Adaptive Combinatorial Search for e-Science

Youssef Hamadi (MSR)

and

Marc Schoenauer (INRIA)

MICROSOFT RESEARCH INRIA
JOINT CENTRE

The “Adapt” Project

Goal: Improve the search performance through the automated tuning of solvers’ parameters

Expected result: Black-box, **GO** button

e.g. Improved usability, e-scientists, etc.

Technically:

A Meta-Learning problem

- Before the run → off-line learning
- During the run → on-line learning

Staff

- **Researchers**:
  - Anne Auger (TAO, INRIA Saclay)
  - Youssef Hamadi (MSR)
  - Nikolaus Hansen (Joint lab → TAO, INRIA Saclay)
  - Marc Schoenauer (TAO, INRIA Saclay)
  - Michèle Sebag (TAO, CNRS LRI)
- **Post-doc**:
  - Said Jabbour [→ Associate Pr. CRIL/CNRS]
- **PhD students**:
  - Alejandro Arbelaez (planned defense in May 2011)
  - Alvaro Fialho (defended Dec. 22, 2010)
    [→ post-doc OSD Microsoft CNRS Chaire]

Impact

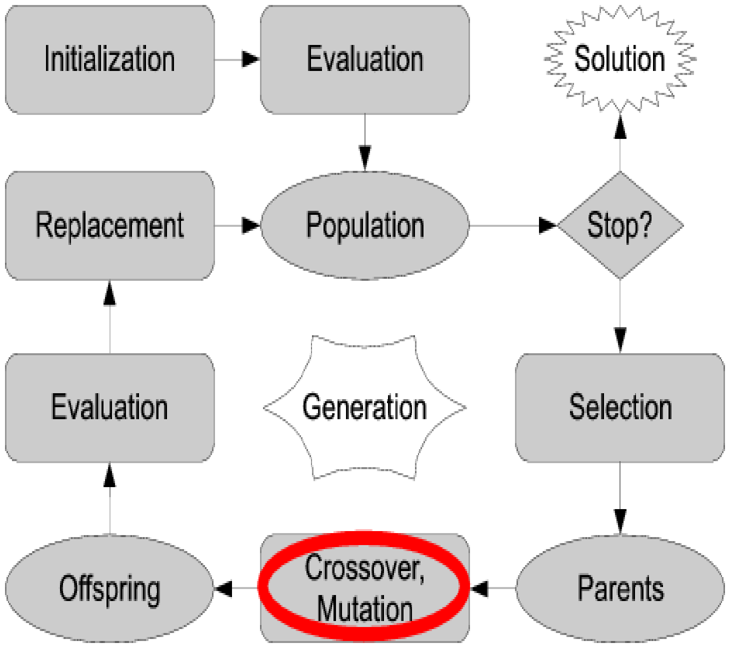
- **Publications**
  - 12 journal papers
  - 1 book + 6 book chapters
  - 34 intl. conferences
  - 4 PhD/HDR
  - 10 invited conf. talks
  - 2 Workshops
  - 7 TRs (+ BBOB)
- **Awards**
  - 4 Best papers (LION'09, ICTAI'09, EvoBIO'09, ACM-GECCO'10)
  - Silver Humies (ACM-GECCO'10)
  - ManySAT
    - Gold Medal SAT-Race 2008
    - Gold, Silver, Bronze SAT-Competition 2009
    - Silver, Bronze, SAT-Race 2010
- **Free software**
  - **GUIDE** – GUI for Evolutionary Algorithms
  - **CMA-ES** – Covariance Matrix Adaptation Evolution Strategy + variants
  - **DomFD** – add-on to Gecode (public domain CP solver)
  - **ManySAT** – parallel SAT solver
- **ManySAT→ Microsoft's Z3**
  - Internal, SAGE, Spec#/Boogie, Pex, Yogi, SLAM, F7, VS3, FORMULA, HAVOC and VCC
  - External, Dassault Aviation, Airbus, Synopsis, etc.

Agenda

- **Context**
  - Stochastic Search / Optimisation
  - Heuristics and Meta-Heuristics
  - Combinatorial, discrete, and continuous settings
- **Overview**
  - Evolutionary Algorithms and Invariance Properties
    * CMA-ES: Variants and Surrogates
    * Adaptive Operator Selection
  - Instance-based tuning for Constraint Programming
    * Static Tuning (Protein Structure Prediction, AI Planning)
    * Continuous Search

Evolutionary Algorithms

- **Stochastic Optimization**
  - “Generate and test” stochastic algorithms
  - Blind Variations + Survival of the fittest
- **Variation Operators**
  - Determine the trajectory of the search
  - Problem-dependent (with initialisation … and evaluation)



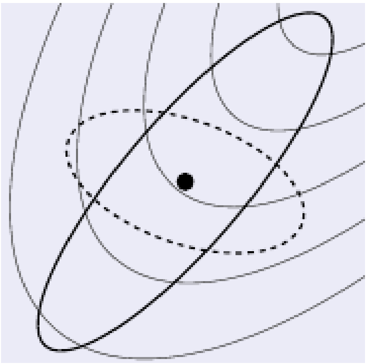
Invariance and Search

- **Source of Robustness**
  - Same parameter setting for whole equivalence class
- **Comparison-based algorithms**
  - Invariance by monotonous transformations of the objective function
  - Most Evolutionary Optimizers
  - Generally lost with 'goodies'
    Surrogate Models, Adaptive Operators Selection, ...
- **Adaptive encoding**
  - CMA-ES base principle
  - Invariance w.r.t. coordinate tranformations

Agenda

- **Context**
  - Stochastic Search / Optimisation
  - Heuristics and Meta-Heuristics
  - Combinatorial, discrete, and continuous settings
- **Overview**
  - Evolutionary Algorithms and Invariance Properties
    * CMA-ES: Variants and Surrogates
    * Adaptive Operator Selection
  - Instance-based tuning for Constraint Programming
    * Static Tuning (Protein Structure Prediction, AI Planning)
    * Continuous Search

Covariance Matric Adaptation

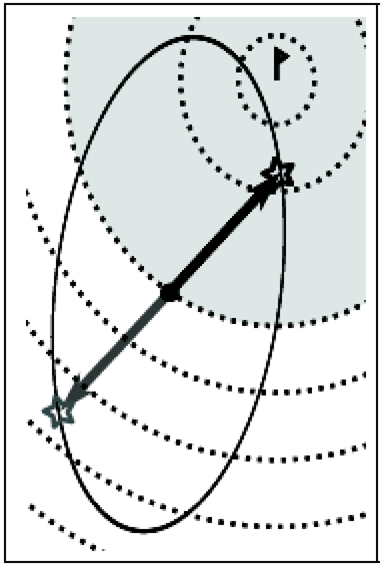


- Stochastic Evolutionary Search in continuous domain [Hansen et al. 94-01]
- Multi-variate Gaussian Mutation that adapts based on search path
- Won the most comprehensive benchmarking contests (vs all types of optimizers) **[ACM-GECCO'09-10]**
- Hundreds of applications in industry

Comparison-based and coordinate-independent
→ almost parameterless [Hansen & Ostermeier 2001]

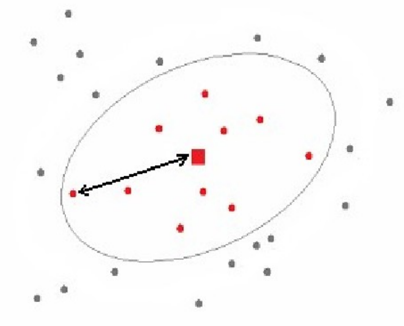
CMA-ES: Contributions

- **Generic Adaptive Encoding**
  - → make any algorithm coordinate-independent
  - Proof of concept with Cauchy mutations **[PPSN'08]**
  - Adaptive Coordinate Descent **[ACM-GECCO'11]**
- **Active CMA-ES** **[ACM-GECCO'10] Best paper**
  - Use unsuccessful trials too
- **Mirrored Sampling and Sequential Selection** **[PPSN'10, FOGA'11]**
  - Try two oppposite directions
  - Stop sampling if success



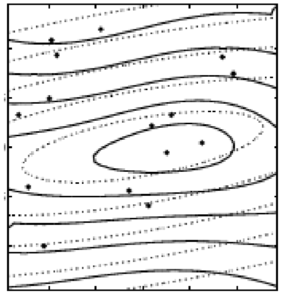
Surrogates with CMA-ES

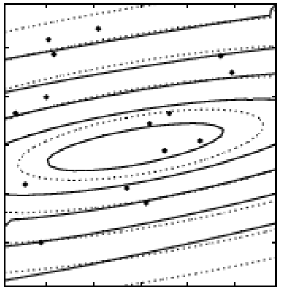
- **nlmm-CMA-ES** **[EvoNum'10; ACM-GECCO'11]**
  - Build a quadratic **l**ocal **m**eta-**m**odel around the point to evaluate
  - MM acceptance criterion based on approximate ranking
  - Can exploit partial separability



- **ACM-ES** **[PPSN'10]**
  - Rank-SVM for ordinal regression
  - With adaptive kernel

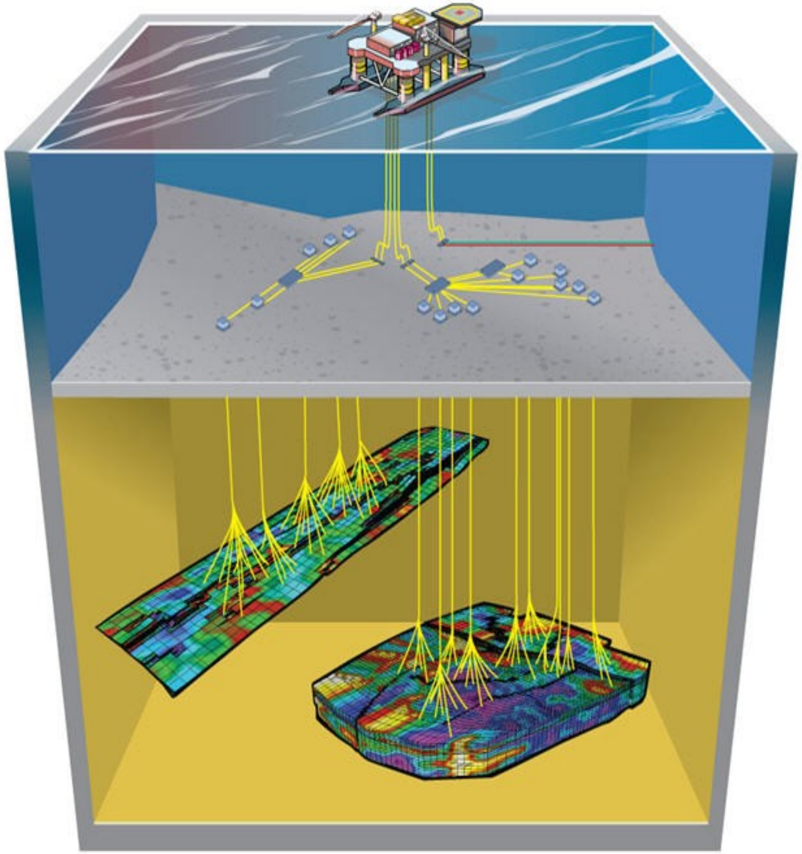
  → comparison-based surrogate



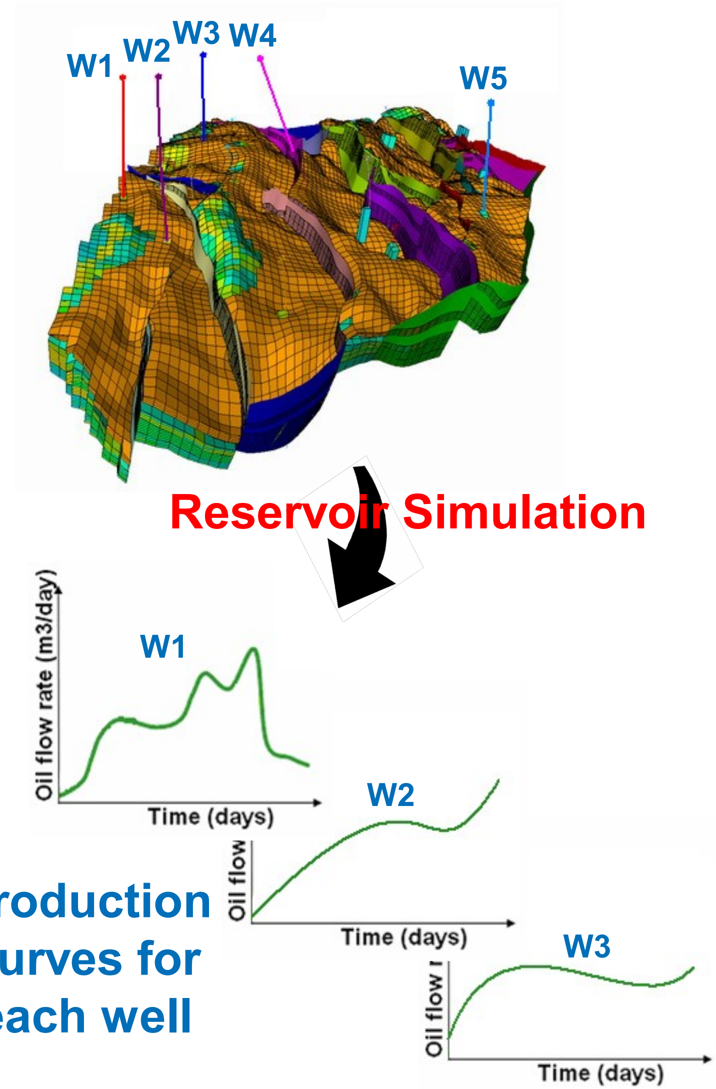


Well placement

coll. IFP **[ACM-GECCO'11]**

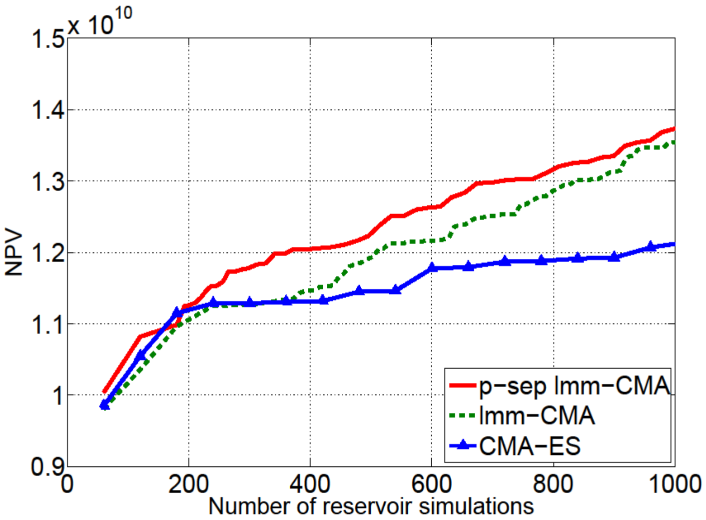


20mn per simulation



Production curves for each well

Results



- **Using partially-separable nlmm-CMA-ES**

  →

- **60% increase production**

- **20% less simulations**

- **All proposed configurations are feasible (or close to feasible domain).**

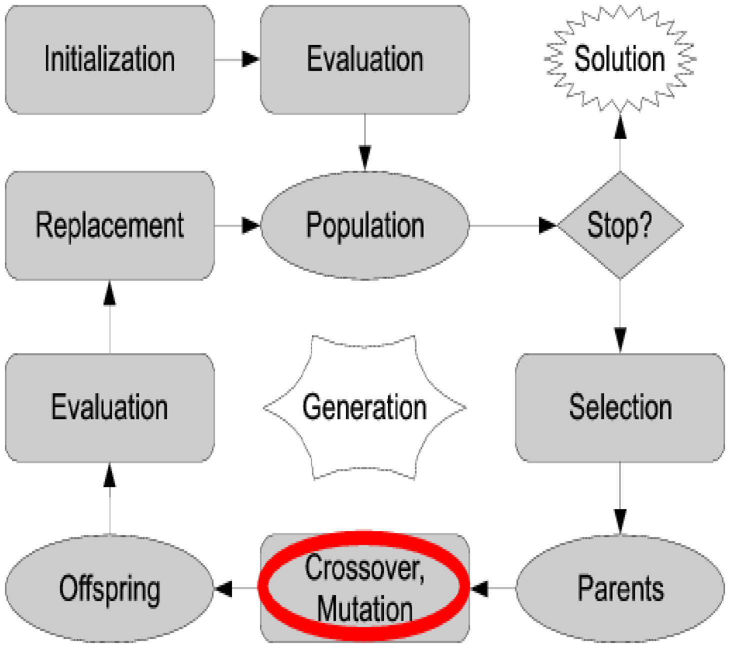
Agenda

- **Context**
  - Stochastic Search / Optimisation
  - Heuristics and Meta-Heuristics
  - Combinatorial, discrete, and continuous settings
- **Overview**
  - Evolutionary Algorithms and Invariance Properties
    * CMA-ES: Variants and Surrogates
    * Adaptive Operator Selection
  - Instance-based tuning for Constraint Programming
    * Static Tuning (Protein Structure Prediction, AI Planning)
    * Continuous Search

Evolutionary Algorithms

- **Stochastic Optimization**
  - “Generate and test» algorithms
  - Blind Variations + Survival of the fittest

- **Adaptive Operator Selection**
  - On-line choice of variation operator
  - Based on results of previous applications



AOS: Contributions

- **Bandit-based** Selection Rules
  - Choose operator j maximizing

$$\hat{p}_{j,t} + C\sqrt{\frac{2\log\sum_k n_{k,t}}{n_{j,t}}}$$

- **Dynamic context**
  - Page-Hinckley change detection test **[ACM-GECCO'08]**
  - Sliding window **[AMAI 2010]**
- **Extreme statistics**
  - Extreme events more important than averages **[PPSN'08]**

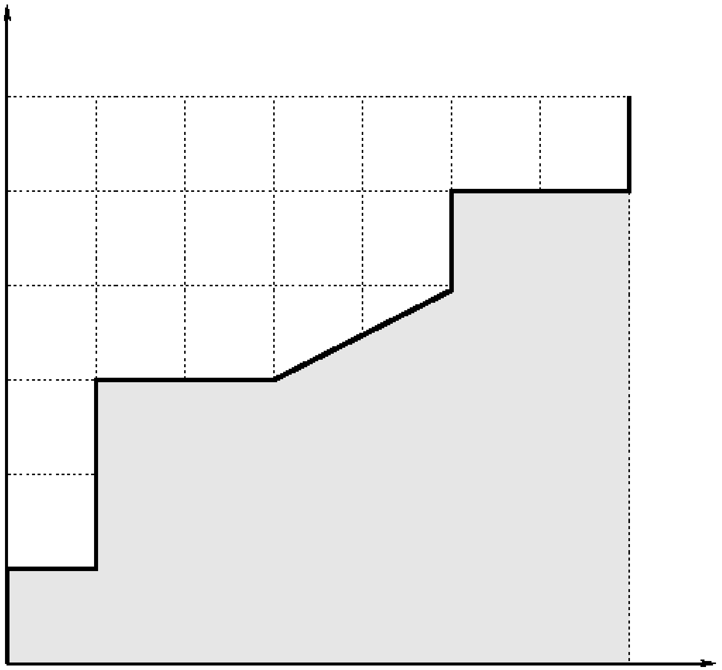
Outperforms previous AOS **[ACM-GECCO'09, CEC'09, LION'09]**

Very sensitive to fitness-scaling hyper-parameters **C**

AUC-based AOS

- Rank the rewards
  sliding window
- Compute the Area Under the Curve
  (1 op vs all others)
- Directly use in UCB

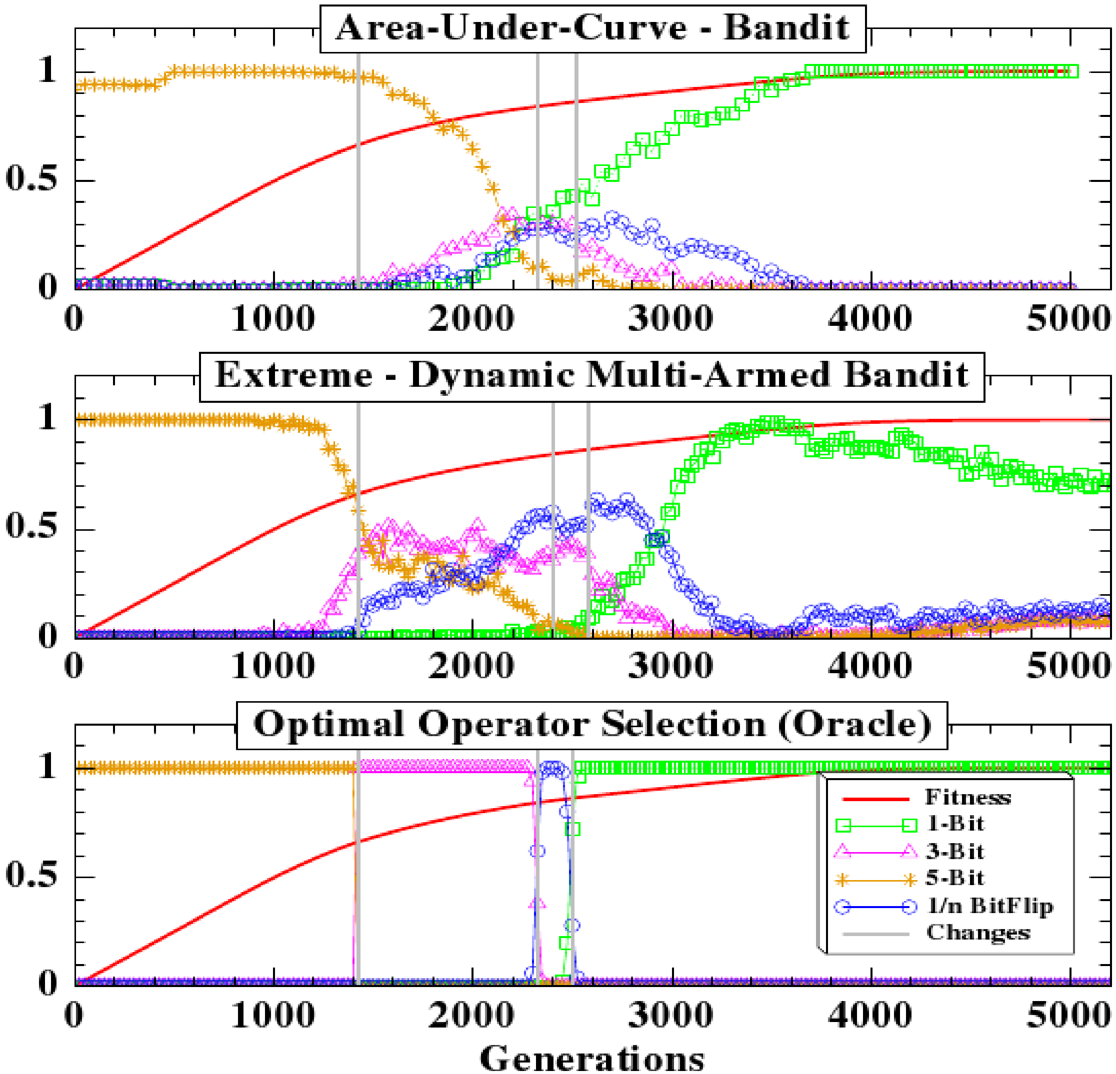
$$AUC_{j,t} + C\sqrt{\frac{2\log\sum_k n_{k,t}}{n_{j,t}}}$$



[ACM-GECCO'10]

Ranked list: 212211[112]2112

Results - OneMax



AOS	#Gens.
Naive	7813 ± 708
Best Naive	6793 ± 625
Optimal	**5059 ± 203**
AUC - MAB	**5069 ± 292**
Ext - DMAB	**5254 ± 313**
Avg - DMAB	7494 ± 611
Ext - MAB	9044 ± 840
Avg - MAB	8198 ± 683

Robustness - OneMax

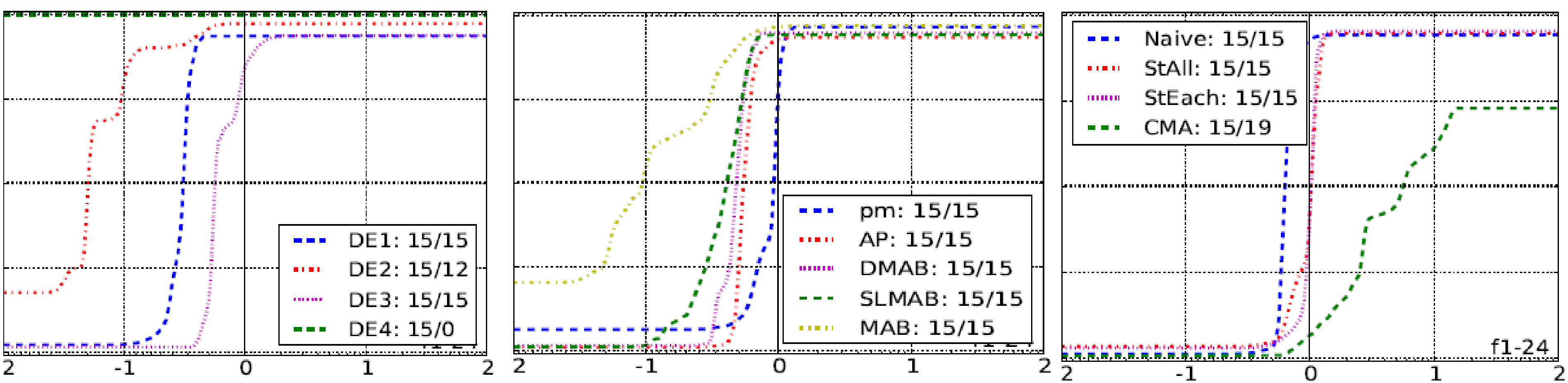
- Replace F with Log(F), exp(F), F^2

(h-l)	$\mathcal{F} = \sum b_i$	$\log(\mathcal{F})$	$\exp(\mathcal{F})$	$\mathcal{F}^2$	AOS tech.
485	**5103/427**	**5195/430**	**5562/950**	**5588/950**	AUC-MAB
807	**5123/218**	5431/223	5930/334	5792/382	Ext-AP
0	5726/399	5726/399	5726/399	5726/399	FAUC-MAB
2591	5376/285	7967/718	7722/2151	6138/516	Ext-DMAB
6971	6059/667	8863/694	13030/3053	12136/949	Ext-SLMAB
7052	9044/840	7947/1267	14999/0	14999/0	Ext-MAB

Results – DE on BBOB

[PPSN'10]

- **Differential Evolution**
  - Implicitly self-adaptive population-based EA [Storn & Price 95]
  - NP=100*dim; CR=1; F=0.5
  - 4 possible mutation strategies
- **Black-Box Optimization Benchmarking**
  - 24 functions x 15 instances with controlled difficulties
  - Several dimensions (2, 3, 5, 10, <u>20</u>, 40)

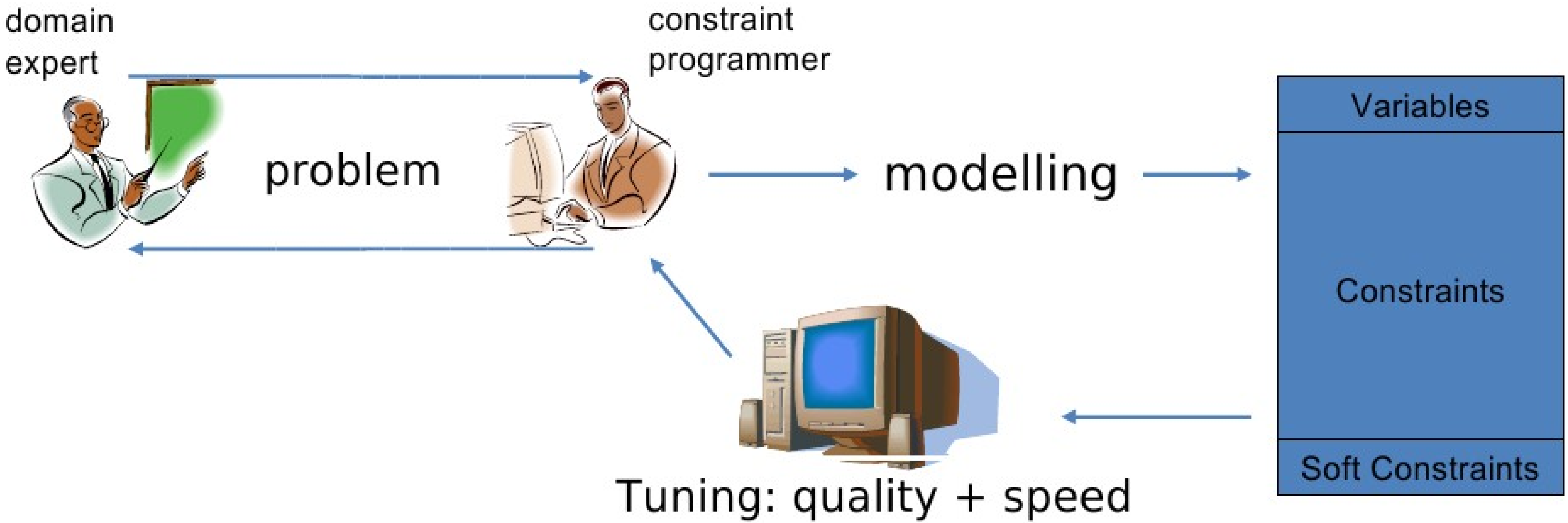


Agenda

- **Context**
  - Stochastic Search / Optimisation
  - Heuristics and Meta-Heuristics
  - Combinatorial, discrete, and continuous settings
- **Overview**
  - Evolutionary Algorithms and Invariance Properties
    * CMA-ES: Variants and Surrogates
    * Adaptive Operator Selection
  - Instance-based tuning for Constraint Programming
    * Static Tuning (Protein Structure Prediction, AI Planning)
    * Continuous Search

Constraint Programming

Does not scale:



- Learn a single function that maps instance features and parameter settings to runtime

Machine Learning for Adaptive Search

- **Goal**: Predict the best algorithm (or parameter set) for each instance
- **Examples** (actual runs) to learn the mapping

(Instance, algo/param) → Performance

or

Instance → Best algo/param

- **Issues**

- **Features** (describe the instances)

- **Performance**

Time to given value vs Value in given time

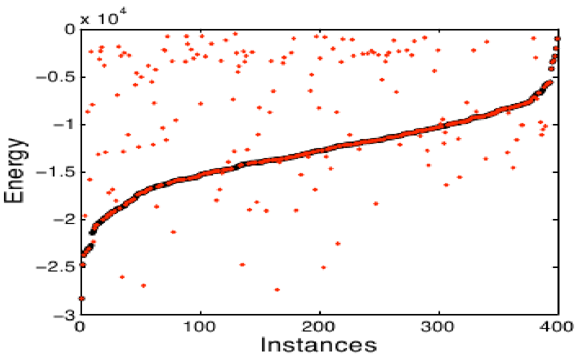
Agenda

- **Context**
  - Stochastic Search / Optimisation
  - Heuristics and Meta-Heuristics
  - Combinatorial, discrete, and continuous settings
- **Overview**
  - Evolutionary Algorithms and Invariance Properties
    * CMA-ES: Variants and Surrogates
    * Adaptive Operator Selection
  - Instance-based tuning for Constraint Programming
    * Static Tuning (Protein Structure Prediction, AI Planning)
    * Continuous Search

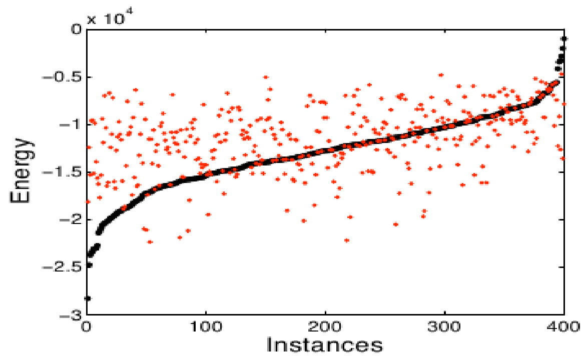
Protein Structure Prediction

[WCBMB'10]

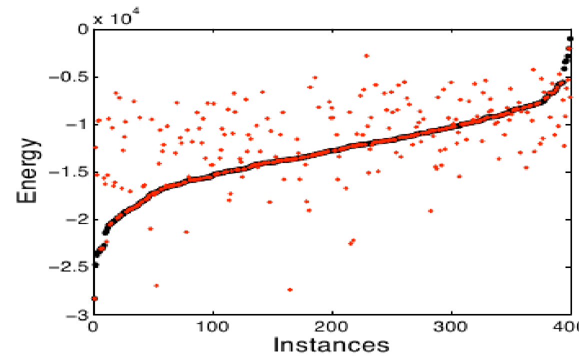
- A Constraint Programming formulation: Minimize energy for lattice configurations
- 4 efficient heuristics to choose from
- 105 domain-based features from literature
- Minimize energy after 5mn CPU time
- Supervised learning with decision trees (J48)
- Outperforms fixed strategies and SATzilla (on avg)



(a) wdeg^{+} vs J48



(b) domFD^{+} vs J48



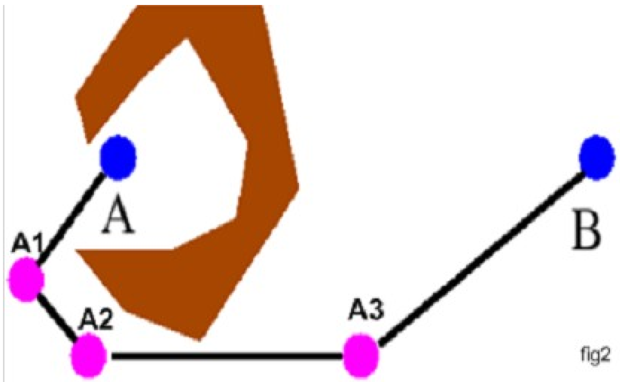
(c) SATzilla vs J48

Evolutionary AI Planning

coll. Thalès - ONERA

Divide-and-Evolve

- Decompose a planning problem into a sequence of (simpler?) sub-problems
- Each sub-problem is solved by a standard planner
- Evolution to optimize the sequence of intermediate states:
  - Planning Heuristics used for
    * Initialisation
    * Variation operators



Silver medal, Humies Awards, ACM-GECCO 2010

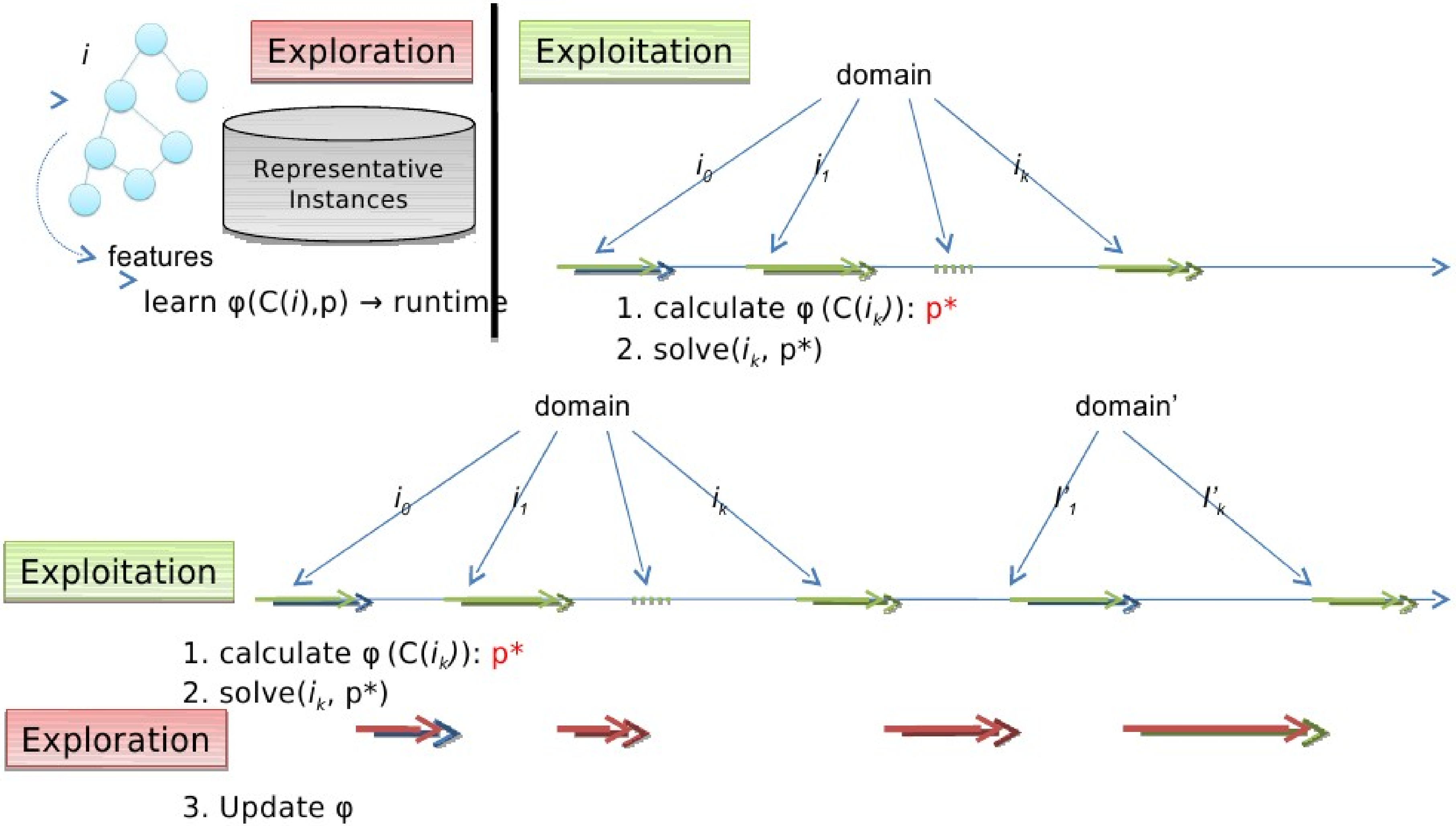
Parameter Tuning for DaE

- Performance: normalized makespan in 30mn CPU
- **Standard Racing** **[ACM-GECCO'10]**
  - Computed for one instance, one domain, or across several domains
  - Results as expected
- **ML-based Learn-and-Optimize** **[ACM-GECCO'11]**
  - 14 features to describe an instance
  - Learn the (Instance → best parameters) mapping
  - Using an Artificial Neural Network
  - Significant improvement over the default values
  - Marginaly worse than best per-instance values

Agenda

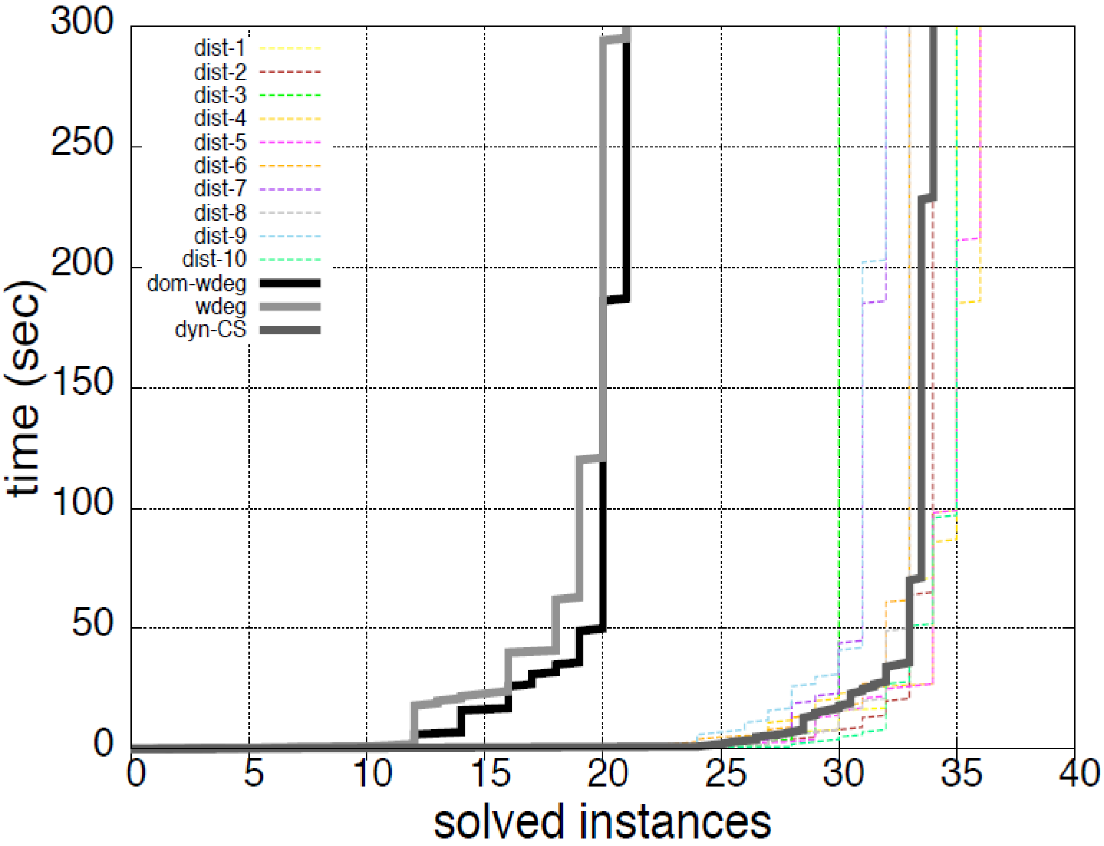
- **Context**
  - Stochastic Search / Optimisation
  - Heuristics and Meta-Heuristics
  - Combinatorial, discrete, and continuous settings
- **Overview**
  - Evolutionary Algorithms and Invariance Properties
    * CMA-ES: Variants and Surrogates
    * Adaptive Operator Selection
  - Instance-based tuning for Constraint Programming
    * Static Tuning (Protein Structure Prediction, AI Planning)
    * Continuous Search

Continuous Search in CP



Continuous Search in CP

Langford numbers



Continuous Search in CP

Experiments

Problem	*dom-wdeg*			*wdeg*			*dyn-CS*		
	#sol	time(h)	avg-time(m)	#sol	time(h)	avg-time(m)	#sol	time(h)	avg-time(m)
nsp	68	3.9	2.34	**88**	**2.6**	**1.56**	77	2.9	1.74
bibd	**68**	**1.8**	**1.37**	68	1.8	1.37	65	2.0	1.44
js	**76**	**4.9**	**2.26**	73	5.1	2.35	73	5.2	2.4
lfn	21	5.2	3.75	21	5.3	3.83	**33**	**4.1**	**2.96**
geom	64	3.9	2.34	27	6.8	4.08	**67**	**3.3**	**1.98**
Total	297	19.7	2.39	274	21.6	2.61	**315**	**17.5**	**2.11**

Table 1. Total solved instances

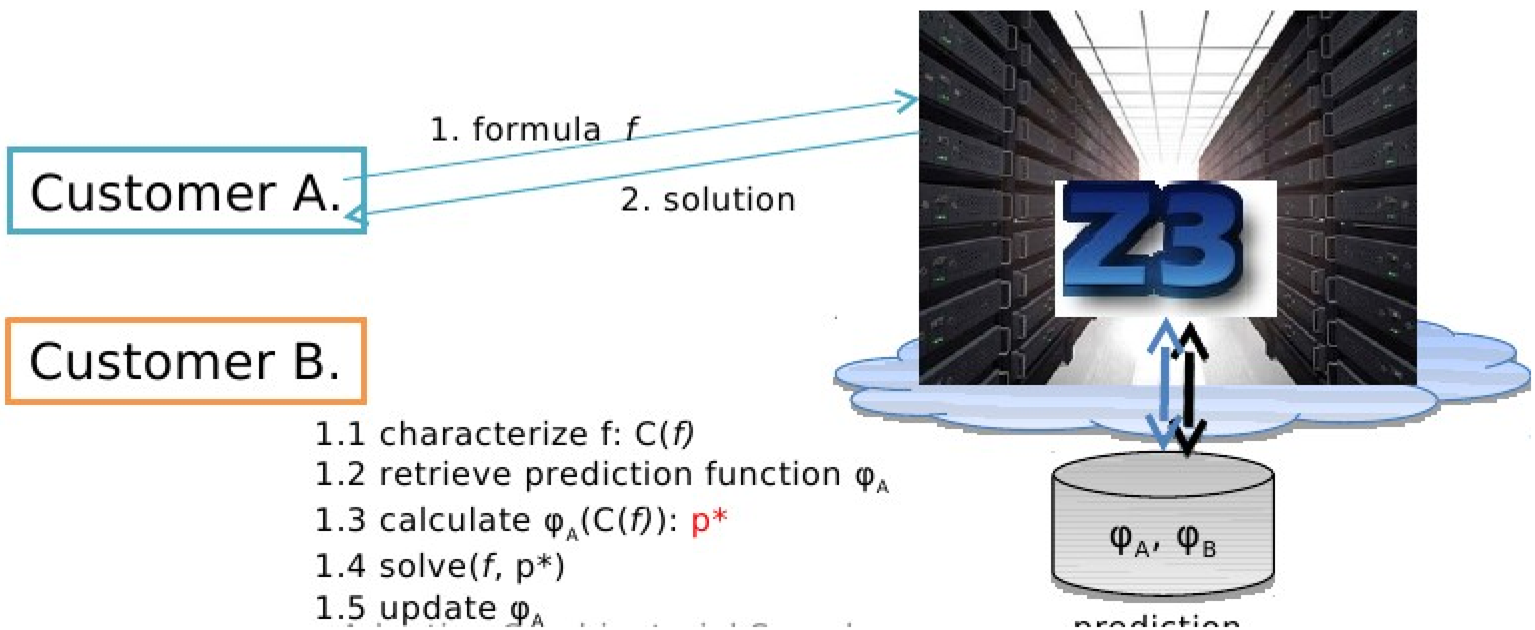
Problem	#Sol	time (h)
lfn-bibd‡	23	5.3
lfn-bibd†	63	2.3

Symbolic Reasoning in the Cloud

Z3: MSR SMT's solver

several reasoning engines, >256 parameters

Continuous Search: customer-based self tuning



Perspectives

- **Algorithm → Automatic tuning**
  - parameters
- **Components → Algorithm Design**
  - Operators
  - Neighborhoods
  - Data Structures
- **Components, System → Autonomic Computing**
  - Resources (very large scale)
  - Self-healing (fault tolerance)