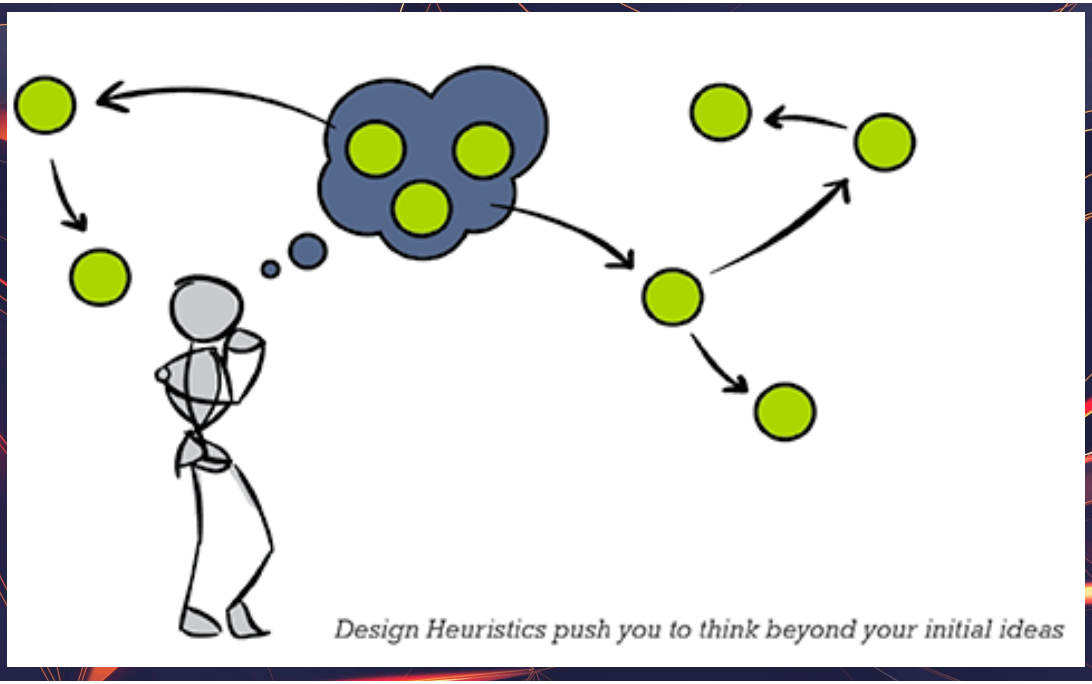


COMBINATORIAL OPTIMIZATION AND METAHEURISTICS

Prof. Eduardo Pécora

CALENDAR

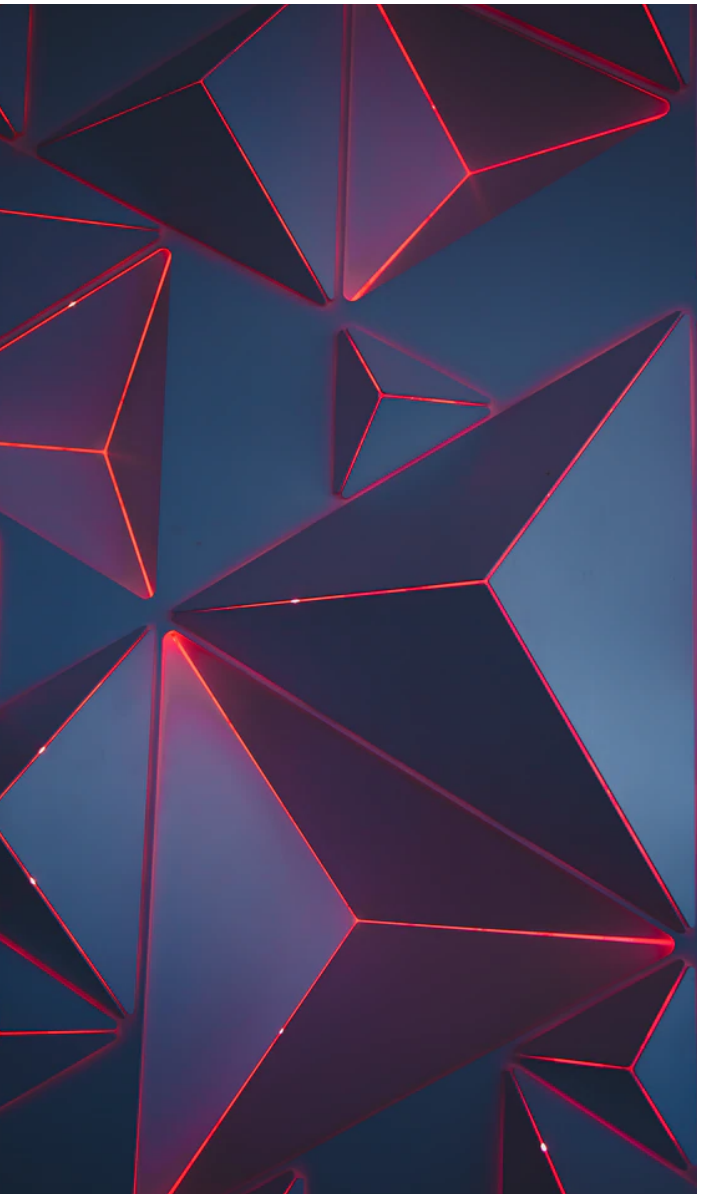
Lecture	Date	Theme
1	13/07/2021	Introduction
2	20/07/2021	Single Solution Metaheuristic methods for combinatorial problems
3	27/07/2021	Local search heuristic methods and the local optimal trap (VNS, ILS, GRASP)
4	03/08/2021	Tabu Search
5	10/08/2021	Performance Analysis
6	17/08/2021	Population Based Metaheuristics: Genetic Algorithms
7	24/08/2021	Ants Colony, Particle Cloud, BRKGA
8	31/08/2021	Parameter Settings - Irace
	07/09/2021	Holliday
9	14/09/2021	Very Large Neighborhood Search, Granular Tabu Search, Other Metaheuristics
10	21/09/2021	Hybridization: Heuristic and Exact methods working together
11 and 12	28/09/2021	Presentations of the Final Works

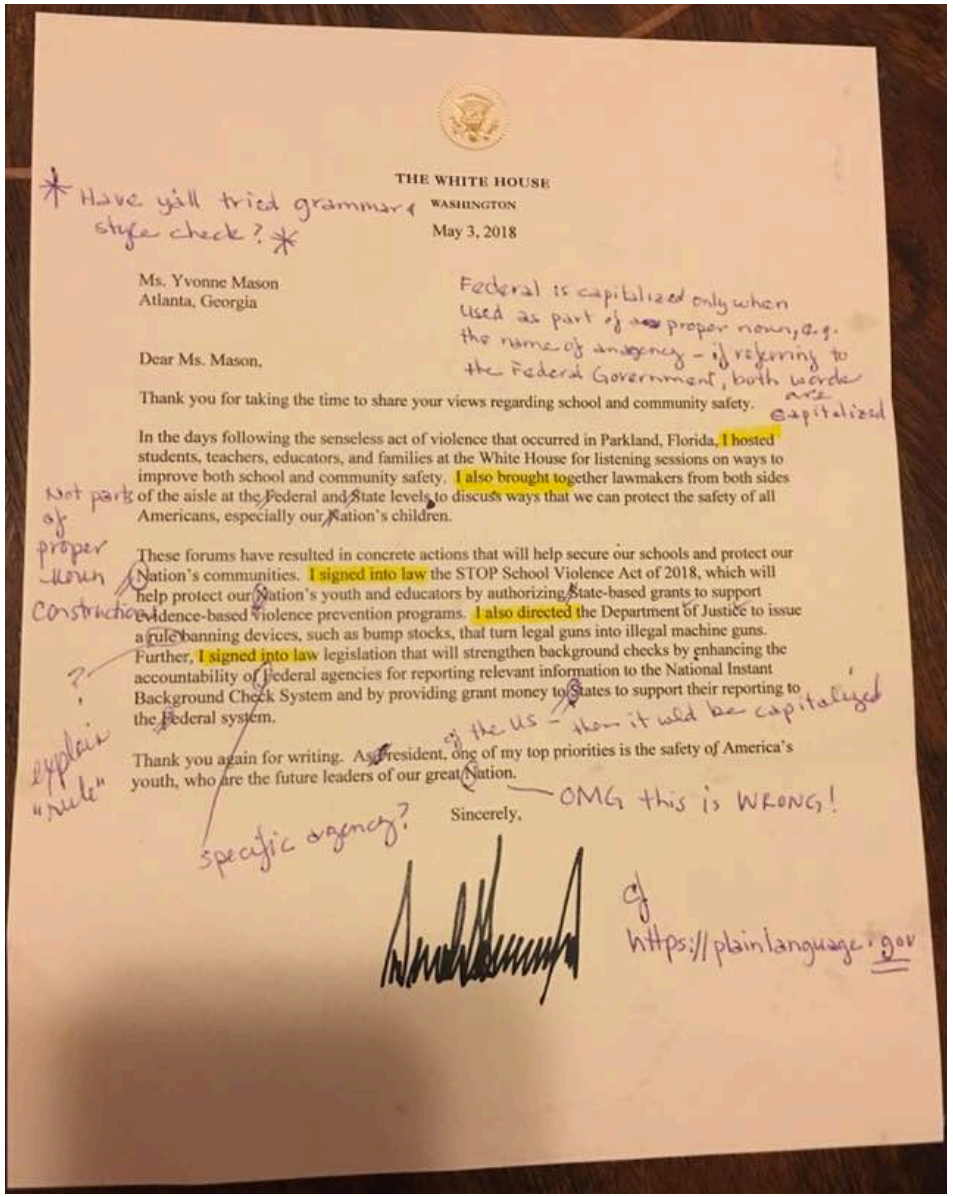
EVALUATION

Evaluation	Percentage
LABs	40%
Final Project (article)	30%
Presentations	30%

TEACHING CONTRACT

Prof. Eduardo Pécora





THE WHITE HOUSE
WASHINGTON

May 3, 2018

Ms. Yvonne Mason
Atlanta, Georgia

Dear Ms. Mason,

Thank you for taking the time to share your views regarding school and community safety.

In the days following the senseless act of violence that occurred in Parkland, Florida, I hosted students, teachers, educators, and families at the White House for listening sessions on ways to improve both school and community safety. I also brought together lawmakers from both sides of the aisle at the Federal and State levels to discuss ways that we can protect the safety of all Americans, especially our Nation's children.

These forums have resulted in concrete actions that will help secure our schools and protect our Nation's communities. I signed into law the STOP School Violence Act of 2018, which will help protect our Nation's youth and educators by authorizing State-based grants to support evidence-based violence prevention programs. I also directed the Department of Justice to issue a rule banning devices, such as bump stocks, that turn legal guns into illegal machine guns. Further, I signed into law legislation that will strengthen background checks by enhancing the accountability of Federal agencies for reporting relevant information to the National Instant Background Check System and by providing grant money to States to support their reporting to the Federal system.

Thank you again for writing. As President, one of my top priorities is the safety of America's youth, who are the future leaders of our great Nation.

Sincerely,

* Have y'all tried grammar + style check? *

Federal is capitalized only when used as part of a proper noun, e.g. the name of an agency - if referring to the Federal Government, both words are capitalized

Not part of proper noun construction?

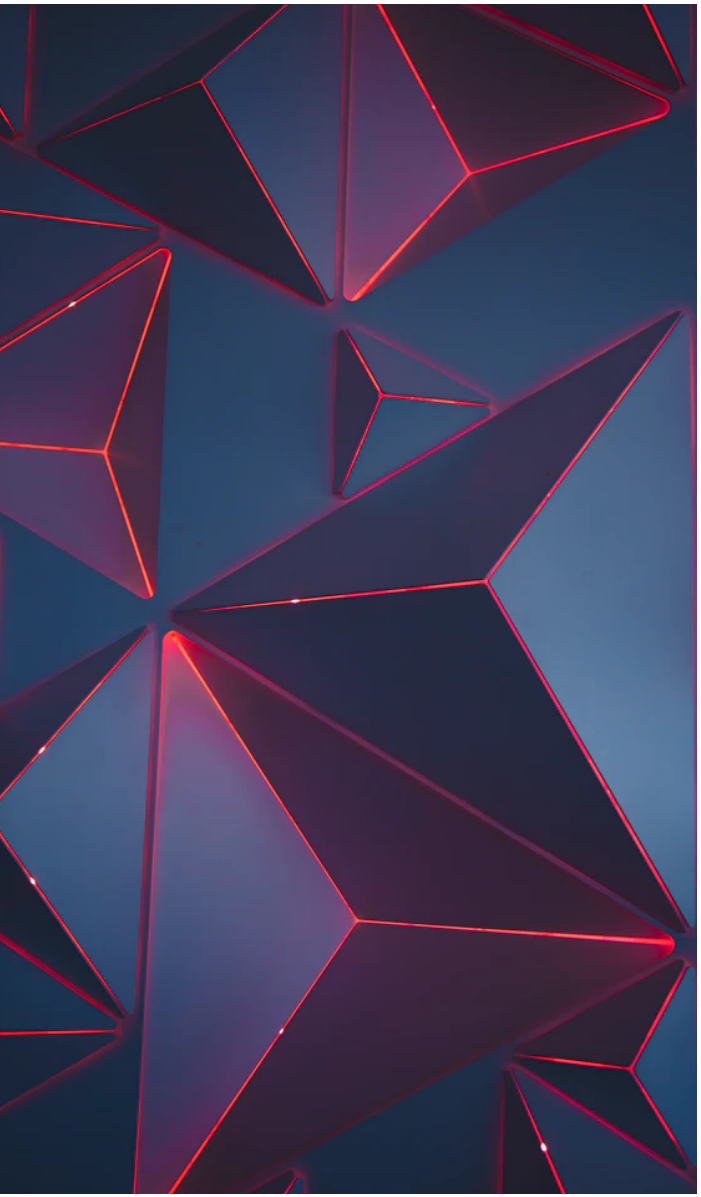
explain "rule"

specific agency?

of the US - then it wld be capitalized

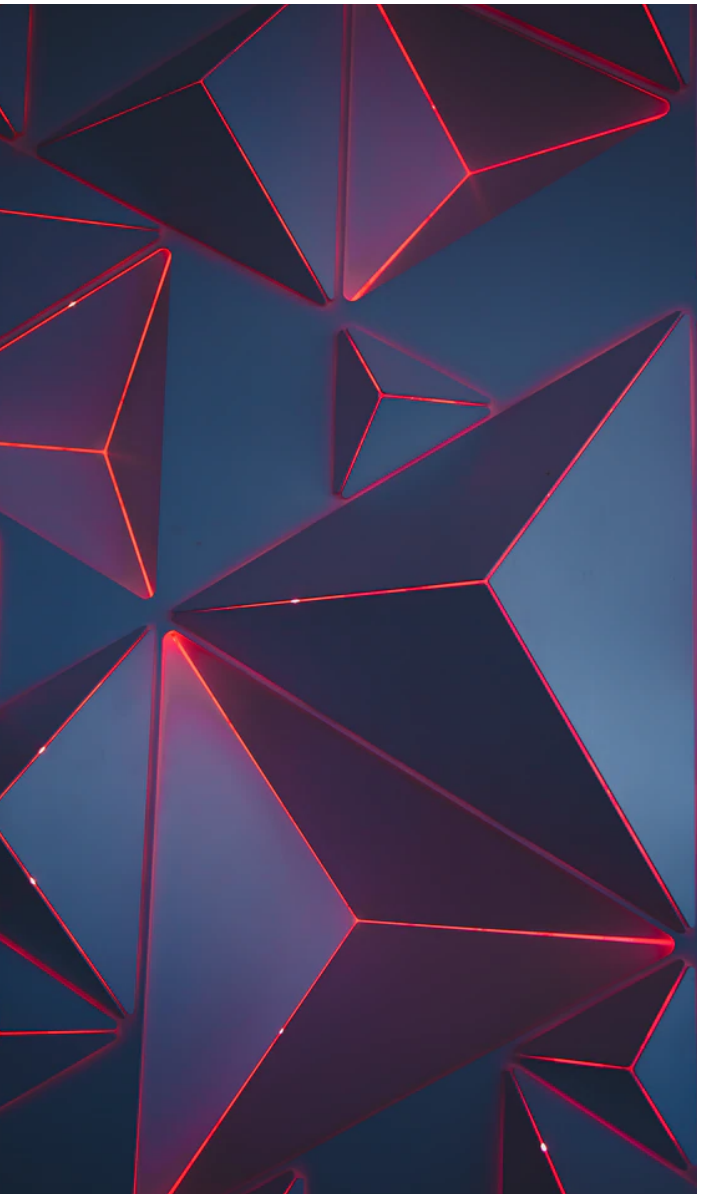
OMG this is WRONG!

cf https://plainlanguage.gov



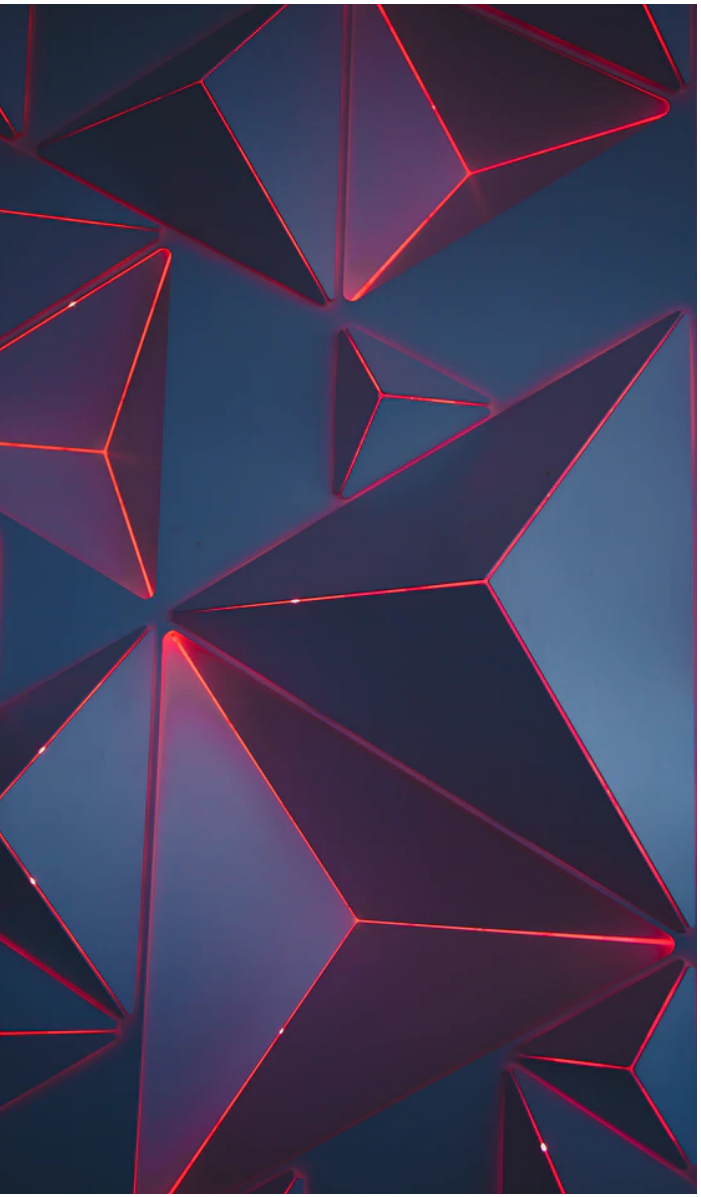
USEFUL LINKS

- English Test: https://www.efset.org/
- skell
  - https://skell.sketchengine.co.uk/run.cgi/skell
- Dictionary -
  - https://www.merriam-webster.com/
- Thesaurus -
  - http://www.thesaurus.com/
- Google Translator
- DeepL



GOOD HABITS

- Keep an open mind! Interact with your classmates and professor!
- This class will help to improve your English! Let's learn together!
- Discuss how (and if) you want to be corrected!
- Keep a notepad with the vocabulary you do not know.
  - At the end of the class, look for their meaning or ask a colleague or the professor.
  - The primary focus of this class is to learn Metaheuristics.

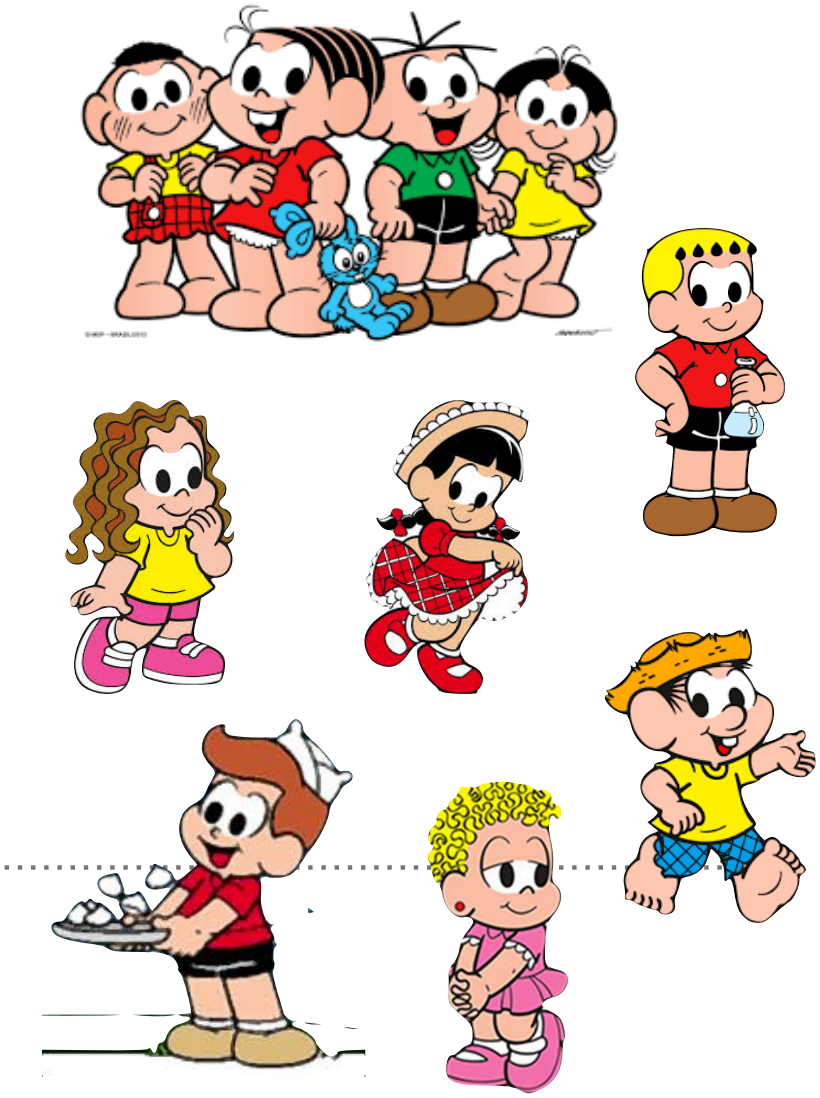


CONTRACT

- Create a notepad for unknown words
- Be pro-active
- Teams

COMBINATORIAL OPTIMIZATION

Prof. Eduardo Pécora



ACTIVITY 1.0

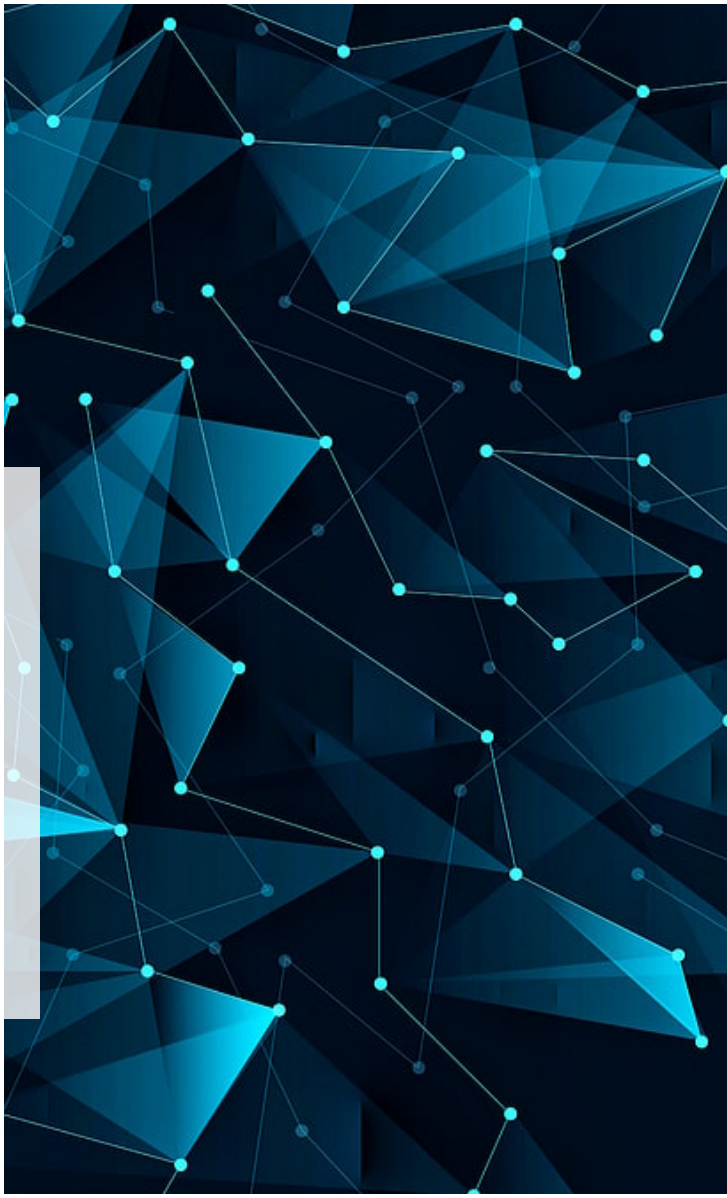
S:

S->SS:

SS<->SS:

SS->C

	Chico Bento	Cascão	Cebolinha	Quinzinho	Franjinha
Mônica	3	7	1	2	3
Cascuda	9	8	8	3	2
Rosinha	8	5	6	1	8
Magali	6	9	4	3	5
Marina	8	5	5	7	3



	Chico Bento	Cascão	Cebolinha	Quinzinho	Franjinha
Mônica	3	7	1	2	3
Cascuda	9	8	8	3	2
Rosinha	8	5	6	1	8
Magali	6	9	4	3	5
Marina	8	5	5	7	3

BIG NUMBERS

- With 5 males and 5 females we may have 120 different combinations of couples.
- And if we had 70 males and 70 females. How many possible combinations will we have?

There are 70! (seventy factorial) possible different combinations

How big is this number?

Just a hint!

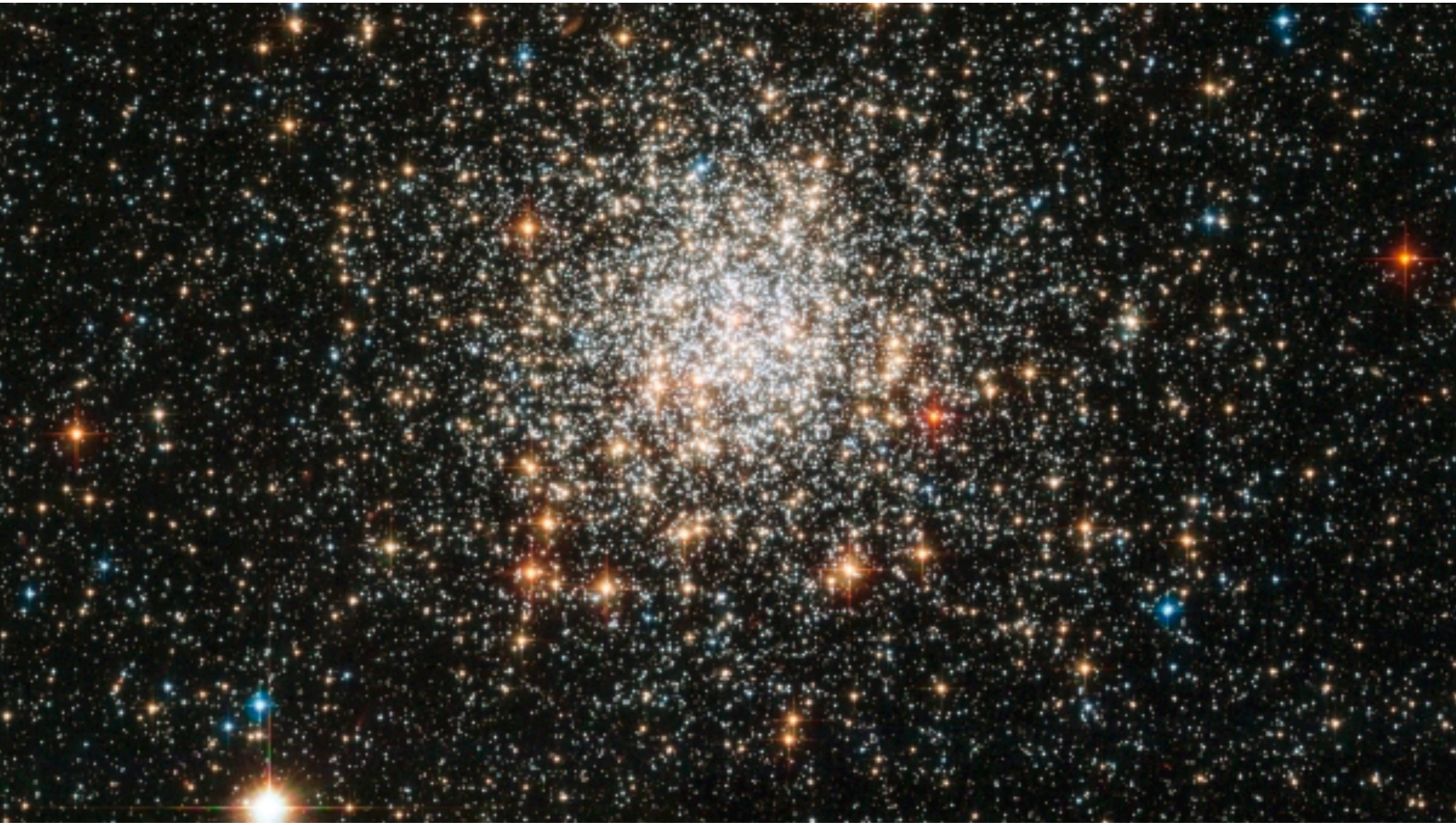
$$70! > 10^{100}$$

BIG NUMBERS

Do you know this constant from physics?

$$7.1 \times 10^{79}$$

$$70! > 10^{100} > 7.1 \times 10^{79}$$



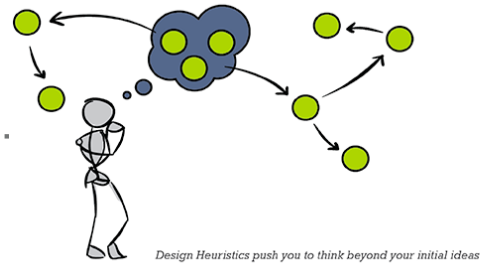
BASICS

WHAT IS A HEURISTIC?

- Heuristic
  - : involving or serving as an aid to learning, discovery, or problem-solving by experimental and especially trial-and-error methods <heuristic techniques> <a heuristic assumption>; also
  - : of or relating to **exploratory problem-solving techniques** that utilize self-educating techniques (as the evaluation of feedback) to improve performance <a heuristic computer program>

Source: http://www.merriam-webster.com/dictionary/heuristic

HEURISTIC METHODS



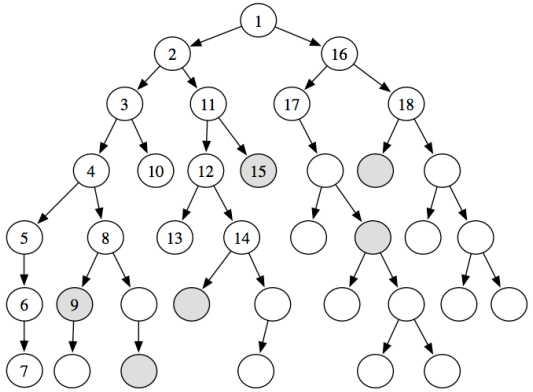
+ empirically proven to find good solutions in a short time

+ exploit the problem characteristics for better performance

- do not guarantee the optimality

- may leave unexplored regions of the space

EXACT METHODS



+ systematic exploration of the space

+ provides the optimal solution

+ very efficient for small sized instances

- very expensive computationally

- neither efficient nor fast enough for large scale problems

SOLVING APPROACH TREE

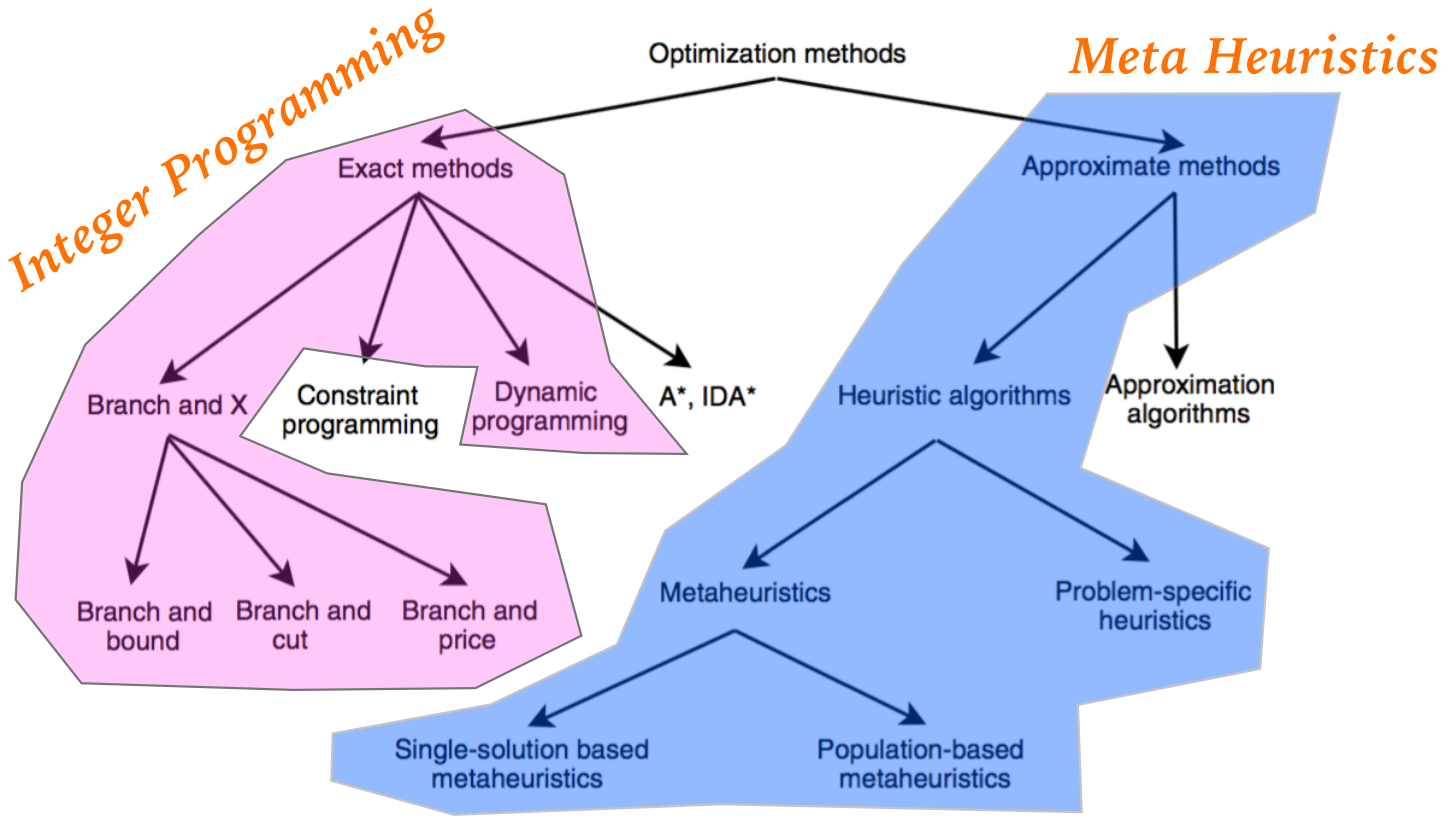


FIGURE 1.7 Classical optimization methods.

DESIGN OF METAHEURISTICS

Conception	*Am I Exploring the problem's structure?* *Are there any constraints which may be relaxed?*
Programming	*Define Programming Language*
Small test	*Generate a small data instance* *Solve it*
Setting	*Are the parameters at their best values?* *May I prove it? Even empirically?*
Computational experiments	*Generate a large data instance* *Validate them* *Are my instances set large enough?* *How long takes to solve them?*

DEFINITION

Formal

- Definition 1.1: Global optimum. A solution $s* \in S$ is a global optimum if it has a better objective function than all solutions of the search space, that is, $\forall s \in S$, $f(s*) \leq f(s)$.

Casual

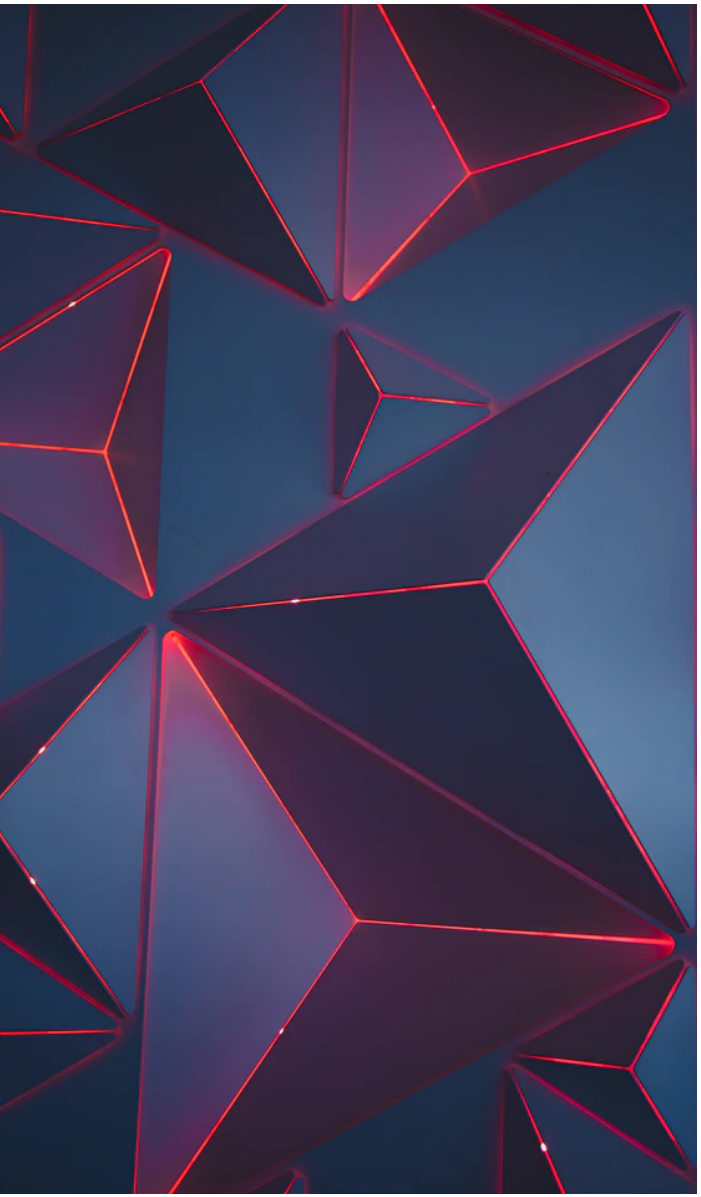
- The best of all solutions ever!
- Nobody is better than me!
- I am so good!

COMPLEXITY OF AN ALGORITHM

COMPLEXITY OF ALGORITHM

- An algorithm needs two important resources to solve a problem: time and space.
  - The time complexity of an algorithm is the number of steps required to solve a problem of size n. The complexity is generally defined in terms of the worst-case analysis.
  - The goal in the determination of the computational complexity of an algorithm **is not to obtain an exact step** count but an asymptotic bound on the step count.
  - The Big-O notation makes use of asymptotic analysis. It is one of the most popular notations in the analysis of algorithms.

SS<->SS;



SORT THIS LIST

12

45

34

13

78

01

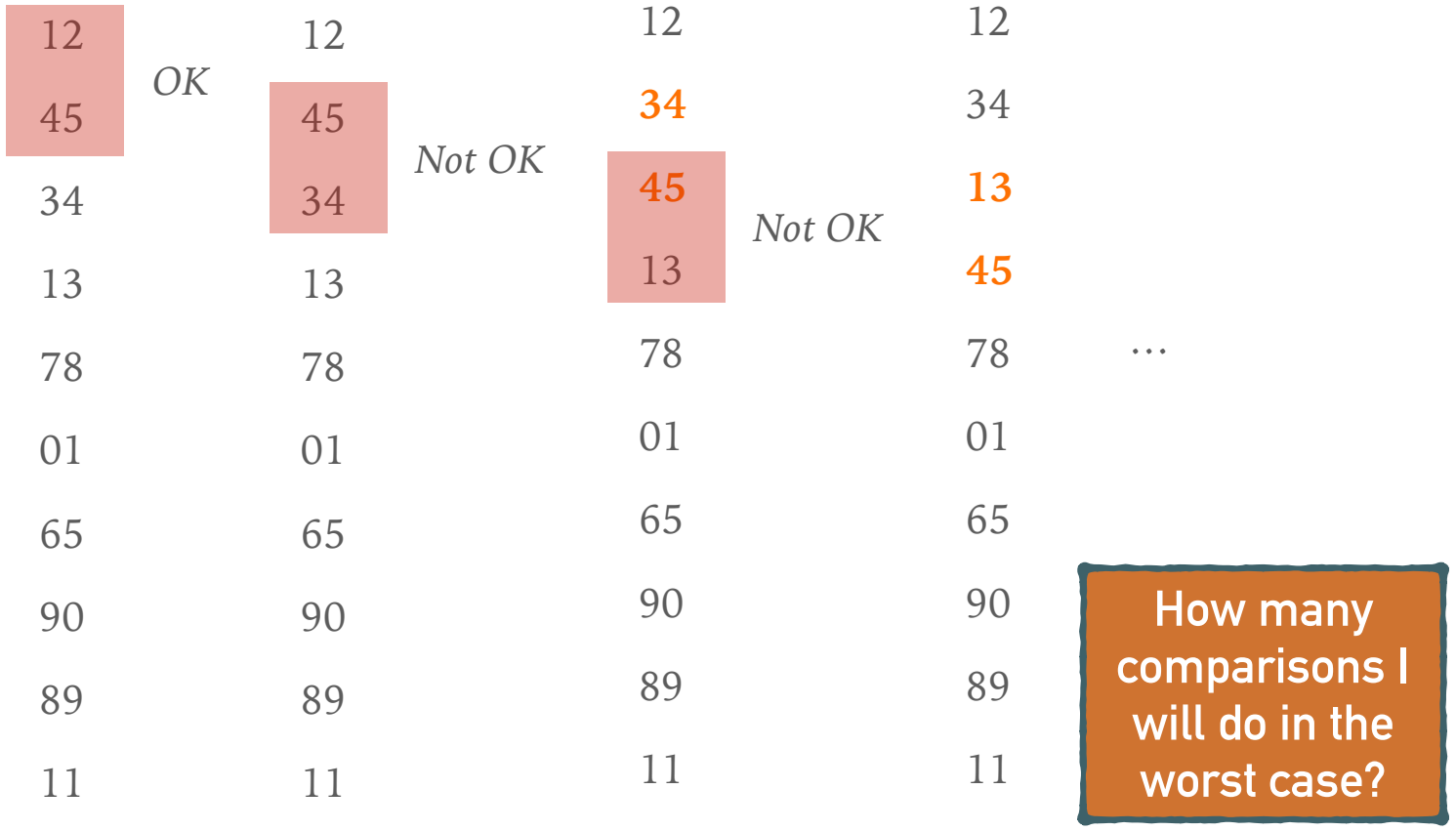
65

90

89

11

ALGORITHM 01



ALGORITHM 02 – QUICK SORT

12

45

34

13

78

01

65

90

89

11

How many comparisons I will do in the worst case?

DEFINITION

Formal

- Definition 1.2: Big-O notation. An algorithm has a complexity $f(n) = O(g(n))$ if there exist positive constants n0 and c such that $\forall n > n0$, $f(n) \leq c \cdot g(n)$.

Casual

- f (n) is upper bounded by the function g(n)
- From a certain point f(n) always loses

DEFINITION

Formal

- **Definition 1.3 Polynomial-time algorithm.** *An algorithm is a polynomial-time* algorithm if its complexity is $O(p(n))$, where $p(n)$ is a polynomial function of n.
- **Definition 1.4 Exponential-time algorithm.** An algorithm is an exponential-time algorithm if its complexity is $O(c^n)$, where c is a real constant strictly superior to 1.

Casual

- Polynomial - piece of cake
- Exponential - OH My GOD! I have no memory, no time …

EXAMPLE EXPONENTIAL ALGORITHM

TABLE 1.3 Search Time of an Algorithm as a Function of the Problem Size Using Different Complexities (from [299])

Complexity	Size = 10	Size = 20	Size = 30	Size = 40	Size = 50
$O(x)$	0.00001 s	0.00002 s	0.00003 s	0.00004 s	0.00005 s
$O(x^2)$	0.0001 s	0.0004 s	0.0009 s	0.0016 s	0.0025 s
$O(x^5)$	0.1 s	0.32 s	24.3 s	1.7 mn	5.2 mn
$O(2^x)$	0.001 s	1.0 s	17.9 mn	12.7 days	35.7 years
$O(3^x)$	0.059 s	58.0 mn	6.5 years	3855 centuries	2×10^8 centuries

M. Garey and D. Johnson. *Computers and Intractability: A Guide to the Theory on NP-Completeness*. W. H. Freeman and Co. Publishers, New York, 1979.

COMPLEXITY OF A PROBLEM

SS<->SS;

WORK IN PAIRS AND ANSWER THESE QUESTIONS!

- What is a P class?
- What is a Non Deterministic Algorithm?
  - Can you give me an example?
- What is a NP class?
- What is a NP-HARD class?
- What is a NP-complete Class?

COMPLEXITY OF A PROBLEM

- **The complexity of a problem is equivalent to the complexity of the best algorithm TO SOLVE that problem.**
- A problem is tractable (or easy) if there exists a polynomial-time algorithm to solve it.
- A problem is intractable (or difficult) if no polynomial-time algorithm exists to solve the problem.

- Attention!!! This is only a definition!!
- An algorithm that run in time **10^7 n** is nor efficient neither easy, although it is linear
- The same reasoning applies to O(n^10) algorithms

COMPLEXITY OF A PROBLEM

- P - Problems which may be solved by a deterministic polinomial algorithm
- NP - which may be solved by a non-deterministic polinomial algorithm
- NP-HARD: X is an NP-HARD problem if every problem in NP is polynomially deductible to X ("at least as hard as the hardest problems in NP")
- NP-COMPLETE: X is NP-COMPLETE if
  - (1) X belongs to NP and
  - (2) X is NP-HARD

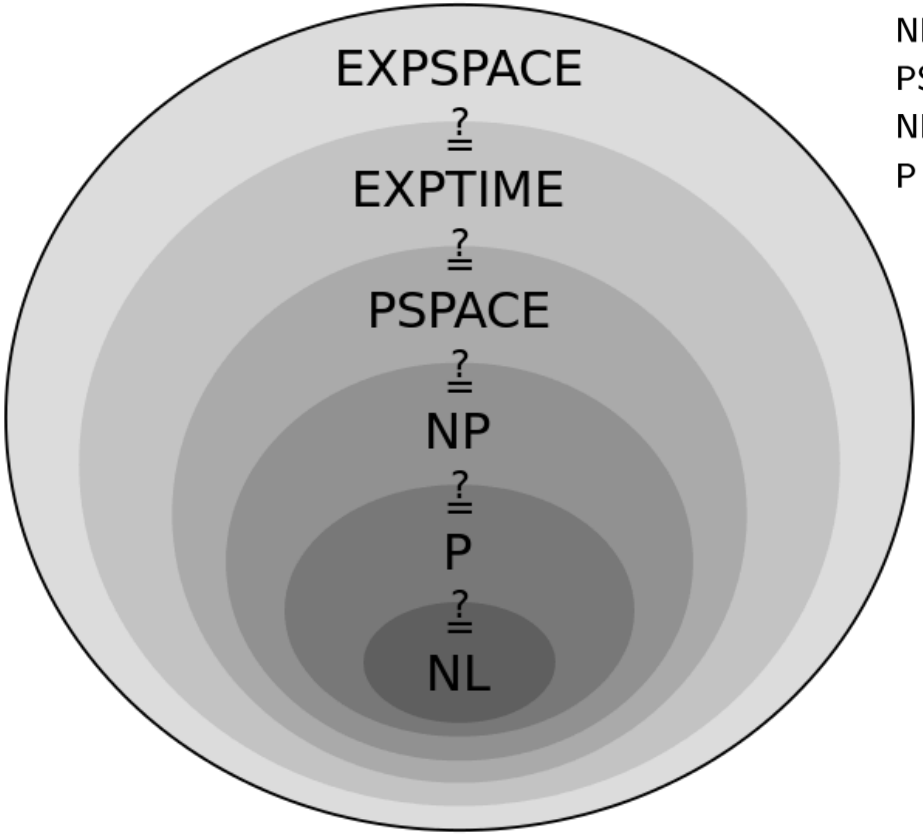
NON DETERMINISTIC ALGORITHM

Given a set of N objects. Each object O has a specified weight and a specified value. Given a capacity, which is the maximum total weight of the knapsack, and a quota, which is the minimum total value that one wants to get. The 0–1 knapsack decision problem consists in finding a subset of the objects whose total weight is at most equal to the capacity and whose total value is at least equal to the specified quota.

Algorithm 1.1 Nondeterministic algorithm for the knapsack problem.

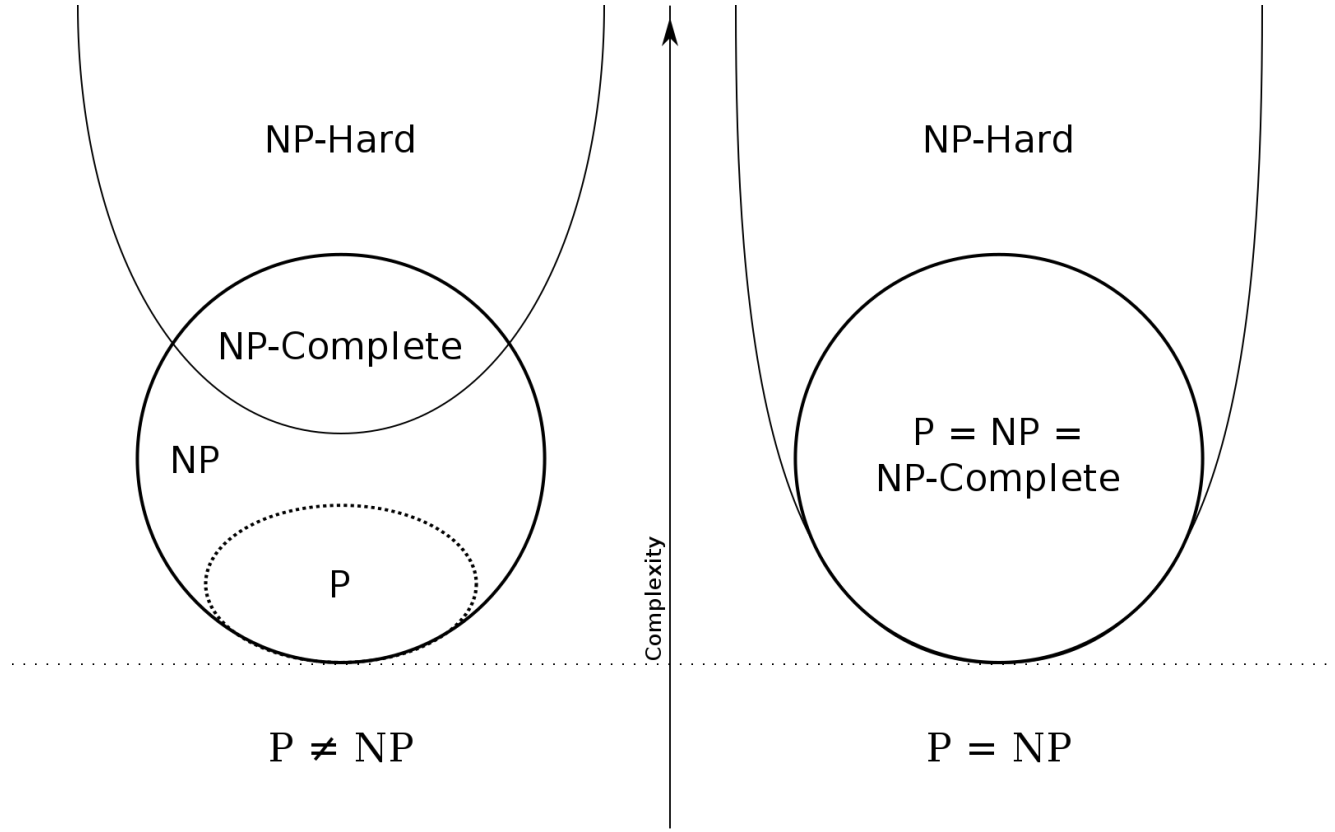
```
Input OS : set of objects ; QUOTA : number ; CAPACITY : number.
Output S : set of objects ; FOUND : boolean.
  S = empty ; total_value = 0 ; total_weight = 0 ; FOUND = false ;
  Pick an order L over the objects ;
  Loop
    Choose an object O in L ; Add O to S ;
    total_value = total_value + O.value ;
    total_weight = total_weight + O.weight ;
    If total_weight > CAPACITY Then fail
    Else If total_value ≥ QUOTA
         FOUND = true ;
         succeed ;
    Endif Endif
    Delete all objects up to O from L ;
  Endloop
```

COMPLEXITY OF A PROBLEM

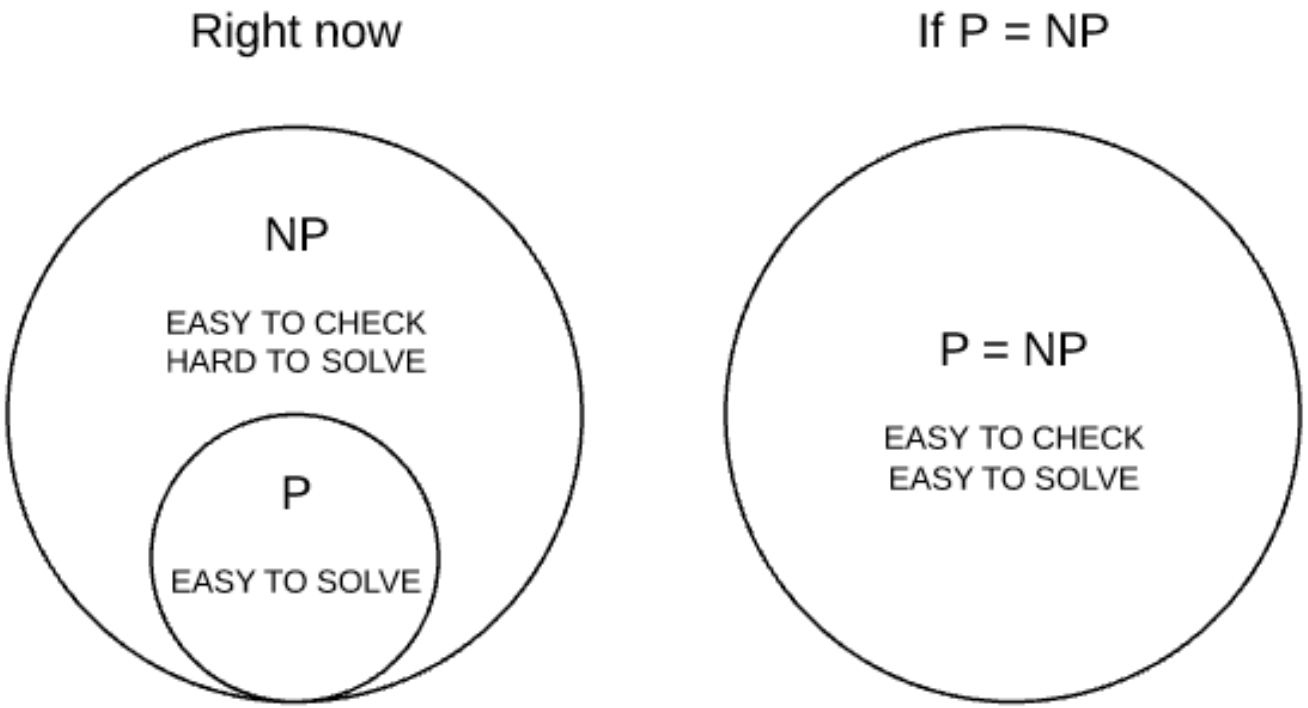


$NL \subseteq P \subseteq NP \subseteq PH \subseteq PSPACE$
$PSPACE \subseteq EXPTIME \subseteq EXPSPACE$
$NL \subsetneq PSPACE \subsetneq EXPSPACE$
$P \subsetneq EXPTIME$

COMPLEXITY OF A PROBLEM



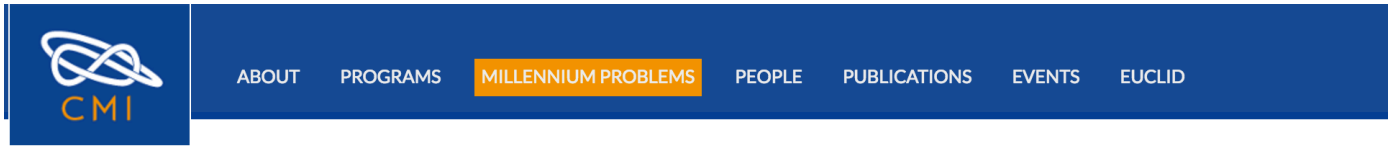
COMPLEXITY OF A PROBLEM



EXEMPLES OF PROBLEM'S COMPLEXITY

- Class P.
  - Some classical problems belonging to class P are minimum spanning tree, shortest path problems, maximum flow network, maximum bipartite matching, and linear programming continuous models.
- Class NP complete:
  - Sequencing and scheduling problems such as flow-shop scheduling, job-shop scheduling, or open-shop scheduling.
  - Assignment and location problems such quadratic assignment problem(QAP), generalized assignment problem (GAP), location facility, and the p-median problem.
  - Grouping problems such as data clustering, graph partitioning, and graph coloring.
  - Routing and covering problems such as vehicle routing problems (VRP), set cov- ering problem (SCP), Steiner tree problem, and covering tour problem (CTP).
  - Knapsack and packing/cutting problems, and so on.

P = NP ?



Millennium Problems

Yang–Mills and Mass Gap

Experiment and computer simulations suggest the existence of a "mass gap" in the solution to the quantum versions of the Yang-Mills equations. But no proof of this property is known.

Riemann Hypothesis

The prime number theorem determines the average distribution of the primes. The Riemann hypothesis tells us about the deviation from the average. Formulated in Riemann's 1859 paper, it asserts that all the 'non-obvious' zeros of the zeta function are complex numbers with real part 1/2.

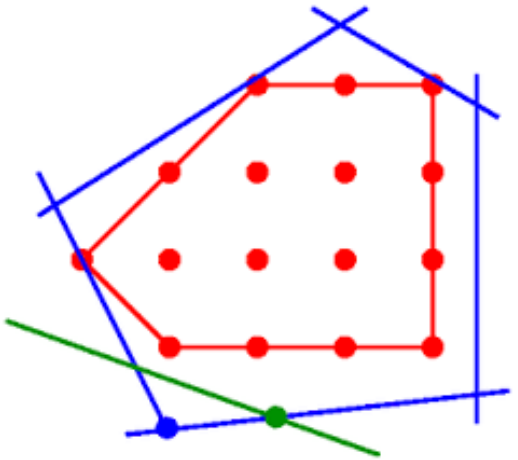
P vs NP Problem

If it is easy to check that a solution to a problem is correct, is it also easy to solve the problem? This is the essence of the P vs NP question. Typical of the NP problems is that of the Hamiltonian Path Problem: given N cities to visit, how can one do this without visiting a city twice? If you give me a solution, I can easily check that it is correct. But I cannot so easily find a solution.

Navier–Stokes Equation

This is the equation which governs the flow of fluids such as water and air. However, there is no proof for the most basic questions one can ask: do

INTEGER PROGRAMMING



- Integer programming?
  - P or NP?

A matrix A is totally unimodular if the determinant of each square submatrix of A is 0, 1, or +1.

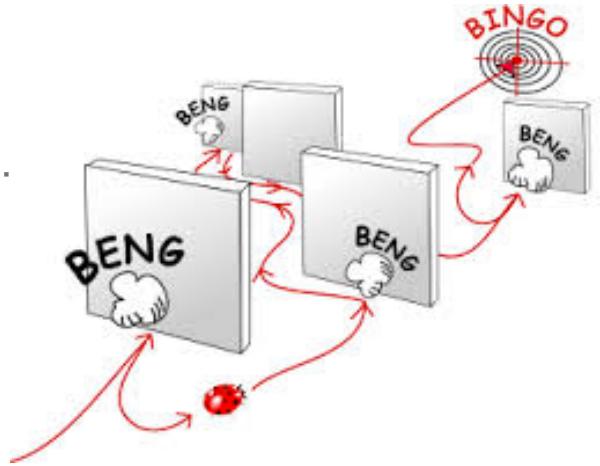
Theorem 1: If A is totally unimodular, then every vertex solution of $Ax >= b$ is integer.

- Conclusion ?
- !&#¥©ßð¬ˆø
- Search for theory of Matroids too …

HEURISTICS' CLASSIFICATION

BASIC HEURISTICS

BALL, M., and M. MAGAZINE (1981)

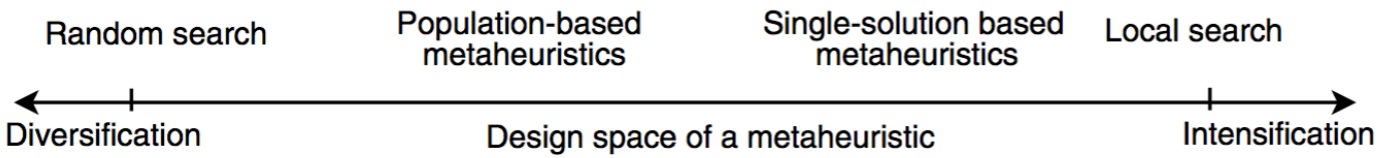


- CONSTRUCTION
- IMPROVEMENT
- MATHEMATICAL PROGRAMMING
- DECOMPOSITION
- PARTITIONING
- RESTRICTION OF FEASIBLE SOLUTION SPACE
- RELAXATION OF FEASIBLE SOLUTION SPACE

SS<->T; YOU READ THE ARTICLE. CAN YOU EXPLAIN THEM TO ME?

INTENSIFICATION – DIVERSIFICATION PARADIGM

- In designing a metaheuristic, two contradictory criteria must be taken into account: exploration of the search space (diversification) and exploitation of the best solutions found (intensification).
- Promising regions are determined by the obtained "good" solutions. In intensification, the promising regions are explored more thoroughly in the hope to find better solutions. In diversification, nonexplored regions must be visited to be sure that all regions of the search space are evenly explored and that the search is not confined to only a reduced number of regions.
- Random search Population-based Single-solution based Local search metaheuristics metaheuristics.



CLASSIFICATION OF METAHEURISTICS

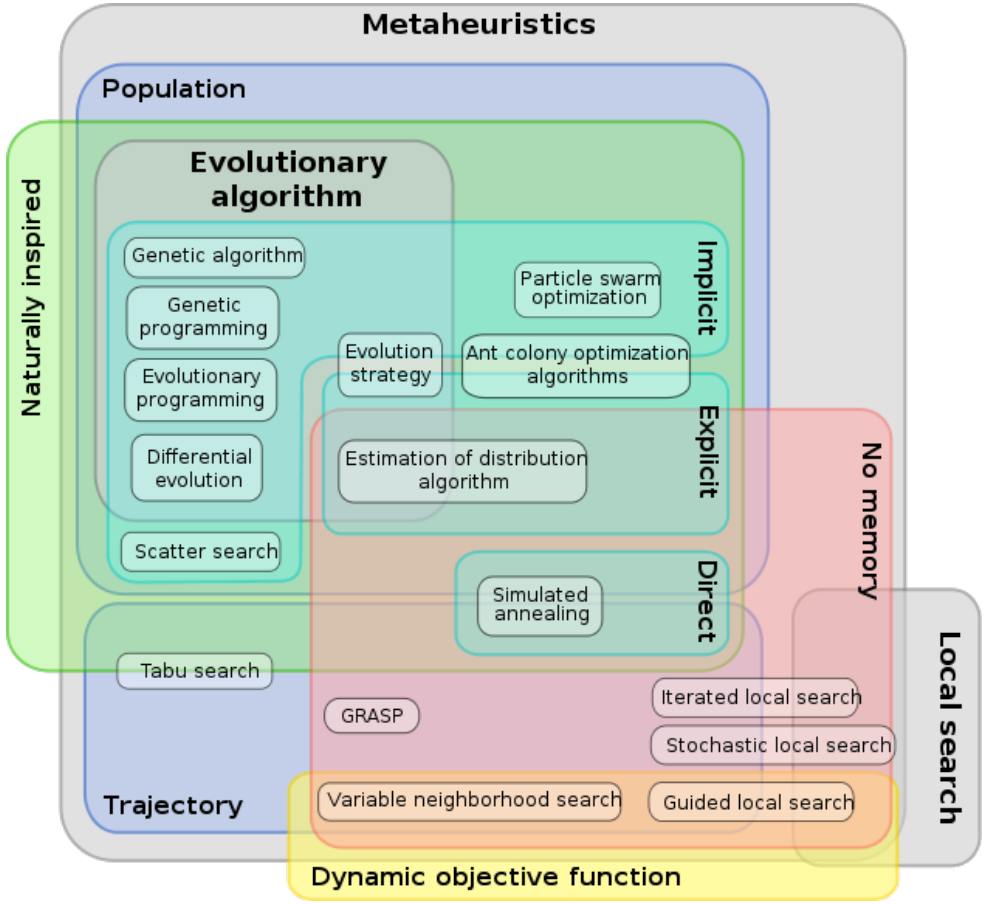
- **Nature inspired versus non nature inspired:** Many metaheuristics are inspired by natural processes: evolutionary algorithms and artificial immune systems from biology; ants, bees colonies, and particle swarm optimization from swarm intelligence into different species (social sciences); and simulated annealing from physics.
- **Memory usage versus memoryless methods:** Some metaheuristic algorithms are memoryless; that is, no information extracted dynamically is used during the search. Some representatives of this class are local search, GRASP, and simulated annealing. While other metaheuristics use a memory that contains some information extracted online during the search. For instance, short-term and long-term memories in tabu search.

CLASSIFICATION OF METAHEURISTICS

- **Deterministic versus stochastic:** A deterministic metaheuristic solves an optimization problem by making deterministic decisions (e.g., local search, tabu search). In stochastic metaheuristics, some random rules are applied during the search (e.g., simulated annealing, evolutionary algorithms). In deterministic algorithms, using the same initial solution will lead to the same final solution, whereas in stochastic metaheuristics, different final solutions may be obtained from the same initial solution.
- **Iterative versus greedy:** In iterative algorithms, we start with a complete solution (or population of solutions) and transform it at each iteration using some search operators. Greedy algorithms start from an empty solution, and at each step a decision variable of the problem is assigned until a complete solution is obtained.

CLASSIFICATION OF METAHEURISTICS

- **Population-based search versus single-solution based search:** Single-solution based algorithms (e.g., local search, simulated annealing) manipulate and transform a single solution during the search while in population-based algorithms (e.g., particle swarm, evolutionary algorithms) a whole population of solutions is evolved.



I picked this image from wikipedia!! At least I referenced it!

(S)
HOME WORK

LAB 1

Chose three of the following problems:

- Knapsack Problem
- Single Machine
- Parallel Machines
- Travelling Salesman Problem
- Location and Allocation Problem

(1) Provide a Mathematical formulation;
(2) Identify the complexity of the problem;
(3) Propose a constructive heuristic;
(4) Classify your method into diversification and intensification;

(Deliver before Class 3).