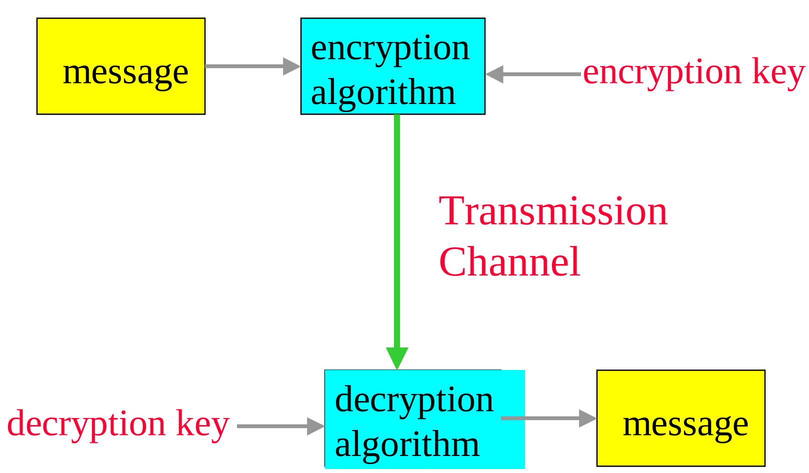
Hard Problems

- Some problems are hard to solve.
  - No polynomial time algorithm is known.
  - E.g., NP-hard problems such as machine scheduling, bin packing, 0/1 knapsack.
- Is this necessarily bad?
- Data encryption relies on difficult to solve problems.

Cryptography



Public Key Cryptosystem (RSA)

- A public encryption method that relies on a public encryption algorithm, a public decryption algorithm, and a public encryption key.
- Using the public key and encryption algorithm, everyone can encrypt a message.
- The decryption key is known only to authorized parties.
- Asymmetric method.
  - Encryption and decryption keys are different; one is not easily computed from the other.

Public Key Cryptosystem (RSA)

- p and q are two prime numbers.
- $n = pq$
- $m = (p-1)(q-1)$
- a is such that $1 < a < m$ and $\gcd(m,a) = 1$.
- b is such that $(ab) \bmod m = 1$.
- a is computed by generating random positive integers and testing $\gcd(m,a) = 1$ using the extended Euclid's gcd algorithm.
- The extended Euclid's gcd algorithm also computes b when $\gcd(m,a) = 1$.

RSA Encryption And Decryption

- Message $M < n$.
- Encryption key $= (a,n)$.
- Decryption key $= (b,n)$.
- Encrypt $\Rightarrow E = M^a \bmod n$.
- Decrypt $\Rightarrow M = E^b \bmod n$.

Breaking RSA

- Factor n and determine p and q, $n = pq$.
- Now determine $m = (p-1)(q-1)$.
- Now use Euclid's extended gcd algorithm to compute $gcd(m,a)$. b is obtained as a byproduct.
- The decryption key (b,n) has been determined!

Security Of RSA

- Relies on the fact that prime factorization is computationally very hard.
- Let q be the number of bits in the binary representation of n.
- No algorithm, polynomial in q, is known to find the prime factors of n.
- Try to find the factors of a 100 bit number.

Elliptic Curve Cryptography (ECC)

- Asymmetric Encryption Method
  - Encryption and decryption keys are different; one is not easily computed from the other.
- Relies on difficulty of computing the discrete logarithm problem for the group of an elliptic curve over some finite field.
  - Galois field of size a power of 2.
  - Integers modulo a prime.
- 1024-bit RSA ~ 200-bit ECC (cracking difficulty).
- Faster to compute than RSA?

Data Encryption Standard

- Used for password encryption.
- Encryption and decryption keys are the same, and are secret.
- Relies on the computational difficulty of the satisfiability problem.
- The satisfiability problem is NP-hard.

Satisfiability Problem

- The permissible values of a boolean variable are true and false.
- The complement of a boolean variable x is denoted $\bar{x}$.
- A literal is a boolean variable or the complement of a boolean variable.
- A clause is the logical or of two or more literals.
- Let x_1, x_2, x_3, …, x_n be n boolean variables.

Satisfiability Problem

- Example clauses:
  - $x_1 + x_2 + x_3$
  - $x_4 + \bar{x}_7 + x_8$
  - $x_3 + x_7 + x_9 + x_{15}$
  - $x_2 + x_5$
- A boolean formula (in conjunctive normal form CNF) is the logical and of m clauses.
- $F = C_1 C_2 C_3 \ldots C_m$

Satisfiability Problem

- $F = (x_1 + x_2 + x_3)(x_4 + \bar{x}_7 + x_8)(x_2 + x_5)$
- F is true when x_1, x_2, and x_4 (for e.g.) are true.

Satisfiability Problem

- A boolean formula is satisfiable iff there is at least one truth assignment to its variables for which the formula evaluates to true.
- Determining whether a boolean formula in CNF is satisfiable is NP-hard.
- Problem is solvable in polynomial time when no clause has more than 2 literals.
- Remains NP-hard even when no clause has more than 3 literals.

Other Problems

- Partition
  - Partition n positive integers s_1, s_2, s_3, …, s_n into two groups A and B such that the sum of the numbers in each group is the same.
  - [9, 4, 6, 3, 5, 1,8]
  - A = [9, 4, 5] and B = [6, 3, 1, 8]
- NP-hard.

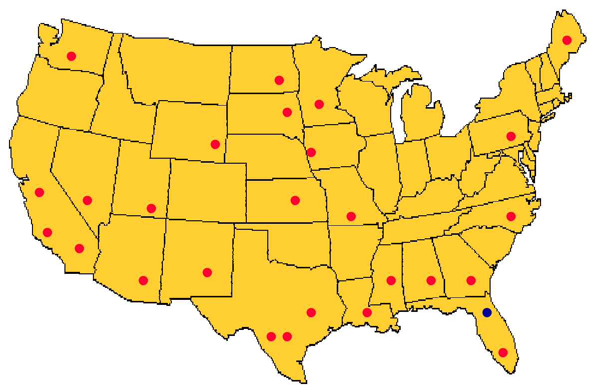
Subset Sum Problem

- Does any subset of n positive integers s_1, s_2, s_3, …, s_n have a sum exactly equal to c?
- [9, 4, 6, 3, 5, 1,8] and c = 18
- A = [9, 4, 5]
- NP-hard.

Traveling Salesperson Problem (TSP)

- Let G be a weighted directed graph.
- A tour in G is a cycle that includes every vertex of the graph.
- TSP => Find a tour of shortest length.
- Problem is NP-hard.

Applications Of TSP [redacted]



- Home city
- Visit city

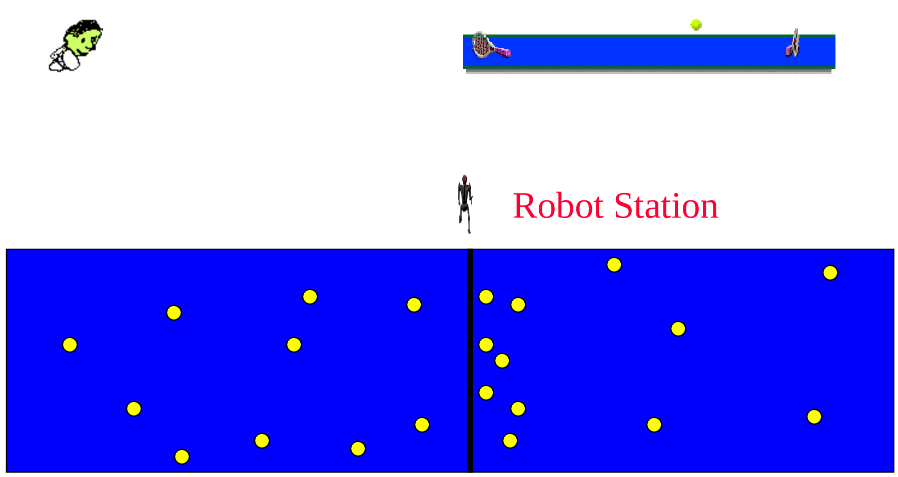
Applications Of TSP

- Each vertex represents a city that is in Joe’s sales district.
- The weight on edge (u,v) is the time it takes Joe to travel from city u to city v.
- Once a month, Joe leaves his home city, visits all cities in his district, and returns home.
- The total time he spends on this tour of his district is the travel time plus the time spent at the cities.
- To minimize total time, Joe must use a shortest-length tour.

Applications Of TSP

- Tennis practice.
- Start with a basket of approximately 200 tennis balls.
- When balls are depleted, we have 200 balls lying on and around the court.
- The balls are to be picked up by a robot (more realistically, the tennis player).
- The robot starts from its station visits each ball exactly once (i.e., picks up each ball) and returns to its station.

Applications Of TSP

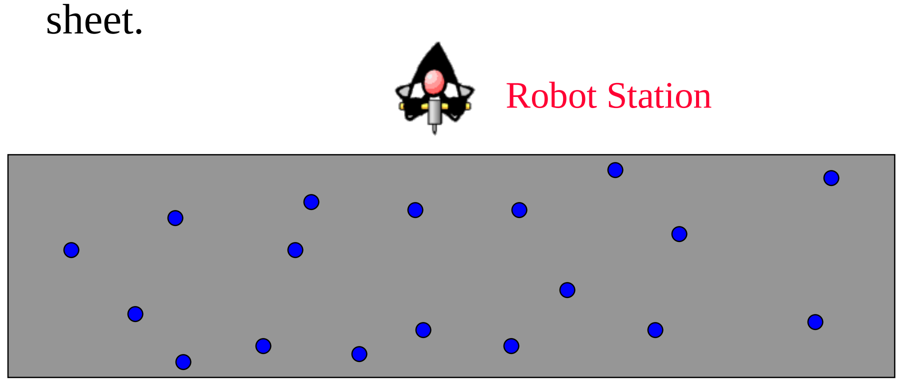


Applications Of TSP

- 201 vertex TSP.
- 200 tennis balls and robot station are the vertices.
- Complete directed graph.
- Length of an edge (u,v) is the distance between the two objects represented by vertices u and v.
- Shortest-length tour minimzes ball pick up time.
- Actually, we may want to minimize the sum of the time needed to compute a tour and the time spent picking up balls using the computed tour.

Applications Of TSP

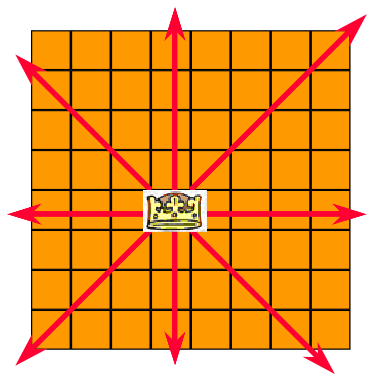
- Manufacturing.
- A robot arm is used to drill n holes in a metal sheet.



n+1 vertex TSP.

n-Queens Problem

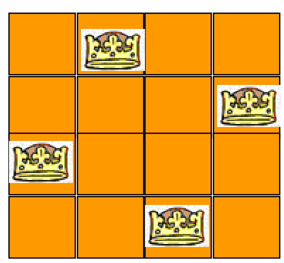
A queen that is placed on an n x n chessboard, may attack any piece placed in the same column, row, or diagonal.



8x8 Chessboard

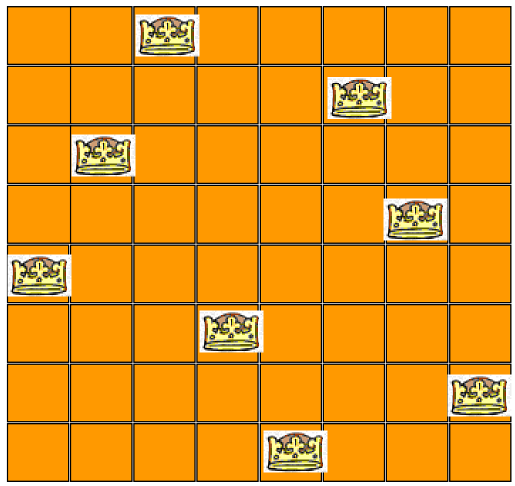
n-Queens Problem

Can n queens be placed on an n x n chessboard so that no queen may attack another queen?



4x4

n-Queens Problem



8x8

Difficult Problems

- Many require you to find either a subset or permutation that satisfies some constraints and (possibly also) optimizes some objective function.
- May be solved by organizing the solution space into a tree and systematically searching this tree for the answer.

Subset Problems

- Solution requires you to find a subset of n elements.
- The subset must satisfy some constraints and possibly optimize some objective function.
- Examples.
  - Partition.
  - Subset sum.
  - 0/1 Knapsack.
  - Satisfiability (find subset of variables to be set to true so that formula evaluates to true).
  - Scheduling 2 machines.
  - Packing 2 bins.

Permutation Problems

- Solution requires you to find a permutation of n elements.
- The permutation must satisfy some constraints and possibly optimize some objective function.
- Examples.
  - TSP.
  - n-queens.
    - Each queen must be placed in a different row and different column.
    - Let queen i be the queen that is going to be placed in row i.
    - Let c_i be the column in which queen i is placed.
    - $c_1, c_2, c_3, \ldots, c_n$ is a permutation of $[1,2,3, \ldots, n]$ such that no two queens attack.

Solution Space

- Set that includes at least one solution to the problem.
- Subset problem.
  - n = 2, {00, 01, 10, 11}
  - n = 3, {000, 001, 010, 100, 011, 101, 110, 111}
- Solution space for subset problem has 2^n members.
- Nonsystematic search of the space for the answer takes $O(p2^n)$ time, where p is the time needed to evaluate each member of the solution space.

Solution Space

- Permutation problem.
  - n = 2, {12, 21}
  - n = 3, {123, 132, 213, 231, 312, 321}
- Solution space for a permutation problem has n! members.
- Nonsystematic search of the space for the answer takes O(pn!) time, where p is the time needed to evaluate a member of the solution space.