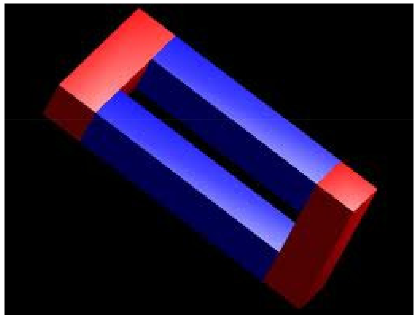
Parallel Experiences

Multiplicity Computing and SBSE



University of Málaga, SPAIN

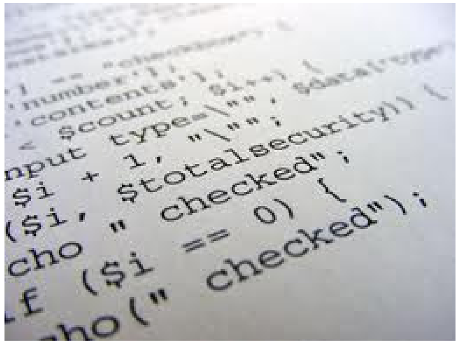
Enrique Alba

Parallel Experiences

Index

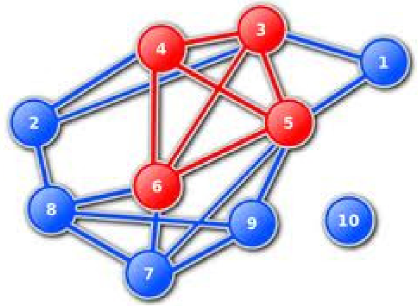
Introduction: SEP's Modeled as Search Problems

THE PROBLEM

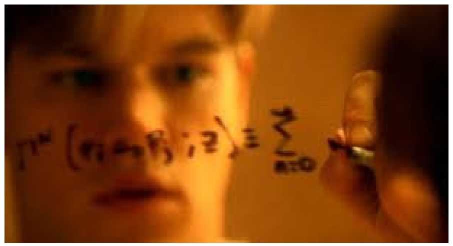


MODELING PHASE

D/C SEARCH



ELEMENTS



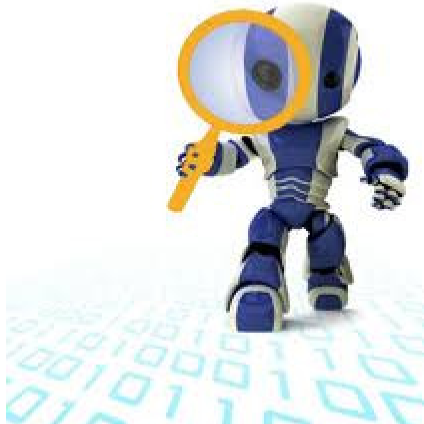
KNOWLEDGE

$$\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}$$

$$f(x) = a_0 + \sum_{n=1}^{\infty}\left(a_n \cos\frac{n\pi x}{L} + b_n \sin\frac{n\pi x}{L}\right)$$

$$x = \frac{-b \pm \sqrt{b^2-4ac}}{2a}$$

QUANTIFICATION



SOLVER

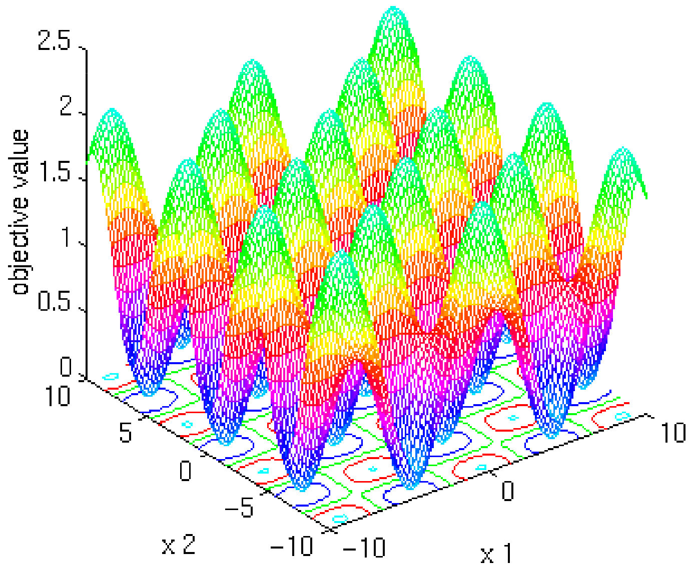
Parallel Experiences

Introduction: THE Problem

Index

- **Objective of a global search problem:**

$$f(\vec{x}) \rightarrow max: \quad find\ a\ vector \quad \vec{x}^*$$
$$such\ that\ \forall \vec{x} \in M : f(\vec{x}) \leq f(\vec{x}^*) := f^*$$



- **Minimizing is also possible**
- **Vectors can map to other data structures**

Parallel Experiences

Index

Introduction: Taxonomy of Search Algorithms

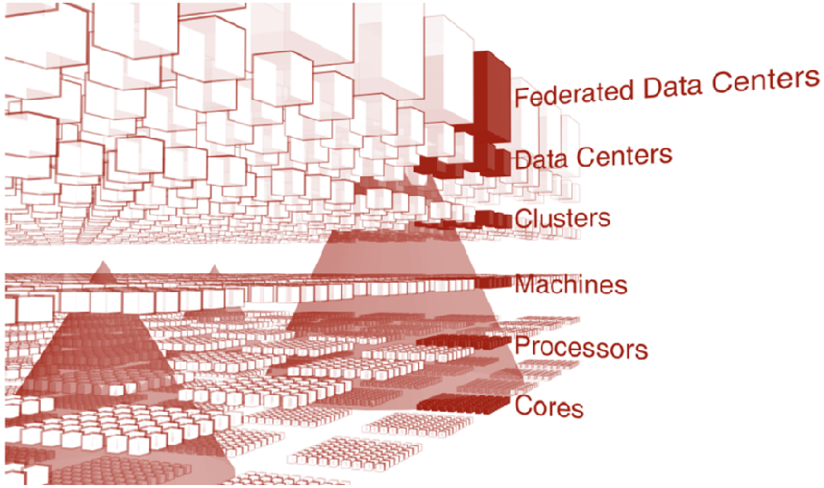
- Optimization Algorithms
  - Exact
    - Calculus
      - Direct
        - Newton
        - Greedy
      - Indirect
    - Enumerative
      - DP
      - B&B
  - Ad-hoc Heuristic
  - Metaheuristic
    - Trajectory
      - SA
      - VNS
      - TS
    - Population
      - EA
      - ACO
      - PSO

• nature inspired in red

Parallel Experiences

Index

Introduction: Multiplicity Computing



The idea: multiple levels ot technology

The resources: computation, communication, storage

The approach: take all of them into account in research

The requirements: be measurable and predictable

The goals: improved scalability, quality, dependability, security

How does this relate to parallel solvers?

Parallel Experiences

Introduction: Parallel Solvers for Complex Problems

- **Basic components of parallel Solvers are:**
  - **Nodes performing separate search**
  - **Communication pattern among the nodes**
  - **Policy of the search (start, end, solution…)**

Ex1: Parallel ACO
- ACO nodes
- All to all
- Exchange pheromone matrix

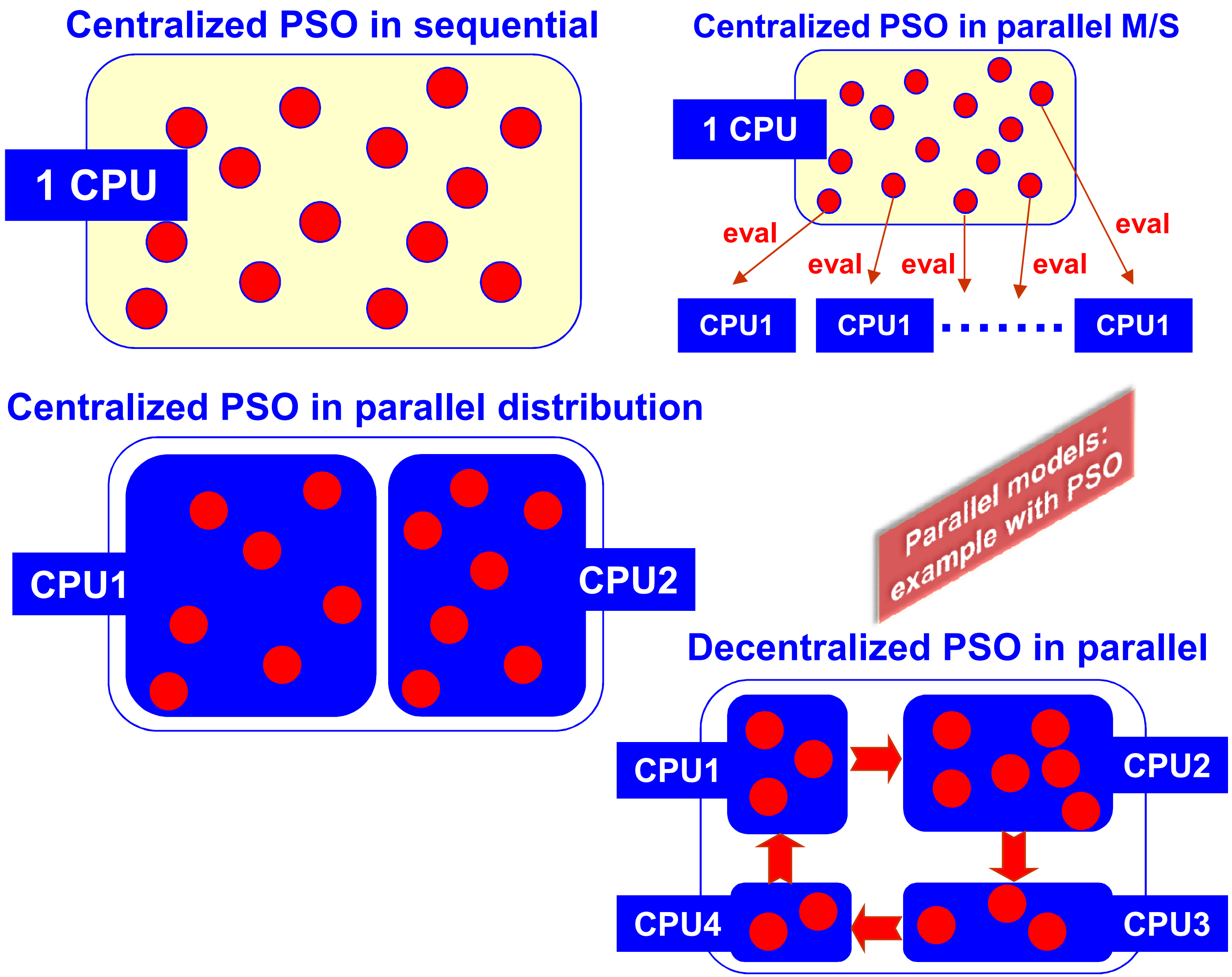
Ex3: Parallel EA
- Subpops. of partial solutions
- Static ring
- Exchange random selected sols.

Ex2: Parallel SA
- SA nodes
- Random target
- Exchange actual best solution

Parallel Experiences

Index

Introduction: Parallel Solvers for Search

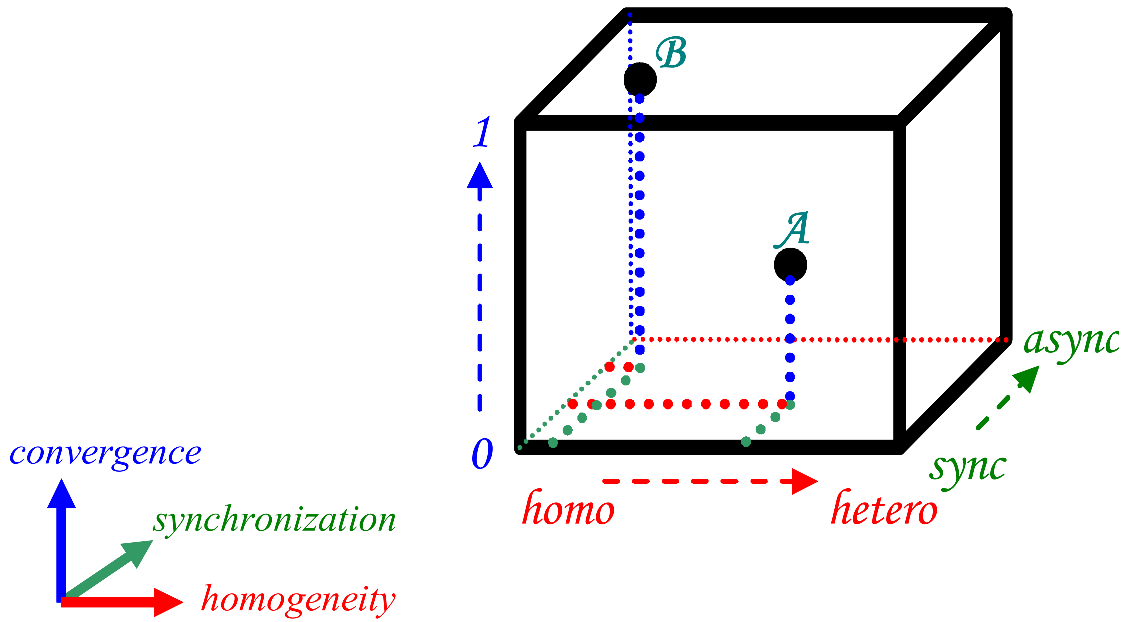


Parallel Experiences

Index

Introduction: Taxonomy

- **Three features:**



- **Other important issues are:**
  - **Node granularity**
  - **Central memory**
  - **Static/Dynamic features**
  - **…**

Parallel Experiences

Index

Research with Parallel Solvers

Working with parallel solvers is far from trivial:

1. Must know on concurrency and parallel theory
2. Must know on parallel software languages and tools
3. Must know on hardware realizations for communication
4. Must know on communication protocols and networks
5. Programming is more error prone than in sequential
6. A parallel solver is hard to analyze
7. After all the work, maybe you are not gaining in time
8. Many people just don’t want to know on parallel issues
9. ...

Parallel Experiences

Index

Research with Parallel Solvers

...but it is worthwile!

Problems not solved before become now solvable by using parallel solvers

Parallel Experiences

Index

Facts in Research with Parallel Solvers

Facts in parallel search:

❶ **Model & implementation are different**

❷ **Metrics need a revision**

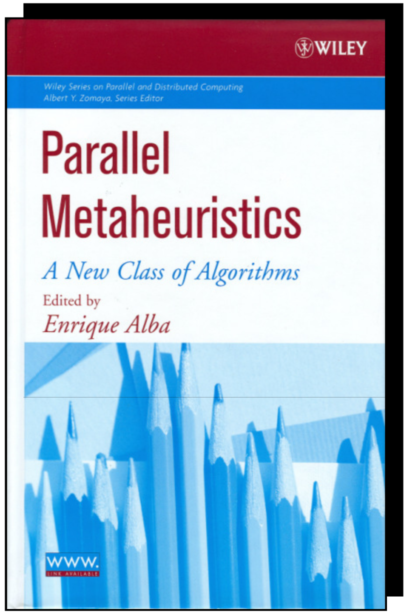
❸ **Superlinear speedup is a fact**

❹ **Heterogeneity is a must nowadays**

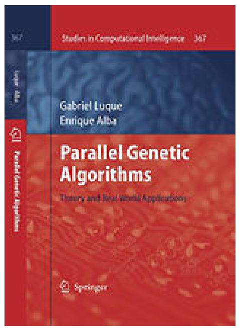
❺ **The experimental setup is important**

❻ **Algorithms are Software**

❼ **Other facts**



E. Alba
Parallel Metaheuristics: A New Class of Algorithms
Wiley, ISBN 0-471-67806-6, July 2005



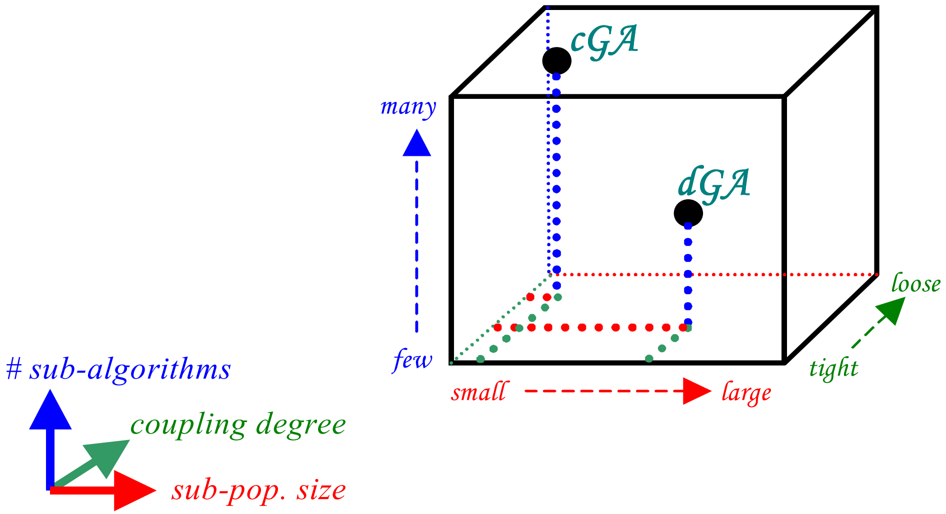
G. Luque, E. Alba
Parallel Genetic Algorithms: Theory and RW Applications
Springer-Verlag, ISBN 978-3-642-22083-8, July 2011

Parallel Experiences

Fact 1: Model and Implementation are Different

Node in a decentralized EA

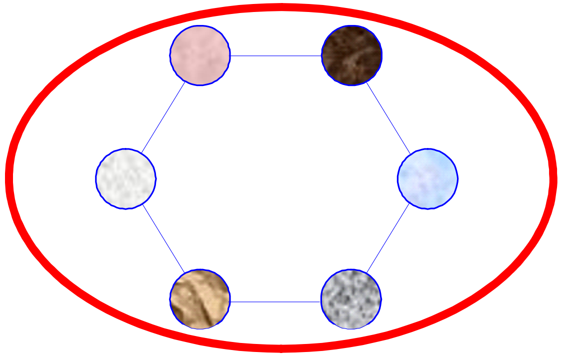
```
① Generate initial population
② Evaluate present evaluation
③ While not stop criterion do:
       ③① Select partners
       ③② Apply variation operators
       ③③ Communication with neighbors
       ③④ Replace old solutions by the new ones
       ③⑤ Compute statistics and performances
```



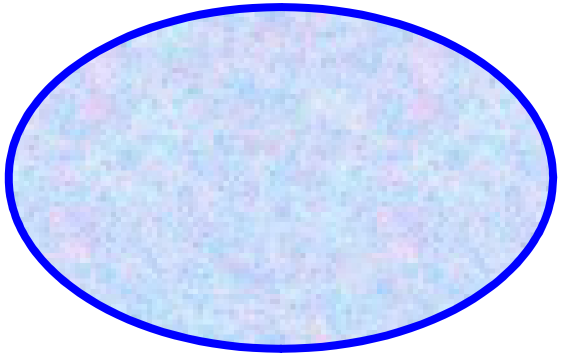
Parallel Experiences

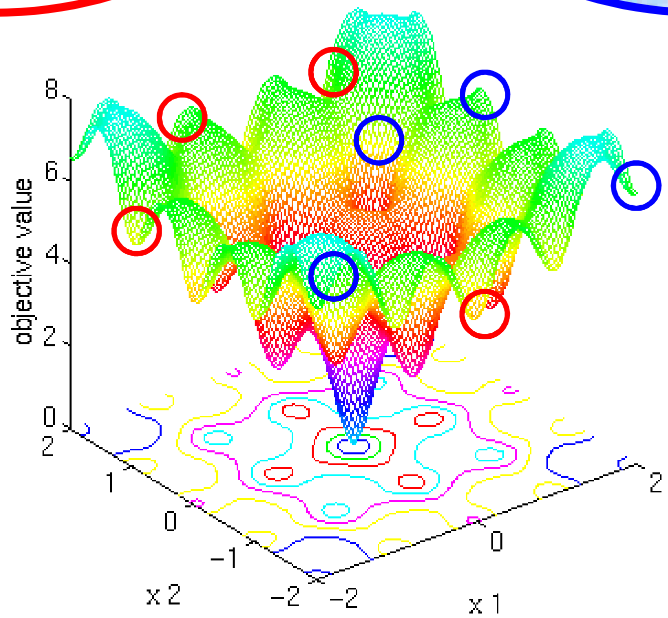
Fact 1: Model and Implementation are Different

Decentralized Model



Centralized Model



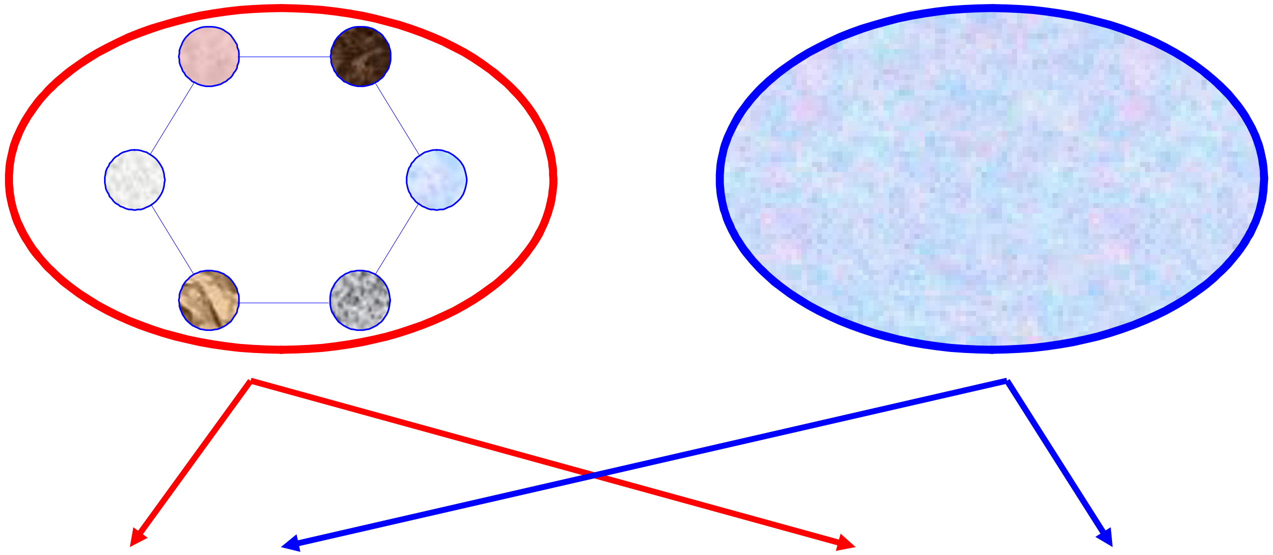


Parallel Experiences

Index

Fact 1: Model and Implementation are Different

Decentralized Model **Centralized Model**



IMPLEMENTATION-IMPLEMENTATION-IMPLEMENTATION

1 CPU

Manycore Cluster

Fact 2: Metrics Need a Revision

- **Exact and approximate algorithms become different algorithms when run in parallel**
- **Specially, parallel nondeterministic algorithms can stop at solutions of very different quality**
- **Comparing times against the sequential version could be meaningless if the two algorithms are not exactly the same or the final solution quality is different**
- **A clear example is the speedup (efficiency)**

A taxonomy is needed

Alba E. (2002) "Parallel EAs Can Achieve SuperLinear Performance". *Information Processing Letters*, Elsevier, 82(1):7-13

I. Strong Speedup

II. Weak Speedup

A. Speedup with solution-stop

1. Versus Panmixia

2. Orthodox

B. Speedup with predefined effort

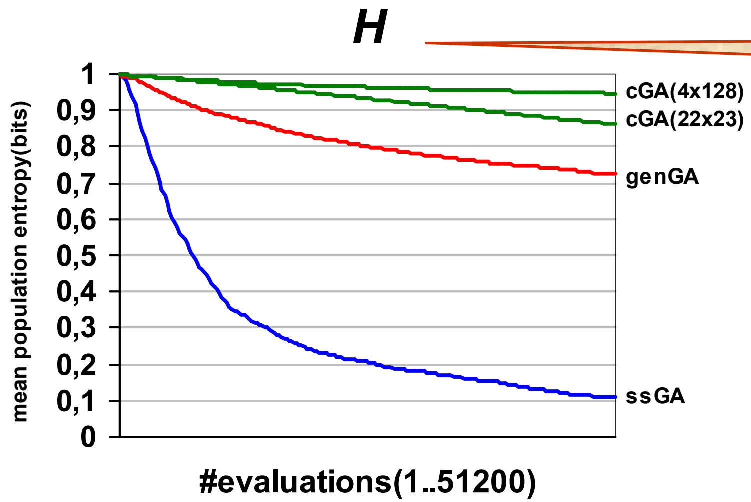
$$s_m = \frac{\overline{T_1}}{\overline{T_m}}$$

Parallel Experiences

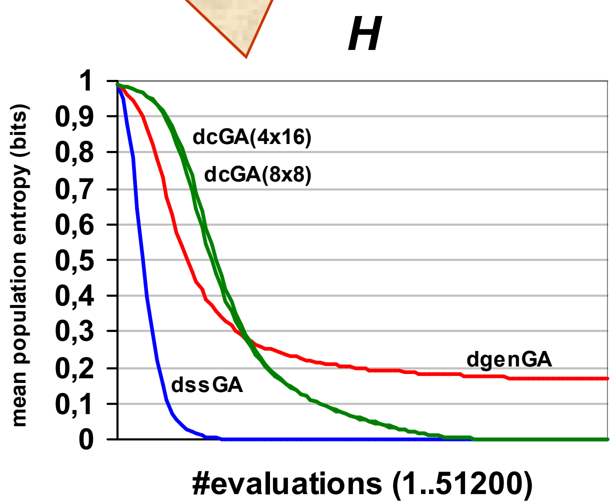
Fact 2: Advanced Metrics (Entropy)

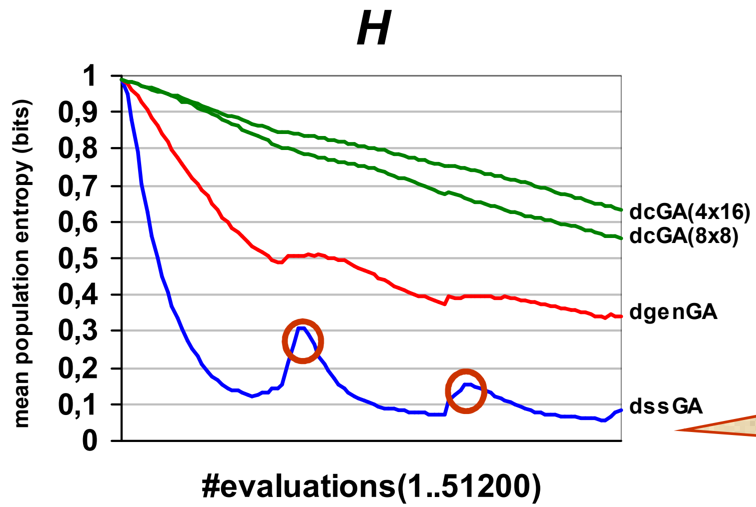
SPH16-32
8×64 individuals
DPX1(p_c=1.0)
MUT(p_m=1/l)

Non Distributed



Distributed Sync (ζ=1)





Distributed Sync (ζ=32)

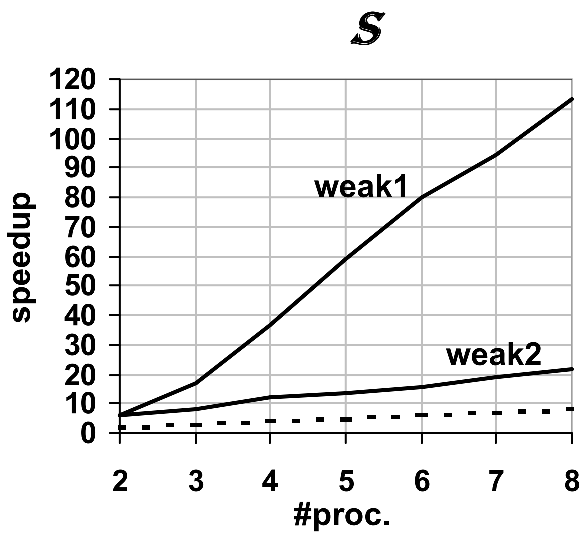
Parallel Experiences

Index

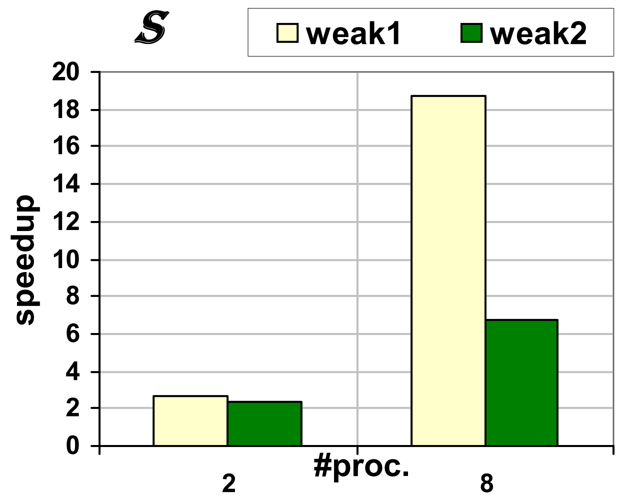
Fact 3: Superlinear Speedup is a Fact

$$s(n_{proc}) = \frac{\overline{T}_1}{\overline{T}_{n_{proc}}}$$

weak1: against panmixia (1 proc)
weak2: only changing n_{proc}



SPH16-32



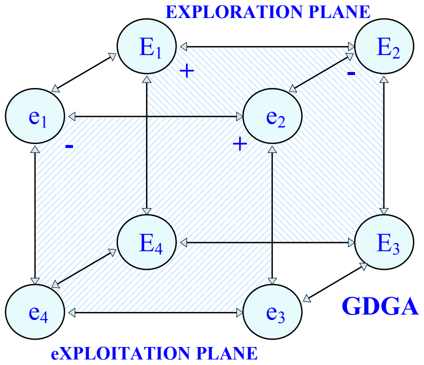
SSS128

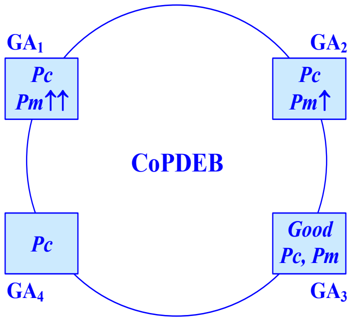
Parallel Experiences

Index

Fact 4: Heterogeneity is a Must Nowadays

Algorithmic components could be heterogeneous



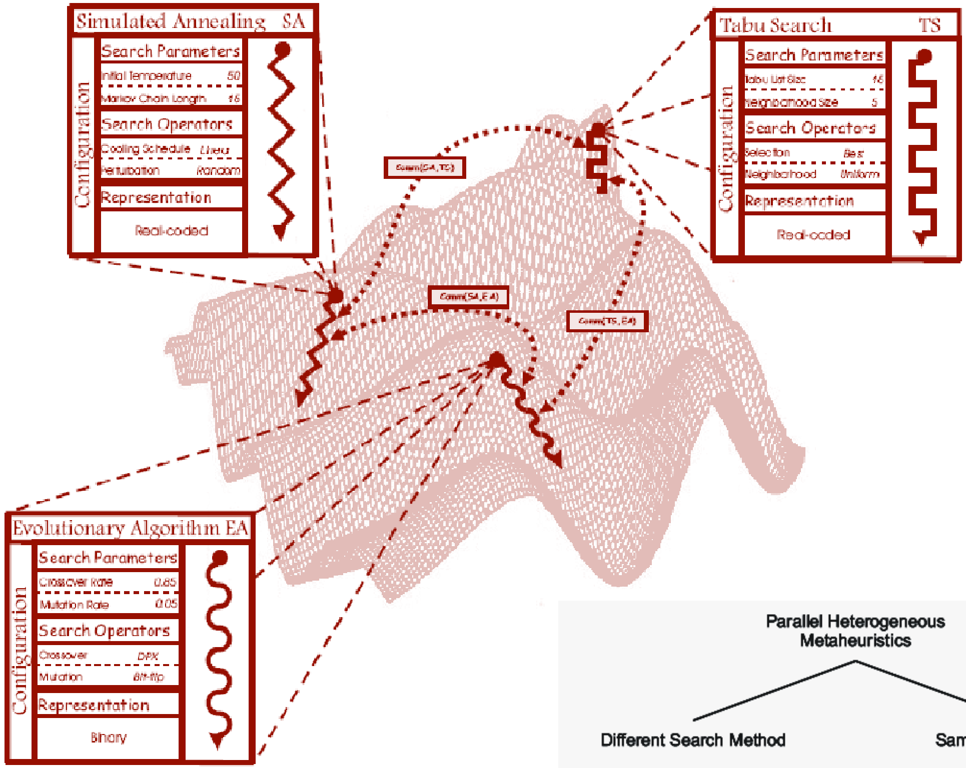


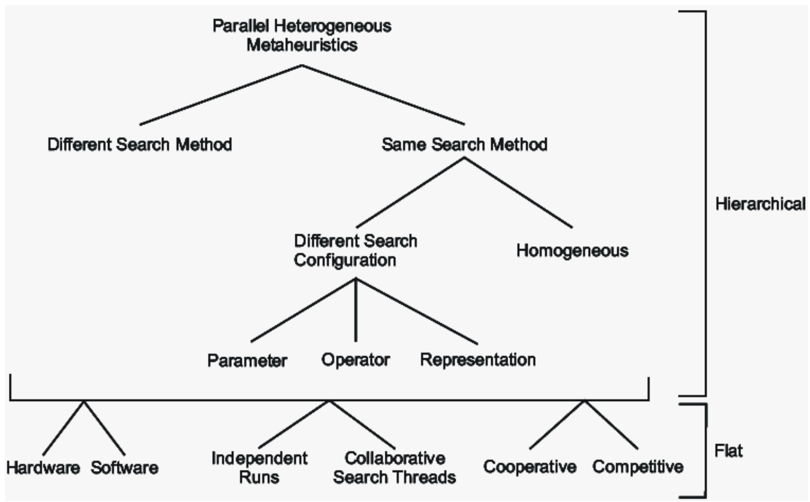
Parallel hardware could be heterogeneous

Parallel Experiences

Index

Fact 4: Heterogeneity is a Must Nowadays





Parallel Experiences

Index

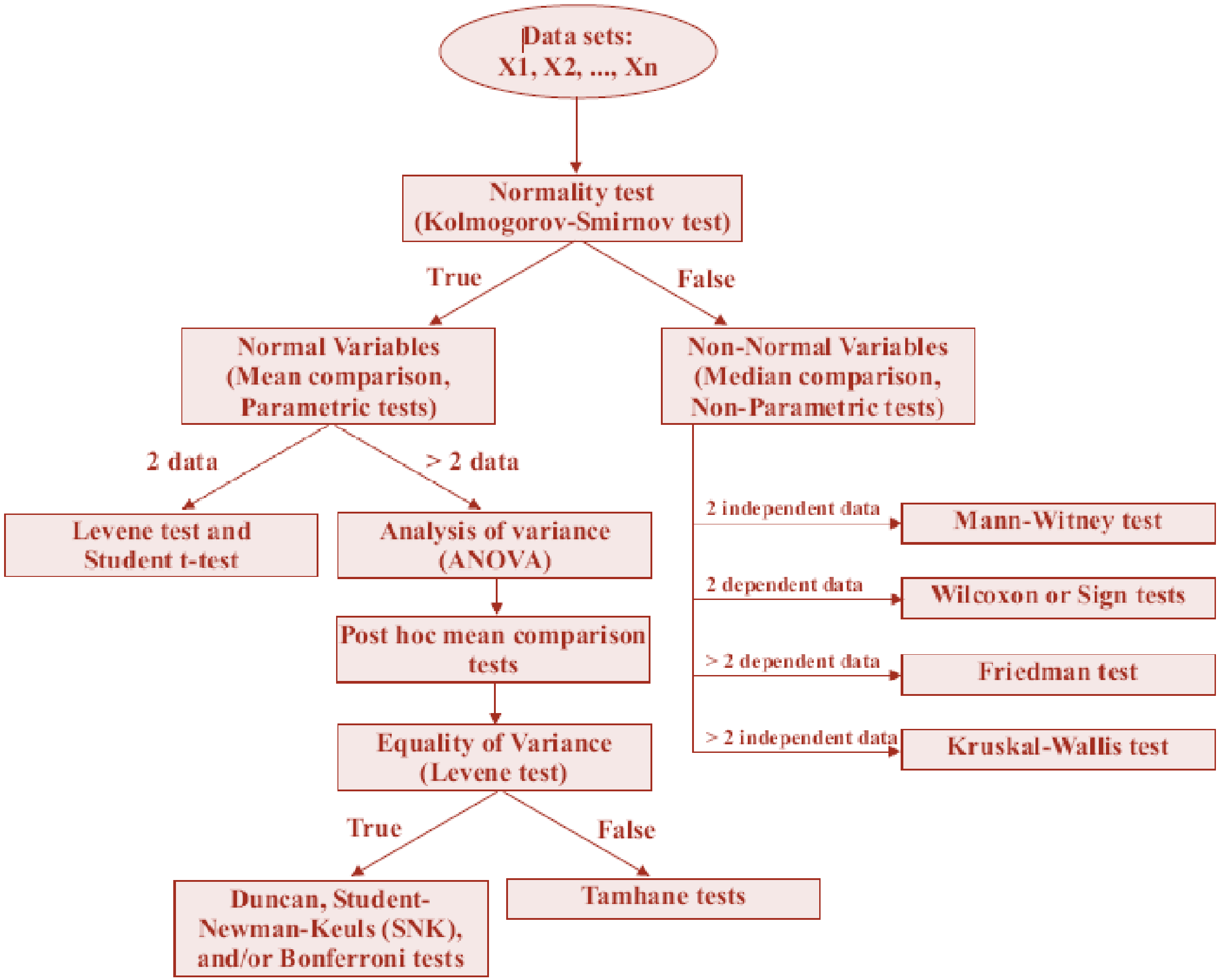
Fact 5: The Experimental Setup is Important

Parallel heuristic/exact methods have often nondeterministic behaviors, so:

1. **Multiple independent runs are necessary**
2. **Statistical hypothesis tests must be used: Student *t*-test, ANOVA, Wilcoxon, …**
3. **Average, max and min values need to be reported, but: is that all? Run times?**
4. **Advises: more than 30 ind. runs, give all the parameters in one table, always report on times, give hardware and software used, …**

Parallel Experiences

Fact 5: The Experimental Setup is Important: Guidelines



Parallel Experiences

Index

Fact 6: Algorithms Are Software

We should worry about the design of algorithms

- **At the end, they are software pieces (!)**
- **Take care of the design and take care of the documentation**
- **Efficiency is usually the more important issue**

Traditional tools

- **Fortran, C/C++, Java, Haskell, ...**
- **Generalization is in conflict with efficiency**
- **Apply well-accepted recomms. on GOTO, global vars, ...**

Object Orientation

- **Present best issue (long term development and design)**
- **Efficiency can be modulated**
- **Parallel software tools are developed nowadays**
- **Java versus C++**

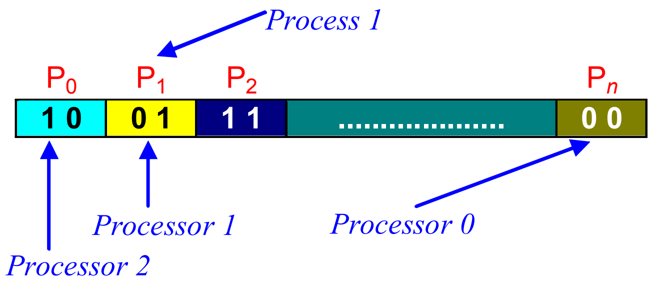
Alba E., Troya J.M., Gaining New Fields of Application for OOP: the Parallel Evolutionary Algorithm Case, *The Journal of Object-Oriented Programming*, December 2001

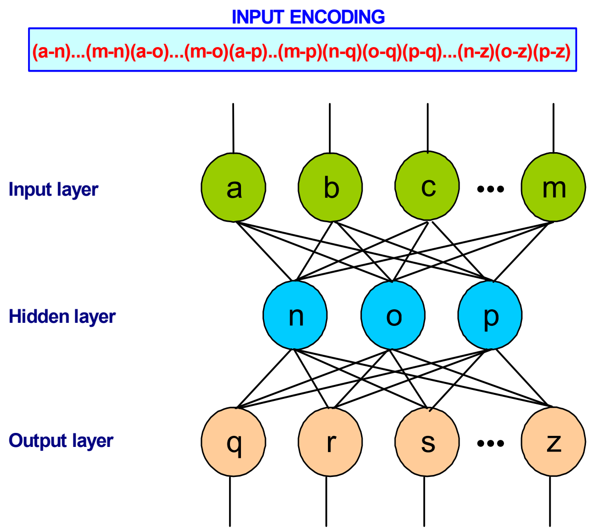
Parallel Experiences

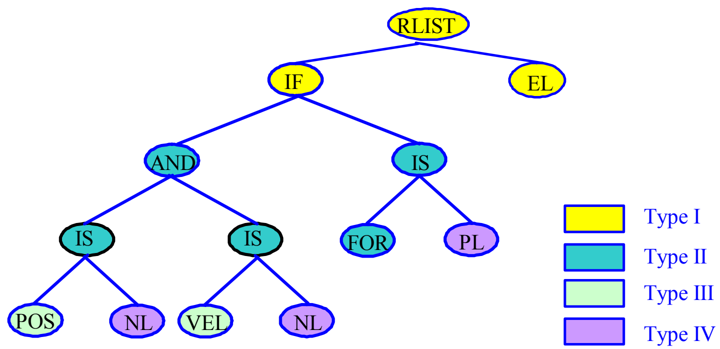
Index

Fact 6: Algorithms Are Software

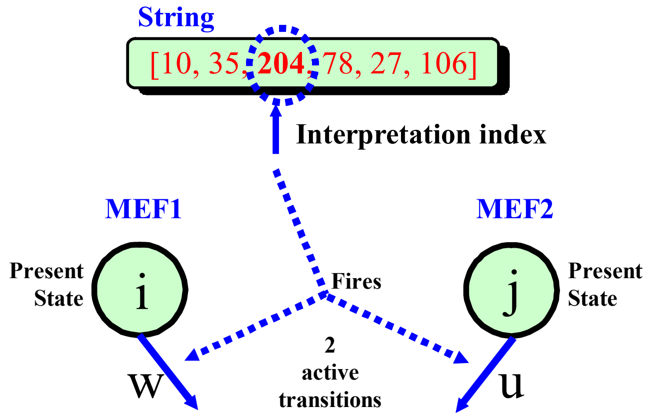
Complex Data Structures for Complex Problems







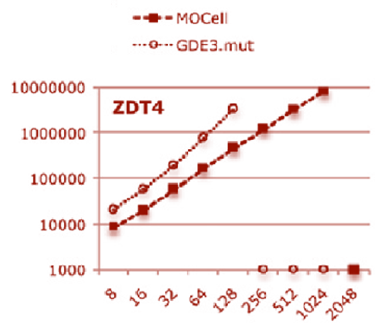
IF *pos* IS NL AND *vel* IS NL THEN *for* IS PL

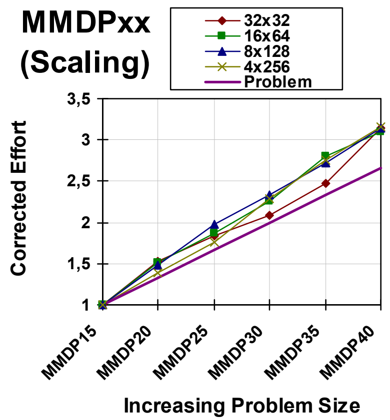


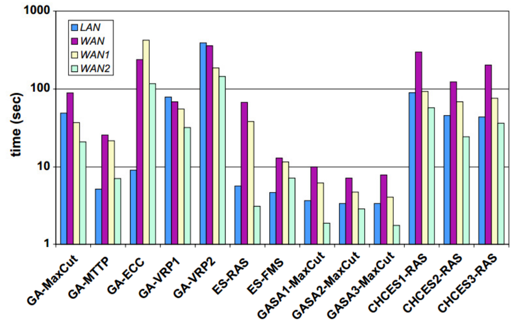
Parallel Experiences

Relevant Facts for Multiplicity Computing

- **Scalability:** Numerical ………….. vs. Physical

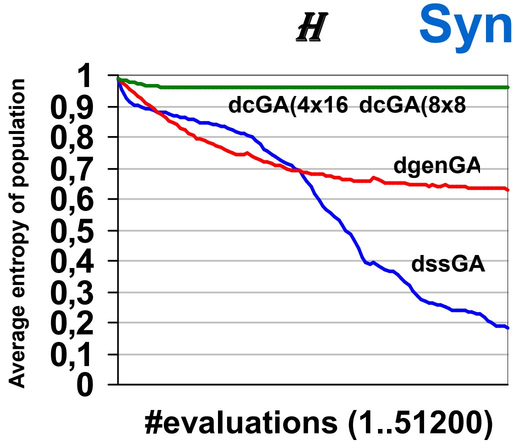


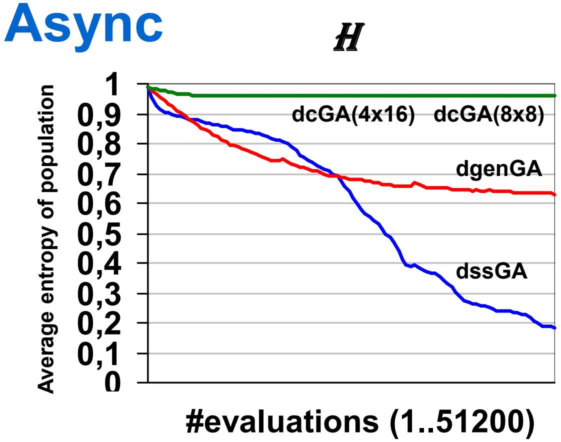




- **Asynchronism:**

Sync vs. Async





Parallel Experiences

Index

Other Facts

- **Theory is hard but important:** convergence, time complexity, landscape theory, math oracles, …

- **Knowledge exchange between fields helps:** developing common frameworks for grid algorithms, exact plus heuristic issues, …

- **Parallel algorithms are not always better:** communication overhead, numerical search could progress slowly, …

- **Be always ready for new facts!**

Parallel Experiences

Index

Software for Parallelism

There exist multiple software tools to develop parallel applications:

- **Low level tools: Sockets, …**
- **High Level Comm Libraries: PVM, MPI, …**
- **Language Embedded: Java RMI, …**
- **Middleware: CORBA, MANIFOLD, MR, …**
- **Metacomputing Systems: Globus, Condor, BOINC, …**
- **Related to Internet: .NET, SOAP, XML, …**
- **Plateform specific: CUDA, OpenCL, Handle C, …**
- **Others: OpenMP, HPF, …**

Parallel Experiences

Our Software

http://neo.lcc.uma.es/software/index/

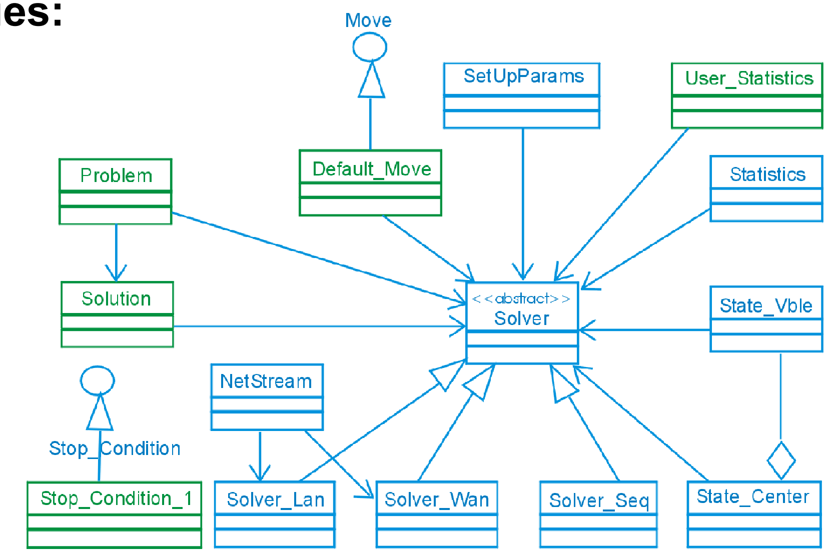
Parallel Experiences

Index

MALLBA

- **Library for complex optimization problems**
  - **3 types of techniques:**
    - **Exact**
    - **Heuristic**
    - **Hybrid**
  - **3 implementations:**
    - **Sequential**
    - **LAN**
    - **WAN**

UML design of SA in MALLBA



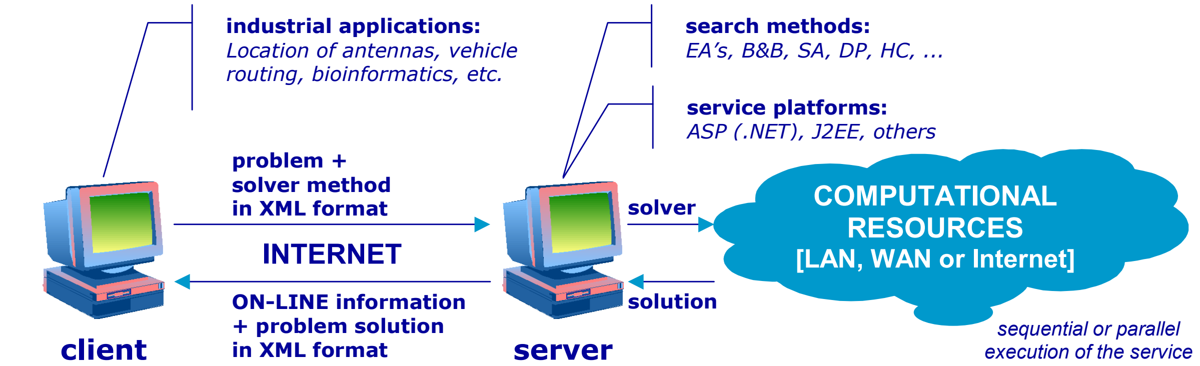
- **Goals:**
  - **Wide genericiy but low effort of instantiation**
  - **Simplified utilization (no parallel skills needed)**
  - **Geographically distributed computing systems**

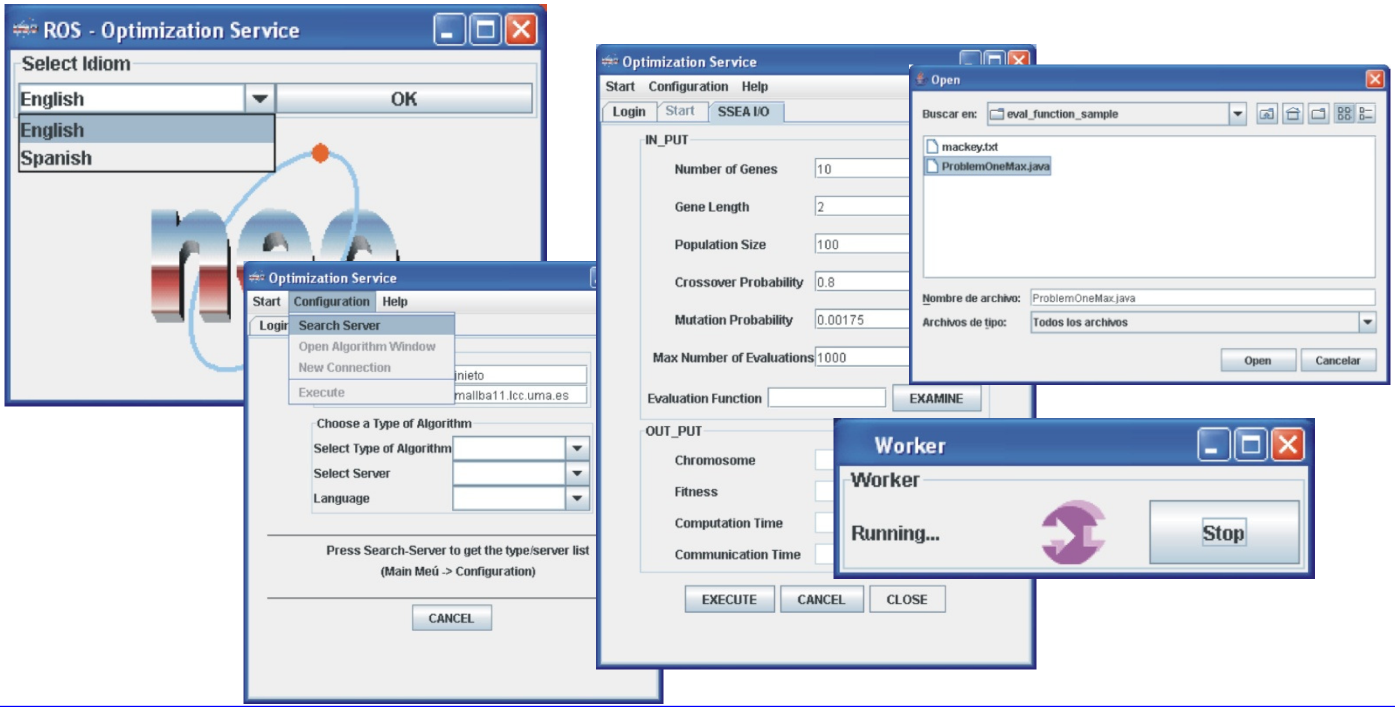
E. Alba, et al., MALLBA: A Library of Skeletons for Combinatorial Optimisation, *Proceedings of the Euro-Par*, Paderborn (GE), LNCS 2400, pp. 927-932, 2002

Parallel Experiences

Index

R.O.S.





Parallel Experiences

New Technologies Are Revitalizing this Field

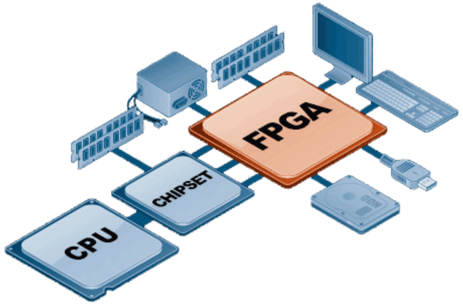
Index

Research in either separate or join levels of parallelism is welcome in companies and journals

Grid and cloud computing

GPU

FPGA



multicore

Parallel Experiences

Index

Summary of Challenges

1. Parallel solvers are not just fast versions of sequential solvers: they are new algorithms
2. Create standard benchmarking for parallel search
3. Develop a uniform experimental setup methodology in which metrics and statistics have a niche
4. Connect the solver field to the application field, and to the software domain: metrics for OO, parallelization of compilers, automatic testing, optimized software protocols, scheduling…
5. Tackle new issues in cloud computing
6. New research fields: GPGPU, Multicores, programming
7. Create a body of knowledge in parallel algorithmics independently of their applications

Parallel Experiences

Index

End of Presentation

Málaga