Backtracking search: look-back

ICS 275
Winter 2016
Look-back:
Backjumping / Learning

- **Backjumping:**
  - In deadends, go back to the most recent culprit.

- **Learning:**
  - constraint-recording, no-good recording.
  - good-recording

Winter 2016
Backjumping

- (X1=r,x2=b,x3=b,x4=b,x5=g,x6=r,x7={r,b})
- (r,b,b,b,g,r) conflict set of x7
- (r,-,b,b,g,-) c.s. of x7
- (r,-,b,-,-,-,-) minimal conflict-set
- Leaf deadend: (r,b,b,b,g,r)
- Every conflict-set is a no-good
Backjumping

- \((X1=r, x2=b, x3=b, x4=b, x5=g, x6=r, x7={r,b})\)
- \((r,b,b,b,g,r)\) conflict set of \(x7\)
- \((r,-,b,b,g,-)\) c.s. of \(x7\)
- \((r,-,-,-,-,-,-)\) minimal conflict-set
- Leaf deadend: \((r,b,b,b,g,r)\)
- Every conflict-set is a no-good
Flavor of Gaschnig’s jumps only at leaf-dead-ends

Internal dead-ends: dead-ends that are non-leaf

Example 6.3.1 In Figure 6.4, all of the backjumps illustrated lead to internal dead-ends, except for the jump back to \((x_1, \text{green}), (x_2, \text{blue}), (x_3, \text{red}), (x_4, \text{blue})\), because this is the only case where another value exists in the domain of the culprit variable.

Winter 2016
Gaschnig jumps only at leaf-dead-ends
Internal dead-ends: dead-ends that are non-leaf

Example 6.3.1 In Figure 6.4, all of the backjumps illustrated lead to internal dead-ends, except for the jump back to \((x_1, \text{green}), (x_2, \text{blue}), (x_3, \text{red}), (x_4, \text{blue})\), because this is the only case where another value exists in the domain of the culprit variable.

Winter 2016
Backjumping styles

- **Jump at leaf only** (Gaschnig 1977)
  - Context-based

- **Graph-based** (Dechter, 1990)
  - Jumps at leaf and internal dead-ends, graph information

- **Conflict-directed** (Prosser 1993)
  - Context-based, jumps at leaf and internal dead-ends
Conflict Analysis

- Conflict set
- Leaf deadend
- Nogood
- Safe jump
Gaschnig’s backjumping: Culprit variable

Definition 6.2.1 (culprit variable) Let $\vec{a}_i = (a_1, ..., a_i)$ be a leaf dead-end. The culprit index relative to $\vec{a}_i$ is defined by $b = \min\{j \leq i | \vec{a}_j \text{ conflicts with } x_{i+1}\}$. We define the culprit variable of $\vec{a}_i$ to be $x_b$.

- If $a_i$ is a leaf deadend and $x_b$ its culprit variable, then $a_b$ is a safe backjump destination and $a_j$, $j < b$ is not.
- The culprit of $x_7$ $(r,b,b,b,g,r)$ is $(r,b,b) \rightarrow x_3$
Gaschnig’s backjumping implementation [1979]

- Gaschnig uses a marking technique to compute culprit.
- Each variable $x_j$ maintains a pointer ($latest_j$) to the latest ancestor incompatible with any of its values.
- While forward generating $a_i$, keep array $latest_i$, $1 \leq j \leq n$, of pointers to the last value conflicted with some value of $x_j$. The algorithm jumps from a leaf-dead-end $x_{i+1}$ back to $latest_{i+1}$ which is its culprit.
Gaschnig's backjumping

```plaintext
procedure GASCHNIG'S-BACKJUMPING
Input: A constraint network $\mathcal{R} = (X, D, C)$
Output: Either a solution, or a decision that the network is inconsistent.

\[ i \leftarrow 1 \] (initialize variable counter)
\[ D'_i \leftarrow D_i \] (copy domain)
\[ latest_i \leftarrow 0 \] (initialize pointer to culprit)

while \( 1 \leq i \leq n \)
  instantiate \( x_i \leftarrow SELECTVALUE-CBJ \)
  if \( x_i \) is null (no value was returned)
    \[ i \leftarrow latest_i \] (backjump)
  else
    \[ i \leftarrow i + 1 \]
    \[ D'_i \leftarrow D_i \]
    \[ latest_i \leftarrow 0 \]
  end while
if \( i = 0 \)
  return "inconsistent"
else
  return instantiated values of \( \{x_1, \ldots, x_n\} \)
end procedure

procedure SELECTVALUE-CBJ

while \( D'_i \) is not empty
  select an arbitrary element \( a \in D'_i \) and remove \( a \) from \( D'_i \)
  \( \text{consistent} \leftarrow \text{true} \)
  \[ k \leftarrow 1 \]
  while \( k < i \) and \( \text{consistent} \)
    if \( k > latest_i \)
      \[ latest_i \leftarrow k \]
    if not CONSISTENT(\( a \), \( x_i = a \))
      \( \text{consistent} \leftarrow \text{false} \)
    else
      \[ k \leftarrow k + 1 \]
  end while
if \( \text{consistent} \)
  return \( a \)
end while
return null (no consistent value)
end procedure
```

Winter 2016

Figure 6.3: Gaschnig's backjumping algorithm.
Example 6.2.3 Consider the problem in Figure 6.1 and the order $d_1$. At the dead-end for $x_7$ that results from the partial instantiation ($< x_1, \text{red} >, < x_2, \text{blue} >, < x_3, \text{blue} >, < x_4, \text{blue} >, < x_5, \text{green} >, < x_6, \text{red} >$), $\text{latest}_7 = 3$, because $x_7 = \text{red}$ was ruled out by $< x_1, \text{red} >$, $x_7 = \text{blue}$ was ruled out by $< x_3, \text{blue} >$, and no later variable had to be examined. On returning to $x_3$, the algorithm finds no further values to try ($D'_3 = \emptyset$). Since $\text{latest}_3 = 2$, the next variable examined will be $x_2$. Thus we see the algorithm’s ability to backjump at leaf dead-ends. On subsequent dead-ends, as in $x_3$, it goes back to its preceding variable only. An example of the algorithm’s practice of pruning the search space is given in Figure 6.2.
Properties

- Gaschnig’s backjumping implements only safe and maximal backjumps in leaf-deadends.
Gaschnig jumps only at leaf-dead-ends

Internal dead-ends: dead-ends that are non-leaf

**Example 0.3.1** In Figure 0.4, all of the backjumps illustrated lead to internal dead-ends, except for the jump back to $(\langle x_1, green \rangle, \langle x_2, blue \rangle, \langle x_3, red \rangle, \langle x_4, blue \rangle)$, because this is the only case where another value exists in the domain of the culprit variable.
Gaschnig jumps only at leaf-dead-ends

Internal dead-ends: dead-ends that are non-leaf

Example 6.3.1 In Figure 6.4, all of the backjumps illustrated lead to internal dead-ends, except for the jump back to \((x_1, \text{green}), (x_2, \text{blue}), (x_3, \text{red}), (x_4, \text{blue})\), because this is the only case where another value exists in the domain of the culprit variable.

Winter 2016
Graph-based backjumping scenarios

Internal deadend at X4

- Scenario 1, deadend at x4:
- Scenario 2: deadend at x5:
- Scenario 3: deadend at x7:
- Scenario 4: deadend at x6:
Graph-based backjumping

- Uses only graph information to find culprit
- Jumps both at leaf and at internal dead-ends
- Whenever a deadend occurs at x, it jumps to the most recent variable y connected to x in the graph. If y is an internal deadend it jumps back further to the most recent variable connected to x or y.
- The analysis of conflict is approximated by the graph.
- Graph-based algorithm provide graph-theoretic bounds.
Ancestors and parents

- $\text{anc}(x_7) = \{x_5, x_3, x_4, x_1\}$
- $p(x_7) = 5$
- $p(r, b, b, b, g, r) = x_5$

**Definition 6.3.2 (ancestors, parent)** Given a constraint graph and an ordering of the nodes $d$, the ancestor set of variable $x$, denoted $\text{anc}(x)$, is the subset of the variables that precede and are connected to $x$. The parent of $x$, denoted $p(x)$, is the most recent (or latest) variable in $\text{anc}(x)$. If $\tilde{a}_i = (a_1, ..., a_i)$ is a leaf dead-end, we equate $\text{anc}(\tilde{a}_i)$ with $\text{anc}(x_{i+1})$, and $p(\tilde{a}_i)$ with $p(x_{i+1})$.  

Winter 2016
Internal deadends analysis

Definition 6.3.5 (session) We say that backtracking visits $x_i$ if it processes $x_i$ coming from a variable earlier in the ordering. The session of $x_i$ starts upon the invising of $x_i$ and ends when retracting to a variable that precedes $x_i$. At a given state of the search where variable $x_i$ is already instantiated, the current session of $x_i$ is the set of variables processed by the algorithm since the most recent invist to $x_i$. The current session of $x_i$ includes $x_i$ and therefore the session of a leaf dead-end variable has a single variable.

Definition 6.3.6 (relevant dead-ends) The relevant dead-ends of $x_i$'s session are defined recursively as follows. The relevant dead-ends of a leaf dead-end $x_i$, denoted $r(x_i)$, is $x_i$. If $x_i$ is variable to which the algorithm retracted from $x_i$, then the relevant-dead-end set of $x_i$, denoted $r_{i}(x_i)$, is $x_i$. Otherwise, the relevant-dead-end set of $x_i$ is the union of all relevant-ends of variables that precede $x_i$. Formally, $r_{i}(x_i) = \bigcup_{i'=1}^{i-1} P_{i'}(Y)$. The induced parent of $x_i$ relative to $Y$, $P_{i}(Y)$, is the latest variable in $I_{i}(Y)$. We call $P_{i}(Y)$ the graph-based culprit of $x_i$.
Graph-based backjumping scenarios

Internal deadend at X4

- Scenario 1, deadend at x4:
- Scenario 2: deadend at x5:
- Scenario 3: deadend at x7:
- Scenario 4: deadend at x6:

Figure 6.1: A modified coloring problem.
Graph-based backjumping algorithm, but we need to jump at internal deadends too.

**procedure** GRAPH-BASED-BACKJUMPING

**Input:** A constraint network $\mathcal{R} = (X, D, C)$

**Output:** Either a solution, or a decision that the network is inconsistent.

compute $\text{anc}(x_i)$ for each $x_i$ (see Definition 6.3.2 in text)

$i \leftarrow 1$ (initialize variable counter)

$D'_i \leftarrow D_i$ (copy domain)

$I_i \leftarrow \text{anc}(x_i)$ (copy of anc() that can change)

while $1 \leq i \leq n$

instantiate $x_i \leftarrow \text{SELECTVALUE}$

if $x_i$ is null (no value was returned)

$qpres \leftarrow i$

$i \leftarrow \text{latest index in } I_i$ (backjump)

$I_i \leftarrow I_i \cup \{qpres \} - \{x_i\}$

else

$i \leftarrow i + 1$

$D'_i \leftarrow D_i$

$I_i \leftarrow \text{anc}(x_i)$

end while

if $i = 0$

return “inconsistent”

else

return instantiated values of $\{x_1, \ldots, x_n\}$

end procedure

**procedure** SELECTVALUE (same as BACKTRACKING’s)

while $D'_i$ is not empty

select an arbitrary element $a \in D'_i$, and remove $a$ from $D'_i$

if CONSISTENT($\overline{a}_{i-1}, x_i = a$)

return $a$

end while

return null (no consistent value)

end procedure

When not all variables $X_i$ are relevant deadends?

In the session above, read example 6.6

Figure 6.5: The graph-based backjumping algorithm.
Properties of graph-based backjumping

- Algorithm graph-based backjumping jumps back at any deadend variable as far as graph-based information allows.
- For each variable, the algorithm maintains the induced-ancestor set $l_i$ relative the relevant dead-ends in its current session.
- The size of the induced ancestor set is at most $w^*(d)$. 
Conflict-directed backjumping
(Prosser 1990)

- Extend Gaschnig’s backjump to internal dead-ends.
- Exploits information gathered during search.
- For each variable the algorithm maintains an induced jumpback set, and jumps to most recent one.
- Use the following concepts:
  - An ordering over variables induced a strict ordering between constraints: \( R_1 < R_2 < \ldots R_t \)
  - Use earliest minimal conflict-set (emc\((x_{i+1})\)) of a deadend.
  - Define the jumpback set of a deadend
Example 6.4.5 Consider the problem of Figure 6.1 using ordering $d_1 = (x_1, \ldots, x_7)$. Given the dead-end at $x_7$ and the assignment $\vec{a}_6 = (\text{blue, green, red, red, blue, red})$, the emc set is $(<x_1, \text{blue}>, <x_3, \text{red}>)$, since it accounts for eliminating all the values of $x_7$. Therefore, algorithm conflict-directed backjumping jumps to $x_3$. Since $x_3$ is an internal dead-end whose own var – emc set is $\{x_1\}$, the jumpback set of $x_3$ includes just $x_1$, and the algorithm jumps again, this time back to $x_1$. \[\square\]
Properties

- Given a dead-end $\vec{a}_i$, the latest variable in its jumpback set $J_i$ is the earliest variable to which it is safe to jump.
- This is the culprit.
- Algorithm conflict-directed backtracking jumps back to the latest variable in the dead-ends’s jumpback set, and is therefore safe and maximal.
Conflict-directed backjumping

procedure CONFLICT-DIRECTED-BACKJUMPING
Input: A constraint network $\mathcal{R} = (X, D, C)$.
Output: Either a solution, or a decision that the network is inconsistent.

\begin{verbatim}
i ← 1 (initialize variable counter)
\end{verbatim}

\begin{verbatim}
D'_i ← D_i (copy domain)
J_i ← ∅ (initialize conflict set)
\end{verbatim}

\begin{verbatim}
while $1 \leq i \leq n$
\begin{verbatim}
    instantiate $x_i ← SELECTVALUE-CBJ$
    if $x_i$ is null (no value was returned)
        ipree ← i
        $i ←$ index of last variable in $J_i$ (backjump)
        $J_i ← J_i \cup J_{ipree} - \{x_i\}$ (merge conflict sets)
    else
        $i ← i + 1$ (step forward)
        $D'_i ← D_i$ (reset mutable domain)
        $J_i ← ∅$ (reset conflict set)
    end while
\end{verbatim}

if $i = 0$
\begin{verbatim}
    return “inconsistent”
else
\end{verbatim}
\begin{verbatim}
    return instantiated values of $\{x_1, \ldots, x_n\}$
end procedure
\end{verbatim}

subprocedure SELECTVALUE-CBJ

\begin{verbatim}
while $D'_i$ is not empty
    select an arbitrary element $a ∈ D'_i$, and remove $a$ from $D'_i$
    \begin{verbatim}
        consistent ← true
    \end{verbatim}
    \begin{verbatim}
        $k ← 1$
    \end{verbatim}
    \begin{verbatim}
        while $k < i$ and consistent
            if CONSISTENT($\bar{a}_k, x_i = a$)
                \begin{verbatim}
                    $k ← k + 1$
                \end{verbatim}
            else
                let $R_S$ be the earliest constraint causing the conflict
                add the variables in $R_S$’s scope $S$, but not $x_i$, to $J_i$
                \begin{verbatim}
                    consistent ← false
                \end{verbatim}
            end while
        end if
    \end{verbatim}
\end{verbatim}

end while
\begin{verbatim}
if consistent
\end{verbatim}
\begin{verbatim}
    return $a$
end while
\end{verbatim}

return null (no consistent value)
\end{verbatim}

Winter 2016

Figure 6.7: The conflict-directed backjumping algorithm.
Example 6.5.1 Consider, once again, the CSP in Figure 6.1. A DFS ordering \( d_2 = (x_1, x_7, x_4, x_5, x_6, x_2, x_3) \) and its corresponding DFS spanning tree are given in Figure 6.6c,d. If a dead-end occurs at node \( x_3 \), the algorithm retreats to its DFS parent, which is \( x_7 \).
Graph-based Backjumping on DFS ordering

- Example: \( d = x_1, x_2, x_3, x_4, x_5, x_6, x_7 \)
- Constraints: \((6,7)(5,2)(2,3)(5,7)(2,7)(2,1)(2,3)(1,4)3,4\)
- Rule: go back to parent. No need to maintain parent set

![Graphs](Image)

Figure 6.6: Several ordered constraint graphs of the problem in Figure 6.1: (a) along ordering \( d_1 = (x_1, x_3, x_2, x_4, x_5, x_6, x_7) \), (b) the induced graph along \( d_1 \), (c) along ordering \( d_2 = (x_1, x_7, x_4, x_5, x_6, x_3) \), and (d) a DFS spanning tree along ordering \( d_2 \).

**Theorem 6.5.2** Given a DFS ordering of the constraint graph, if \( f(x) \) denotes the DFS parent of \( x \), then, upon a dead-end at \( x \), \( f(x) \) is \( x \)'s graph-based earliest safe variable for both leaf and internal dead-ends.
Complexity of Graph-based Backjumping

- \( T_i = \) number of nodes in the AND/OR search space rooted at \( x_i \) (level \( m-i \))
- Each assignment of a value to \( x_i \) generates subproblems:
  - \( T_i = k b T_{i-1} \)
  - \( T_0 = k \)
- Solution: \( T_m = b^m k^{m+1} \)

**Theorem 6.5.3** When graph-based backjumping is performed on a DFS ordering of the constraint graph, the number of nodes visited is bounded by \( O((b^m k^{m+1})) \), where \( b \) bounds the branching degree of the DFS tree associated with that ordering, \( m \) is its depth and \( k \) is the domain size. The time complexity (measured by the number of consistency checks) is \( O(e k (bk)^m) \), where \( e \) is the number of constraints.
DFS of graph and induced graphs

Spanning-tree of a graph;
DFS spanning trees, Pseudo-tree
Pseudo-tree is a spanning tree that does not allow arcs across branches.
Complexity of backjumping uses pseudo-tree analysis

Simple: always jump back to parent in pseudo tree
Complexity for csp: $\exp(\text{tree-depth})$
Complexity for csp: $\exp(w \log n)$
Complexity of backjumping

Theorem 6.5.5 If \( d \) is a DFS ordering of \( (G^*, d_1) \) for some ordering \( d_1 \), having depth \( m_d^* \), then the complexity of graph-based backjumping using ordering \( d \) is \( O(\exp(m_d^*)) \).
Look-back: No-good Learning

Learning means recording conflict sets used as constraints to prune future search space.

- \((x_1=2, x_2=2, x_3=1, x_4=2)\) is a dead-end

- Conflicts to record:
  - \((x_1=2, x_2=2, x_3=1, x_4=2)\) 4-ary
  - \((x_3=1, x_4=2)\) binary
  - \((x_4=2)\) unary
Learning, constraint recording

- Learning means recording conflict sets
- An opportunity to learn is when deadend is discovered.
- Goal of learning is to not discover the same deadends.
- Try to identify small conflict sets
- Learning prunes the search space.
Nogoods explain deadends

Learning means recording explanations to conflicts
They are implied constraints

- Conflicts to record are explanations
  - \((x_1=2,x_2=2,x_3=1,x_4=2)\) 4-ary
  - \((x_1=2,x_2=2,x_3=1,x_4=2)\) \(\Rightarrow\) \((x_5 \neq 1)\) and
  - \((x_3=1,x_4=2)\) \(\Rightarrow\) \((x_5 \neq 1)\)
  - \((x_4=2)\) \(\Rightarrow\) \((x_5 \neq 1)\)
Learning example

Figure 6.9: The search space explicated by backtracking on the CSP from Figure 6.1, using the variable ordering \((x_6, x_3, x_4, x_2, x_7, x_1, x_5)\) and the value ordering \((blue, red, green, teal)\). Part (a) shows the ordered constraint graph, part (b) illustrates the search space. The cut lines in (b) indicate branches not explored when graph-based learning is used.
Learning issues

- Learning styles
  - Graph-based or context-based
  - i-bounded, scope-bounded
  - Relevance-based
- Non-systematic randomized learning
- Implies time and space overhead
- Applicable to SAT
Graph-based learning algorithm

procedure GRAPH-BASED-BACKJUMP-LEARNING

    instantiate \( x_i \leftarrow \text{SELECTVALUE} \)
    if \( x_i \) is null (no value was returned)
        record a constraint prohibiting \( \tilde{a}_{i-1}[I_i] \).
    \( iprev \leftarrow i \)
    (algorithm continues as in Fig. 6.5)

Figure 6.10: Graph-based backjumping learning, modifying CBJ
Deep learning

- Deep learning: recording all and only minimal conflict sets
- Example:
- Although most accurate, overhead can be prohibitive: the number of conflict sets in the worst-case:

\[
\binom{r}{r/2} = 2^r
\]
Jumpback learning

- Record the jumpback assignment

Example: For the problem and ordering of Example at the first dead-end, jumpback learning will record the no-good \((x_2 = \text{green}, x_3 = \text{blue}, x_7 = \text{red})\), since that tuple includes the variables in the jumpback set of \(x_1\).

\[
\text{procedure} \text{ CONFLICT-DIRECTED-BACKJUMP-LEARNING}
\]

\[
\text{instantiate } x_i \leftarrow \text{SELECTVALUE-CBJ}
\]

\[
\text{if } x_i \text{ is null} \quad \text{(no value was returned)}
\]

\[
\text{record a constraint prohibiting } \vec{a}_{i-1}[J_i] \text{ and corresponding values}
\]

\[
i_{prev} \leftarrow i
\]

\[
\text{(algorithm continues as in Fig. 6.7)}
\]
Bounded and relevance-based learning

Bounding the arity of constraints recorded.

- When bound is \( i \): \( i \)-ordered graph-based, \( i \)-order jumpback or \( i \)-order deep learning.
- Overhead complexity of \( i \)-bounded learning is time and space exponential in \( i \).

Definition 6.7.3 (i-relevant) A no-good is \( i \)-relevant if it differs from the current partial assignment by at most \( i \) variable-value pairs.

Definition 6.7.4 (\( i \)'th order relevance-bounded learning) An \( i \)'th order relevance-bounded learning scheme maintains only those learned no-goods that are \( i \)-relevant.
Complexity of backtrack-learning for CSP

- The complexity of learning along d is time and space exponential in $w^*(d)$:

  The number of dead-ends is bounded by $O(nk^{w^*(d)})$
  Number of constraint tests per dead-end are $O(e)$

  Space complexity is $O(nk^{w^*(d)})$
  Time complexity is $O(n^2 \cdot k^{w^*(d)+1})$

m- depth of tree, e- number of constraints
Good caching:
Moving from one to all or counting

Winter 2016
Summary:
Time-space for constraint processing (consistency and counting)

- **Constraint-satisfaction**
  - Search with backjumping
    - Space: linear, Time: $O(nk\log nw^*)$
  - Search with learning no-goods
    - time and space: $O(nk^{w*})$
  - Variable-elimination
    - time and space: $O(nk^{w*})$

- **Counting, enumeration**
  - Search with backjumping
    - Space: linear, Time: $O(k^n)$
  - Search with no-goods caching only
    - space: $O(\exp(w^*))$  Time: $O(\exp(n))$
  - Search with goods and no-goods learning
    - Time and space: $O(\exp(path-width), O(\exp(\log n w^*)))$
  - Variable-elimination
    - Time and space: $O(\exp(w^*))$
Non-Systematic Randomized Learning

- Do search in a random way with interrupts, restarts, unsafe backjumping, **but record conflicts**.
- Guaranteed completeness.
Look-back for SAT

A partial assignment is a set of literals: $\sigma$

A jumpback set if a J-clause:

Upon a leaf deadend of $x$ resolve two clauses, one enforcing $x$ and one enforcing $\neg x$ relative to the current assignment

A clause forces $x$ relative to assignment $\sigma$ if all the literals in the clause are negated in $\sigma$.

Resolving the two clauses we get a nogood.

If we identify the earliest two clauses we will find the earliest conflict.

The argument can be extended to internal deadends.
Integration of algorithms

procedure FC-CBJ
Input: A constraint network $\mathcal{R} = (X, D, C)$.
Output: Either a solution, or a decision that the network is inconsistent.

\begin{align*}
i &\leftarrow 1 \quad \text{(initialize variable counter)} \\
\text{call } &\text{SELECT VARIABLE} \quad \text{(determine first variable)} \\
D'_i &\leftarrow D_i \text{ for } 1 \leq i \leq n \quad \text{(copy all domains)} \\
J_i &\leftarrow \emptyset \quad \text{(initialize conflict set)} \\
\text{while } &1 \leq i \leq n \\
\text{instansiate } &x_i \leftarrow \text{SELECTVALUE-FC-CBJ} \\
\text{if } &x_i \text{ is null} \quad \text{(no value was returned)} \\
\quad &\text{iprev } \leftarrow i \\
\quad &i \leftarrow \text{latest index in } J_i \quad \text{(backjump)} \\
\quad &J_i \leftarrow J_i \cup J_{\text{iprev}} - \{x_i\} \\
\quad &\text{reset each } D'_k, k > i, \text{ to its value before } x_i \text{ was last instantiated} \\
\quad \text{else} \quad \text{(step forward)} \\
\quad &i \leftarrow i + 1 \\
\quad \text{call } &\text{SELECT VARIABLE} \quad \text{(determine next variable)} \\
\quad D'_i &\leftarrow D_i \\
\quad J_i &\leftarrow \emptyset \\
\text{end while} \\
\text{if } &i = 0 \\
\quad &\text{return } \text{ "inconsistent"} \\
\text{else} \\
\quad &\text{return instantiated values of } \{x_1, \ldots, x_n\} \\
\text{end procedure}
\end{align*}

Winter 2016

Figure 6.13: The main procedure of the FC-CBJ algorithm.
subprocedure SELECTVALUE-FC-CBJ

while $D'_i$ is not empty
    select an arbitrary element $a \in D'_i$, and remove $a$ from $D'_i$
    \[ empty-domain \leftarrow false \]
    for all $k$, $i < k \leq n$
        for all values $b$ in $D'_k$
            if not CONSISTENT($\overline{a}_{i-1}, x_i=a, x_k=b$)
                let $R_S$ be the earliest constraint causing the conflict
                add the variables in $R_S$'s scope $S$, but not $x_k$, to $J_k$
                remove $b$ from $D'_k$
            endfor
        endif
        if $D'_k$ is empty \( (x_i=a \text{ leads to a dead-end}) \)
            \[ empty-domain \leftarrow true \]
        endif
    endfor
    if $empty-domain$ \( (\text{don't select } a) \)
        reset each $D'_k$ and $J_k$, $i < k \leq n$, to status before $a$ was selected
    else
        return $a$
    end while
return null \( (\text{no consistent value}) \)
end subprocedure

Figure 6.14: The SelectValue subprocedure for FC-CBJ.
Winter 2016
Relationships between various backtracking algorithms
Empirical comparison of algorithms

- Benchmark instances
- Random problems
- Application-based random problems
- Generating fixed length random k-sat \((n,m)\) uniformly at random
- Generating fixed length random CSPs
- \((N,K,T,C)\) also arity, \(r\).
The Phase transition (m/n)
Some empirical evaluation

- Sets 1-3 reports average over 2000 instances of random csps from 50% hardness. Set 1: 200 variables, set 2: 300, Set 3: 350. All had 3 values:
- Dimacs problems

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Set 1</th>
<th>Set 2</th>
<th>Set 3</th>
<th>ssa 038</th>
<th>ssa 158</th>
</tr>
</thead>
<tbody>
<tr>
<td>FC</td>
<td>207</td>
<td>68.5</td>
<td>-</td>
<td>46</td>
<td>14.5</td>
</tr>
<tr>
<td>FC+AC</td>
<td>40</td>
<td>55.4</td>
<td>1</td>
<td>0.6</td>
<td>1</td>
</tr>
<tr>
<td>FCr-CBJ</td>
<td>189</td>
<td>69.2</td>
<td>222</td>
<td>119.3</td>
<td>182</td>
</tr>
<tr>
<td>FC-CBJ+LVO</td>
<td>167</td>
<td>73.8</td>
<td>132</td>
<td>86.8</td>
<td>119</td>
</tr>
<tr>
<td>FC-CBJ+LRN</td>
<td>186</td>
<td>63.4</td>
<td>32</td>
<td>15.6</td>
<td>1</td>
</tr>
<tr>
<td>FC-CBJ+LRN+LVO</td>
<td>160</td>
<td>74.0</td>
<td>26</td>
<td>14.0</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 6.16: Empirical comparison of six selected CSP algorithms. See text for explanation. In each column of numbers, the first number indicates the number of nodes in the search tree, rounded to the nearest thousand, and final 000 omitted; the second number is CPU seconds.
All Solutions and Counting

- For all solutions and counting we will see
  - The additional impact of Good learning
  - BFS makes sense with good learning
  - BFS and DFS time and space $\exp(\text{path-width})$
  - Good-learning doesn’t help consistency task
#CSP – OR Search Tree
### #CSP – OR Search Tree

**Diagram:**

- Nodes: C, A, F, B, E, D
- Edges: C→A, A→F, B→E, D→B, E→D

**Tables:**

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>R&lt;sub&gt;ABC&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>R&lt;sub&gt;BCD&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>R&lt;sub&gt;ABE&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

|   | A | E | F | R<sub>AEF</sub> |
|---|---|---|-----------------|
| 0 | 0 | 0 | 0 | 0               |
| 1 | 0 | 1 | 1 | 0               |
| 1 | 0 | 1 | 1 | 0               |
| 1 | 0 | 1 | 1 | 0               |
| 1 | 1 | 0 | 1 | 0               |
| 1 | 1 | 0 | 1 | 0               |
| 1 | 1 | 0 | 1 | 0               |
| 1 | 1 | 0 | 1 | 0               |
#CSP - OR Search Tree

14 solutions
#CSP - Tree DFS Traversal

Value of node = number of solutions below it
#CSP - OR Search Tree
#CSP - Searching the Graph by Good Caching

A  context(A) = [A]
B  context(B) = [AB]
C  context(C) = [ABC]
D  context(D) = [ABD]
E  context(E) = [AE]
F  context(F) = [F]
#CSP - Searching the Graph by Good Caching

A context(A) = [A]

B context(B) = [AB]

C context(C) = [ABC]

D context(D) = [ABD]

E context(E) = [AE]

F context(F) = [F]
#CSP - Searching the Graph by Good Caching

(A=0, E=0) is good
V(A=0, E=0) = 1
#CSP - Searching the Graph by Good Caching

No caching: $O(\exp(n))$

Good-caching: $O(\exp(pw))$
Summary: Time-Space for Constraint Processing

- Constraint-satisfaction, one solution
  - Naive backtracking
    - Space: $O(n)$,
    - Time: $O(\exp(n))$
  - Backjumping
    - Space: $O(n)$,
    - Time: $O(\exp(\log n \ w^*))$
  - Learning no-goods
    - Space: $O(\exp(w^*))$
    - Time: $O(\exp(w^*))$
  - Variable-elimination
    - Space: $O(\exp(w^*))$
    - Time: $O(\exp(w^*))$

- Counting, enumeration
  - Backtracking, backjumping
    - Space: $O(n)$,
    - Time: $O(\exp(n))$
  - Learning no-goods
    - Space: $O(\exp(w^*))$
    - Time: $O(\exp(n))$
  - Search with goods and no-goods learning
    - Space: $O(\exp(pw^*))$
    - Time: $O(\exp(pw^*))$, both, $O(\exp(w^*\log n))$
  - Variable-elimination
    - Space: $O(\exp(w^*))$
    - Time: $O(\exp(w^*))$
  - BFS is time and space $O(\exp(pw^*))$
Summary: search principles

- DFS is better than BFS search
- Constraint propagation (i.e., bounded inference) prunes search space
- Constraint propagation yields good advise for how to branch and where to go
- Backjumping and no-good learning helps prune search space and revise problem.
- Good learning revise problem but helps only counting, enumeration