# Efficient Propagators for Global Constraints

#### **Claude-Guy Quimper**

Supervisor: Alejandro López-Ortiz





#### Outline

- My first contact with constraint programming
- The all-different constraint
- The global cardinality constraint
- The inter-distance constraint
- Post-doctoral work

# My first contact with constraint programming

- I took Peter's course in Constraint Programming
- The field requires efficient algorithms that are executed gazillions of times.
- Project: To implement Thiel and Mehlhorn's alldiff propagator.



Peter van Beek

- Scheduling: We want execution times to be all different.
- Encoding permutations.
- Sometimes, one simply wants things to be different!

нцццц

#### The All-Different Constraint

All-Different $(X_1, \ldots, X_n) \iff X_i \neq X_j$ 

Régin '94	Domaín	$O(n^{1.5}d)$







Régin '94	Domaín	$O(n^{1.5}d)$

1

Régin '94	Domaín	$O(n^{1.5}d)$
Leconte '96	Range	$O(n^2)$
Puget '98	Bounds	O(nlogn)
Mehlhorn & Thiel	Bounds	0(n)
López-Ortíz, Químper, Tromp, S Van Beek	Bounds	0(n)







Régin '94	Domaín	0(n <sup>1.5</sup> d)
Leconte '96	Range	$O(n^2)$
Puget '98	Bounds	O(nlogn)







HH

Régin '94	Domaín	0(n <sup>1.5</sup> d)
Leconte '96	Range	0(n <sup>2</sup> )
Puget '98	Bounds	O(nlogn)
Mehlhorn & Thiel	Bounds	O(n)

1

Régin '94	Domaín	$O(n^{1.5}d)$
Leconte '96	Range	$O(n^2)$
Puget '98	Bounds	O(nlogn)
Mehlhorn & Thiel	Bounds	0(n)
López-Ortíz, Químper, Tromp, S Van Beek	Bounds	0(n)

 $dom(X_1) = [3, 4]$  $dom(X_2) = [3, 4]$  $dom(X_3) = [1, 3]$ 

 $dom(X_1) = [3, 4]$  $dom(X_2) = [3, 4]$  $dom(X_3) = [1, 3]$ 

A Hall interval is an interval of k values that contains the domains of k variables.

 $dom(X_1) = [3,4]$   $dom(X_2) = [3,4]$  Hall interval  $dom(X_3) = [1,3]$ 

A Hall interval is an interval of k values that contains the domains of k variables.

 $dom(X_1) = [3,4]$   $dom(X_2) = [3,4]$  Hall interval  $dom(X_3) = [1,2]$ 

A Hall interval is an interval of k values that contains the domains of k variables.

- $\operatorname{dom}(X_1) = [2,3]$
- $\operatorname{dom}(X_2) = [2,3]$
- $\operatorname{dom}(X_3) = [3,4]$
- $\operatorname{dom}(X_4) = [2, 6]$



- $dom(X_1) = [2,3]$ 
  - $\operatorname{dom}(X_2) = [2,3]$
  - $\operatorname{dom}(X_3) = [3,4]$
  - $\operatorname{dom}(X_4) = [2, 6]$



#### A Propagator for the **Bounds Consistency** $dom(X_1) = [2,3]$ $dom(X_2) = [2,3]$ $dom(X_3) = [3, 4]$ $dom(X_4) = [2, 6]$



#### A Propagator for the **Bounds Consistency** $dom(X_1) = [2,3]$ $dom(X_2) = [2,3]$ $dom(X_3) = [3, 4]$ $dom(X_4) = [2, 6]$

1 2 3 4 5 6

~

- $dom(X_1) = [2,3]$
- $dom(X_2) = [2,3]$ 
  - $dom(X_3) = [3, 4]$
  - $dom(X_4) = [2, 6]$

1 2 3 4 5 6 ~

1 2 3 4 5 6

~

- $dom(X_1) = [2,3]$
- $dom(X_2) = [2,3]$ 
  - $dom(X_3) = [3, 4]$
  - $dom(X_4) = [2, 6]$

1 2 3 4 5 6

- $dom(X_1) = [2,3]$
- $dom(X_2) = [2,3]$ 
  - $dom(X_3) = [3, 4]$
  - $dom(X_4) = [2, 6]$

6

1 2 3 4 5

Hall

- $dom(X_1) = [2,3]$
- $dom(X_2) = [2,3]$ 
  - $dom(X_3) = [3, 4]$
  - $dom(X_4) = [2, 6]$

- $\operatorname{dom}(X_1) = [2,3]$
- $\operatorname{dom}(X_2) = [2,3]$
- $dom(X_3) = [3,4]$ 
  - $\operatorname{dom}(X_4) = [2, 6]$



- $\operatorname{dom}(X_1) = [2,3]$
- $\operatorname{dom}(X_2) = [2,3]$
- $dom(X_3) = [3,4]$ 
  - $\operatorname{dom}(X_4) = [2, 6]$



6

1 2 3 4 5

Hall

- $dom(X_1) = [2,3]$
- $dom(X_2) = [2,3]$
- $\longrightarrow$  dom $(X_3) = [4, 4]$ 
  - $dom(X_4) = [2, 6]$

- $dom(X_1) = [2,3]$
- $dom(X_2) = [2,3]$
- $\longrightarrow$  dom $(X_3) = [4, 4]$ 
  - $dom(X_4) = [2, 6]$

1 2 3 4 5 6 Hall
6

1 2 3 4 5

Hall

- $\operatorname{dom}(X_1) = [2,3]$
- $\operatorname{dom}(X_2) = [2,3]$
- $dom(X_3) = [4,4]$ 
  - $\operatorname{dom}(X_4) = [2, 6]$
- [4, 4][2, 6]

- $\operatorname{dom}(X_1) = [2,3]$
- $\operatorname{dom}(X_2) = [2,3]$
- $\operatorname{dom}(X_3) = [4,4]$
- $dom(X_4) = [2,6]$





 $dom(X_1) = [2,3]$ 

- $dom(X_2) = [2,3]$
- $\operatorname{dom}(X_3) = [4,4]$
- $\longrightarrow \operatorname{dom}(X_4) = [2,6]$

### A Propagator for the **Bounds Consistency** $dom(X_1) = [2,3]$ $dom(X_2) = [2,3]$ $\operatorname{dom}(X_3) = [4,4]$ $\longrightarrow \operatorname{dom}(X_4) = [5,6]$ 2 3 4 5 1 6 ~ 1 ~ Hall

2 3 4 5

Hall

6

1

 $dom(X_1) = [2,3]$  $dom(X_2) = [2,3]$  $dom(X_3) = [4,4]$  $dom(X_4) = [5,6]$ 

# 

## Analysis of the Algorithm

version	Complexity	Note
Fírst version	$O(n^2 + m)$	6 línes of C code!

# Analysis of the Algorithm

ЦЦ

Version	Complexity	Note
Fírst version	$O(n^2 + m)$	6 línes of C code!
uníon-find data structure	O(nlogn)	The fastest in practice

# Analysis of the Algorithm

Version	Complexity	Note
Fírst version	$O(n^2 + m)$	6 línes of C code!
Union-find data structure	O(nlogn)	The fastest in practice
Balanced union-find data structure	$O(n \alpha(n))$	Slightly slower than the previous version

# Analysis of the Algorithm

1

Version	Complexity	Note
Fírst version	$O(n^2 + m)$	6 línes of C code!
Union-find data structure	O(nlogn)	The fastest in practice
Balanced union-find data structure	$O(n \alpha(n))$	Slightly slower than the previous version
Gabow and Tarjan's data structure	0(n)	Slower and pages of code

- $GCC([X_1, \ldots, X_n], l, u) \iff \forall v \ l_v \le |\{i \mid X_i = v\}| \le u_v$ 
  - $\Box \quad A \text{ value } \vee \text{ must be taken at least } I_v \text{ times}$ and at most  $u_v \text{ times}$ .

- $GCC([X_1, \ldots, X_n], l, u) \iff \forall v \ l_v \le |\{i \mid X_i = v\}| \le u_v$ 
  - A value v must be taken at least l<sub>v</sub> times and at most u<sub>v</sub> times.
  - Scheduling: No more than 2 tasks can be executed at a given time.

- $GCC([X_1, \dots, X_n], l, u) \iff \forall v \ l_v \le |\{i \mid X_i = v\}| \le u_v$ 
  - $\Box \quad A \text{ value } \vee \text{ must be taken at least } I_v \text{ times}$ and at most  $u_v \text{ times}$ .
  - Scheduling: No more than 2 tasks can be executed at a given time.
  - Sequencing: We want to restrict the number of occurrences of an event in a sequence.

 $GCC([X_1, \ldots, X_n], l, u) \iff \forall v \ l_v \le |\{i \mid X_i = v\}| \le u_v$ 

IRégin '96] gives a propagator achieving domain consistency.

 $GCC([X_1, \ldots, X_n], l, u) \iff \forall v \ l_v \le |\{i \mid X_i = v\}| \le u_v$ 

IRégin '96] gives a propagator achieving domain consistency.

There were no propagators for bounds consistency.

### 

## **Decomposing the GCC**



## **Decomposing the GCC**

The upper bound constraint (ubc)

Each value ís assigned to at most 2 variables.



## **Decomposing the GCC**



## **Decomposing the GCC**



All values must be assigned to at most 2 variables.

$$\begin{array}{cccc} X_1:\{1 & 2 & \} \\ X_2:\{1 & & \} \\ X_3:\{1 & 2 & \} \\ X_4:\{ & 2 & \} \\ X_5:\{1 & 2 & 3\} \end{array}$$

All values must be assigned to at most 2 variables.



All values must be assigned to at most 2 variables.



Upper capacity:  $\lceil S \rceil = 2 + 2 = 4$ 

All values must be assigned to at most 2 variables.



Upper capacity:  $\lceil S \rceil = 2 + 2 = 4$ 

All values must be assigned to at most 2 variables.



Upper capacity:  $\lceil S \rceil = 2 + 2 = 4$ 

All values must be assigned to at most 2 variables.



Upper capacity:  $\lceil S \rceil = 2 + 2 = 4$ 

Hall Interval

An interval containing as many domains as its upper capacity.  $\lceil S \rceil = |\{i \mid \operatorname{dom}(X_i) \subseteq S\}|$ 

### A Propagator for the UBC

Símílar to the one for the all-dífferent
 Constraínt.



### A Propagator for the UBC

- Símílar to the one for the all-dífferent
  Constraínt.
- values can have more than one bucket.



### A Propagator for the UBC

- Símílar to the one for the all-dífferent
  Constraínt.
- Values can have more than one bucket.



All values must be assigned to at least 1 variable.



All values must be assigned to at least 1 variable.

 $X_{1}: \{1 \}$   $X_{2}: \{ 4 \}$   $X_{3}: \{1 \}$   $X_{4}: \{1 \}$   $X_{4}: \{1 \}$   $X_{5}: \{ 2 \}$   $S = \{2, 3\}$ 

All values must be assigned to at least 1 variable.

 $S = \{2, 3\}$ Lower capacity:  $\lfloor S \rfloor = 1 + 1 = 2$ 

All values must be assigned to at least 1 variable.

 $X_{1}: \{1 \\ X_{2}: \{ \\ 4 \} \\ X_{3}: \{1 \\ 4 \} \\ X_{4}: \{1 \\ 2 \\ 3 \\ 4 \} \cap S \neq \emptyset \\ X_{5}: \{ 2 \\ 3 \\ 4 \} \cap S \neq \emptyset \\ S = \{2, 3\} \\ \text{Lower capacity: } \lfloor S \rfloor = 1 + 1 = 2$ 

All values must be assigned to at least 1 variable.

 $X_{1} : \{1 \} \\ X_{2} : \{ 4 \} \\ X_{3} : \{1 \} \\ X_{4} : \{ 2 \} \\ X_{5} : \{ 2 \} \cap S \neq \emptyset \\ S = \{2, 3\} \\ \text{Lower capacity: } \lfloor S \rfloor = 1 + 1 = 2$ 

All values must be assigned to at least 1 variable.

 $X_{1} : \{1 \} \\ X_{2} : \{ 4 \} \\ X_{3} : \{1 \} \\ X_{4} : \{2 \} \} \cap S \neq \emptyset \\ X_{5} : \{2 \} \cap S \neq \emptyset \\ S = \{2, 3\} \\ \text{Lower capacity: } \lfloor S \rfloor = 1 + 1 = 2$ 

Unstable Set

A set intersecting as many domains as its lower capacity.  $|S| = |\{i \mid \operatorname{dom}(X_i) \cap S \neq \emptyset\}|$ 

### A Propagator for the LBC

We adapted the algorithm for the All-different constraint

Detects <u>unstable sets</u> rather than <u>Hall intervals</u>.

□ Time complexity: O(n)




### The Global Cardinality Constraint



# The Global Cardinality Constraint



### Theorem:

A value has a support in the GCC iff it has a support in the UBC and the LBC.

### Proof:

Based on the relationship between Hall sets and unstable sets.

### Note:

Holds for domaín, range, and bounds consistency

# A Propagator for the GCC



# A Propagator for the GCC



<u>Theorem</u>: This algorithm never loops! <u>Proof</u>: Based on the relationship between Hall sets and unstable sets.

### Note:

Holds for domaín, range, and bounds consistency

## **Extended GCC**

### $\Box$ EGCC( $[X_1, \ldots, X_n], [C_1, \ldots, C_m]$ ) is satisfied when $\lor$ is taken $C_{\lor}$ times.

### **Extended GCC**

### $\Box$ EGCC( $[X_1, \ldots, X_n], [C_1, \ldots, C_m]$ ) is satisfied when v is taken $C_v$ times.

### Theorem

When domains are <u>sets</u>, testing the satisfiability of EGCC is NP-Hard.

### **Extended GCC**

 $\Box$  EGCC( $[X_1, \ldots, X_n], [C_1, \ldots, C_m]$ ) is satisfied when v is taken  $C_v$  times.

### Theorem

When domains are <u>sets</u>, testing the satisfiability of EGCC is NP-Hard.

### Theorem

When domains are <u>intervals</u>, filtering EGCC takes linear time.

Katriel & Thiel

ALL-DIFFERENT $(X_1, \ldots, X_n) \iff X_i \neq X_j$ 

ALL-DIFFERENT $(X_1, \ldots, X_n) \iff X_i \neq X_j$ 

Variables could be sets, multi-sets, or tuples.

ALL-DIFFERENT $(X_1, \ldots, X_n) \iff X_i \neq X_j$ 

- Variables could be sets, multi-sets, or tuples.
- Sets, multi-set, and tuple variables often have large domains.

ALL-DIFFERENT $(X_1, \ldots, X_n) \iff X_i \neq X_j$ 

- Variables could be sets, multi-sets, or tuples.
- Sets, multi-set, and tuple variables often have large domains.

 $\Box \{\} \subseteq X \subseteq \{1, \dots, u\} \Rightarrow |X| = 2^u$ 

ALL-DIFFERENT $(X_1, \ldots, X_n) \iff X_i \neq X_j$ 

- Variables could be sets, multi-sets, or tuples.
- Sets, multi-set, and tuple variables often have large domains.
- $\Box \{\} \subseteq X \subseteq \{1, \dots, u\} \Rightarrow |X| = 2^u$
- $\square$  We adapted the propagator to obtain a polynomial complexity:  $O(n^{2.5}+n^2u)$

INTER-DISTANCE $([X_1, \ldots, X_n], p) \iff |X_i - X_j| \ge p$ 

□ There must be a gap of p between each variable.

### INTER-DISTANCE( $[X_1, \ldots, X_n], p$ ) $\iff |X_i - X_j| \ge p$

There must be a gap of p between each variable.
 When p = 1, we obtain the All-Different
 Constraint.

INTER-DISTANCE $([X_1, \ldots, X_n], p) \iff |X_i - X_j| \ge p$ 

- □ There must be a gap of p between each variable.
- $\Box$  When p = 1, we obtain the All-Different Constraint.
- Scheduling: Execution times must be punits of time apart.

INTER-DISTANCE $([X_1, \ldots, X_n], p) \iff |X_i - X_j| \ge p$ 

- □ There must be a gap of p between each variable.
- $\Box$  When p = 1, we obtain the All-Different Constraint.
- Scheduling: Execution times must be punits of time apart.

Radio frequency allocation problem.

IRégin '97I introduces the global minimum distance constraint.

- IRégin '97I introduces the global minimum distance constraint.
- □ [Artionchine & Baptiste '05]

- IRégin '97I introduces the global minimum distance constraint.
- □ [Artionchine & Baptiste '05]
  - prove the constraint is NP-Hard when variables are sets.

- IRégin '971 introduces the global
  minimum distance constraint.
- □ [Artionchine & Baptiste '05]
  - prove the constraint is NP-Hard when variables are sets.
  - achieve bounds consistency in cubic time.

Place two blocks of size 4 on the axis without overlapping them.



Place two blocks of size 4 on the axis without overlapping them.

No block can have its left end inside a red zone.

### Internal Adjustment Intervals

Artiouchine & Baptiste '05

















Place the 3 blocks on the axis such that the blue blocks are in the box.



Place the 3 blocks on the axis such that the blue blocks are in the box.



Place the 3 blocks on the axis such that the blue blocks are in the box.



Place the 3 blocks on the axis such that the blue blocks are in the box.



The green box cannot have its left end inside a red zone.

### Parenthesis

If you place n blue blocks of size one inside a box of size n, you obtain a red zone of n elements.



□ This is a Hall interval!

Place the 3 blocks on the axis such that the blue blocks are in the box.



The green box cannot have its left end inside a red zone
Artiouchine & Baptiste '05



## 0 1 2 3 4 5 6 7 8 9 10 11 12















 $O(n^2) \times O(n) = O(n^3)$ 

## $O(n^2) \times O(n) = O(n^3)$

## Number of intervals [l, u]

 $O(n^2) \times O(n) = O(n^3)$ 

Number of red zones produced per ínterval

Number of intervals [l, u]

,  $O(n^2) \times O(n) = O(n^3)$ 

Number of red zones produced per ínterval

Number of intervals [l, u]

Total number of red zones

## Number of Adjustment Intervals Complexity of Number of red zones Artiouchine & Baptiste's produced per interval propagator , $O(n^2) \times O(n) = O(n^3)$ Number of intervals [l, u] Total number of red zones

## **Dominance** 0 1 2 3 4 5 6 7 8 9 10 11 12

и

## Dominance

## 0 1 2 3 4 5 6 7 8 9 10 11 12 l

## Dominance 0 1 2 3 4 5 6 7 8 9 10 11 12

## Dominance 0 1 2 3 4 5 6 7 8 9 10 11 12

## Dominance



## Dominance



Theorem

Only O(n<sup>2</sup>) red zones needs to be computed to achieve bounds consistency.

## Propagator

Uses a special data structure to store the adjustment intervals

## Propagator

- Uses a special data structure to store the adjustment intervals
- $\Box$  Time complexity:  $O(n^2)$

Bounds consistency for the
All-Different Constraint.

- Bounds consistency for the All-Different Constraint.
- Generalization of Hall's marriage theorem for the GCC.

- Bounds consistency for the All-Different Constraint.
- Generalization of Hall's marriage theorem for the GCC.
- D Extension to non-integer domains

- Bounds consistency for the All-Different Constraint.
- Generalization of Hall's marriage theorem for the GCC.
- Extension to non-integer domains
- Quadratic propagator for the Inter-Distance.

## 

## Life after the PhD



## Life after the PhD



## Microsoft® Research

## Life after the PhD



## Microsoft<sup>®</sup> **Research**



## Life after the PhD



## **Special Thanks to a Special Supervisor**



Alex López-Ortiz

## **Special Thanks to Special Collaborators**



Peter van Beek



Toby Walsh

## Thanks to my Thesis Committee



# Thanks to the ACP

