Problem Solving and Search in Artificial Intelligence

Basic Concepts

Nysret Musliu
Database and Artificial Intelligence Group
Institut für Informationssysteme, TU-Wien

Introduction

- Where to begin?
- You have to create the plan for generating of solution
- Always consider all of the available data
- Can you make connection between the goal and what is given

Puzzle: Ages of Three Sons

- Product of ages of three sons is 36
- Sum of the son's ages is the same as the number of windows in building
- The oldest son has blue eyes

How old is each of sons?

Puzzle: Ages of Three Sons

Product of ages of three sons is 36

Sum of the son's ages is the same as the number of windows in building

Age of son 1	Age of son 2	Age of son 3
36	1	1
18	2	1
12	3	1
9	4	1
9	2	2
6	6	1
6	3	2
4	3	3

Possible Sums
38
21
16
14
13
13
11
10

The oldest son has blue eyes

Why are some problems difficult to solve?

- The number of possible solutions in the search space is to large
- The problem is too complicated: we have to use such simplified models of the problem that any result is essentially useless
- Evaluation function is noisy or varies with time
- The possible solutions are to heavily constrained

Search space

- SAT problem:
  - Make a a compound statement of Boolean variables to evaluate to true
  - Example: Formula of 100 variables given in conjunctive normal form:

$$F(x) = (x_{17} \vee \bar{x}_{37} \vee x_{73}) \wedge (\bar{x}_{11} \vee \overline{x_{12}}) \wedge \ldots \wedge (\bar{x}_{2} \vee x_{43} \vee x_{22})$$

  - Find the truth assignment for each variable x_i such that *F(x)=TRUE*

Search space

- For the SAT problem in the previous example:
  - There are 100 variables
  - Each variable can take two values (0,1)
- The size of the search space is:

$$|S| = 2^{100}$$

- How should be the evaluation function?

Traveling Salesman Problem

- Are given:
  - N cities
  - Distance between each pair of cities
- Problem:
  - Travelling salesman must visit every city exactly once and than return home covering the shortest distance
- Size of search space:

$$|S| = n!/2n$$

Modeling the problem

- We are finding solution to a model of the problem
- Model is a simplification of real world
- Process of model solving consist of two separate general steps:
  - Creating of model of the problem
  - Using that model to generate a solution:

    Problem => Model => Solution

Change over time

- Change of problem while we are deriving the solution
- Travelling Salesman Problem:
  - Travel time between two cities depends from many factors
    - Make all the green (or red!) traffic lights along the way
    - Get stuck in the road, weather, road conditions…

Change over time

- Example: You are owner of Supermarket chain and have to decide where to place new store
  - Calculate the cost of construction in each possible location, the existing competition etc.
  - Formulate fitness function
  - As you are deciding where to put new story the competition have their own story to construct
  - The competition tries to find the best place to their new store so as to minimize your success

Constraints

- Almost all practical problems pose constraints
- Two types of constraint
  - Hard constraints
  - Soft constraints
- Constraints make the search space smaller, but
  - It is hard to create operators that will act on feasible solution and generate in turn new feasible solutions that are an improvement of previous solution
  - The geometry of search space gets tricky

Constraints

Example: Timetable of the classes at a college in one semester

- We are given
  - List of courses that are offered
  - List of students assigned to each class
  - Professors assigned to each class
  - List of available classrooms, and information for size and other facilities that each offer

University Timetabling

- Construct timetables that fulfil hard constraints:
  - Each class must be assigned to an available room that has enough seats and requisite facilities
  - Students who are enrolled in more than one class can not have their classes held at the same time on the same day
  - Professors can not be assigned to teach courses that overlap in time

University Timetabling

- Soft Constraints:
  - Courses that meets twice a week should preferably be assigned to Mondays and Wednesdays or Tuesdays and Thursdays
  - Courses that meets three times per week should preferably be assigned to Mondays, Wednesdays, and Fridays
  - Course time should be assigned so that students do not have to take final exams for multiple courses without any break in between
  - If more than one room satisfies the requirements for a course and is available at the designated time, the course should be assigned to the room with the capacity that is closest to the class size
  - ...

University Timetabling

- Any timetable that meets the hard constraints is feasible
- The timetable has to be optimized in the light of soft constraints
- Each soft constraint has to be quantified
- We can evaluate two candidate assignments and decide that one is better than other

The problem of proving things

- Find solution to a problem easier than proving something
- Example:
  - You will take four hours to fill a pool using a large pipe.
  - You will take six hours if you use the small pipe.
  - How long would it take if you used both pipes?
- Possible problem formulations:
  - Prove that the amount of time to fill the pool with two pipes is less than *a* -> harder
  - Find value of *x* (amount of time required to fill the pool using both pipes) ->easy

Basic definitions

- Representation
  - Encodes alternative candidate solutions for manipulation
- Objective Function
  - Describes the purpose to be fulfilled
- Evaluation Function
  - Indicates the quality of any particular solution given the representation

Representation

- Satisfiability problem:
  - Candidate solution is represented with a binary string of length n (number of varaibles)

x1 x2 x3 x4 x5 ...

0 1 1 0 1 ...

- Travelling Salesman Problem:
  - Solution is represented as a permutation of natural numbers 1,..,n

1 3 5 6 7 2 4 8 ...

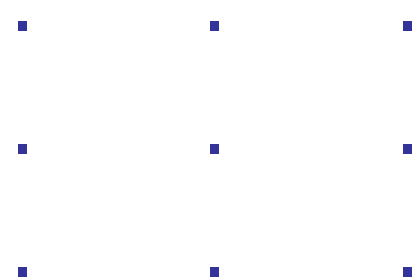
- The size of search space is not determined by the problem; it‘s determined by your representation and the manner in which you handle this encoding

Representation

- Example: There are six matches on the table and the task is construct four equilateral triangles where the length of each side is equal to the length of a match

Representation: 9 dots puzzle

Connect the 9 dots below with three lines without ever lifting your pencil from the paper



The Objective

- Mathematical statement of the task to be achieved
- Travelling Salesman Problem (TSP):
  - Objective: Minimize the total distance traversed by the salesman subject to constraint of visiting each city once and only once and returning to the starting city.

$$\min \sum dist(x,y)$$

The evaluation function

- Mapping from the space of possible candidate solutions under the chosen representation to a set of numbers
- TSP:
  - Possible evaluation function may map each tour to its corresponding problem's total distance
- Question: Which evaluation function could be used SAT problem?
- The Evaluation function is not given to you with the problem; You have to chose the evaluation function

Definition of search problem

- Given e search space S together with its feasible part $F \subseteq S$, find $x \in F$ such that

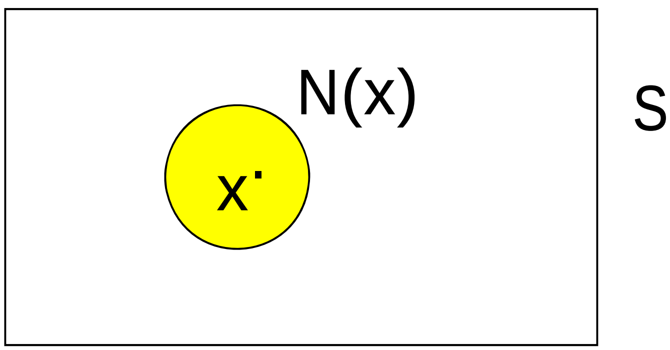
  $eval(x) \leq eval(y)$

  For all $y \in F$

- x that satisfies the above condition is called global optimum

Neighborhood and local optima

- Region of the search space that is near particular point in the space



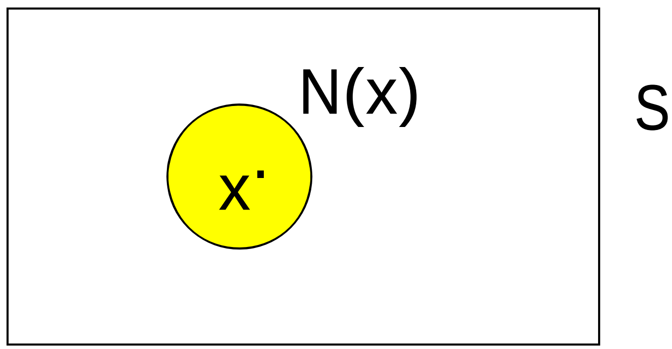
- A potential solution $x \in F$ ia a local optimum with respect to the neighborhood N, if and only if

  $eval(x) \leq eval(y)$,

  for all $y \in N(x)$

Local Search Techniques

- Are based on the neighbourhood of the current solution



- The solution is changed iteratively with so called neighbourhood relations (moves) until an acceptable or optimal solution is reached

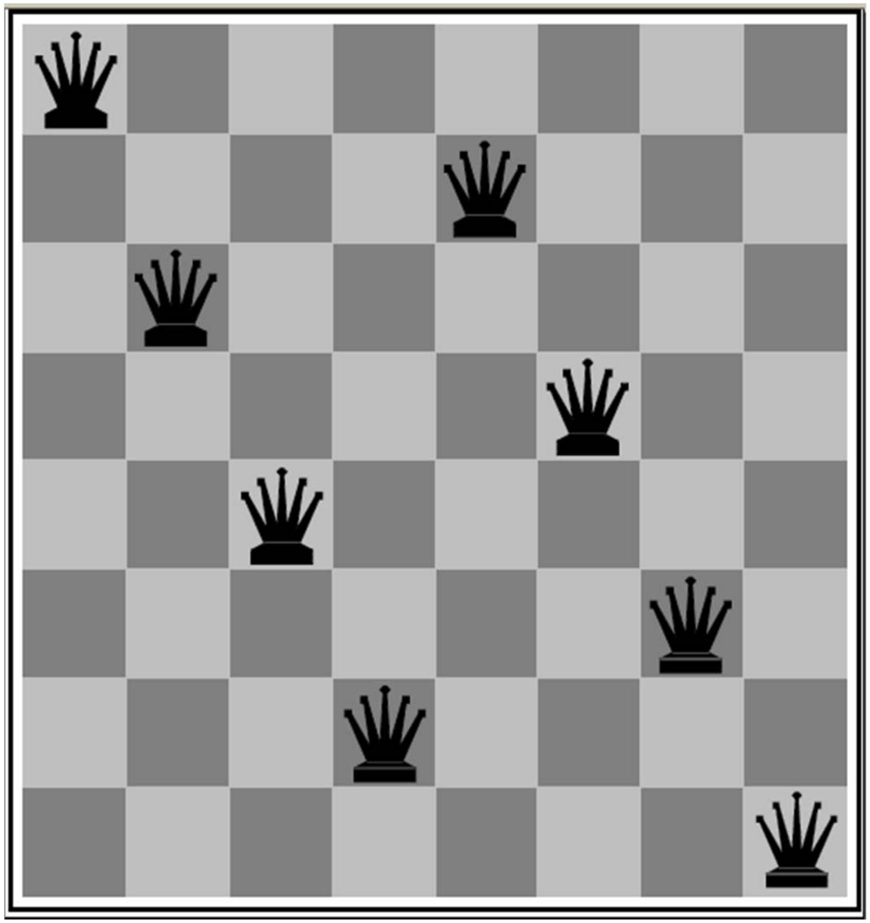
Local Search Techniques

- Construct the initial solution s,
- Generate neighbourhood N(s) of solution s
- Select from the neighbourhood the descendant of the current solution
- Go to step 2

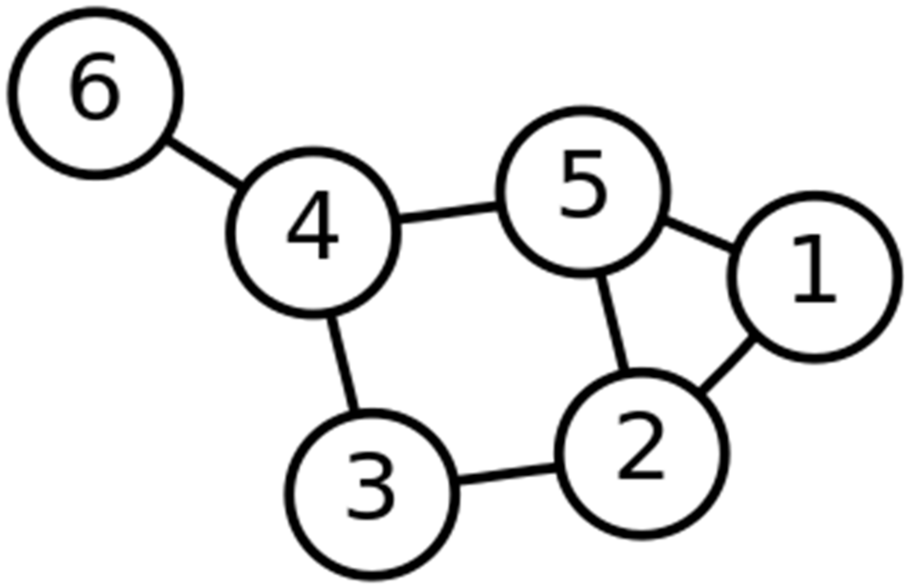
Exercises

- Analyse for the problems in the next slides:
  - Problem representation
  - Solution representation
  - Search space
  - Evaluation function

8 Queens problem



Graph Coloring



Sudoku

5	3			7				
6			1	9	5			
	9	8					6	
8				6				3
4			8		3			1
7				2				6
	6					2	8	4
			4	1	9	6	3	5
				8		1	7	9

Rotating Workforce Scheduling

Length of schedule: If the schedule is cyclic the total length of a planning period will be: NumberOfEmployees*7

	Mo	Tu	We	Th	Fr	Sa	Su
A	D	D	D			D	D
B	D	D	D	D			
C	A	A	A	A			A
D	A	A	A	A	A		
E	D	D	A	A	A		
F	A	A	N	N	N		
G	N	N	N	N	N		
H		N	N	N	N	N	
I			D	D	D	A	A
J				D	D	N	N
K	N				A	A	N
L	N	N			D	D	D

Number of employees

Employees working shifts:

D: Day shift ; A: Afternoon shift ,

N: Night shift; Day off

Constraints

	Mo	Tu	We	Th	Fr	Sa	Su
A	D	D	D			D	D
B	D	D	D	D			
C	A	A	A	A			A
D	A	A	A	A	A		
E	D	D	A	A	A		
F	A	A	N	N	N		
G	N	N	N	N	N		
H		N	N	N	N	N	
I			D	D	D	A	A
J				D	D	N	N
K	N				A	A	N
L	N	N			D	D	D

Not allowed sequences of shifts:

N - D
N D
A D
N A
N - A
A - D

Maximum and minimum length of periods of successive shifts.

e.g.: N: 2-5, D: 2-6

Temporal requirements: required number of employees in shift *i* during day *j*

Monday (Mo): D: 3, N: 3, A: 3

Maximum and minimum length of work days and days-off blocks

e.g.: days-off block: 2-4

work block: 2-6

Objective

Find a cyclic schedule (assignment of shifts to employees) that satisfies the temporal requirement, and all other constraints

Possible soft constraints:

- Optimization of free weekends (weekends off)
- …

Literature

- Z. Michalewicz and D. B. Fogel. How to Solve It: Modern Heuristics
  - Chapter 1,2