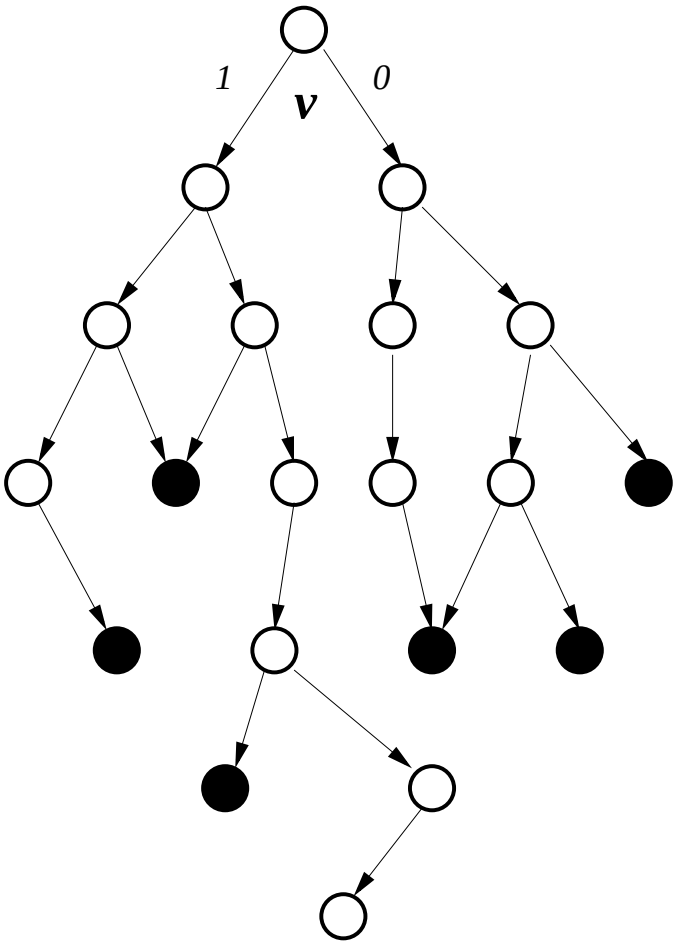
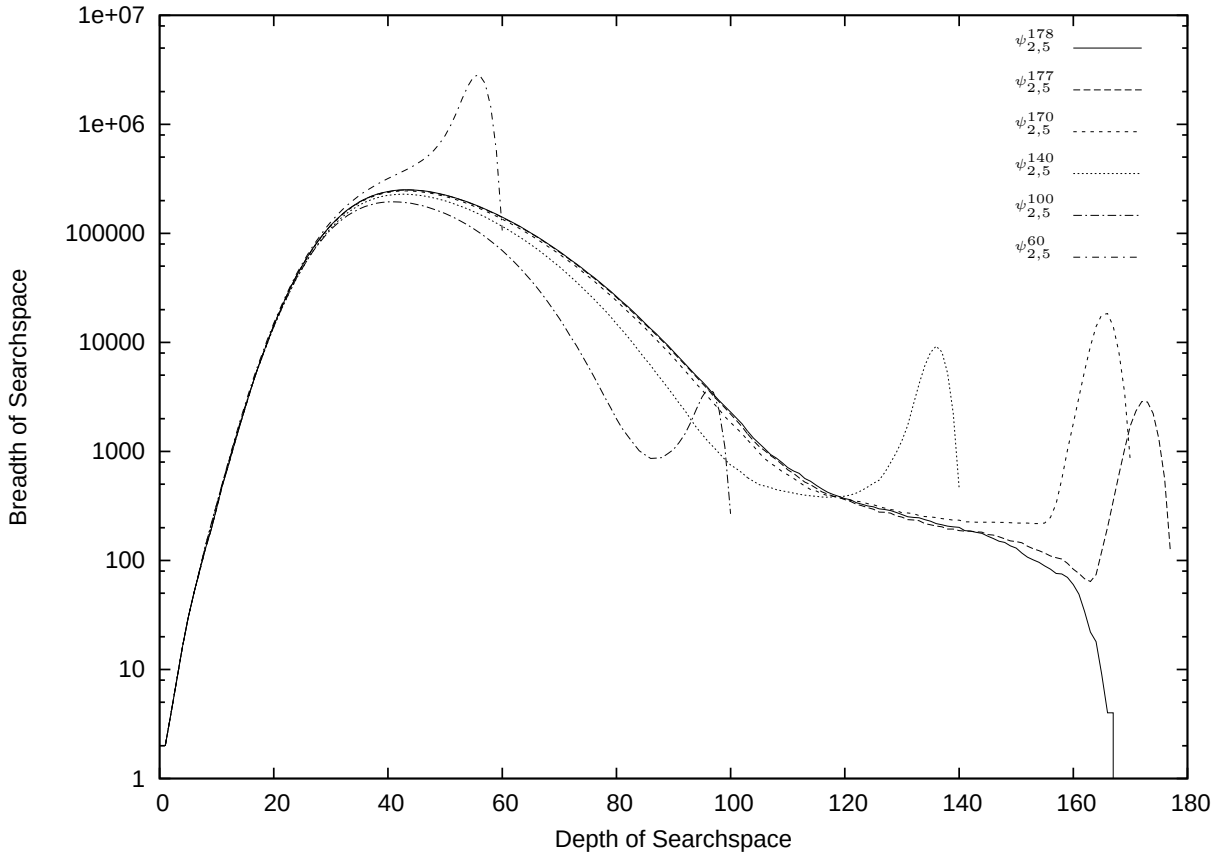
Resolution Tunnels for Improved SAT Solver Performance

A Search Space



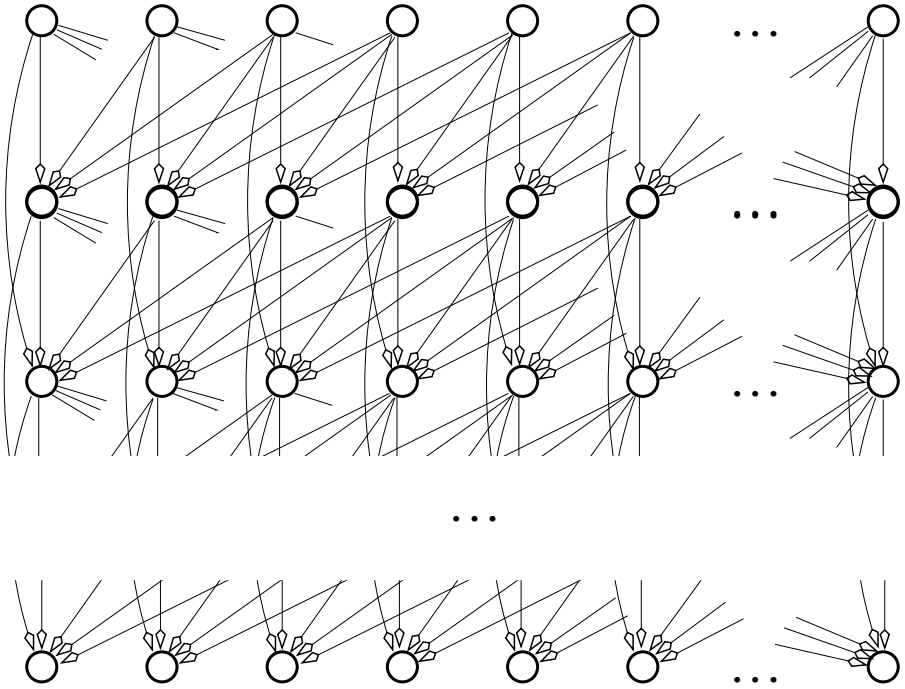
Search Profile



b

Typical SAT solver performance on a "hard" CNF formula.

Hard Problems That Fit This Profile



For example, Bounded Model Checking problems

Required to Reduce Search

- **To reduce search: need to learn forced (inferenced) constraints (clauses and values) earlier**
- **If formula is unsatisfiable, and is sparse, then must infer exponentially many large constraints before smaller constraints and values can be inferred**

Speed up search

Reduce Search Space:

- **Reduce breadth of search space by initially adding constraints that have no business being there**
- **At some search depth, when breadth is decreasing, remove the constraints**
- **Continue to the end**

Speed up search

Reduce Search Space:

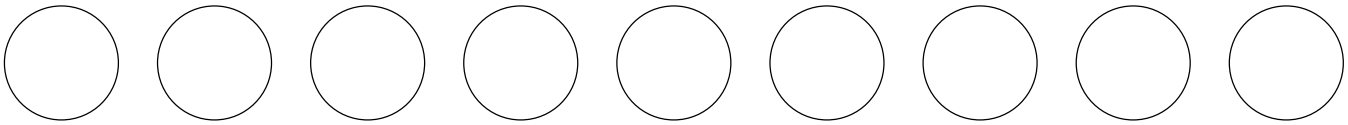
- **Reduce breadth of search space by initially adding constraints that have no business being there**
- **At some search depth, when breadth is decreasing, remove the constraints**
- **Continue to the end**

Notes:

- **Obtain constaints by looking for patterns in *Solutions* to smaller instances of the same family (downplay formula structure)**
- **May fail to find an existing solution, cannot be applied to unsat formulas, but**

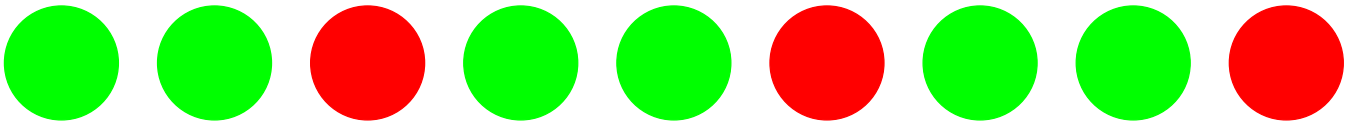
Example - Van der Waerden Formulas

Can you color 9 disks with red and green so that no three disks separated by 0, 1, 2, ... disks are the same color?



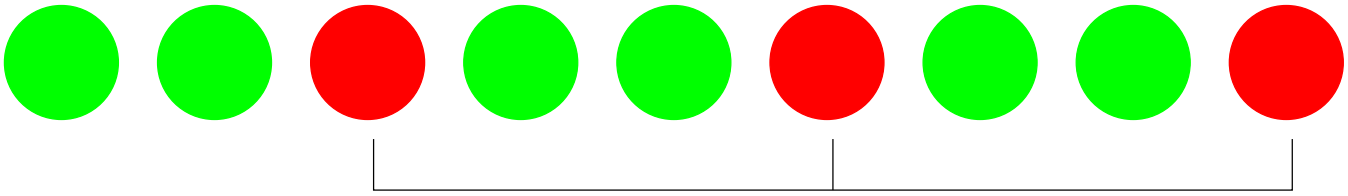
Example - Van der Waerden Formulas

Can you color 9 disks with red and green so that no three disks separated by 0, 1, 2, ... disks are the same color?



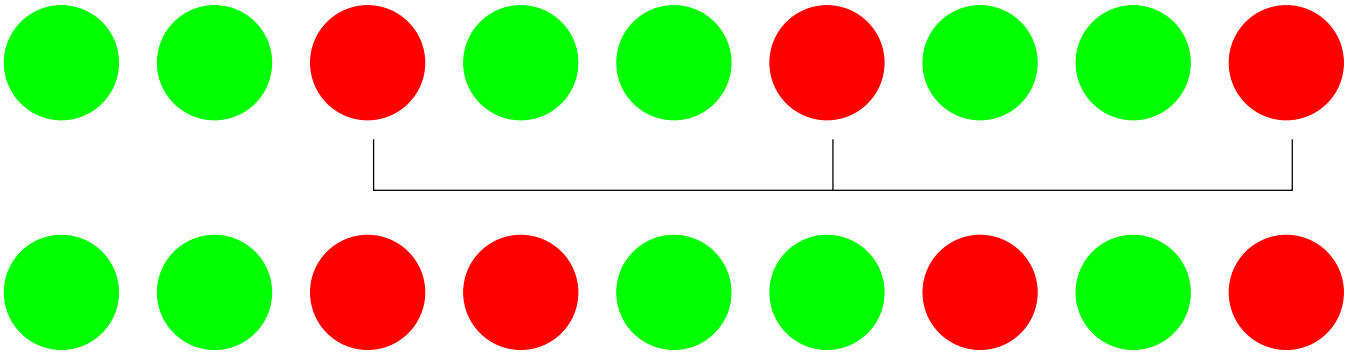
Example - Van der Waerden Formulas

Can you color 9 disks with red and green so that no three disks separated by 0, 1, 2, ... disks are the same color?



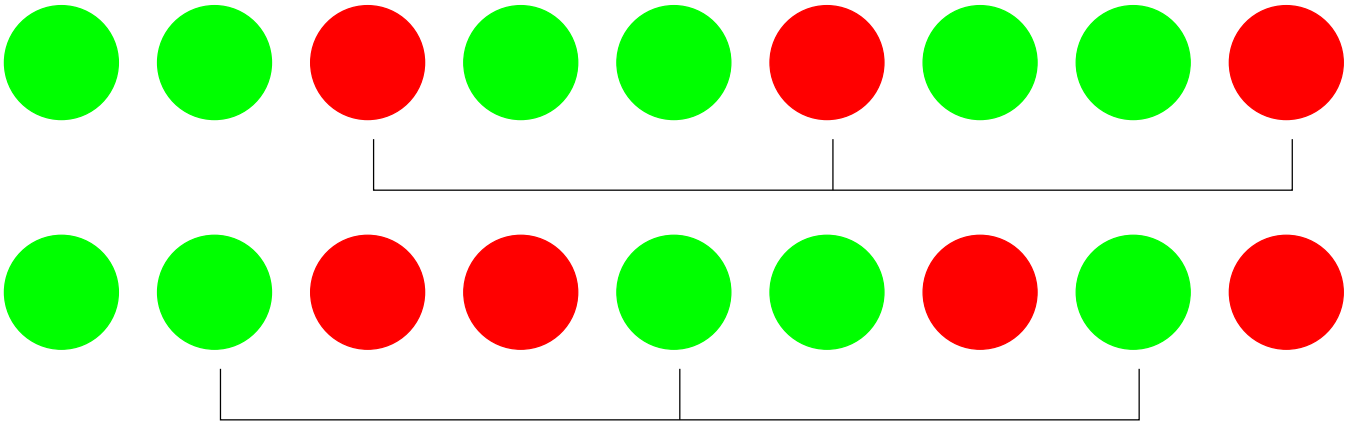
Example - Van der Waerden Formulas

Can you color 9 disks with red and green so that no three disks separated by 0, 1, 2, ... disks are the same color?



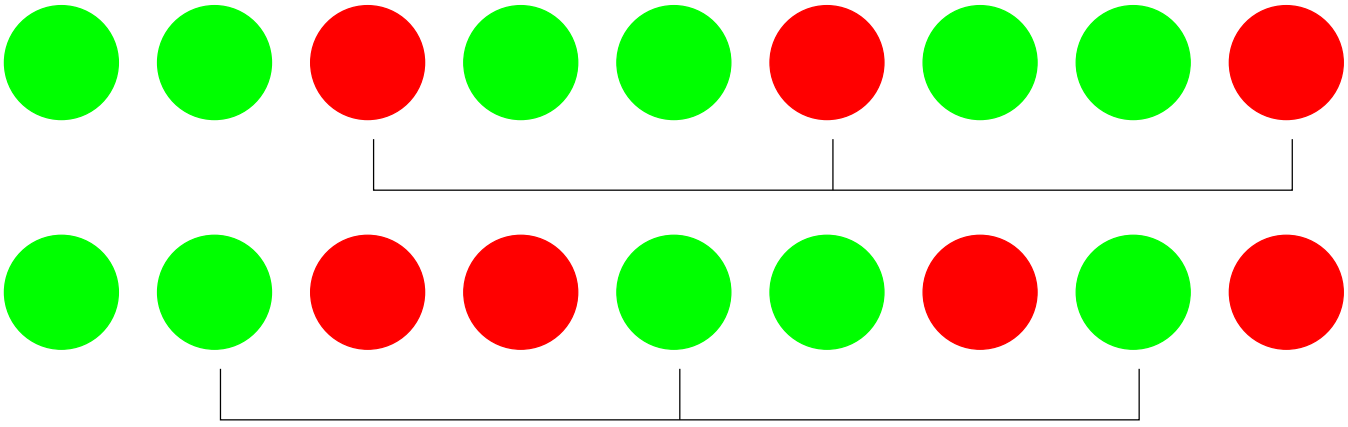
Example - Van der Waerden Formulas

Can you color 9 disks with red and green so that no three disks separated by 0, 1, 2, ... disks are the same color?



Example - Van der Waerden Formulas

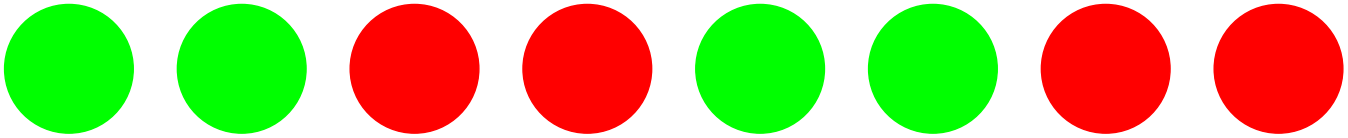
Can you color 9 disks with red and green so that no three disks separated by 0, 1, 2, ... disks are the same color?



Cannot do it!

Example - Van der Waerden Formulas

Can you color 8 disks with red and green so that no three disks separated by 0, 1, 2, ... disks are the same color?



Example - Van der Waerden Numbers

Partition the set $S_n = \{1, ...n\}$ of the first n positive consecutive integers into k classes. Let $P_{n,k}(l)$ be a proposition that is $True$ if and only if all partitions of S_n into k classes contain at least one arithmetic progression of length l in at least one class. The k, l Van der Waerden number, denoted $W(k, l)$, is the minimum n for which $P_{n,k}(l)$ is $True$.

Example: $k = 2$, $l = 3$, $n = 9$.

Not possible

Example: $k = 2$, $l = 3$, $n = 8$.

$\{\{1, 2, 5, 6\}\{3, 4, 7, 8\}\}$

Example - Van der Waerden Numbers

There is no known closed form expression for $W(k, l)$ and all but five of the first few numbers are unknown.

Table below shows all the known Van der Waerden numbers. In 1979 $W(4, 3)$ became the most recent addition to this table.

$k \setminus l$	**3**	**4**	**5**
2	**9**	**35**	**178**
3	**27**		
4	**76**		

Previous Bounds, Van der Waerden Numbers

$k \setminus l$	**3**	**4**	**5**	**6**	**7**	**8**
2	**9**	**35**	**178**	> 695	> 3702	> 7483
3	**27**	> 291	> 1209	> 8885	> 43854	> 161371
4	**76**	> 1047	> 10436	> 90306	> 262326	
5	> 125	> 2253	> 24044	> 177955		
6	> 206	> 3693	> 56692			

Example - Van der Waerden Formulas

Variables	Subscript Range	Meaning
$v_{i,j}$	$1 \le i \le n, 1 \le j \le k$	$v_{i,j} \equiv 1$ **iff** $i \in C_j$

Clauses	Subscript Range	Meaning
$\{\bar{v}_{i,r}, \bar{v}_{i,s}\}$	$1 \le i \le n, 1 \le r < s \le k$	i **is in at most one class**
$\{v_{i,1}, \ldots, v_{i,k}\}$	$1 \le i \le n$	i **is in at least one class**
$\{\bar{v}_{r,j}, \bar{v}_{r+1,j}, \ldots, \bar{v}_{r+l-1,j}\}$ $\{\bar{v}_{r,j}, \bar{v}_{r+2,j} \ldots, \bar{v}_{r+2(l-1),j}\}$ **...** $\{\bar{v}_{r,j}, \bar{v}_{r+t,j} \ldots, \bar{v}_{r+t(l-1),j}\}$	$1 \le r \le n - l + 1$ $1 \le j \le k$ **...** $t = \lfloor (n-r)/(l-1) \rfloor$	**no arithmetic progression of length** l **in** C_j

Previous Bounds, Van der Waerden Formulas

$k \setminus l$	**3**	**4**	**5**	**6**	**7**	**8**
2	**9**	**35**	**178**	> 341	> 614	> 1322
3	**27**	> 193	> 676	> 2236		
4	**76**	> 416				
5	> 125	> 880				
6	> 194					

Formulas

$k \setminus l$	**3**	**4**	**5**	**6**	**7**	**8**
2	**9**	**35**	**178**	> 695	> 3702	> 7483
3	**27**	> 291	> 1209	> 8885	> 43854	> 161371
4	**76**	> 1047	> 10436	> 90306	> 262326	
5	> 125	> 2253	> 24044	> 177955		
6	> 206	> 3693	> 56692			

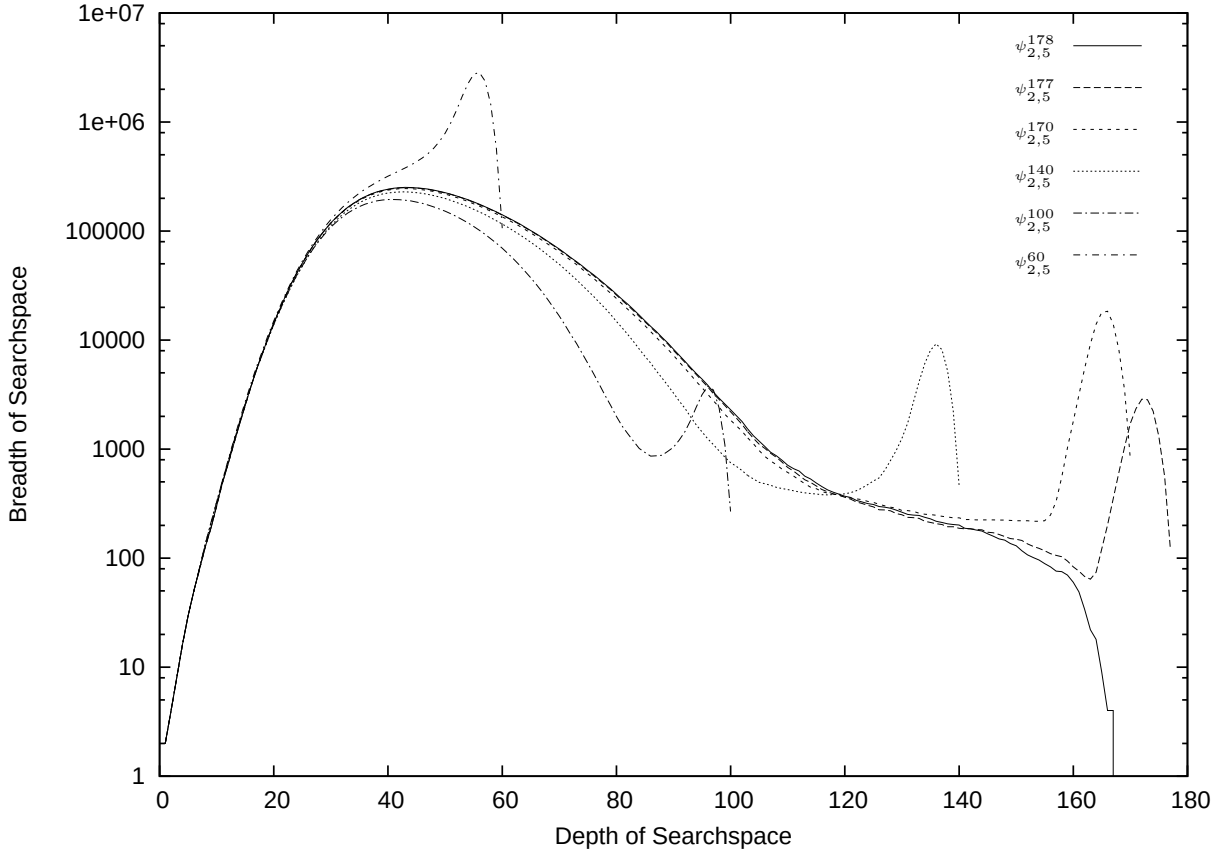
Analysis

Target W(2,6)

Variables	Subscript Range	Meaning
v_i	$-n/2 < i \leq n/2$	$v_i \equiv 1$ **if** $i + n/2 \in C_1$ $v_i \equiv 0$ **if** $i + n/2 \in C_2$

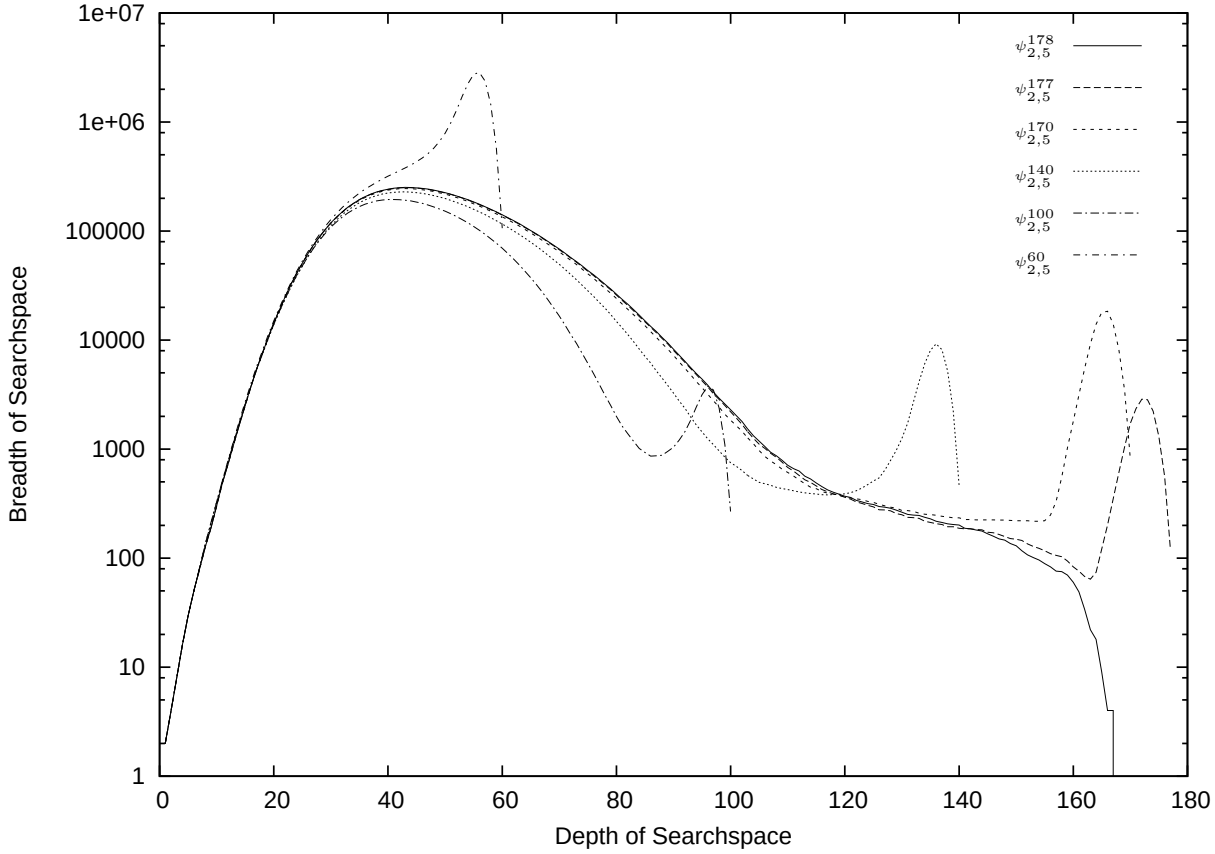
Clauses	Subscript Range	Meaning
$\{\bar{v}_i, \bar{v}_{i+1}, \ldots, \bar{v}_{i+5}\}$ $\{\bar{v}_i, \bar{v}_{i+2} \ldots, \bar{v}_{i+10}\}$ **...** $\{\bar{v}_i, \bar{v}_{i+t} \ldots, \bar{v}_{i+5t}\}$	$-n/2 < i \leq n/2 - 5$ **...** $t = \lfloor (n/2 - i)/5 \rfloor$	**no arithmetic progression of length 6 in** C_1
$\{v_i, v_{i+1} \ldots, v_{i+5}\}$ $\{v_i, v_{i+2} \ldots, v_{i+10}\}$ **...** $\{v_i, v_{i+t} \ldots, v_{i+5t}\}$	$-n/2 < i \leq n/2 - 5$ **...** $t = \lfloor (n/2 - i)/5 \rfloor$	**no arithmetic progression of length 6 in** C_2

Motivation - Analysis



***Observation 1*: Greatest maximum at about $W(2,l)/(2(l-1))$, approx the same value for $n > W(2,l)/(l-1)$, orders of magnitude greater than search breadth at depth $W(2,l)/(l-1)$.**

Motivation - Analysis

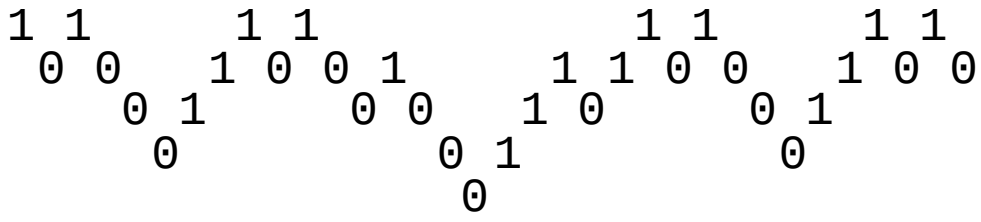


***Observation 2*: $W(2,l) \approx l * W(2,l-1)$, at least for small l.**

Solver Outline

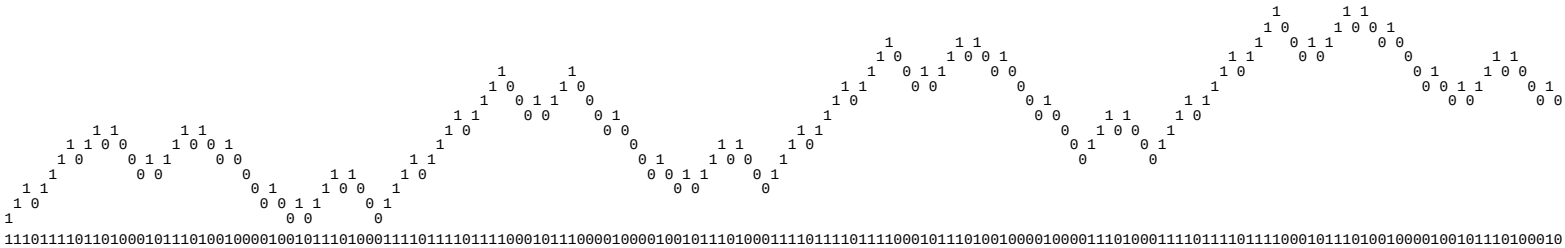
- **Choose small n formula (say around 210 variables)**
- **Add crazy constraints having no business there**
- **Solve until get through the mountain (check search breadth)**
- **Hopefully there is a solution left. Remove the crazy constraints.**
- **Finish solving. Hopefully there is a solution left.**
- **Increase n (add some "normal" clauses)**
- **Finish solving. Repeat until cannot get a solution.**

Tunnel 1



10100011101001000111011010001110 10

Analysis of solutions to $\psi_{2,4}$



11101111011010001011101001000010010111010001111011110111100010111000010000100101110100011110111101111000101110100100001000011101000111101111011110001011101001000010010111010001 0

Analysis of solutions to $\psi_{2,5}$

Limited length patterns of reverse symmetry

Tunnel 1

Conjecture 0.1 ***For every*** $\psi_{2,l}^{W(2,l)-1}$ ***there exists a solution that contains at least one reflected pattern of length*** $W(2,l)/((l-1)*2)$ ***with the middle positioned somewhere between*** $W(2,l)/(l-1)$ ***and*** $W(2,l)*(l-2)/(l-1)$**.**

Tunnel 1 is designed as a filter for consecutive variable assignment patterns that are not reverse symmetric.

Tunnel Clauses	Subscript Range	Meaning
$\{v_{-i}, v_{i+1}\}, \{\bar{v}_{-i}, \bar{v}_{i+1}\}$	$0 \leq i < s/2$	**force** $v_{-i} \equiv \bar{v}_{i+1}$**.**

Tunnel 2

- **Some small assignment patterns *do not* occur in solutions.**
- **The second tunnel filters those patterns.**
- **This action is opposite to that of *forcing* patterns to occur which is the objective of the first tunnel.**

Tunnel Clauses	Subscript Range	Filters
$\{v_i, \bar{v}_{i+t}, v_{i+2t}, \bar{v}_{i+3t}, v_{i+4t}, \bar{v}_{i+5t}\}$	$-n/2 < i \leq n/2 - 5t$	**010101**
$\{\bar{v}_i, v_{i+t}, \bar{v}_{i+2t}, v_{i+3t}, \bar{v}_{i+4t}, v_{i+5t}\}$	$1 \leq t \leq 20$	**101010**
$\{v_i, v_{i+t}, \bar{v}_{i+2t}, \bar{v}_{i+3t}, v_{i+4t}, \bar{v}_{i+5t}, \bar{v}_{i+6t}, v_{i+7t}\}$		**00110110**
$\{\bar{v}_i, \bar{v}_{i+t}, v_{i+2t}, v_{i+3t}, \bar{v}_{i+4t}, v_{i+5t}, v_{i+6t}, \bar{v}_{i+7t}\}$		**11001001**
$\{v_i, \bar{v}_{i+t}, \bar{v}_{i+2t}, v_{i+3t}, \bar{v}_{i+4t}, \bar{v}_{i+5t}, v_{i+6t}, v_{i+7t}\}$		**01101100**
$\{\bar{v}_i, v_{i+t}, v_{i+2t}, \bar{v}_{i+3t}, v_{i+4t}, v_{i+5t}, \bar{v}_{i+6t}, \bar{v}_{i+7t}\}$	$-n/2 < i \leq n/2 - 7t$	**10010011**
$\{v_i, v_{i+t}, \bar{v}_{i+2t}, \bar{v}_{i+3t}, v_{i+4t}, v_{i+5t}, v_{i+6t}, \bar{v}_{i+7t}\}$	$1 \leq t \leq 20$	**00111001**
$\{\bar{v}_i, \bar{v}_{i+t}, v_{i+2t}, v_{i+3t}, v_{i+4t}, \bar{v}_{i+5t}, \bar{v}_{i+6t}, v_{i+7t}\}$		**11000110**
$\{\bar{v}_i, v_{i+t}, v_{i+2t}, \bar{v}_{i+3t}, \bar{v}_{i+4t}, \bar{v}_{i+5t}, v_{i+6t}, v_{i+7t}\}$		**10011100**
$\{v_i, \bar{v}_{i+t}, \bar{v}_{i+2t}, v_{i+3t}, v_{i+4t}, v_{i+5t}, \bar{v}_{i+6t}, \bar{v}_{i+7t}\}$		**01100011**

Tunnel 3

- **Analytic solutions to $\psi_{2,6}^{n}$ have been found for various values of n including 565 and 695. For solution to $\psi_{2,6}^{565}$, re-index the assigned variables**

  $$v_{-282}, ..., v_0, v_1, ...v_{282}$$

  to

  $$v_{-564}, v_{-562}, ..., v_0, v_2, ..., v_{562}, v_{564}$$

  and add unassigned variables

  $$v_{-565}, v_{-563}, ..., v_1, ..., v_{563}, v_{565}.$$

- **This operation *will not introduce any arithmetic progression* among the even indexed variables.**
- **The assignment to the even indexed variables is the third tunnel.**

Results

- **Any of the tunnels works - improves performance of off-the-shelf solvers (modified) by orders of magnitude**
- **All tunnels give the same result - bound of 1132 on $W(2, 6)$ (compare with 696 and 342)**
- **Probably broke the sound barrier on this particular instance**

Future

- **Want to extend the results to Bounded Model Checking and other circuit problems.**
- **Similarities in formula construction**
- **But generally we want to show BMC formulas are not satisfiable. Maybe we can actually introduce "bugs," then see if they are found. If all bugs are found and no solution is found for the original formula, we have a pretty good feeling that the formula was not satisfiable in the first place.**