

Final Project of Combinatorial Decision Making and Optimization, Module 1

SAT Model

Federico Cichetti - federico.cichetti@studio.unibo.it

February 27, 2022

Contents

1	Model	2
1.1	Order Encoding	2
1.1.1	Ordering Integrity	2
1.2	Variables	3
1.3	Constraints	3
1.3.1	Keeping Circuits in Bounds	3
1.3.2	Bidimensional No-Overlap	4
1.3.3	Implied Constraints	5
1.3.4	Symmetry Breaking Constraints	5
1.3.5	Other Reductions	6
2	Search	6
2.1	From Satisfiability to Optimality	7
3	Rotation Extension	7
4	Implementation	7
5	Results	8

Introduction

This report describes the choices behind the implementation of different solutions for the Very Large Scale Integration (VLSI) problem, presented as a project work for the course of Combinatorial Decision Making and Optimization, Module 1.

This document in particular focuses on how the problem can be modeled using *propositional logic* and solved efficiently using a *SAT solver*. Some of the concepts used in this solution have been presented in the report describing the *CP* model and are therefore not repeated here. Furthermore, our main source for the ideas presented in this solution is [1], which provides an impressive SAT model and discusses many techniques for improving search efficiency.

1 Model

Our SAT model is more powerful than the model we presented for CP, but also more complex due to the difficulty to express high level concepts with the limited language of propositional logic.

1.1 Order Encoding

The core idea we adopted from [1] is that of starting from a limited but sufficient amount of constraints expressed in the language of constraint programming and *translating them* into propositional logic using *order encoding*.

Order encoding [2] is a simple but effective way to *translate CSP constraints into boolean variables*. In practice, a constraint of the form $x \leq c$ with x an integer variable and c an integer value, is encoded into a set of boolean variables $p_{x,c}$ for all meaningful c .

Example 1. We have variable $x_1 \in \{0, 1, 2, 3\}$. We can express $x_1 = 1$ using the following set of boolean variables:

$$\neg p_{x_1,0}, \quad p_{x_1,1}, \quad p_{x_1,2}, \quad p_{x_1,3}$$

From this set of constraints, we can retrieve the correct assignment $x_1 = 1$ by ordering the variables and picking the value of the first true assignment (value i for which $p_{x_1,i-1}$ is false and for all $k \geq i$, $p_{x_1,k}$ is true).

Example 2. We have variables $x_1, x_2 \in \{0, 1, 2, 3\}$. The constraint $x_1 + 1 \leq x_2$ can be decomposed into this set of *primitive comparisons*:

$$\neg(x_2 \leq 0), \quad (x_1 \leq 0) \vee \neg(x_2 \leq 1)^1, \quad (x_1 \leq 1) \vee \neg(x_2 \leq 2), \quad (x_1 \leq 2)$$

This set of constraints is expressed using the following set of boolean variables:

$$\neg p_{x_2,0}, \quad p_{x_1,0} \vee \neg p_{x_2,1}, \quad p_{x_1,1} \vee \neg p_{x_2,2}, \quad p_{x_1,2}$$

1.1.1 Ordering Integrity

Order encoding requires that an internal *ordering integrity* is maintained: for instance if $x_2 \leq 0$, the constraints $x_2 \leq 1, x_2 \leq 2, \dots$ are always true. Therefore, $p_{x_2,i}$ must be true $\forall i > 0$. The order integrity constraint is usually maintained with a constraint like the following:

$$\neg p_{x_2,0} \vee p_{x_2,1}, \quad \neg p_{x_2,1} \vee p_{x_2,2}, \dots$$

expressing the impossibility to have $p_{x_2,i}$ true and $p_{x_2,i+1}$ false.

¹Personally, an intuitive way to understand the *or-not* comparisons is to think of them as *if-then* comparisons on conditions that are not respected: for instance, *if $x_2 \leq 1$ then $x_1 \leq 0$* . Also, *if $x_1 > 0$, then $x_2 > 1$* . The logic expression also adds to the meaning the fact that both the expressions can be true at the same time, and that if one is true we don't really care about the other's truth value as far as the constraint's satisfiability is concerned.

1.2 Variables

Given any instance of the problem, the values we have access to are:

- W : The width of the plate, which is fixed.
- n : The number of circuits to be placed.
- The measures of the circuits, that we save into two arrays: cw for widths and ch for heights. The set of circuits is C .
- The initial solution, obtained with the algorithm we have described in the CP report. In SAT there are no warm start mechanisms, so we use the initial solution just to have an *higher bound* on the height of the circuit: $maxh$.
- Like for the CP model, we can also provide a lower bound to the height by summing all circuits' areas and dividing by W . This lower bound is referred to as *lowh* and is formally defined as $lowh = \lfloor \frac{\sum_i ch_i cw_i}{W} \rfloor$.

The boolean variables we instantiate for the problem are the following:

- $px_{c,e}$ for all circuits $c \in \{0, \dots, n-1\}$, for all possible x-axis positions $e \in \{0, \dots, W-1\}$. Each one of these variables indicates whether circuit c is placed at an x-position $\leq e$.
- $py_{c,f}$ for all circuits $c \in \{0, \dots, n-1\}$, for all possible y-axis positions $f \in \{0, \dots, maxh-1\}$. Each one of these variables indicates whether circuit c is placed at an y-position $\leq f$.

1.3 Constraints

The SAT encoding shares with the CP model the principles behind the main constraints: the two rules of interest in the positioning of the blocks are:

- The circuits should be placed *entirely* within the W and $maxh$ boundaries of the board.
- The circuits should not overlap.

The problem is that we are working within the *order encoding* framework, which, in general, makes posing constraints related to the positioning of the circuits quite complex because, as we discussed in 1, the only way to know the exact position of a circuit is to order variables and select the first true one, which is something that can't be done dynamically while the solver is running. The main constraints have to be re-designed keeping this shift from *absolute positioning* coordinates to *relative positioning* coordinates into account.

1.3.1 Keeping Circuits in Bounds

The idea is that circuit c with dimensions $cw_c \times ch_c$ cannot be placed closer than cw_c to the W limit, nor closer than ch_c to the $maxh$ limit. We can express this constraint by asking that *at least one* of the variables between $px_{c,0}$ and $px_{c,W-cw_c}$ and between $py_{c,0}$ and $py_{c,maxh-ch_c}$ is assigned true. This in turn means that all circuits respect $px_c \leq W - cw_c$ and $py_c \leq maxh - ch_c$, which is the objective of the first requirement.

$$\bigwedge_{c \in C} \bigvee_{0 \leq e < W - cw_c} px_{c,e}$$

$$\bigwedge_{c \in C} \bigvee_{0 \leq f < maxh - ch_c} py_{c,f}$$

Being positional encodings, the arrays px and py also respect the ordering integrity constraints we mentioned in Section 1.1.1:

$$\bigwedge_{c \in C} \bigwedge_{0 \leq e < W - cw_c} (\neg px_{c,e} \vee px_{c,e+1})$$

$$\bigwedge_{c \in C} \bigwedge_{0 \leq f < maxh - ch_c} (\neg py_{c,f} \vee py_{c,f+1})$$

1.3.2 Bidimensional No-Overlap

The second requirement is a two-dimensional no-overlap constraint, which, using the notation we used for the CP report, can be decomposed into:

$$\forall ci, cj \in C, ci < cj, \\ (xpos_{ci} + ws_{ci} \leq xpos_{cj}) \vee (xpos_{cj} + ws_{cj} \leq xpos_{ci}) \vee (ypos_{ci} + hs_{ci} \leq ypos_{cj}) \vee (ypos_{cj} + hs_{cj} \leq ypos_{ci})$$

Unfortunately, we don't have *global constraints* like in CP, so we aim for a more direct translation of this decomposition into our framework.

In [1], the authors propose to add two arrays of variables to better express this complex constraint:

- $lr_{i,j}$ for all circuits $i, j \in \{0, \dots, n-1\}$, indicating whether circuit c_i is placed to the left of circuit c_j .
- $ud_{i,j}$ for all circuits $i, j \in \{0, \dots, n-1\}$, indicating whether circuit c_i is placed below circuit c_j .

The meaning of these variables can be visualized in Figure 1

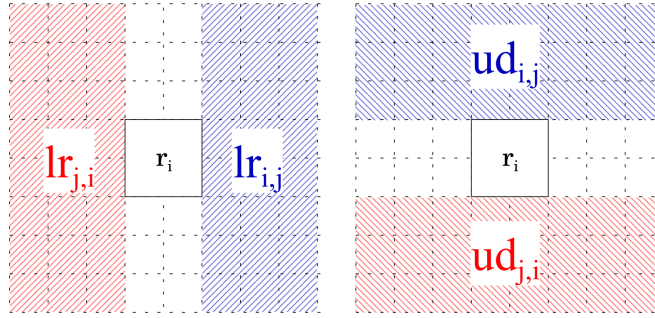


Figure 1: The meaning of variables $lr_{i,j}, lr_{j,i}, ud_{i,j}, ud_{j,i}$ between two circuits r_i and r_j . Source.

The no-overlap constraint can then be expressed by first defining what the relations *lr* and *ud* imply for the positioning of the circuits and then requiring that for each pair of circuits at least one between the possible relations $lr_{i,j}, lr_{j,i}, ud_{i,j}, ud_{j,i}$ is true.

The relations definition is quite tricky. The main idea is that given that a circuit ci is on the left or above circuit cj , if ci is placed *after* position e , then cj cannot be placed before position $e + cw_{ci}$, or position $e + ch_{ci}$ on the vertical case. A similar concept is valid for all placements of all pairs of circuits, so we can add the following constraints to the solver:

$$\bigwedge_{ci, cj \in C, ci < cj} \bigwedge_{0 \leq e < W - cw_{ci}} (\neg lr_{ci, cj} \vee px_{ci, e} \vee \neg px_{cj, e + cw_{ci}}) \\ \bigwedge_{ci, cj \in C, ci < cj} \bigwedge_{0 \leq e < W - cw_{cj}} (\neg lr_{cj, ci} \vee px_{cj, e} \vee \neg px_{ci, e + cw_{cj}}) \\ \bigwedge_{ci, cj \in C, ci < cj} \bigwedge_{0 \leq f < maxh - ch_{ci}} (\neg ud_{ci, cj} \vee py_{ci, f} \vee \neg py_{cj, f + ch_{ci}}) \\ \bigwedge_{ci, cj \in C, ci < cj} \bigwedge_{0 \leq f < maxh - ch_{cj}} (\neg ud_{cj, ci} \vee py_{cj, f} \vee \neg py_{ci, f + ch_{cj}})$$

To this set of constraints, we add:

$$\bigwedge_{ci, cj \in C, ci < cj} (lr_{ci, cj} \vee lr_{cj, ci} \vee ud_{ci, cj} \vee ud_{cj, ci})$$

which ensures that all pairs of circuits have at least one of the previously defined relations active. This large set of constraints overall ensures that circuits do not overlap and satisfy the second requirement for producing a solution.

1.3.3 Implied Constraints

We also add a constraint that proved successful in our CP model, that is the requirement of *exactly one* circuit to be placed in position (0,0). We have seen in class that the constraint *exactly one* can be decomposed into *at least one*, which is an \vee of all boolean variables involved in the constraint, and *at most one*, which poses $\neg(b_1 \wedge b_2)$ constraints between all pairs of boolean variables b_1 and b_2 . Therefore, we can add a constraint like the following to the problem:

$$\bigwedge_{c \in C} \text{exactly_one}(px_{c,0} \wedge py_{c,0})$$

Furthermore, we can add the following constraints: if circuit ci is placed on the left of or below circuit cj , cj cannot be placed in the first cw_{ci} or ch_{ci} positions, because there wouldn't be enough space to fit both circuits. Formally:

$$\begin{aligned} & \bigwedge_{ci, cj \in C, ci < cj} \bigwedge_{0 \leq e < cw_{ci}} (\neg lr_{ci, cj} \vee \neg px_{cj, e}) \\ & \bigwedge_{ci, cj \in C, ci < cj} \bigwedge_{0 \leq e < cw_{cj}} (\neg lr_{cj, ci} \vee \neg px_{ci, e}) \\ & \bigwedge_{ci, cj \in C, ci < cj} \bigwedge_{0 \leq f < ch_{ci}} (\neg ud_{ci, cj} \vee \neg py_{cj, f}) \\ & \bigwedge_{ci, cj \in C, ci < cj} \bigwedge_{0 \leq f < ch_{cj}} (\neg ud_{cj, ci} \vee \neg py_{ci, f}) \end{aligned}$$

1.3.4 Symmetry Breaking Constraints

As already highlighted in the CP report, the model contains 3 very obvious kinds of symmetries that we need to eliminate: vertical, horizontal and diagonal flippings.

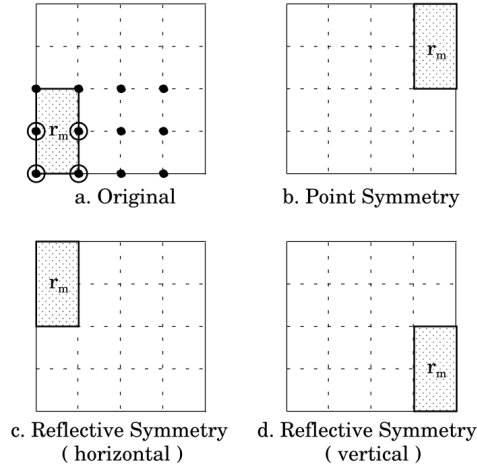


Figure 2: Effect of domain reduction and the symmetries it breaks. The original domain of the large circuit cm is represented with dots, while the reduced domain with circles.

Within the order encoding framework, it becomes quite difficult to express symmetry breaking constraints imposing a lexicographic ordering of the circuits like we did for CP, because we only know position relations, not exact coordinates for the circuits.

In [1], the authors propose to reduce the domain of the largest circuit's possible x and y coordinates: if circuit cm is the largest out of all the circuits, its position must be (a, b) with $0 \leq a \leq \lfloor \frac{W - cw_{cm}}{2} \rfloor$ and $0 \leq b \leq \lfloor \frac{maxh - ch_{cm}}{2} \rfloor$. In this way, we constrain the largest circuit to be placed in the *bottom-left* area of the board, thus making complete horizontal, vertical or diagonal flippings impossible. Figure 2 (from the original paper) shows the effect of the constraint on a large circuit.

1.3.5 Other Reductions

We also implemented other types of search space reductions proposed by the paper:

1. **Same Rectangle Reduction:** If two rectangles have the same dimension, we can decide the positional relation between them so that the solver will only be able to try a subset of the possible placements and will never swap the two circuits. We set $lr_{cj,ci} = \perp$ and then require that either ci is at the left or below cj .

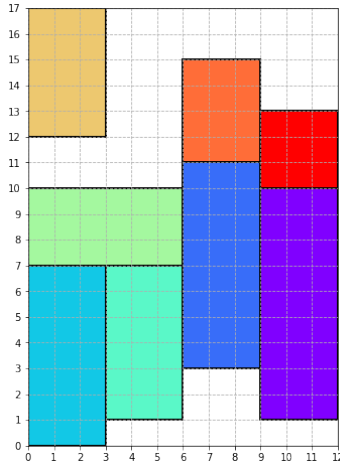
$$\bigwedge_{ci,cj \in C, ci < cj, ch_{ci}=ch_{cj}, cw_{ci}=cw_{cj}} (\neg lr_{cj,ci} \wedge (lr_{ci,cj} \vee \neg ud_{cj,ci}))$$

2. **Large Rectangle Reduction:** For each pair of rectangles $ci, cj \in C, ci < cj$, if the sum of their widths is greater than W ($cw_{ci} + cw_{cj} > W$) we cannot pack them horizontally, meaning that they can never be on the left or on the right of one another, but only above or below. Therefore, we can set $lr_{ci,cj} = \perp$, $lr_{cj,ci} = \perp$ and remove all constraints which contain $\neg lr_{ci,cj}$ or $\neg lr_{cj,ci}$ (i.e. half of the constraints in Sections 1.3.2 and 1.3.3) because they will always be satisfied. The same can be said on the vertical dimensions: $(ch_{ci} + ch_{cj} > maxh) \implies (ud_{ci,cj} = \perp \wedge ud_{cj,ci} = \perp)$ and the removal of all constraints containing these variables.

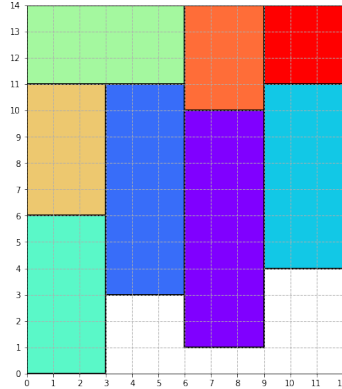
ideally
the sec-
ond part
of the
domain
reduction

2 Search

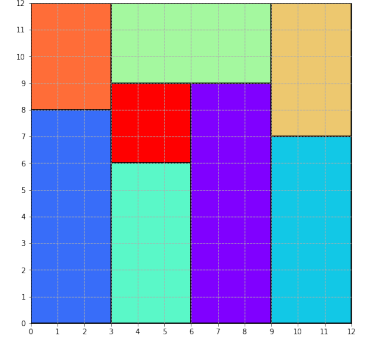
A model with the constraints we presented above is only able to provide a solution placing circuits in a $W \times maxh$ board, like the one presented in Figure 3a. Of course, we would like to *minimize* the height, so obtain a solution like the one in Figure 3c. SAT solving has no mechanisms to deal with *optimization problems* directly, therefore we adopt the indirect, but powerful method presented in [1], implemented using the same order encoding framework as all other variables.



(a) Non-optimal solution, max allowed height set to $maxh$.



(b) Sub-optimal solution, max allowed height set to 14.



(c) Optimal solution of the instance.

Figure 3: Solutions of the same instance with different cumulative heights.

First of all, we add yet another set of boolean variables to the model: ph_o for $o \in \{lowh, \dots, maxh\}$ indicating whether all circuits can be packed with no overlaps within a total height that is $\leq o$. Given that we have an initial solution at height $maxh$, ph_{maxh} is always true. Since the height boundary is expressed in order encoding, we have the usual ordering integrity constraint:

$$\bigwedge_{o \in \{lowh, \dots, maxh-1\}} (\neg ph_o \vee ph_{o+1})$$

By definition, this set of variables should affect the placement of the circuits in the board so that if ph_o is true for some o , it means that all circuits $c \in C$ must have been placed before $o - ch_c$: no circuit must extend above o . We translate this requirement into the following constraint:

$$\bigwedge_{o \in \{lowh, \dots, maxh-1\}} \bigwedge_{c \in C} (\neg ph_o \vee py_{c, o-ch_c})$$

The direct consequence of adding these variables to the problem is that we have the power to *decide* the maximum height to be reached by the circuits. Indeed, all it takes to keep circuits within the height limit o is to add $ph_o = \top$ as a constraint into the problem: if the problem is satisfiable at that height level, thanks to our constraints, the solver will assign y-coordinates that respect that boundary.

2.1 From Satisfiability to Optimality

Even though we have a powerful mechanism to set the maximum height boundary of the board, our model is missing an automated way to obtain the *optimal solution*. In most of the provided instances, the optimal h is equal to $lowh$ because there is always at least one way to place circuits so that there are no empty spaces in between. The problem is that we cannot make this assumption for the general case, so we need an algorithm to *test solvability* at many different boundaries and *choose* the best one as our optimal h .

In [1], the *bisection method* is proposed as an efficient way to obtain the best solution. The idea is to try and solve the problem using as height boundary $o = l + \lfloor (m - l)/2 \rfloor$ (so, setting $ph_o = \top$), where l and m are initially set to $lowh$ and $maxh$. If the problem is SAT, m is reduced to o , while if it is not SAT, l is increased to o . Then, we restore the previous solver state (retract $ph_o = \top$ and all learnt clauses), we update o with the new boundaries and repeat the process until $l = m$. When this happens, if the problem is not SAT with $ph_l = \top$, we declare the problem unsolvable, otherwise l is the best possible height boundary for the problem.

Solving multiple SAT problems can of course be quite time-consuming and does not scale well to large instances, but for most of the provided instances the solver is already quite fast and the sequence of problems is solved in just a few seconds (see Table 1). Furthermore, Z3 efficiently implements and *incremental solving* mechanism which allows us to *save* (push) the status of a solver and *restore* (pop) it later. We heavily employ this mechanism for adding and retracting the constraints relative to ph variables.

3 Rotation Extension

4 Implementation

The model has been implemented in Python using the Z3 Python API for expressing the variables and constraints. Similarly to the CP program, we have built a *launcher* which contains the code for loading the instances specified by the user, for instantiating the solver, the variables and constraints and for saving the solution as well as optionally displaying it graphically. The algorithms that provide the initial solution and sort the circuits by their size have been re-used from the CP program implementation.

Differently from CP, the output of the solving process is a mapping $var \rightarrow v \in \{\top, \perp\}$ from each of the boolean variables instanced in the model to the value that has been assigned to it by the SAT solver. To obtain a clear solution, we apply a *post-processing step* where we take the set of *true* variables and transform it into the sequence of horizontal and vertical coordinates by taking the first variable in order for each circuit, as explained in Example 1.

During the solution, we keep track of the elapsed time with different timers: in particular, we noticed that the actual solving of the problem (`solver.check()` call) is of course the heaviest part of the program, but the post-processing step also requires some time due to the amount of variables, string filtering and such.

Since the post-processing step is fundamental for the coherence of the solution, we decided to add these two timers together to express the final time seen in Table 1.

5 Results

We managed to solve 35 of the provided instances within the 5 minutes limit (depending on randomness of search and hardware), so 4 instances more with respect to the CP model.

Table 1 shows the results and solution times we obtained for each of the 40 instances using our models.

References

- [1] Takehide Soh et al. “A SAT-based Method for Solving the Two-dimensional Strip Packing Problem”. In: *Fundam. Inform.* 102 (Jan. 2010), pp. 467–487. DOI: 10.3233/FI-2010-314.
- [2] Naoyuki Tamura et al. “Compiling finite linear CSP into SAT”. In: vol. 14. Sept. 2006. ISBN: 978-3-540-46267-5. DOI: 10.1007/11889205_42.

	Solving Time (s)	Solved
1	0.003	True
2	0.001	True
3	0.001	True
4	0.000999	True
5	0.003001	True
6	0.004	True
7	0.004004	True
8	0.007	True
9	0.012	True
10	0.023004	True
11	0.157001	True
12	0.040999	True
13	0.106001	True
14	0.246993	True
15	0.126002	True
16	2.052551	True
17	0.678999	True
18	1.752	True
19	11.680127	True
20	0.837998	True
21	4.499001	True
22	15.56	True
23	0.715002	True
24	1.479997	True
25	87.786002	True
26	8.306998	True
27	0.569	True
28	1.079	True
29	1.442411	True
30	13.829572	True
31	1.645	True
32	300.058168	False
33	0.377997	True
34	1.181	True
35	4.612996	True
36	10.829999	True
37	300.04	False
38	300.056001	False
39	300.104998	False
40	301.203068	False

Table 1: Solving times and optimality status achieved by our model on the 40 instances of the problem. Times don't include the overhead of the Python launcher nor the time for finding an initial solution, but in most cases they are negligible.