

Today's Class

• What's a Constraint Satisfaction Problem (CSP)?

· A.K.A., Constraint Processing / CSP paradigm

• How do we solve them?

· Algorithms for CSPs

· Search Terminology

Constraint (n): A relation ... between the values of one or more mathematical variables (e.g., x>3 is a constraint on x).

**Constraint satisfaction** assigns values to variables so that all constraints are true.

- http://foldoc.org/constraint

1

Constraint Satisfaction

• Con-straint /kən strānt/, (noun):

- Something that limits or restricts someone or something.<sup>1</sup>
- \* A relation ... between the values of one or more mathematical variables (e.g., x>3 is a constraint on x), that...
- Assigns values to variables so that all constraints are true.<sup>2</sup>
- In search, constraints exist on?
- General Idea
  - View a problem as a set of variables
  - To which we have to assign values
  - That satisfy a number of (problem-specific) constraints

[1] Merriam-Webster onlin The Free Online Computing Dictionar

3

Overview

Constraint satisfaction: a problem-solving paradigm

 Constraint programming, constraint satisfaction problems (CSPs), constraint logic programming...

- Algorithms for CSPs
- · Backtracking (systematic search)
- Constraint propagation (k-consistency)
- Variable and value ordering heuristics
- · Backjumping and dependency-directed backtracking

4

#### Search Vocabulary

- We've talked about caring about goals (end states) vs. paths
- · These correspond to...
  - · Planning: finding sequences of actions
  - · Paths have various costs, depths
  - · Heuristics to guide, frontier to keep backup possibilities
  - Examples: chess moves; 8-puzzle; homework 2
  - Identification: assignments to variables representing unknowns
  - The goal itself is important, not the path
  - $\bullet$  Examples: Sudoku; map coloring; N queens, scheduling, planning
- CSPs are specialized for identification problems

Slightly Less Informal Definition of CSP

**CSP** = Constraint Satisfaction Problem

- Given:
  - . A finite set of variables
  - Each with a **domain** of possible values they can take (often finite)
  - 3. A set of **constraints** that limit the values the variables can take on
- **Solution**: an assignment of values to variables that satisfies all constraints.



5

#### **CSP** Applications

- · Decide if a solution exists
- · Find some solution
- · Find all solutions
- Find the "best solution"
  - · According to some metric (objective function)
  - Does that mean "optimal"?



• Given a 2D map, it is always possible to color it using three colors (we'll use red, green, blue)

Such that:

- No two adjacent regions are the same color
- Start thinking: What are the values, variables, constraints?



#### Slightly Less Informal

- Constraint satisfaction problems (CSPs): a special subset of search problems
- **State** is defined by variables *Xi* with values from a domain D
- · D may be finite
- Sometimes D depends on i
- **Goal test** is a set of constraints specifying allowable combinations of values for variables



9

### Example: N-Queens (1)

- Formulation 1:
  - Variables:  $X_{ij}$
  - Domains: {0, 1}
  - · Constraints:

 $\forall i, j, k \ (X_{ij}, X_{ik}) \in \{(0, 0), (0, 1), (1, 0)\}$ 

 $\forall i, j, k \ (X_{ij}, X_{kj}) \in \{(0,0), (0,1), (1,0)\}$ 

 $\forall i, j, k \ (X_{ij}, X_{i+k,j+k}) \in \{(0,0), (0,1), (1,0)\}$ 

 $\forall i, j, k \ (X_{ij}, X_{i+k,j-k}) \in \{(0,0), (0,1), (1,0)\}$ 

Example: SATisfiability

Special case!

 $\sum X_{ij} = N$ 

10

#### Example: N-Queens (2)

- Formulation 2:
  - Variables:  $Q_k$
  - Domains:  $\{1, 2, 3, \dots N\}$
  - Actually, tuples of {(1-N, 1-N)}

  - Constraints: Implicit:  $\forall i, j$  non-threatening $(Q_i, Q_j)$

Explicit:  $(Q_1, Q_2) \in \{(1, 3), (1, 4), \ldots\}$ 

- Given a set of propositions containing variables, find an assignment of the variables to  $\{folse, frue\}$  that satisfies them.
- For example, the clauses:
- $^{\circ}$  (A  $\vee$  B  $\vee$   $\neg$ C)  $\wedge$  (  $\neg$ A  $\vee$  D)
- (equivalent to  $(C \rightarrow A) \lor (B \land D \rightarrow A)$ )

#### are satisfied by

- A = false
- B = true C = false
- D = false

11 12

 $Q_2$ 

Q<sub>3</sub> ▮ **±** 

# Real-World Problems - Scheduling - Temporal reasoning - Building design - Planning - Optimization/satisfaction - Vision - Graph layout - Network management - Natural language processing - Molecular biology / genomics - VLSI design

13

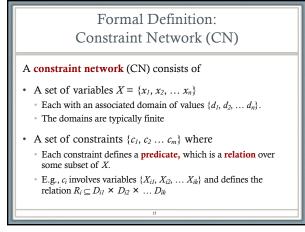
Exercise: Map Coloring II

Variables:
Domains:
Constraints:
One solution:

E
D
A
B
C
B
C

14

16



15

Formal Definition of a CN (cont.)
An instantiation is an assignment of a value d<sub>x</sub> ∈ D to some subset of variables S.
Any assignment of values to variables
Ex: Q<sub>2</sub> = {2,3} ∧ Q<sub>3</sub> = {1,1} instantiates Q<sub>2</sub> and Q<sub>3</sub>
An instantiation is legal iff it does not violate any constraints
A solution is an instantiation of all variables
A correct solution is a legal instantiation of all variables

Typical Tasks for CSP

• Solutions:
• Does a solution exist?
• Find one solution
• Find all solutions
• Given a partial instantiation, can we do these?
• Transform the CN into an equivalent CN that is easier to solve

17 18

# Binary CSP: all constraints are binary or unary Can convert a non-binary CSP → binary CSP by: Introducing additional variables

- One variable per constraint
  One binary constraint for each pair of original constraints that share variables
- "Dual graph construction"

Binary CSPs: Why?

• Can always represent a binary CSP as a constraint graph with:

• A node for each variable

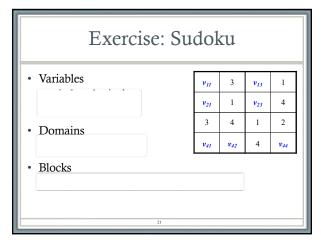
• An arc between two nodes iff there is a constraint on the two variables

• Unary constraint appears as a self-referential arc

"C can't be green"

20

19



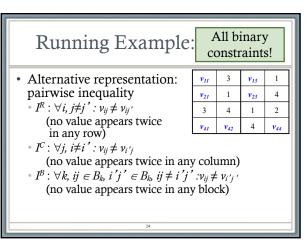
Exercise: Sudoku

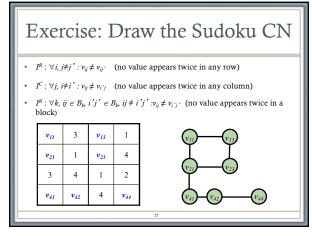
• Variables
•  $v_{i,j}$  is the value in the  $j^{th}$  cell of the  $i^{th}$  row

• Domains
•  $D_{i,j} = D = \{1, 2, 3, 4\}$ • Blocks
•  $B_1 = \{11, 12, 21, 22\}, ..., B_4 = \{33, 34, 43, 44\}$ 

21 22

Running Example: Sudoku • Constraints (implicit or intensional) 4 1  $v_{21}$  $v_{23}$ •  $C^R$ :  $\forall i, \cup_i v_{ii} = D$ 4 1 2 (every value appears 4 v<sub>44</sub>  $v_{42}$ in every row) •  $C^C$ :  $\forall j, \cup_i v_{ij} = D$ (every value appears in every column) •  $C^B$ :  $\forall k, \cup (v_{ii} \mid ij \in B_k) = D$ (every value appears in every block)





Solving Constraint Problems

- 1. Systematic search
  - Generate and test
  - Backtracking
- 2. Constraint propagation (consistency)
- 3. Variable ordering heuristics
- 4. Value ordering heuristics
- 5. Backjumping and dependency-directed backtracking

25 26

#### Generate and Test: Sudoku

• Try every possible assignment of domain elements to variables until you find one that works:

		•		 				_	_	_		
1	3	1	1	1	3	1	1	1	3	1	1	
1	1	1	4	1	1	1	4	1	1	1	4	
3	4	1	2	3	4	1	2	3	4	1	2	
1	1	4	1	1	1	4	2	1	1	4	3	

- Doesn't check constraints until all variables have been instantiated
- Very inefficient way to explore the space of possibilities (4<sup>7</sup> for this trivial Sudoku puzzle, mostly illegal)

Systematic Search: Backtracking

(a.k.a. depth-first search!)

- · Consider the variables in some order
  - 1. Pick an unassigned variable
  - 2. Give it a provisional value
  - 3. That is consistent with all of the constraints
- If no such assignment can be made, we've reached a dead end and need to backtrack to the previous variable
- · Continue this process until:
  - · A solution is found, or
  - We backtrack to the initial variable and have exhausted all possible values

27

28

#### Problems with Backtracking

- Thrashing: keep repeating same failed variable assignments
  - v<sub>21</sub>
     1
     v<sub>23</sub>
     4

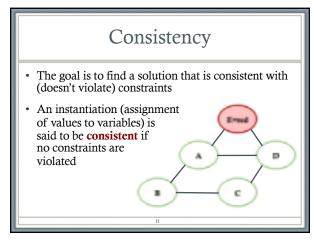
     3
     4
     1
     2
  - Consistency checking can help
  - Intelligent backtracking schemes can also help
- Inefficiency: can spend time exploring areas of search space that aren't likely to succeed
  - Variable ordering can help
  - IF there's a meaningful way to order them

Consistency

The goal is to find a solution that is consistent with (doesn't violate) constraints

An instantiation (assignment of values to variables) is said to be consistent if no constraints are violated

30 31

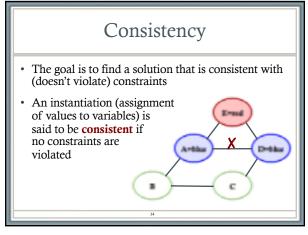


Consistency

• The goal is to find a solution that is consistent with (doesn't violate) constraints

• An instantiation (assignment of values to variables) is said to be consistent if no constraints are violated

32 33



Once the whole graph is consistent, we have a solution (a legal instantiation of values to all variables)
There are multiple kinds of consistency
Different kinds give us different guarantees for performance and correctness

35

34

Node Consistency:

• Node consistency: every value in node X's domain (every value we think it might take) is consistent with X's unary constraints

• A graph is node-consistent if all nodes are node-consistent

• Let's say C can't be green

•  $C = \{red, green, blue\}$ this domain of C makes this node-consistent

Arc Consistency:

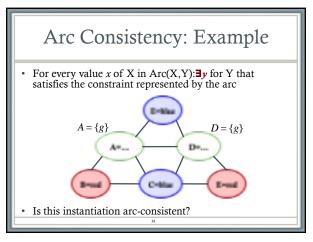
• Arc consistency:

• For every value x of X in Arc(X,Y):

•  $\exists y$  for Y• That satisfies the constraint represented by the arc

• A graph is arc-consistent if all arcs are arc-consistent  $A = \{g, b\}$   $D = \{g, b\}$ 

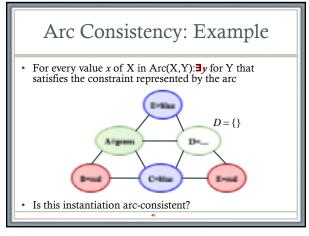
36 37



Arc Consistency: Example

• For every value x of X in Arc(X,Y):  $\exists y$  for Y that satisfies the constraint represented by the arc  $A = \{g\}$ • Is this instantiation arc-consistent? So far, yes!

38 39



Constraint Propagation

• How do we find a set of consistent assignments?

• We perform constraint propagation

• Iteratively reduce the domain of each variable

• Constraints reduce # of legal values for a variable

• Which may then reduce legal values of another variable

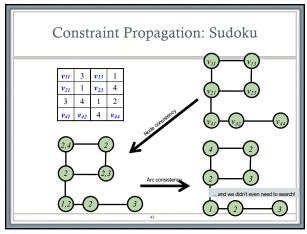
• Key idea: local consistency

• Enforce nearby constraints

• Propagate

40

41



42

Example: Map-Coloring
Variables: WA, NT, Q, NSW, V, SA, T
Domain: D = {red, green, blue}
Constraints: adjacent regions must have different colors
Ex: WA ≠ NT
(WA, NT) ∈ {(red, green), (red, blue), (green, blue), (green, red, (blue, red)}
Solutions are assignments satisfying all constraints, e.g.: {WA = red, NT = green, Q = red, NSW = green, V = red, SA = blue, T = green}

48



Constraint Graphs

· Binary CSP: each constraint relates (at most) two variables

Binary constraint graph: nodes are variables, arcs show constraints

The state of the s

 General-purpose CSP algorithms use the graph structure to speed up search. E.g., Tasmania is an independent subproblem!

50

Standard Search Formulation

- Standard search formulation of CSPs (incremental)
- Let's start with a straightforward, dumb approach, then fix it
- States are defined by the values assigned so far (ex: WA=red, T=red is a state)
  - Initial state: the empty assignment, {}
  - Successor function: assign a value to an unassigned variable
- Goal test: the current assignment is complete and satisfies all constraints

Search Methods

• What does BFS do?

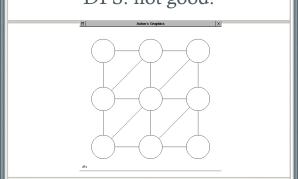
WA SA NSW

· What does DFS do?

51

52

DFS: not good!



**Backtracking Search** 

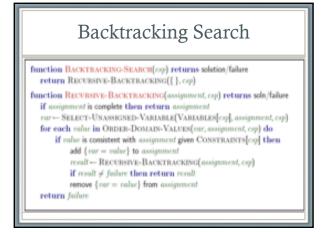
- DFS baaaaad. So how do we improve it?
- Idea 1: Only consider a single variable at each point
  - Variable assignments are commutative, so fix the ordering
  - Ex: [WA = red then NT = green] same as [NT = green then WA = red]
  - Only need to consider assignments to a single variable at each step
  - How many leaves are there now?
- · Idea 2: Only allow fully legal assignments at each point
  - · Consider only values which do not conflict with existing assignments
  - Might have to do some computation to figure out whether a value is ok
  - · "Incremental goal test"

53 54

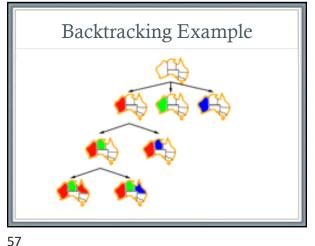
#### **Backtracking Search**

- Idea 1: Only consider a single variable at each point
- Idea 2: Only allow legal assignments at each point
- DFS for CSPs with these two improvements is called backtracking search
  - We backtrack when there's no legal assignment for the next variable
- Backtracking search is the basic uninformed algorithm for CSPs
- Can solve *n*-queens for  $n \approx 25$

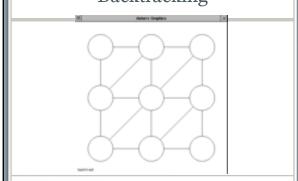
55



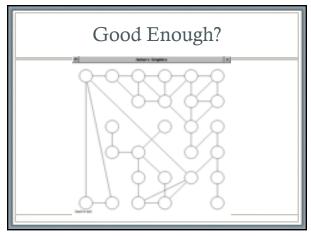
56



Backtracking

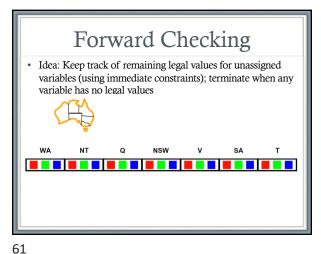


58



Improving Backtracking

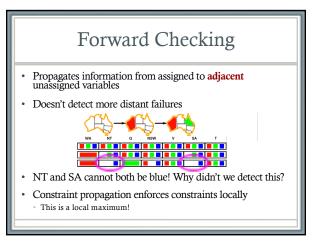
- · General-purpose ideas give huge gains in speed
- 1. Ordering (queueing function ++)
  - · Which variable should be assigned next?
  - In what order should its values be tried?
- 2. Filtering: Can we detect inevitable failure early?
- 3. Structure: Can we exploit the problem structure?



Forward Checking Idea: Keep track of remaining legal values for unassigned variables (using immediate constraints); terminate when any variable has no legal values

62

64



63

**Arc Consistency** • Simplest form of propagation makes each arc consistent  $X \rightarrow Y$  is consistent iff for every value x there is some allowed y · If X loses a value, neighbors of X need to be rechecked! · Arc consistency detects failure earlier than forward checking · What's the downside of arc consistency? Can be run as a preprocessor or after each assignment

K-consistency K-consistency generalizes the notion of arc consistency to sets of more than two variables A graph is **K-consistent** if, for legal values of any K-1 variables in the graph, and for any  $K^{\rm th}$  variable  $V_k,$  there is a legal value for  $V_k$ **Strong** K-consistency = J-consistency for all  $J \le K$ • Node consistency = strong 1-consistency Arc consistency = strong 2-consistency Path consistency = strong 3-consistency

65 66

#### Why Do We Care?

- A strongly N-consistent CSP with N variables can be solved without backtracking
- 2. For any CSP that is strongly K-consistent:
- If we find an appropriate variable ordering (one with "small enough" branching factor)
- We can solve the CSP without backtracking

67

#### Ordered Constraint Graphs

- Select a variable ordering, V<sub>1</sub>, ..., V<sub>n</sub>
- Width of a node in this OCG is the number of arcs leading to *earlier* variables:
- width( $V_i$ ) = count (( $V_i$ ,  $V_k$ ) | k < i)
- Width of the OCG is the maximum width of any node:
  - width(OCG) = max (width (V<sub>i</sub>)),  $1 \le i \le N$
- Width of an unordered CG is the minimum width of all orderings of that graph (best you can do)

68

#### Tree-Structured Constraint Graph

- A constraint tree rooted at V<sub>1</sub> satisfies:
  - There exists an ordering V<sub>1</sub>, ..., V<sub>n</sub> such that every node has zero or
    one parents (i.e., each node only has constraints with at most one
    "earlier" node in the ordering)

 $v_1 = v_2 = v_3 = v_5 = v_5 = v_6 = v_9 = v_1$ 

- Also known as an ordered constraint graph with width 1
- If this constraint tree is also node- and arc-consistent (a.k.a. strongly 2-consistent), it can be solved without backtracking
  - (More generally, if the ordered graph is strongly k-consistent, and has width w < k, then it can be solved without backtracking.)</li>

69

#### **Proof Sketch for Constraint Trees**

- Perform backtracking search in the order that satisfies the constraint tree condition
- Every node, when instantiated, is constrained only by at most one previous node
- Arc consistency tells us that there must be at least one legal instantiation in this case
- (If there are no legal solutions, the arc consistency procedure will collapse the graph some node will have no legal instantiations)
- Keep doing this for all n nodes, and you have a legal solution – without backtracking!

70

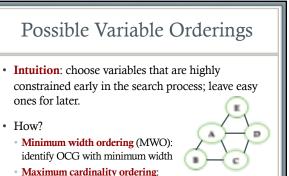
72

## Backtrack-Free CSPs: Proof Sketch

- Given a strongly k-consistent OCG, G, with width w < k:
  - $^{\circ}$  Instantiate variables in order, choosing values that are consistent with the constraints between  $V_{\rm i}$  and its parents
  - Each variable has at most w parents, and k-consistency tells us we can find a legal value consistent with the values of those w parents
- Unfortunately, achieving k-consistency is hard (and can increase the width of the graph in the process!)
- Fortunately, 2-consistency is relatively easy to achieve, so constraint trees are easy to solve
- Unfortunately, many CGs have width greater than one (that is, no equivalent tree), so we still need to improve search

So What If We Don't Have a Tree?

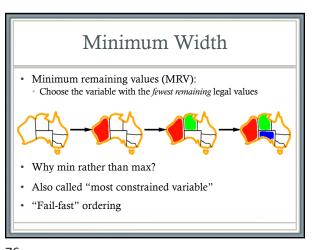
- Answer #1: Try interleaving constraint propagation and backtracking
- Answer #2: Try using variable-ordering heuristics to improve search
- Answer #3: Try using value-ordering heuristics during variable instantiation
- Answer #4: See if iterative repair works better
- Answer #5: Try using **intelligent backtracking** methods



approximation of MWO that's cheaper to compute: order

variables by decreasing cardinality (a.k.a. degree heuristic)

74 75



76

#### Value Ordering

- constrained early on, leaving the most legal values in later variables
  - Maximal options method (a.k.a. least-constraininglegal values for not-yet-instantiated variables
- constrain search space to 'useful' solutions (don't examine more than one symmetric/isomorphic solution)

Possible Variable Orderings • Fail first principle (FFP): choose variable with fewest remaining values • AKA minimum remaining values (MRV)) • Static FFP: use domain size of variables · Dynamic FFP (search rearrangement method): At each choice, select the  $D = \{g,b\}$  $A = \{g\}$ variable with the fewest remaining values

Variable Orderings II

- **Maximal stable set**: find largest set of variables with no constraints between them, save these for 1ast
- **Cycle-cutset tree creation**: Find a set of variables that, once instantiated, leave a tree of uninstantiated variables; solve these, then solve the tree without backtracking
- Tree decomposition: Construct a tree-structured set of connected subproblems

77

- Intuition: Choose values that are the least
  - value heuristic): Choose the value that leaves the most
  - Min-conflicts: For iterative repair search (Coming up)
  - Symmetry: Introduce symmetry-breaking constraints to

#### Iterative Repair

- Start with an initial complete (but probably invalid) assignment
- Repair locally
- Min-conflicts: Select new values that minimally conflict with the other variables
- Use in conjunction with hill climbing or simulated annealing or...
- Local maxima strategies
- Random restart
- Random walk

78 79

#### Min-Conflicts Heuristic

- · Iterative repair method
  - 1. Find some "reasonably good" initial solution
    - E.g., in N-queens problem, use greedy search through rows, putting each queen where it conflicts with the smallest number of previously placed queens, breaking ties randomly
  - 2. Pick a variable in conflict (randomly)
  - 3. Select a new value that *minimizes* the number of constraint violations
    - O(N) time and space
  - 4. Repeat steps 2 and 3 until done

Min-Conflicts Heuristic

- · Iterative repair method
  - 1. Find some "reasonably good" initial solution
  - E.g., in N-queens problem, use greedy search through rows, putting each queen where it conflicts with the smallest number of previously placed queens, breaking ties randomly
  - Pick a variable in
     Select a new valu

Performance depends on quality and informativeness of initial assignment; inversely

constraint violation of constraint violation relations

related to distance to solution

4. Repeat steps 2 and 3 until done

80

#### Intelligent Backtracking

- Backjumping: if V<sub>j</sub> fails, jump back to the variable V<sub>i</sub> with greatest i such that the constraint (V<sub>i</sub>, V<sub>j</sub>) fails (i.e., most recently instantiated variable in conflict with V<sub>j</sub>)
- Backchecking: keep track of incompatible value assignments computed during backjumping
- Backmarking: keep track of which variables led to the incompatible variable assignments for improved backchecking

83

81

82

84

#### Challenges

- What if not all constraints can be satisfied?
  - · Hard vs. soft constraints
  - Degree of constraint satisfaction
  - · Cost of violating constraints
- · What if constraints are of different forms?
  - Symbolic constraints
  - Numerical constraints [constraint solving]
  - · Temporal constraints
  - · Mixed constraints

More Challenges

- · What if constraints are represented intensionally?
  - · Cost of evaluating constraints (time, memory, resources)
- · What if constraints/variables/values change over time?
  - Dynamic constraint networks
  - · Temporal constraint networks
  - · Constraint repair
- What if you have multiple agents or systems involved?
  - Distributed CSPs
- Localization techniques

85