, Xu*
  by detecting interchangeable/symmetrical choices
- Problem specification & solving *Glaubius, Hammack*
  Incremental strategies to remove constraints, define new constraints, force choices, etc.

Benefits

1. Challenging honors thesis, deals with a real-world application:
   → advanced aspects of modeling and processing
2. Service to CSE: help administrators and students
3. New research directions for more theses in ConSystLab
4. Enhance teaching
   CSCE476/876, Introduction to Artificial Intelligence
   CSCE421/821, Foundations of Constraint Satisfaction

A win-win-win-win project..