Constraint Patterns

Toby Walsh

Cork Constraint Computation Center, University College Cork, Ireland. tw@4c.ucc.ie

Abstract. Constraint models contain a number of common patterns. For example, many constraint models involve an array of decision variables with symmetric rows and/or columns. By documenting such constraint patterns, we can share modelling expertise. Constraint solvers can also be extended to exploit such patterns. For example, we can develop specialized methods like the global lexicographical ordering constraint for breaking such row and column symmetry.

1 Introduction

Many patterns occur in constraint programs. In this paper, I argue that we need to identify, formalize and document these patterns in a similar way to the design patterns identified by the software engineering community. The result will be a systematic and comprehensive methodology for modelling an informal problem. Such a methodology will permit us to tackle the modelling “bottleneck” that hinders the uptake of constraint programming. This is an ambitious project - modelling is not a task which lends itself to a piecemeal approach as even the smallest modelling decision can have far reaching consequences. However, we have made some progress and I describe some of the more interesting constraint patterns which have already been identified. I also discuss the different ways that we can exploit such patterns. For example, one way to exploit common constraint patterns is to extend the constraint language.

2 Design patterns

Patterns are an approach to design that started in architecture [1], which has since spread to many other areas including software engineering [2]. A pattern describes not only the context of a problem and its solution, but also the rationale behind the solution. Patterns are a valuable mechanism for describing best practice and good design. Patterns therefore have a useful role to play in software engineering. Designing software is hard. Designing good software is even harder. Fortunately, well engineered code exhibits many common patterns that support extensibility, modularity, and performance. By documenting these patterns, we can support and encourage good software engineering.

I can illustrate this by means of an analogy. Consider how you could become an America’s Cup match racing helmsman. First, you learn the sailing rules
(e.g. a boat on starboard tack has right of way over a boat on port). Then you learn the basic principles of racing (e.g. when you have the lead, you cover tack to protect that lead). Finally, you study past match races to learn the winning patterns of others (e.g. on a downwind leg, an expert helmsman in a trailing boat often rides an approaching gust, blankets the leading boat’s sail, and overtakes on the windward side). Becoming an expert software engineer is little different. First, you learn the rules (e.g. algorithms, and data structures). Then you learn the basic principles (e.g. data abstraction helps code be modular and extensible). Finally, you study expert software engineers to learn valuable patterns (e.g. a good software engineer will often construct iterator methods so that elements of an aggregate object can be accessed without exposing the underlying representation).

Patterns are, by their very nature, not formal objects. They are therefore usually documented in natural language. A pattern description typically includes the following category headings (as well as others that may be more domain specific).

<table>
<thead>
<tr>
<th>Pattern Name</th>
<th>A meaningful name for the pattern.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context</td>
<td>The circumstances in which the problem the pattern solves occurs.</td>
</tr>
<tr>
<td>Problem</td>
<td>The specific problem that is solved.</td>
</tr>
<tr>
<td>Forces</td>
<td>The often opposing considerations that must be taken into account when choosing a solution.</td>
</tr>
<tr>
<td>Solution</td>
<td>The proposed solution to the problem.</td>
</tr>
<tr>
<td>Example</td>
<td>An example of the problem and its resolution using the proposed solution.</td>
</tr>
</tbody>
</table>

Fig. 1. The typical categories used in specifying a pattern.

3 Constraint patterns

Why should we identify and document patterns in constraint programming? Constraint programming is programming and so many of the usual software engineering issues arise. However, constraint programming is also about modelling. There are many recurring patterns in good constraint models. These patterns cannot usually be precisely specified as there are many conflicting interactions in a complex problem. Patterns therefore seem a good vehicle for explicitly capturing the knowledge of expert modellers.

What benefits do constraint patterns bring? First, they help tackle the modelling “bottleneck”. A library of patterns would be a valuable resource for passing on modelling expertise to neophyte constraint programmers. Second, constraint toolkits can be extended to support commonly occurring patterns. As I argue in Section 4.2, a constraint pattern can motivate the development of a new global
constraint or language feature. Third, a longer term goal would be pattern automation. For example, in Sections 4.4 and 4.5, I discuss how we are trying to automate some common constraint patterns.

What drawbacks do constraint patterns have? First, it is hard work to identify and document good patterns. It requires the efforts of a whole community, not just of one individual. Second, patterns do not eliminate all the art of constraint modelling. Many problems also contain an unique feature or an unique combination of features which necessitates a special solution technique that will not be not of a pattern library. Third, patterns are not executable. However, as I argued above, we can look to automate aspects of them.

4 Some examples

To illustrate what a constraint pattern is, and how it can be useful, I will look at four examples from my own and other people’s research.

4.1 Matrix models

Before I describe the first pattern, I want to define the context for a number of common constraint patterns. A matrix model is a constraint program with one or more matrices of decision variables. For example, a natural model of a sports scheduling problem has a 2-d matrix of decision variables, each of which is assigned a value corresponding to the game played in a given week and period [3]. In this case, the matrix is obvious in the solution to the problem: we need a table of fixtures. However, as we demonstrate in [4,5], many problems that are less obviously defined in terms of matrices can be efficiently represented and effectively solved using a matrix model. Sometimes, the matrix model contains multiple matrices of variables. Channelling constraints are then used to link the different matrices together [6-8].

As an example, consider the matrix model given in [4] for the steel mill slab design problem [9]. We have a number of orders, each with a particular weight and colour, to assign to slabs. Slabs come in a number of different sizes. We want to assign orders to slabs and sizes to slabs so that the total weight of orders assigned to a slab does not exceed the slab capacity, and so that each slab contains at most p colours (usually 2). A 2-d matrix of 0/1 decision variables represents which orders are assigned to which slabs. A second matrix of 0/1 decision variables is used to post the colour constraints. Channelling constraints are used to connect this to the order matrix.

In [5], we demonstrated the prevalence of matrix models by surveying the first 31 models in CSPLib. At least 27 of these had natural matrix models, most of them already published. Matrix models are a very natural way to represent relations (e.g., the relation between orders and slabs in the steel mill slab design problem). Matrix models are also a natural way to represent functions (e.g., the function mapping exams to times in an exam timetabling problem). Finally, matrix models are a natural way to represent partitions (e.g., the partitioning
Matrices:

\[\begin{align*}
\text{\(\rightarrow\) slabs} & \rightarrow s_i & 4 & 3 & 3 & 0 & 0 & 0 & 0 \\
\text{\(\rightarrow\) orders} & \rightarrow O_{ij} & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
& & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\
& \downarrow & O_{ij} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
& & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
\rightarrow \text{colours} & C_{ij} & 0 & 0 & 0 & 1 \\
& & 0 & 0 & 1 & 1 \\
& \downarrow & \text{slabs} & 0 & 1 & 0 & 0 \\
& & 1 & 1 & 0 & 0 \\
\end{align*}\]

Constraints:

\[\begin{align*}
\text{(row weighted sum)} & \forall j \sum_i Weight_i \cdot O_{ij} \leq s_j \\
\text{(column sum)} & \forall i \sum_j O_{ij} = 1 \\
\text{(row sum)} & \forall j \sum_i C_{ij} \leq p \\
\text{(channelling)} & \forall ij O_{ij} = 1 \rightarrow C_{\text{Colour},ij} = 1
\end{align*}\]

Fig. 2. Matrix model for the steel mill slab design problem (taken from [4])

of nodes in a graph colouring problem into different colour classes). Another indication of the prevalence of matrix models is the common use of 2-dimension matrices of decision variables in integer linear programming.

4.2 Row and column symmetry

A common pattern in matrix models is row and column symmetry. For example, the order matrix in the steel mill problem has partial row symmetry since slabs of the same size are indistinguishable and partial column symmetry since orders of the same size and colour are also indistinguishable. We can swap any two slabs of the same size, and any two orders of the same size and colour and obtain an essentially equivalent solution. As a second example, consider generating Balanced Incomplete Block Designs or BIBDs (prob028 in CSPLib). A matrix model for this problem uses a 2-d matrix of 0/1 variables, with constraints on the sum of each row and each column, and on the scalar product between rows. This matrix model has complete row and column symmetry since we can permute the rows and columns freely without affecting any of the constraints.

Symmetry in constraint programs is problematic. It increases the search space dramatically. Row and column symmetry is especially problematic as it occurs often and there is a lot of it. For example, an \(n\) by \(m\) BIBD has \(n!m!\) row and
column symmetries. Even if we are only interested in finding one solution, we may explore many failed and symmetrically equivalent branches. When proving optimality, this can be especially painful. An effective way to break such symmetry is by posting constraints that lexicographically order the rows and columns [10]. We call this double lex ordering the matrix. As long as we order the rows and columns in the same direction, (i.e. the rows and columns must both be lexicographically increasing, or must both be lexicographically decreasing), this will leave a solution if one exists. Unfortunately, it does not break all symmetry. Multiple, symmetric solutions can be left after double lex ordering the matrix. Indeed, a paper in this volume [11] proposes an additional ordering constraint, the all perms constraint, which can be effectively posted on the rows or columns to eliminate some (but still not all) of the remaining symmetry. To support double lex and all perms ordering constraints, we have developed efficient linear time constraint propagation algorithms [12, 11]. This illustrates how constraint solvers can be extended to support commonly occurring patterns like a global symmetry-breaking constraint. Techniques to eliminate all symmetries exist [13–15], but they appear to be more expensive than they are worth for dealing with row and column symmetries. The symmetries eliminated by double lex ordering and all perms ordering appear to offer a good compromise.

<table>
<thead>
<tr>
<th>Pattern Name: MatrixSymmetry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context: A matrix model containing an array of decision variables with (partial) row and column symmetry.</td>
</tr>
<tr>
<td>Problem: Eliminating (much of) that symmetry.</td>
</tr>
<tr>
<td>Forces: Eliminating all symmetry can be too expensive.</td>
</tr>
<tr>
<td>Solution: Post lex ordering constraints on rows and/or cols.</td>
</tr>
<tr>
<td>Example: Balanced Incomplete Block Design generation.</td>
</tr>
</tbody>
</table>

Fig. 3. Constraint pattern for MatrixSymmetry

An alternative to double lex ordering is to post lexicographical ordering constraints on the rows, and multiset ordering constraints on the columns (or vice versa). We have also developed an efficient constraint propagation algorithm for multiset ordering constraints [16]. This volume also contains a paper showing how we can effectively post during search just those symmetry breaking constraints that are not yet broken by the current assignment [17]. These examples illustrate how the solution proposed to a constraint pattern may depend on what is available in our particular solver. Constraint patterns also need to apply to a range of problems. For example, double lex ordering can be applied even to problems like the steel mill slab design problem with partial row and column symmetry. In this case, we just order lexicographically those subsets of rows or columns which are indistinguishable.
4.3 Dual models

The next pattern is documented in a number of papers (e.g., [18, 7, 19]) as well as ILOG's Solver 5.3 User's manual (Volume II). A constraint program defines a set of decision variables, each with an associated domain of values, and a set of constraints defining allowed values for subsets of these variables. The efficiency of a constraint program depends on a good choice for the decision variables, and a careful modeling of the constraints on these variables. Unfortunately, there is often considerable choice even in what to make the variables, and what to make the values. For example, in an exam timetabling problem, the variables could be the exams, and the values could be the times. However, we could take an alternative or dual viewpoint in which the variables are the times, and the values are the exams.

The choice of variables and values is especially evident in permutation problems. In a permutation problem, we have as many values as variables, and each variable takes an unique value. We can therefore easily exchange variables for values. Indeed, it is often beneficial to have both sets of variables with channeling constraints between the primal (or original) model and the dual [18, 7, 19]. Many assignment, scheduling and routing problems are permutation problems. For example, sports tournament scheduling can be modelled as finding a permutation of the games to fit into the time slots, or a permutation of the time slots to fit into the games.

<table>
<thead>
<tr>
<th>Pattern Name: DualModelling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem:</td>
</tr>
<tr>
<td>Forces:</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Solution:</td>
</tr>
<tr>
<td>Example:</td>
</tr>
</tbody>
</table>

**Fig. 4.** Constraint pattern for DualModelling

An alternative or dual viewpoint can be beneficial for a number of reasons. First, we can get different amounts of propagation in a primal, dual or a combined model [7, 19]. Second, certain constraints may be more easily stated (and propagated) on a dual model. If others are more easily stated (and propagated) on the primal, we can decide to use a combined model. Third, branching on the dual variables can be useful. For example, in a permutation problem, dual variables correspond to primal values. Therefore a variable ordering heuristic on dual variables is essentially a value ordering heuristic on the primal model.
Variable ordering heuristics like fail first tend to be cheap and effective. On the other hand, value ordering heuristics tend to be neither. Variable ordering on the dual therefore can be an effective means to get value ordering on the primal.

As a concrete example, consider the Balanced Academic Curriculum Problem (prob030 in CSPLib) [20]. The objective is to assign time slots to courses. A natural matrix model is a 2-d array of 0/1 decision variables indicating if a course is given in a particular time slot. Most of the constraints are easy to post using this array of variables (e.g., the constraint that a limited number of courses occur in any time slot is simply a row sum constraint). However, one type of constraint is not easy to post on this array of variables. This is the course prerequisite constraint: every course must occur after all its prerequisites. An easy way to post this constraint is to take an alternative dual viewpoint with a 1-d array of finite-domain variables. Each variable here takes as its value the time slot for a particular course. A prerequisite constraint is now simply ordering constraints between the course variable and each of the course variables associated with its prerequisites. The other constraints are not as easily specified on this 1-d array. It is therefore beneficial to have both arrays and to channel between them [20].

4.4 Auxiliary variables

The next pattern is documented in [21, 22]. A common method for improving a basic constraint model is to introduce auxiliary variables. Such variables permit propagation to occur between constraints with structure in common. Consider, for example, the problem of finding optimal Golomb rulers (prob006 in CSPLib). A Golomb ruler is a set of \( n \) integer valued ticks, such that the distance between any pair of ticks is different from the distance between any other pair. The objective is to find the optimal or shortest such ruler. Such rulers have practical applications in radio astronomy.

A natural model for the Golomb ruler problem has a finite-domain variable, \( X_i \) for each tick, and this is assigned the position of the tick on the ruler. To break symmetry between the ticks, we can post constraints of the form \( X_i < X_j \) for \( i < j \). To ensure that all inter-tick distances are distinct, we can post quaternary constraints of the form: \( |X_i - X_j| \neq |X_k - X_l| \). Each such constraint computes two inter-tick differences, and each of these differences appears in a quadratic number of other constraints. A better model introduces auxiliary variables for these differences, \( D_{ij} \) and ternary constraints of the form: \( D_{ij} = |X_i - X_j| \).

Introducing auxiliary variables can be advantageous for several reasons. For example, we can get more propagation through the domains of these auxiliary variables. As a second example, we may be able to post simpler constraints on the auxiliary variables. In the case of the Golomb ruler problem, we can replace the large number of quaternary constraints by a single all-different constraint on the auxiliary variables. We can then use an efficient algorithm for enforcing GAC or BC on such a constraint [23, 24]. In [22], we show that modelling the Golomb ruler with such auxiliary variables increases the amount of constraint propagation and reduces runtimes significantly. More recently, we have developed
methods for automatically introducing such auxiliary variables into a constraint model [25, 26].

<table>
<thead>
<tr>
<th>Pattern Name: AuxiliaryVars</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context:</td>
</tr>
<tr>
<td>Problem:</td>
</tr>
<tr>
<td>Forcés:</td>
</tr>
<tr>
<td>Solution:</td>
</tr>
<tr>
<td>Example:</td>
</tr>
</tbody>
</table>

Fig. 5. Constraint pattern for AuxiliaryVars

4.5 Implied constraints

The next pattern is also described in a number of papers (e.g., [27, 22]), as well as ILOG’s Solver 5.3 User’s manual (Volume II). A common method for improving a basic constraint model is to introduce implied constraints. These are constraints which are not logically necessary but which may reduce search. Consider again the problem of finding optimal Golomb rulers. In the last section, we argued for the introduction of auxiliary variables for the inter-tick differences, ternary constraints of the form $D_{ij} = |X_i - X_j|$ and an all-different constraint over $D_{ij}$. By transitivity, as $X_i < X_j$ for $i < j$, we can infer that $D_{ij} < D_{ik}$ for any $j < k$. This implied constraint is not logically necessary. We obtain the same set of solutions with or without it. However, as shown in [22], its inclusion in the model (along with other implied constraints) reduces search and saves runtime.

Not all implied constraints are useful. Constraint propagation on an implied constraint that is very immediate may do no more pruning than constraint propagation on a basic model. In addition, even if an implied constraint reduces search, it adds overhead to the constraint propagation. In [22], we outline two basic criteria for deciding which implied constraints to add. First, implied constraints either should have specialized, efficient and effective constraint propagation algorithms or should be of small arity. This limits the overheads of adding the implied constraint and helps ensure it will propagate. Second, circumstances in which an implied constraint leads to pruning should be obvious and frequent. The hope is that the implied constraint will reduce search sufficiently to justify the overhead.

One way to develop useful implied constraints is to study the search process. Suppose the constraint solver explores an “obviously” futile part of the search tree. The partial assignments considered by the solver cannot be extended to a complete solution, but they satisfy the constraints in the model. Our challenge then is to identify an implied constraint that would have pruned this branch immediately. We are currently developing methods for inferring useful implied
constraints automatically [25, 26]. For example, one of our methods identifies a clique of not-equals constraints (e.g. the constraints \( D_{ij} \neq D_{kl} \) in the Golomb ruler problem) and replaces them by an all-different constraint. Another method performs Gaussian-like elimination (e.g. if we introduce an auxiliary variable for a repeated expression, this method eliminates the repeated expression in favour of the auxiliary variable).

<table>
<thead>
<tr>
<th>Pattern Name: ImpliedConstraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context: A basic constraint model.</td>
</tr>
<tr>
<td>Problem: Search going down obviously futile branches.</td>
</tr>
<tr>
<td>Forces: Overhead of introducing additional constraints.</td>
</tr>
<tr>
<td>Applicability of the new implied constraints.</td>
</tr>
<tr>
<td>Solution: Introduce implied constraints that prune such branches.</td>
</tr>
<tr>
<td>Example: Golomb ruler problem.</td>
</tr>
</tbody>
</table>

**Fig. 6.** Constraint pattern for ImpliedConstraints

5 Related work

Unfortunately, constraint patterns are rarely described in a general way that permits their immediate use in other applications. One exception is a paper by Martin Green and David Cohen [28] that identifies a constraint pattern which is useful for modelling a range of assignment problems. The pattern occurs, for example, in the problem of assigning radio frequencies to pilots in a model aircraft tournament. In a straightforward model, in which the pilots are the variables and the values assigned to these variables are the radio frequencies, pilots assigned the same frequency are symmetric. We can therefore swap any two pilots assigned the same frequency and obtain a symmetric solution. To eliminate such symmetries, Green and Cohen propose an alternative viewpoint similar to the swapping of values for variables.

Hans Schlenker and Georg Ringwelski use a design pattern in POOC [29], a platform for object-oriented constraint programming. POOC provides Java wrappers around commercial and academic constraint solvers. Different constraint solvers can thus be easily compared. In addition, Java programmers can rapidly experiment with constraint solving. The wrappers are designed using the object factory design pattern. This defines an interface for creating an object, whilst allowing subclasses to decide which class to instantiate.

6 Conclusions and Future Work

I have argued that we need to identify, formalize and document patterns in constraint models in a similar way to the patterns identified by the software
engineering community. A library of such patterns will help tackle the modelling “bottleneck” that hinders the uptake of constraint programming. I have described some of the more interesting constraint patterns which have already been identified. I also discussed the different ways that we can exploit such patterns. For example, one way to exploit common constraint patterns is to extend the constraint language.

<table>
<thead>
<tr>
<th>IdentifyActors</th>
<th>DistinguishPrefs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SPECIFY</strong></td>
<td>:</td>
</tr>
<tr>
<td><strong>MODEL</strong></td>
<td>:</td>
</tr>
<tr>
<td><strong>Constraint Patterns</strong></td>
<td><strong>AuxiliaryVars</strong></td>
</tr>
<tr>
<td><strong>Refine</strong></td>
<td>:</td>
</tr>
<tr>
<td><strong>Solve</strong></td>
<td>:</td>
</tr>
</tbody>
</table>

Fig. 7. A sketch of a possible constraint pattern taxonomy

There are a number of important directions to follow. First, more patterns need to be collected. As I argued before, this requires the efforts of the whole community, not just of one individual. Second, the patterns outlined here need more detail and generality. Third, the patterns need to be organized into a pattern taxonomy so that they can be accessed easily. See Figure 7 for a possible start to such a taxonomy. Fourth, we need to collect these patterns into a pattern library. I believe such a library would be a significant asset to the constraint programming community. A first attempt at such a library is taking shape at 4c.ucc.ie/patterns/.

Acknowledgments

The author is supported by the Science Foundation Ireland. He wishes to thank the members of the 4C lab, the APES research group (especially Barbara Smith),
the York AI Group (especially Alan Frisch), the Uppsala ASTRA group, and David Cohen.

References

2. Gamma, E., Helm, R., Johnson, R., Vlissides, J.: Design Patterns: Elements of Reusable Object-Oriented Software. Addison-Wesley (1995)