High Performance Computing Center Stuttgart

Panel 2:

Opportunities and Challenges in Simulation-Driven Research

Alexey Cheptsov

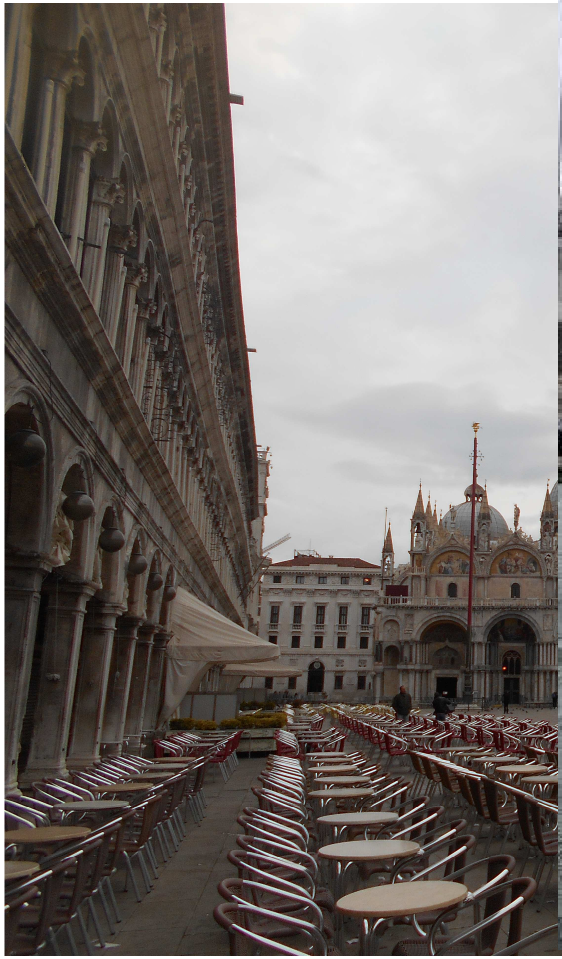
High Performance Computing Center Stuttgart (HLRS)

Simulation-Driven Research

- Panelists:
  - **Yoel Tenne**, Ariel University, Israel
  - **Masaomi Kimura**, Shibaura Institute of Technology, Japan
  - **Dejan Zupan**, University of Ljubljana, Faculty of Civil and Geodetic Engineering, Slovenia
  - **Floriano Scioscia**, Technical University of Bari, Italy
  - **Alexey Cheptsov**, High Performance Computing Center - Stuttgart, Germany

Simulation-Driven Research

- Simulation – can it cover all our needs?



Simulation-Driven Research

- Goals of the Panel:
  - Discuss the modern advances of the Simulation Technology for Science and Industry
  - Analyse the demands of the newest R&D trends on simulation
  - Discuss the emerging application requirements
  - Meet experts in and around the Simulation Technology

Simulation-Driven Research

Problem Statement:
BIG DATA
(in context of Simulation)

Simulation-Driven Research

HPC Center Stuttgart

- First HPC system in Europe (Cray-2, 1986, 4 CPUs, 2GB RAM, approx. 2 GFLOPS)
- German national HPC infrastructure provider since 1995
- EU infrastructure provider since 2005
- **110M core hours delivered to industry in 2014**

HAZEL HEHN (Cray XC40, Intel Haswell [Xeon E5-2680v3] CPU, Aries network)

- **7.712** nodes (24 cores)
- **7.4 PFLOPs** performance
- **128 GB** DDR4 RAM per node
- **10 PB** Disc
- **3000 KW** power consumption / 1.5M Euro

Simulation-Driven Research

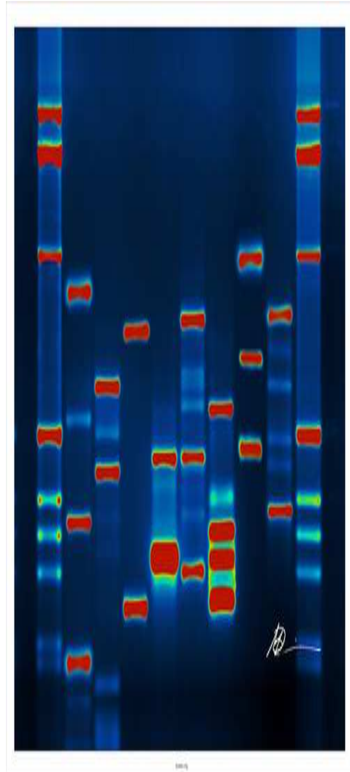
- Evolution of Data Science

continuity: $\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}=0$

x-momentum: $u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}=-\frac{1}{\rho}\frac{\partial p}{\partial x}+\nu\left[\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right]$

y-momentum: $u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}=-\frac{1}{\rho}\frac{\partial p}{\partial y}+\nu\left[\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}\right]$

```
f = figure;
set(f,'Color', [0 0 0]);
sphere(200);
axis vis3d off;
h = findobj('Type','surface');
shading interp;
set(h,'FaceColor',[0.5 0.5 0.5]);
light('Position',[-3 -1 3]);
set(h,'DiffuseStrength',1.0);
set(h,'SpecularStrength',1);
set(h,'SpecularExponent',1);
set(h,'AmbientStrength',0.25);
set(h,'BackFaceLighting','lit')
```



experimental | theory | computation | data-driven

-1000 | -100 | -10 | 2016 Year, A.D.

Simulation-Driven Research

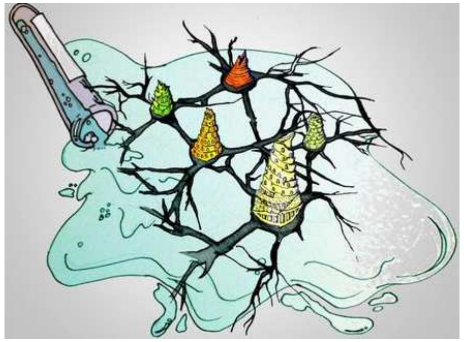
Evolution of Computational Applications	Traditional Computational Sciences	Data-Intensive Sciences
Structure	static	dynamic
Locality	spatially and temporally local	no or little
Volume	fit into memory	do not fit into memory
Arithmetical Complexity	High Precision Arithmetic	variable precision or integer based

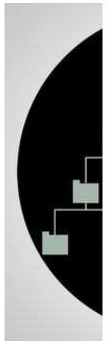
Simulation-Driven Research

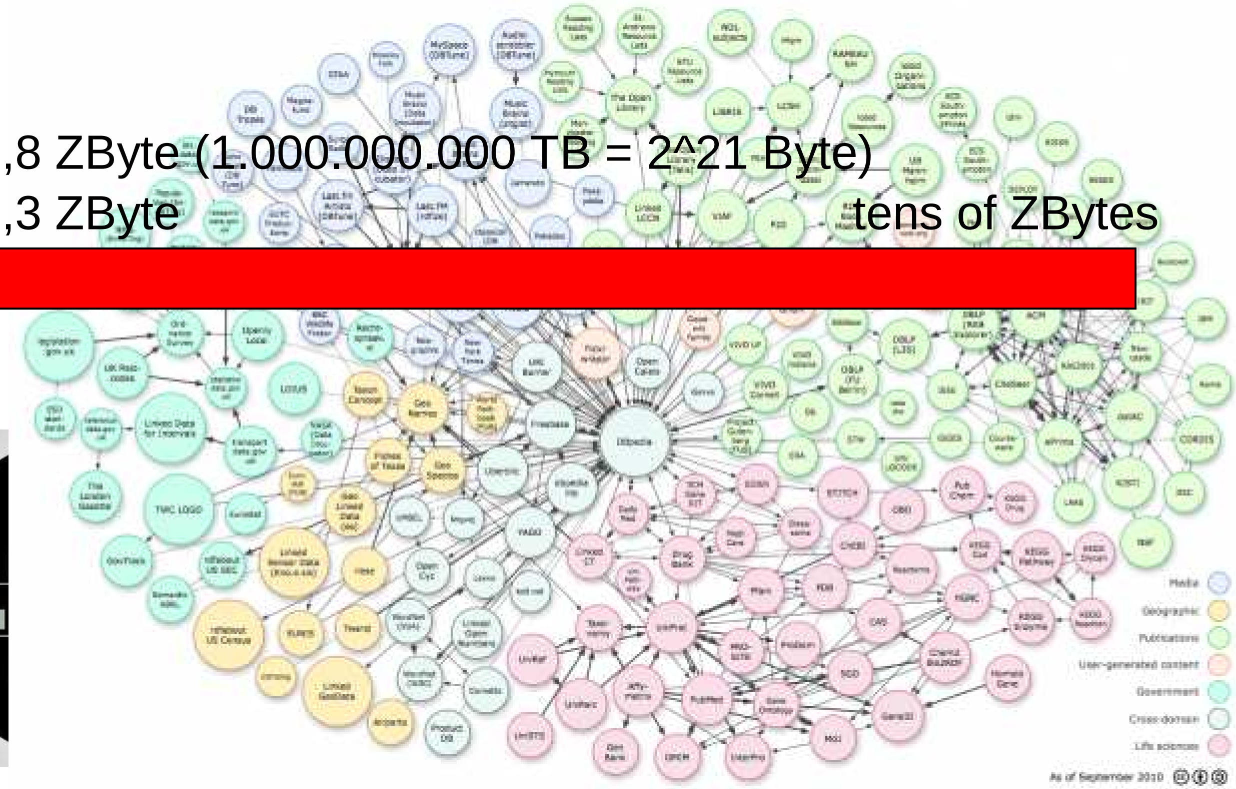
- Let's try the Semantics!

Size of the data universe

2009	0,8 ZByte (1.000.000.000 TB = 2^21 Byte)
2010	1,3 ZByte
2012	tens of ZBytes







structured data — ontologies — linked data — Semantic Web

2000 — **2005** — **2010** — Year, A.D.

Simulation-Driven Research

UPS Study

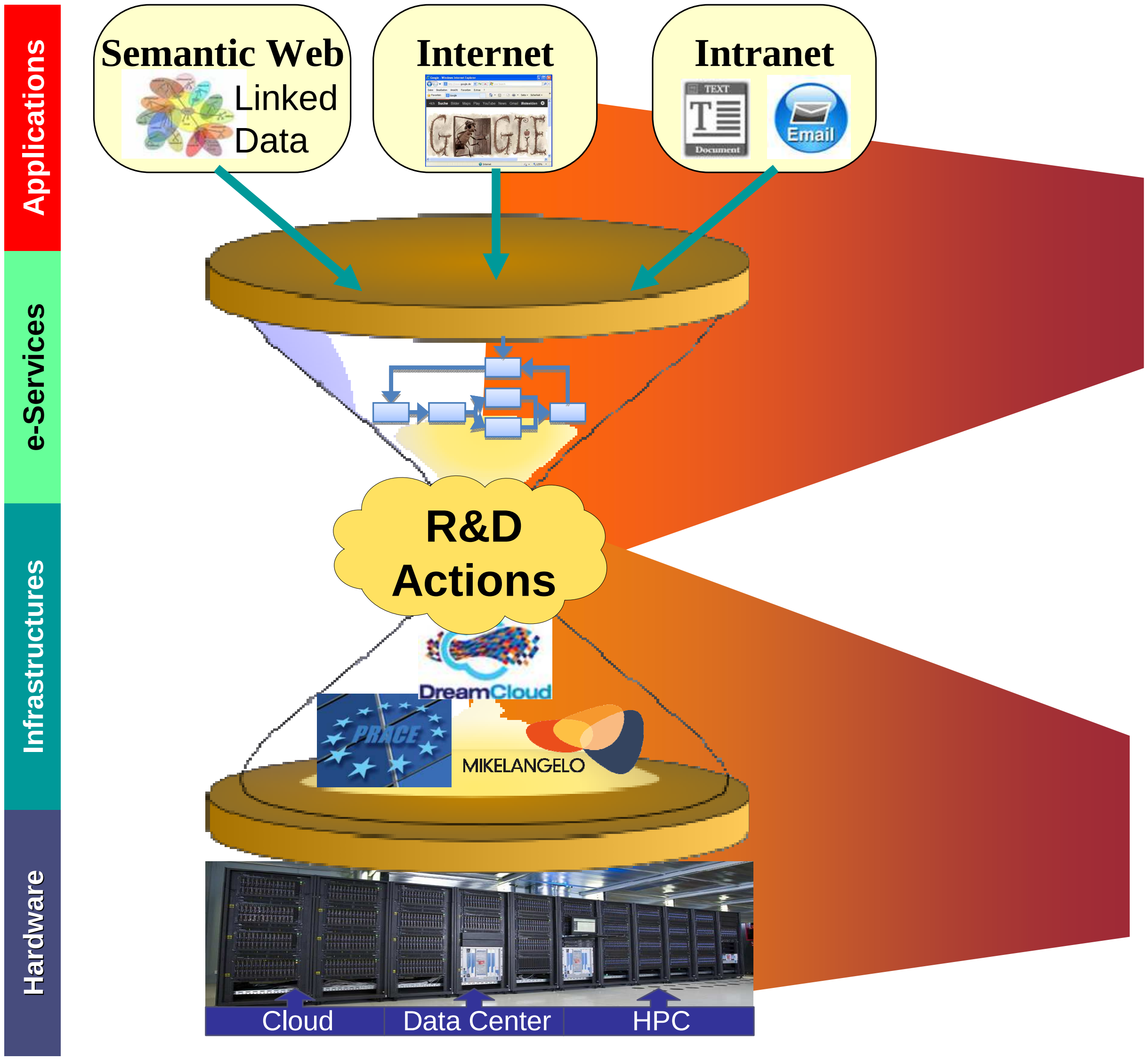
- The world's largest shipping company
- Spending 1 Billion USD on Big Data research
- Very high operational costs
  - costs caused by delivery cars' fall-outs due to traffic accidents are the major point of concern!

- Decision taken:

right-turn only where possible,

regardless on the track's length

Simulation-Driven Research



Applications:

- Programming models
- Parallelisation
- Analysis
- Performance optimisation
- Scalability
- Workflows

Infrastructure:

- Reconfigurability and “on demand” provision
- Distributed platforms
- Efficiency (also energy)
- Middleware (workflows, schedulers)

A challenge of simulation to understand languages

Masaomi Kimura
Shibaura Institute of Technology, JAPAN

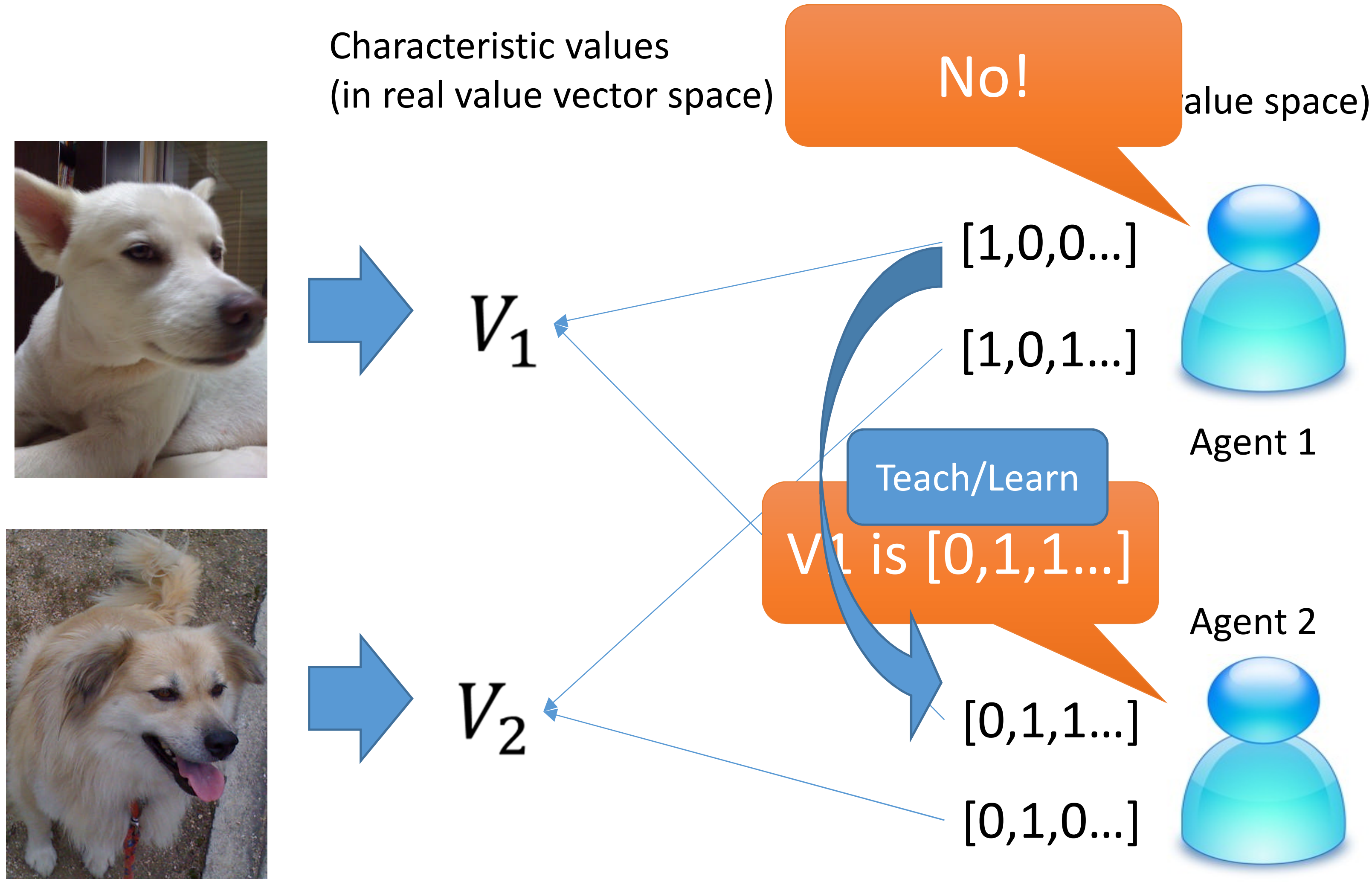
Text mining

We should have
more insights into properties of languages
to realize precise text mining

= My motivation for language simulation
inspired by language emergence

The following is about
the simulation of naming process…

Agent-base simulation



Agent-base simulation

Characteristic values
(in real value vector space)

Names
(in discrete value space)

V_1

[1,0,0...]

[1,0,1...]

Agent 1

Teach/Learn

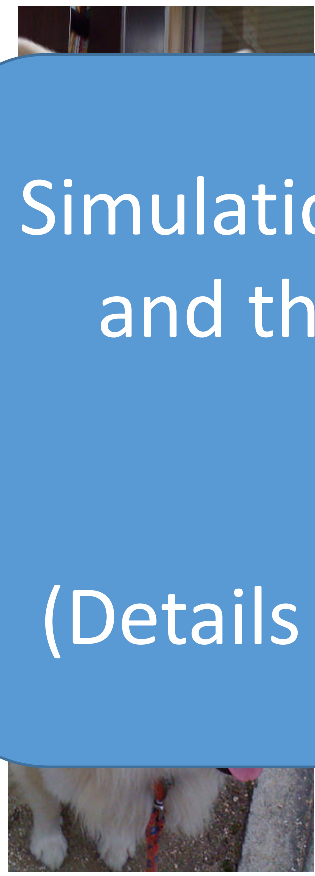
V_2

Agent 2

[1,0,0...]

[0,1,0...]

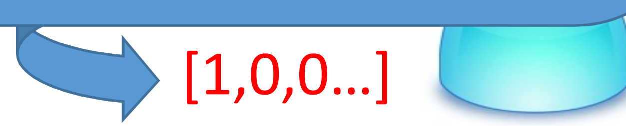
Agent-base simulation



Agent 1

Simulations show that they reach consensus and that a “name” is a label of a cluster.

(Details will be reported in the near future)



Agent 2

Problems

We hope that this simulates a naming process.

But …

- How can we justify the simulation?
  - We do not have a fundamental model or evidences of the process but only the results (=natural languages).

Thank you for your attention

Masaomi Kimura

masaomi@shibaura-it.ac.jp

Panel 2: Discussion Results

Opportunities and Challenges in Simulation-Driven Research

Simulation-Driven Research

- Wrap-Up:
  - Simulation ≠ Algorithms + Software + Hardware
  - + **INTEGRATION**
  - Time for a better consolidation of the simulation techniques
  - Common platform for validation of model
  - Simulation as a service with a pluggable architecture
  - Lack of basic knowledge of the investigated objects - leads to poor model representations
  - Optimization is not only the mathematics

Panel

Opportunities and Challenges in Simulation-driven Research

Floriano Scioscia

Technical University of Bari, Bari (Italy)

The Tenth International Conference on Advances in Semantic Processing (SEMAPRO 2016)
October 9-13, 2016 – Venice, Italy

Why simulation-driven research?

More accurate than purely analytic models

Inspecting individual subsystems

Prototype-less experimental evaluation

“What-if” analysis

Ease of finding **errors** and **performance bottlenecks**

Growing availability of **computing power**

Scalability

Types of simulations

Microscopic

- A single element of a complex system
- Validating behaviors
- Mainly useful for performance profiling and optimization

Mesoscopic

- A small set of homogeneous elements
- Validating interaction patterns: protocols, concurrency, resource management
- Mainly useful for use case tests

Macroscopic

- A large set of heterogeneous elements
- Qualitative and quantitative evaluation of emergent behaviors
- Mainly useful to anticipate or replace expensive prototype development

Some research experiences

ns2 Network Simulator modules

- Bluetooth
- ZigBee
- IEEE 802.11

Radio-Frequency IDentification

- IBM WebSphere RFID Tracking Kit
- Rifidi

Automotive and vehicular networks

- NCTuns
- MATLAB and Simulink

...

Simulations and reproducibility

Reproducibility of research is increasingly important

- EU Horizon 2020
- USA National Science Foundation and National Institute of Health
- Journal publishers

Simulations vs prototypes: easier

- Storage
- Repackaging
- Virtualization (e.g. VMware or Docker)
- Retooling

Issues

- Data-intensive vs computation-intensive simulations
- Performance overhead
- Distributed simulations

Contact

Politecnico di Bari

Floriano Scioscia, Ph.D.

Information Systems Laboratory
Politecnico di Bari, Italy

Tel.: +39 349 5629889

E-mail: floriano.scioscia@poliba.it

Personal Webpage: http://sisinflab.poliba.it/scioscia

Projects Webpage: http://sisinflab.poliba.it/swottools

Github repository: github.com/sisinflab-swot

An Algorithm for Expensive Optimization Problems

Yoel TENNE
Ariel University, Israel

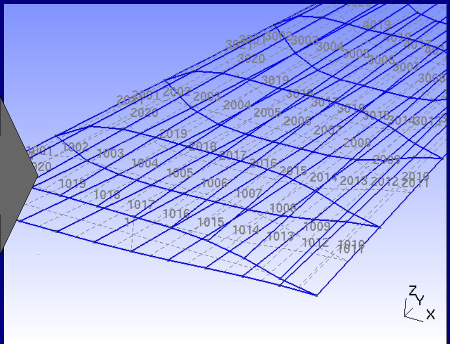
ADVCOMP 2016-10, Venice, Italy

Talk format

- Introduction
- Problem description
- Proposed approach
- Performance analysis
- Summary

Simulation-driven design optimization

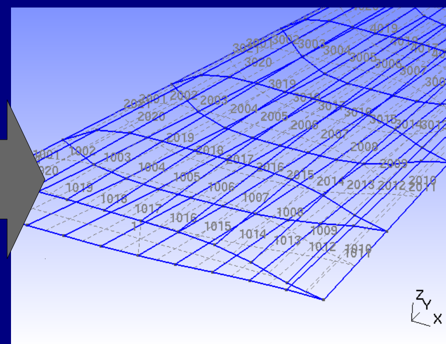
Laboratory experiment



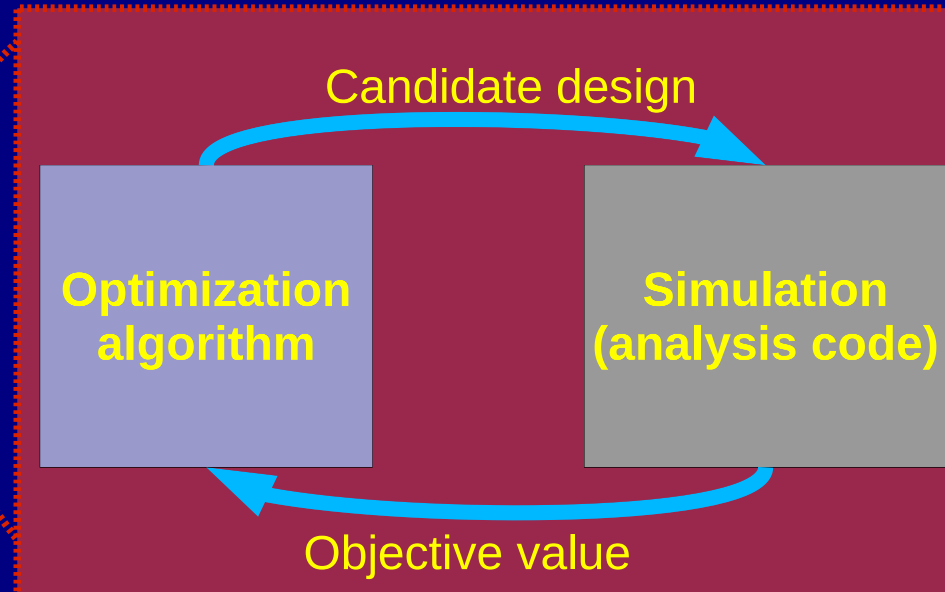
Computer simulation

Simulation-driven design optimization

Laboratory experiment



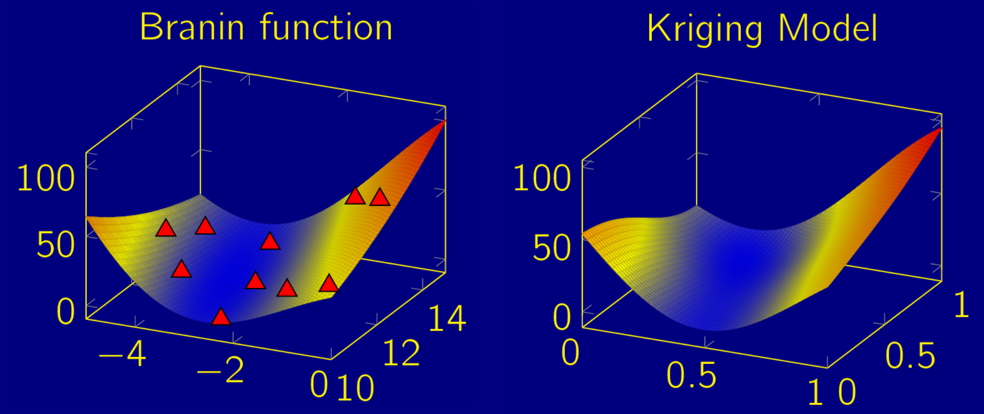
Computer simulation



- No analytic function expression ('Black-box' function).
- 'Expensive' function evaluations.
- Challenging function features (e.g.,multiple optima).

Metamodel-assisted optimization

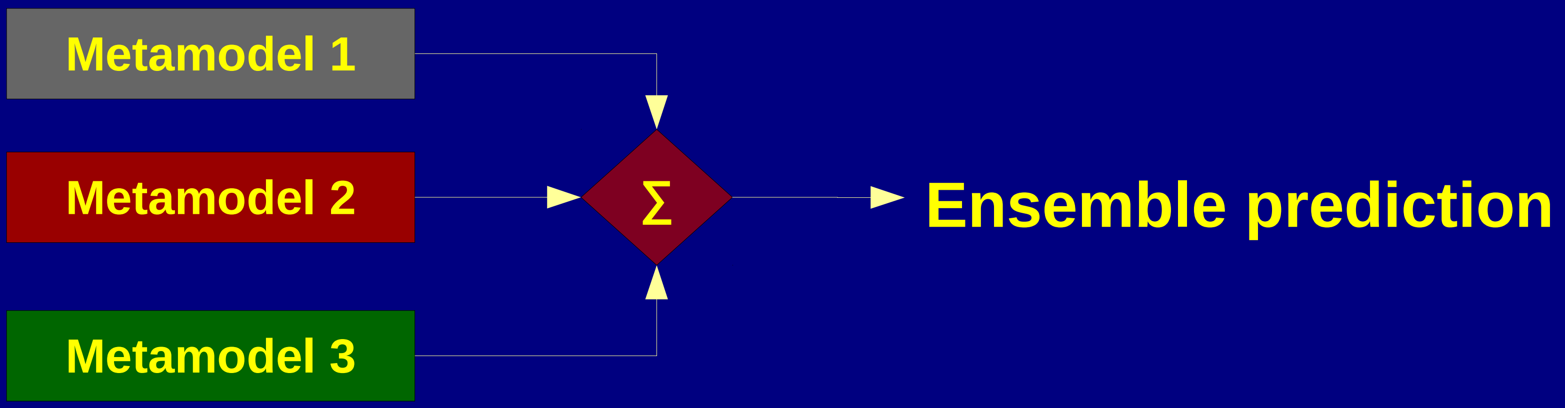
- Step 1: Replace the expensive function (the simulation) with a computationally cheaper approximation (a “metamodel” / “surrogate”).



•Some common variants:
•Polynomials
•Radial basis functions
•Kriging
•Neural networks

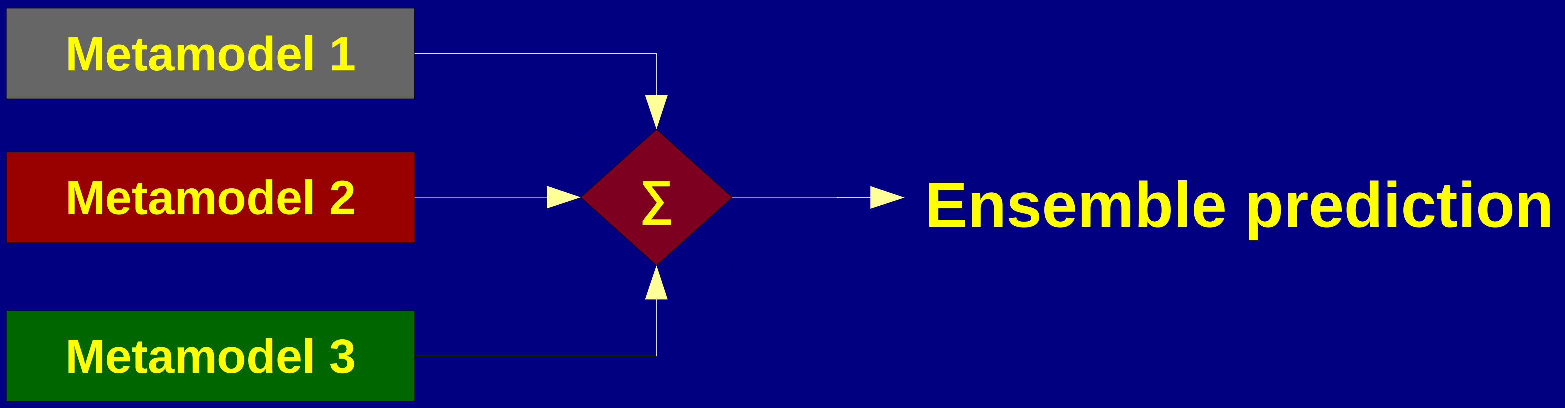
Problem description

- Various metamodel variants exist, but the optimal variant is problem dependent and is typically unknown a-priori.
- To improve the prediction accuracy, *Ensembles* use multiple metamodels and combine their prediction into a single output.



Problem description

- Various metamodel variants exist, but the optimal variant is problem dependent and is typically unknown a-priori.
- To improve the prediction accuracy, *Ensembles* use multiple metamodels and combine their prediction into a single output.



- **Ensemble topology**: which metamodels are incorporated,
- Does the topology affect the prediction accuracy?.

Numerical test

- Comparing 4 different topologies with 3 candidate metamodels: Kriging (K), RBF (R), RBF network (RN).
- Prediction accuracy estimated by the root mean square error.

	Ensemble topology			
Function	R+RN	R+K	RN+K	R+RN+K
Ackley-5D	4.258e-01	3.702e-01	4.151e-01	**2.967e-01**
Rastrigin-10D	1.223e+02	**8.198e+01**	1.312e+02	1.097e+02
Rosenbrock-20D	1.791e+06	1.666e+06	**1.648e+06**	1.693e+06
Schwefel 2.13-30D	**1.882e+06**	2.179e+06	2.343e+06	2.079e+06

R:RBF, RN:RBF neural network, K:Kriging.

- Results show a significant impact of the topology on accuracy.
- Using an unsuitable ensemble can hamper the search.

Existing approaches

- Dynamic selection of a *single* metamdoel (no ensemble):
  Gorrisen (2009), Tenne (2011).
- Using a *fixed* ensemble topology (no selection):

  Regis (2013), Tenne (2014), Muller (2014).

- **In-search selection of the ensemble topology appears to be an open issue.**

Existing approaches

- Dynamic selection of a *single* metamdoel (no ensemble):
  Gorrisen (2009), Tenne (2011).
- Using a *fixed* ensemble topology (no selection):

  Regis (2013), Tenne (2014), Muller (2014).

- **In-search selection of the ensemble topology appears to be an open issue.**

Research goal:

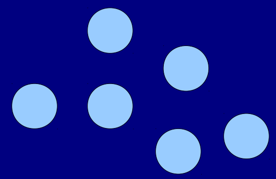
How to select an optimal ensemble topology without prior knowledge on the problem?

Proposed framework

- Goal: Dynamic selection of the optimal ensemble topology.
- Step 1: estimating the prediction accuracy of individual metamodels with cross-validation.

Proposed framework

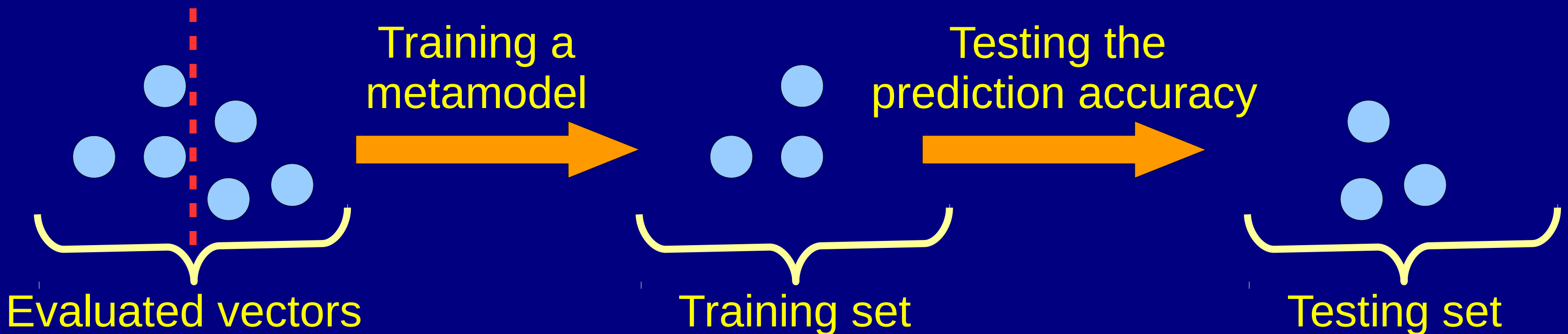
- Goal: Dynamic selection of the optimal ensemble topology.
- Step 1: estimating the prediction accuracy of individual metamodels with cross-validation.



Evaluated vectors

Proposed framework

- Goal: Dynamic selection of the optimal ensemble topology.
- Step 1: estimating the prediction accuracy of individual metamodels with cross-validation.



The root mean square error of the metamodel: $$e_j = \sqrt{\frac{1}{l}\sum_{i=1}^{l}\left[m(x_i) - f(x_i)\right]^2}$$

Proposed framework

- Step 2: Generating the candidate ensemble topologies and corresponding predictions.

  The ensemble prediction is defined as:

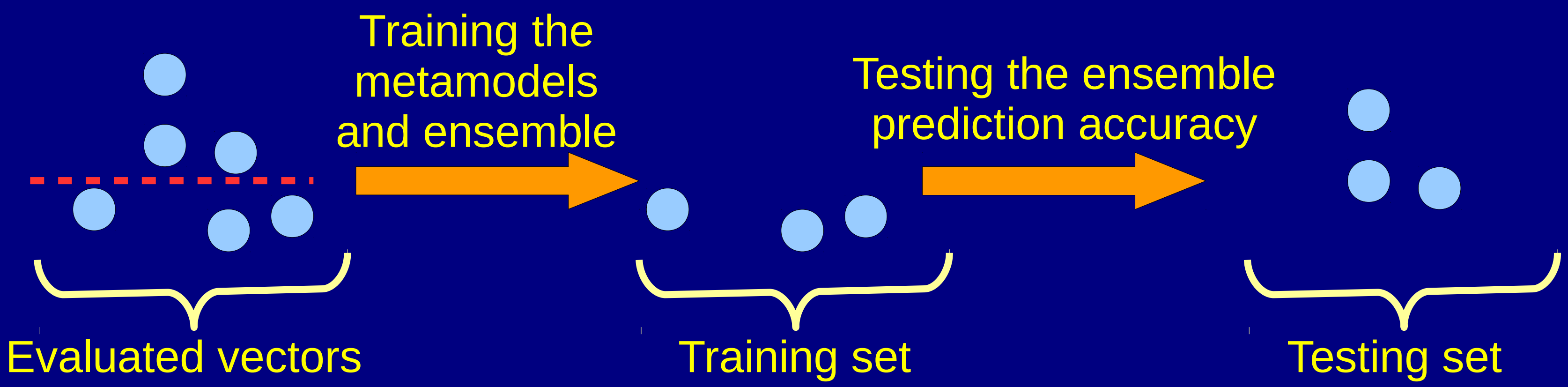
$$\epsilon(x)=\sum_{j=1}^{n} u_j m_j(x) \quad , \qquad u_j=\frac{1/e_j}{\sum_{j=1}^{n} 1/e_j}$$

weight metamodel

The weight (contribution) of each metamodel is inversely proportional to its prediction error (from Step 1).

Proposed framework

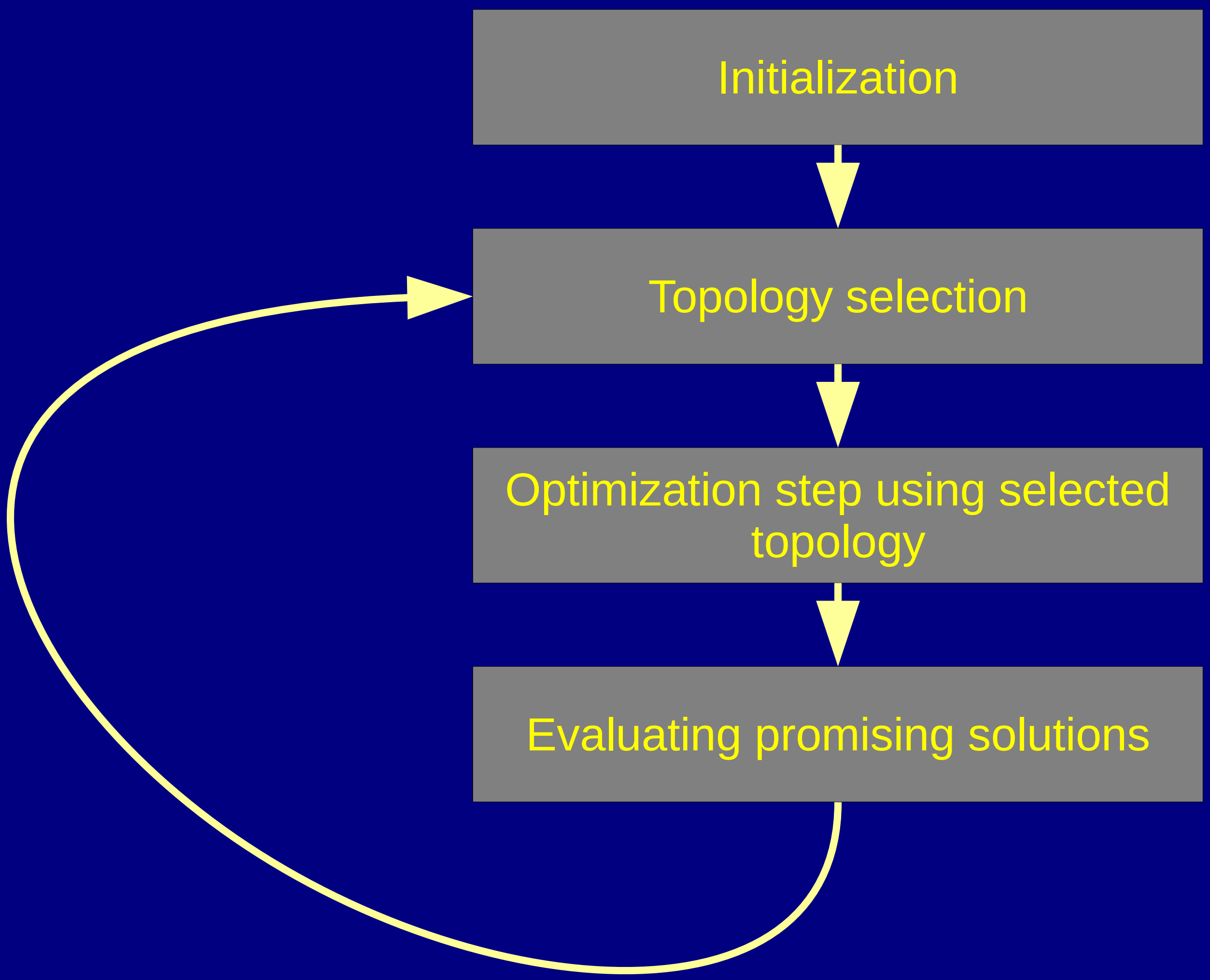
- Step 3: Estimating the prediction accuracy of a candidate ensemble topology with cross-validation:



The root mean square error of topology j: $\hat{e}_j = \sqrt{\frac{1}{l}\sum_{i=1}^{l}[\epsilon(x) - f(x)]^2}$

The topology selected is that having the lowest prediction error.

Proposed framework



Performance Analysis 1

- Using an established set of mathematical test functions.
- Comparing the proposed algorithm to:
  - Variant 1 (V1): Only RBF metamodel, no ensemble.
  - Variant 2 (V2): Fixed ensemble RBF+Kriging+RBF network
  - EA-PS, EI-CMA-ES: Reference algorithms from the literature.
- Limit of 200 evaluations of the true function. 30 repeats.
- This setup was used to check the contribution of the dynamic topology selection.

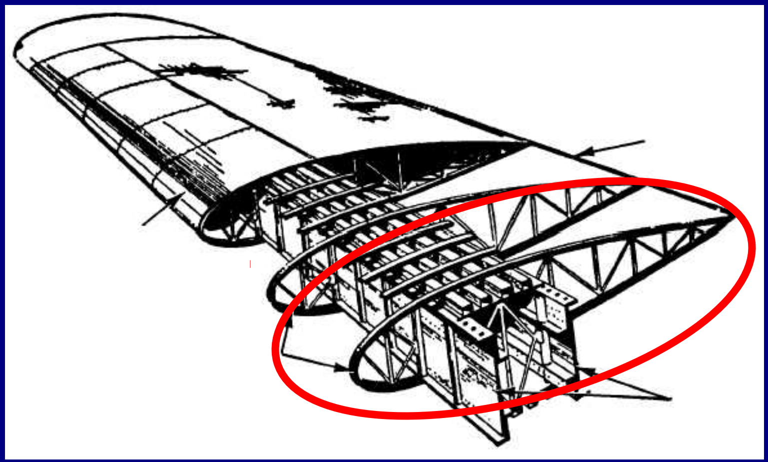
Performance Analysis 1

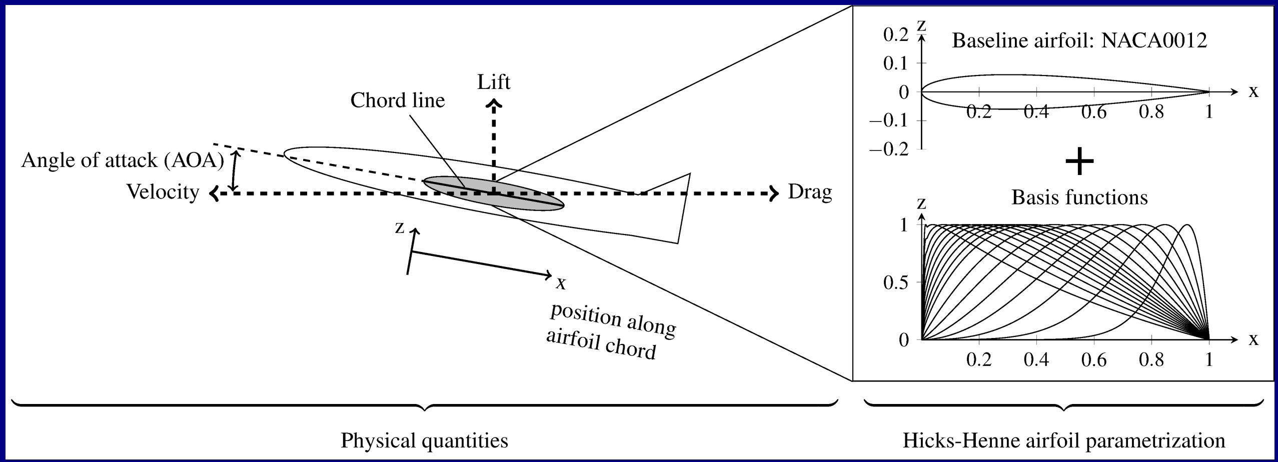
		Proposed	V1	V2	EA–PS	EI–CMA-ES
Ackley-10	Mean	**7.705e+00**	1.455e+01	1.356e+01	5.241e+00	1.796e+01
	SD	8.359e+00	4.649e+00	8.051e+00	5.590e-01	1.529e+00
	Median	**2.314e+00**	1.592e+01	1.908e+01	5.408e+00	1.797e+01
	Min(best)	9.007e-02	2.383e+00	3.457e+00	4.098e+00	1.443e+01
	Max(worst)	1.836e+01	1.825e+01	2.048e+01	6.010e+00	1.988e+01
	α			0.01		0.01
Griewank-10	Mean	**1.304e-01**	1.972e-01	2.078e-01	9.579e-01	9.338e-01
	SD	1.851e-01	1.714e-01	2.213e-01	1.076e-01	2.435e-01
	Median	**7.747e-02**	1.294e-01	1.357e-01	9.862e-01	1.007e+00
	Min(best)	9.350e-03	3.569e-02	2.290e-02	7.146e-01	2.441e-01
	Max(worst)	6.505e-01	5.661e-01	7.601e-01	1.046e+00	1.050e+00
	α				0.01	0.01
Rastrigin-5	Mean	**6.377e+00**	9.360e+00	8.018e+00	7.631e+00	2.131e+01
	SD	3.728e+00	7.852e+00	8.349e+00	4.811e+00	4.890e+00
	Median	5.980e+00	7.464e+00	**4.298e+00**	7.226e+00	2.139e+01
	Min(best)	1.997e+00	1.005e+00	3.369e+00	1.621e+00	1.353e+01
	Max(worst)	1.195e+01	2.787e+01	3.076e+01	1.456e+01	3.006e+01
	α					0.01
Rosenbrock-20	Mean	**5.839e+02**	1.031e+03	8.186e+02	8.435e+02	3.967e+03
	SD	2.094e+02	5.818e+02	3.823e+02	3.012e+02	9.406e+02
	Median	**5.956e+02**	8.665e+02	7.932e+02	7.782e+02	3.685e+03
	Min(best)	2.143e+02	5.483e+02	3.078e+02	4.676e+02	3.141e+03
	Max(worst)	8.905e+02	2.517e+03	1.521e+03	1.439e+03	6.144e+03
	α		0.01		0.05	0.01

The proposed algorithm with dynamic topology selection consistently outperformed the other algorithms.

Performance Analysis 2

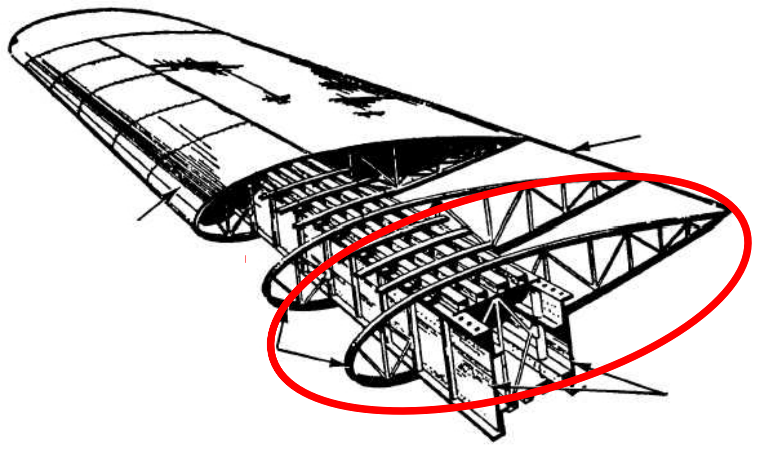
- Using an engineering problem of airfoil shape optimization.

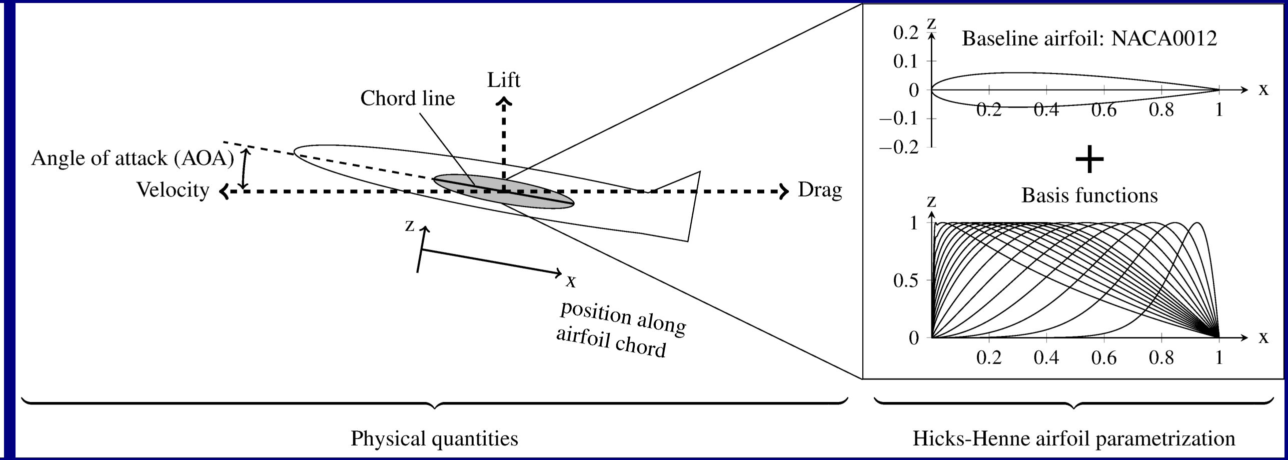




Performance Analysis 2

- Using an engineering problem of airfoil shape optimization.





- 2 cases: 6 or 20 design variables per airfoil.
- Limit of 200 simulation calls.
- Benchmarking against the same algorithms as before.

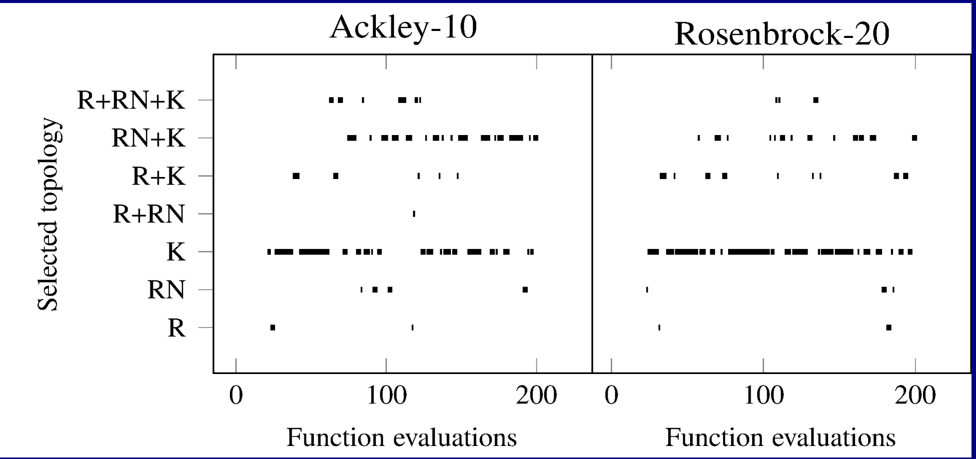
Performance Analysis 2

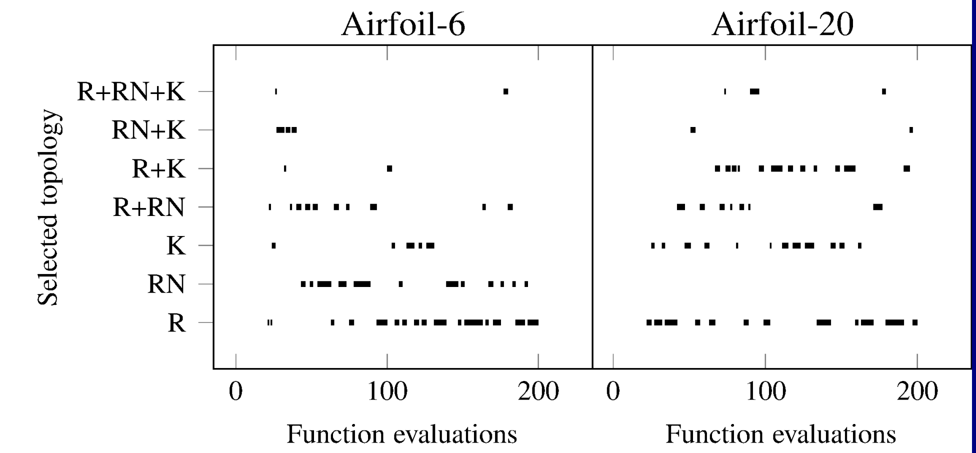
		Proposed	V1	V2	EA–PS	EI–CMA-ES
6D	Mean	**−8.360e+01**	−8.048e+01	−8.203e+01	−7.799e+01	−7.231e+01
	SD	1.320e+01	1.659e+01	2.261e+01	2.250e+00	7.159e−01
	Median	**−7.567e+01**	−7.533e+01	−7.554e+01	−7.831e+01	−7.264e+01
	Min(best)	−1.068e+02	−1.268e+02	−1.436e+02	−8.036e+01	−7.290e+01
	Max(worst)	−7.488e+01	−7.174e+01	−6.405e+01	−7.238e+01	−7.099e+01
	α					0.01
20D	Mean	**−3.247e+00**	−3.202e+00	−3.239e+00	−3.174e+00	−3.212e+00
	SD	6.421e−02	6.991e−02	8.932e−02	8.887e−02	9.405e−02
	Median	**−3.231e+00**	−3.208e+00	−3.206e+00	−3.142e+00	−3.202e+00
	Min(best)	−3.354e+00	−3.303e+00	−3.414e+00	−3.348e+00	−3.327e+00
	Max(worst)	−3.151e+00	−3.098e+00	−3.134e+00	−3.070e+00	−3.036e+00
	α				0.05	

- The proposed algorithm outperformed the other algorithms also in these test problems.

Ensemble updates

- Was the dynamic selection important? how often was the topology updated?





R:RBF, K:Kriging, RN: RBF network

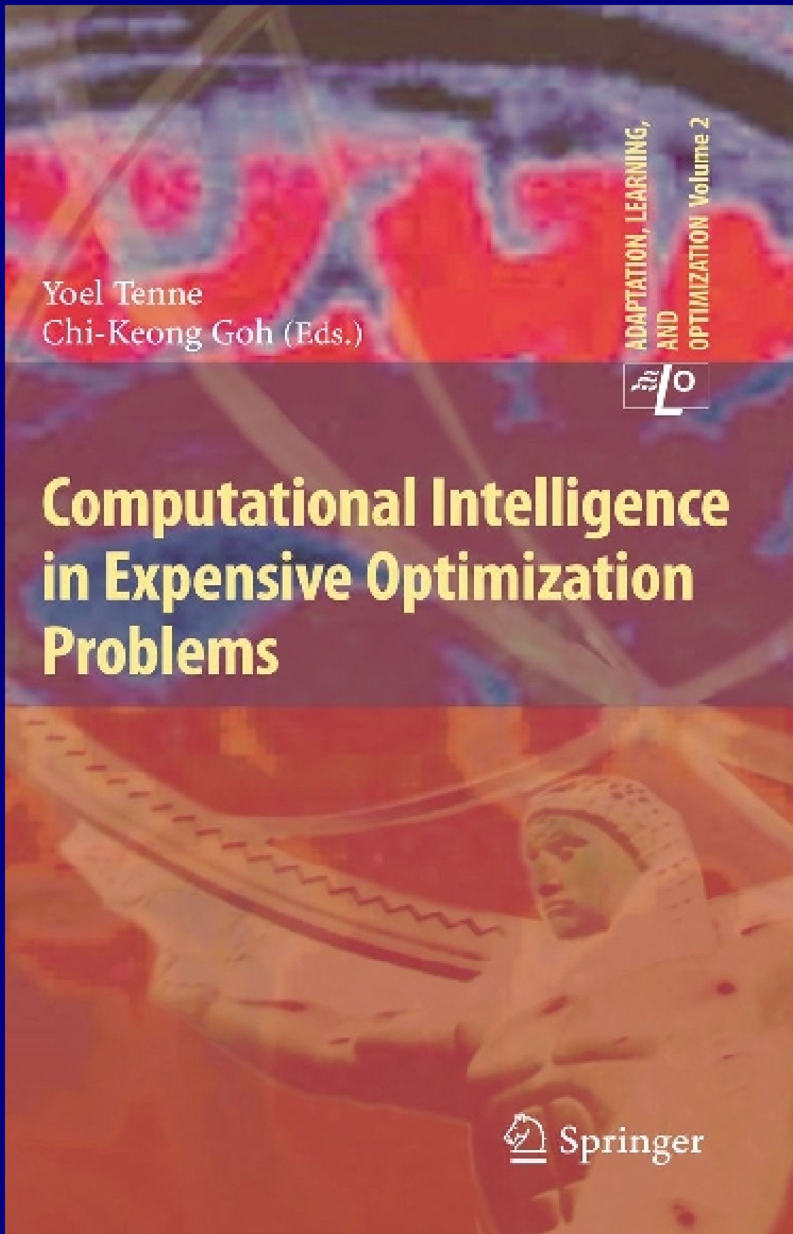
- The optimal topology varied between problems and during the search itself.
- No single topology was the overall optimal.

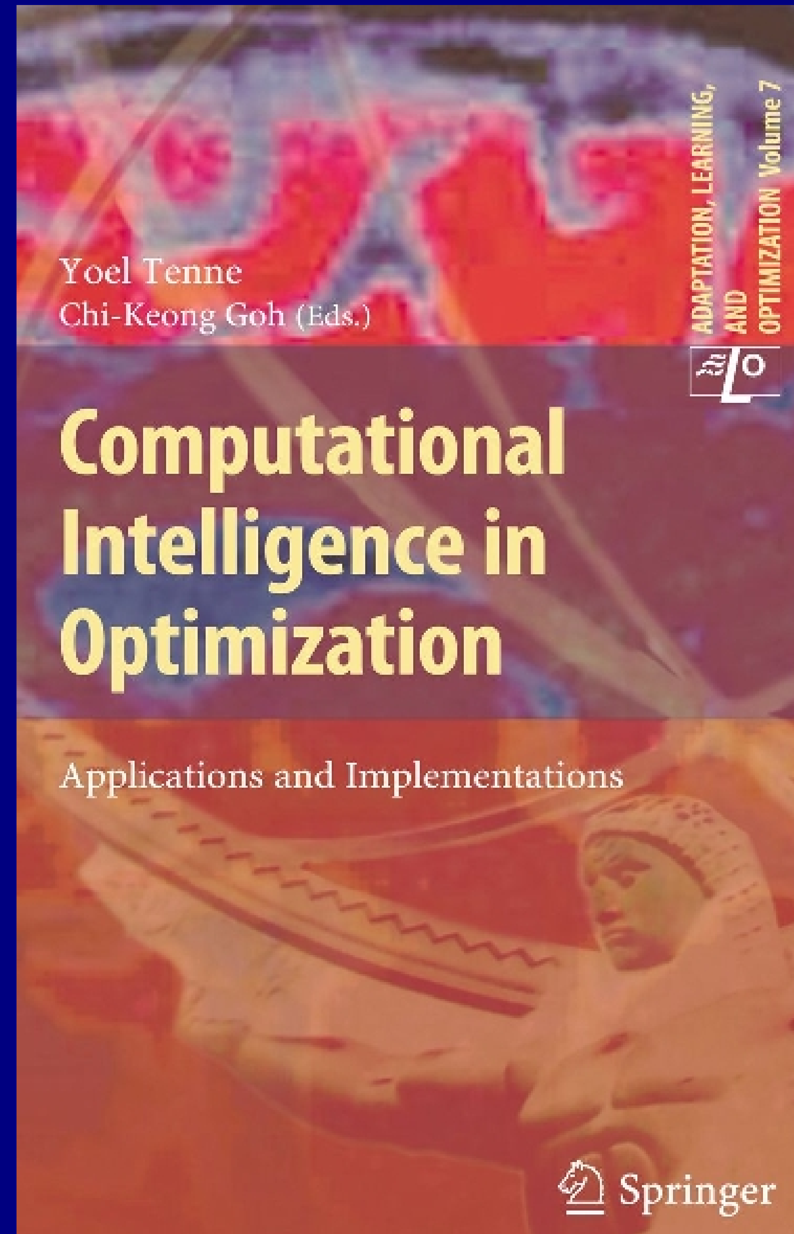
- **The dynamic selection was essential to using an optimal topology.**

Summary

- Ensembles are used to improve the prediction accuracy in simulation-driven optimization.
- This study has proposed a **dynamic topology selection** approach.
- Analysis shows:
  a) an improved search effectiveness, and
  b) that the optimal topology varied dynamically during the search, and so there is no single optimal topology.

Thank you


Yoel Tenne
Chi-Keong Goh (Eds.)
ADAPTATION, LEARNING, AND OPTIMIZATION Volume 2
Computational Intelligence in Expensive Optimization Problems
Springer


Yoel Tenne
Chi-Keong Goh (Eds.)
ADAPTATION, LEARNING, AND OPTIMIZATION Volume 7
Computational Intelligence in Optimization
Applications and Implementations
Springer

Dejan Zupan

Evaluation of Model Parameters: Experiences from Characteristic Value Determination

Panel on "Opportunities and Challenges in Simulation-driven Research"

ADVCOMP 2016

Venice, Italy

October 9 – 13, 2016

*University of Ljubljana, Faculty of Civil and Geodetic Engineering, Slovenia

Simulations in research

- Simulations are a powerful tool when used properly.

Simulations in research

- Simulations are a powerful tool when used properly.
- Quantities we are dealing with are often random variables.

Simulations in research

- Simulations are a powerful tool when used properly.
- Quantities we are dealing with are often random variables.
- We can use simulations to evaluate the parameters of a model.

Simulations in research

- Simulations are a powerful tool when used properly.
- Quantities we are dealing with are often random variables.
- We can use simulations to evaluate the parameters of a model.
- They can be used for confirmation of analytical solutions.

Simulations in research

- Simulations are a powerful tool when used properly.
- Quantities we are dealing with are often random variables.
- We can use simulations to evaluate the parameters of a model.
- They can be used for confirmation of analytical solutions.

Charateristic value

- Let X be a random variable with known CDF $F_X(x)$. The characteristic value of X is such value x_α, that the probability of X being less than x_α equals α:

$$P[X < x_\alpha] = F_X(x_\alpha) = \alpha \quad \longrightarrow \quad x_\alpha = F_X^{-1}(\alpha).$$

Charateristic value

- Let X be a random variable with known CDF $F_X(x)$. The characteristic value of X is such value x_α, that the probability of X being less than x_α equals α:

$$P[X < x_\alpha] = F_X(x_\alpha) = \alpha \quad \longrightarrow \quad x_\alpha = F_X^{-1}(\alpha).$$

- It is common to many practical problems that the parameters of the distribution are not known.

Charateristic value

- Let X be a random variable with known CDF $F_X(x)$. The characteristic value of X is such value x_α, that the probability of X being less than x_α equals α:

$$P[X < x_\alpha] = F_X(x_\alpha) = \alpha \quad \longrightarrow \quad x_\alpha = F_X^{-1}(\alpha).$$

- It is common to many practical problems that the parameters of the distribution are not known.
- The parameters are estimated from the random sample. The characteristic value estimate is itself a random variable, here denoted by $\hat{X}_\alpha$.

Charateristic value

- Let X be a random variable with known CDF $F_X(x)$. The characteristic value of X is such value x_α, that the probability of X being less than x_α equals α:

$$P[X < x_\alpha] = F_X(x_\alpha) = \alpha \quad \longrightarrow \quad x_\alpha = F_X^{-1}(\alpha).$$

- It is common to many practical problems that the parameters of the distribution are not known.
- The parameters are estimated from the random sample. The characteristic value estimate is itself a random variable, here denoted by $\hat{X}_\alpha$.
- For any previously prescribed confidence interval α_λ a characteristic value estimate, $\hat{X}_{\alpha,\lambda}$, should be determined, such that

$$P[\hat{X}_{\alpha,\lambda} < x_\alpha] = 1 - \alpha_\lambda.$$

Solution

- For normal random variables analytical solution can be obtained: ZUPAN, SRPČIČ, TURK: Characteristic value determination from small samples. Structural safety, 2007, vol. 29, no. 4, p. 268-278.

Solution

- For normal random variables analytical solution can be obtained: ZUPAN, SRPČIČ, TURK: Characteristic value determination from small samples. Structural safety, 2007, vol. 29, no. 4, p. 268-278.
- Confirmation of analytical results.

Solution

- For normal random variables analytical solution can be obtained: ZUPAN, SRPČIČ, TURK: Characteristic value determination from small samples. Structural safety, 2007, vol. 29, no. 4, p. 268-278.
- Confirmation of analytical results.
- Results for other distributions.

Solution

- For normal random variables analytical solution can be obtained: ZUPAN, SRPČIČ, TURK: Characteristic value determination from small samples. Structural safety, 2007, vol. 29, no. 4, p. 268-278.
- Confirmation of analytical results.
- Results for other distributions.
- Numerical solution using simulations.

Solution

- For normal random variables analytical solution can be obtained: ZUPAN, SRPČIČ, TURK: Characteristic value determination from small samples. Structural safety, 2007, vol. 29, no. 4, p. 268-278.
- Confirmation of analytical results.
- Results for other distributions.
- Numerical solution using simulations.
- The characteristic value estimate is based on one-parameter formula:

$$\hat{X}_{\alpha,\lambda} = \bar{X} + \lambda S_X^*.$$

Algorithm

Assumption of a distribution and its exact parameters.

Evaluation of exact characteristic value x_α.

Estimation of initial values for λ.

Start of bisection iterations.

  Loop over simulations.

    Generation of a random sample according to the chosen distribution.

    Mean and standard deviation estimation from the sample.

    Calculation of the estimate $\hat{X}_{\alpha,\lambda}$.

  End loop.

  Estimation of probability $P\left[\hat{X}_{\alpha,\lambda} < x_\alpha\right]$.

  Update the value of λ.

Continue bisection iterations until $\left|P\left[\hat{X}_{\alpha,\lambda} < x_\alpha\right] - (1-\alpha_\lambda)\right| \geq \delta$.

Challenges

- Computational demands.

Challenges

- Computational demands.
- Acceptance of the results.

Challenges

- Computational demands.
- Acceptance of the results.
- European building standards.

Challenges

- Computational demands.
- Acceptance of the results.
- European building standards.