Hybrid Tractable Classes of CSPs with Global Cardinality Constraints

Evgenij Thorstensen

Dept. of Computer Science
University of Oxford

January 2012

Global Constraints

Pair $\langle \sigma, \delta \rangle$ with unbounded scope and compact representation.

Many examples

- Sum and Knapsack
- Alldifferent
- Regular

Gen. Global Cardinality Constraints

Very powerful global counting constraint.

Definition (GC constraint)

Pair $\langle \sigma, \pi \rangle$ with a scope σ and $\pi : D \to \mathcal{P}(\mathbb{N})$.

Assigns to every value a set of allowed cardinalities.

Satisfied if every $a \in D$ occurs an allowed number of times.

Generalizes ALLDIFFERENT, interval cardinality...

The problem

I have been studying the following problem.

Definition (GCCP)

A GCCP is a CSP $\langle V, D, C \rangle$ such that every $c \in C$ is a GC constraint.

Modelling examples: Scheduling

Let

- W be a set of workers,
- T a set of tasks,
- $Q : W \to \mathcal{P}(T)$ define who can do what,
- $S : T \to \mathbb{N}$ define numer of workers needed, and
- $C : T \to \mathbb{N}^2$ define lower and upper numbers of each task needed.

Encoding

GCCP with variables W, domain $Q(w)$ for every variable, and single GC constraint $C = \langle W, \pi \rangle$ with

$$\pi(t) = \{S(t) \times i \mid C(t)_1 \leqslant i \leqslant C(t)_2\}$$

Modelling examples: Bin packing

Given item sizes $i_1, \dots, i_n$, a bin size B, find a minimal packing.
Decision version: Given $k \in \mathbb{N}$, find packing with $\leqslant k$ bins.

Encoding

Domain $b_1, \dots, b_k$ (the bins). For every item size i, create i variables.
Post GC constraint on these with $\pi(b) = \{0, i\}$.
Finally, post one GC constraint on everything with $\pi(b) = \{0, 1, \dots, B\}$.

Theoretical analysis

The GCCP is NP-complete (obvious) *for fixed domains*.

Scheduling example had nice structure. Structural classes?

Definition

Structure: Hypergraph formed by scopes. Turns into *primal graph* by relacing every hyperedge with a clique.

Standard structural measures fail: Treewidth of primal graph is trivially unbounded (global constraints).

Hypertreewidth? GCCP NP-complete for htw=1 (reduction from 3COL) due to *lack of extensional representation*.

Hybrid classes

Hybrid classes combine structure and language restrictions.

We will

- exploit the limitation of GC constraints (language) to *build an extensional representation*, and
- also *define structure differently* so that treewidth can be used.

GC limitations

Observation (Bulatov and Marx): Many assignments on a given scope satisfy exactly the same GC constraints.

Definition (Counting function)

Given set of variables V and assignment θ, the counting function $\gamma : D \to \mathbb{N}$ induced by θ is such that for every $d \in D$,

$$\gamma(d) = |\{v \in V \mid \theta(v) = d\}|$$

There are $|D|^{|V|}$ assignments. How many induced counting functions?

Answer: $O(|V|^{|D|})$ (combinatorial argument).

Prioperties of counting functions

Counting functions represent equivalence classes of assignments.

We can sum disjoint counting functions

$$(\gamma + \delta)(d) = \gamma(d) + \delta(d)$$

which relates them to combining partial assignments.

More importantly, we can compute all the counting functions possible on a set of variables V in time $O(|V| + |V|^{|D|})$.

And of course we can check if a counting function satisfies a GC constraint.

Structural part

Observation: Variables that participate in the same set of GC constraints can be lumped together.

Definition (Interaction region)

Given $G = (V, H)$ and a set of hyperedges $K \subseteq H$, the set of vertices $\bigcap K - \bigcup(H - K)$ is an interaction region (IR) of G.

Define hyper/primal graphs on interaction regions (IRs) in the obvious way.

Results

Theorem

The GCCP is tractable for IR primal graphs of bounded treewidth.

Algorithm works on tree decomposition of primal graph.

Key property

For each constraint c there is a node containing all the interaction regions of c.

Algorithm

We precompute the set of counting functions for each IR, and the cross product R of these sets at every node.

Walking up the tree, at every node we do the following:

- For every C enirely contained in this node, remove all tuples that do not satisfy the constraint from R (sum of projection).
- Prune cross product in parent node (semijoin).

Treewidth bound gives polynomial time.

Dichotomy from NP-completeness proof earlier.

More results

What about GC constraints with unbounded number of IRs?

Let's look at incidence graph of IR hypergraph.

Theorem

The GCCP is tractable on acyclic IR incidence graphs.

Key properties

- Every hyperedge h (constraint) has IRs as children and parent.
- Children IRs do not interact with constraints above h.

Algorithm

We precompute set of counting functions Γ for every IR.

At every hyperedge node, we have

- Compute set of sums of children's Γ.
- Make cross product R with parent's Γ.
- Prune R and update parent's Γ.

The sum above has polynomial size.

Unfortunately no dichotomy here.

Open problems

What other global constraints can we use this idea on? SUM and KNAPSACK should work.

Strengthening IR incidence graph?

Capturing this tractable class (may relate to incidence graphs):

Definition (Nested disjoint hypergraph)

A hypergraph is nested disjoint if for every pair of hyperedges h, h', either $h \subset h'$, $h' \subset h$, or $h \cap h' = \emptyset$.

Summary

Two novel hybrid classes of CSPs with GC constraints found by combining

- the limitations of GC expressive power with
- a different way of looking at structure.

Questions?