POLITECNICO
MILANO 1863

Using Competitions and Simulations for Robotics Experiments

Francesco Amigoni

Summary

- Robot competitions as experiments
- Simulated experiments

Robot competitions as experiments

Robot competitions are popular!

Robot competitions serve several (sometimes overlapping) purposes:

- promote education and research
- push the field forward
- entertain general audience
- build community
- ...

Competitions as experiments

In recent years, a point of view considering competitions as experiments has emerged

[Anderson et al., "Recasting robotics challenges as experiments", IEEE Robotics and Automation Magazine, 2011]

"RoCKIn is an EU project that will be run over the next three years, consisting of robot competitions, […] facilitating cognitive and networked robot systems' testing, and streamlining research and development through standardised testbeds and benchmarks"

"euRathlon is a new outdoor robotics competition which invites teams to test the intelligence and autonomy of their robots in realistic mock emergency-response scenarios"

At a first sight…

Competitions (and challenges) and experiments share some similarities… but also have some differences

[Takayama, "Toward a science of robotics: Goals and standards for experimental research", RSS Workshop on Good Experimental Methodology in Robotics, 2009]

Competitions	Experiments	
Precisely definite settings	Controlled conditions	✓
Measuring performance: scoring	Measures and criteria	✓
Evaluating a whole robot system	Evaluating just one (or few) robot ability	✗
Hardly repeatable in the same conditions (e.g., challenges)	Intended to be repeatable in the same conditions	✗

Purpose

I would like to take a deeper critical view on the relationships between competitions and experiments

Spoiler: different competitions map to different types of experiments

- Definition of terms: competition and experiment
- Examples of mapping from competitions to experiments

Definition of terms

Competitions

"The act or process of trying to get or win something (such as a prize or a higher level of success) that someone else is also trying to get or win" (Merriam-Webster)

A robot competition usually involves:

- some **robots**
- a dynamic, but rather controlled, **environment** (arena)
- clear **measures** of success: goals, scoring rules, thresholds, human supervisors, …

Experiments

An experiment is a controlled experience, namely a set of observations and actions, performed in a controlled context, to test a given hypothesis

In robotics, experiments are the *rigorous* empirical practice to gain and check knowledge about a system

Some principles are usually ascribed to experiments: replicability, reproducibility, comparison, generalization, …

How are experiments intended and employed in computing?

[Tedre, The Science of Computing: Shaping a Discipline, CRC Press/Taylor & Francis, 2014]

Types of experiments

Feasibility experiment: empirical demonstration, existence of proof of the ability to build a tool or a system

Trial experiment: evaluation of various aspects of a system using some predefined variables which are often measured in laboratories

Field experiment: evaluating the performances of a system against some measures, outside the laboratory in complex sociotechnical contexts

Comparison experiment: comparing different solutions with the goal of looking for the best solution of a specific problem; comparison is made in some setup and is based on some measures and criteria to assess the performance

Controlled experiment: the golden standard of experimentation of traditional scientific disciplines, refers to the original idea of experiment as controlled experience, where the activity of rigorously controlling the factors that are under investigation is central, while eliminating the confounding factors, and allowing for generalization and prediction

Examples of competitions

RoboCup Middle Size Soccer League

Two robotic teams play against each other in a soccer game

The environment is precisely defined and can be easily reproduced (not true for the opponent team and some conditions, like light and noise)

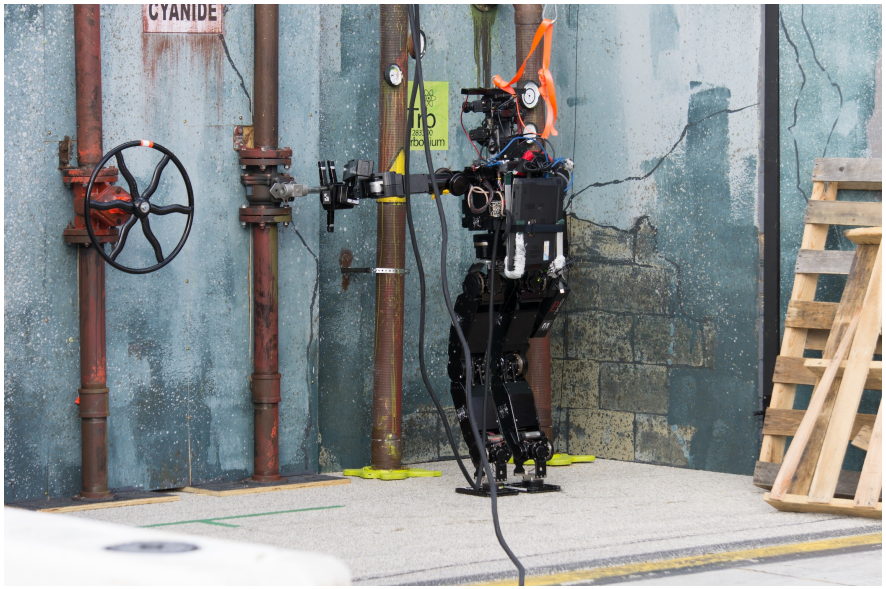
The measures and the criteria according to which the two robotic systems (teams) are compared are clearly defined only for the purposes of the game

This competition can be considered as a feasibility experiment and, partly, as a trial experiment

DARPA Robotics Challenge (DRC)

The DRC consists of tasks related to human assistance in responding to disasters

"It was designed to be extremely difficult"



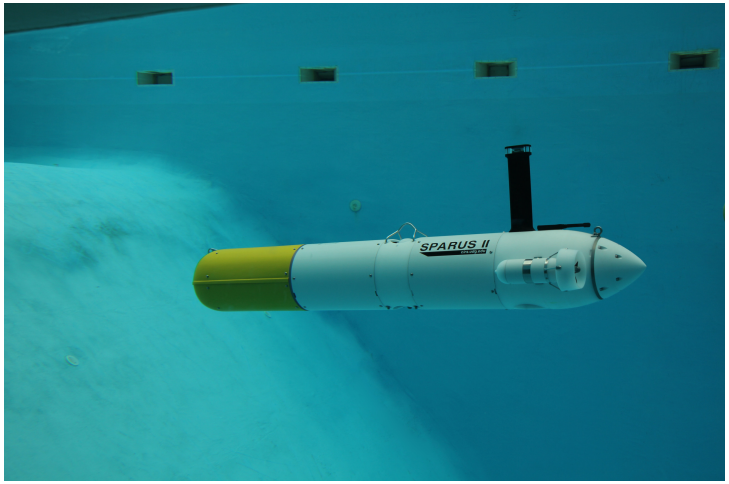
Tasks settings are defined rather precisely

Task completion is evaluated using thresholds (e.g., number of open valves, plus bonus), also time is considered

This competition could work as a field experiment

euRathlon

euRathlon2013 has been about land robotics, euRathlon2014 about sea robotics, and euRathlon2015 about air, land, and sea robotics



Settings and tasks are defined precisely in the rules

"Competitors have the possibility to deal with real life conditions (i.e., limited visibility and salty water)"

Scores are a mix between measured quantities and subjective judgments given by a Judging Team

This competition seems to move from trial to field experiment

ICRA Humanitarian Robotics and Automation Technology Challenge

A robot has to explore an environment performing autonomous landmine detection

The arena and the mines specifications are defined in the rules and can be easily reproduced

Simulation + real world runs (remote, same robot for all teams)

Scoring combines the number of correctly detected (and unexploded!) mines, the area swept, …

This competition could be considered as standing between trial and field experiments

RoCKIn

RoCKIn@Home focuses on domestic service robots that have to perform some socially useful tasks

Home settings are precisely defined in the rules

Task benchmarks: evaluating the performance of integrated robotic systems in performing tasks

Evaluation is performed according to achievements (yes/no)

Functionality benchmarks: evaluating the performance of specific sub-systems (like object recognition and localization)

Evaluation is performed according to specified quantitative measures and criteria

This competition comes closer to comparison experiments

Summary
(not a ranking!)

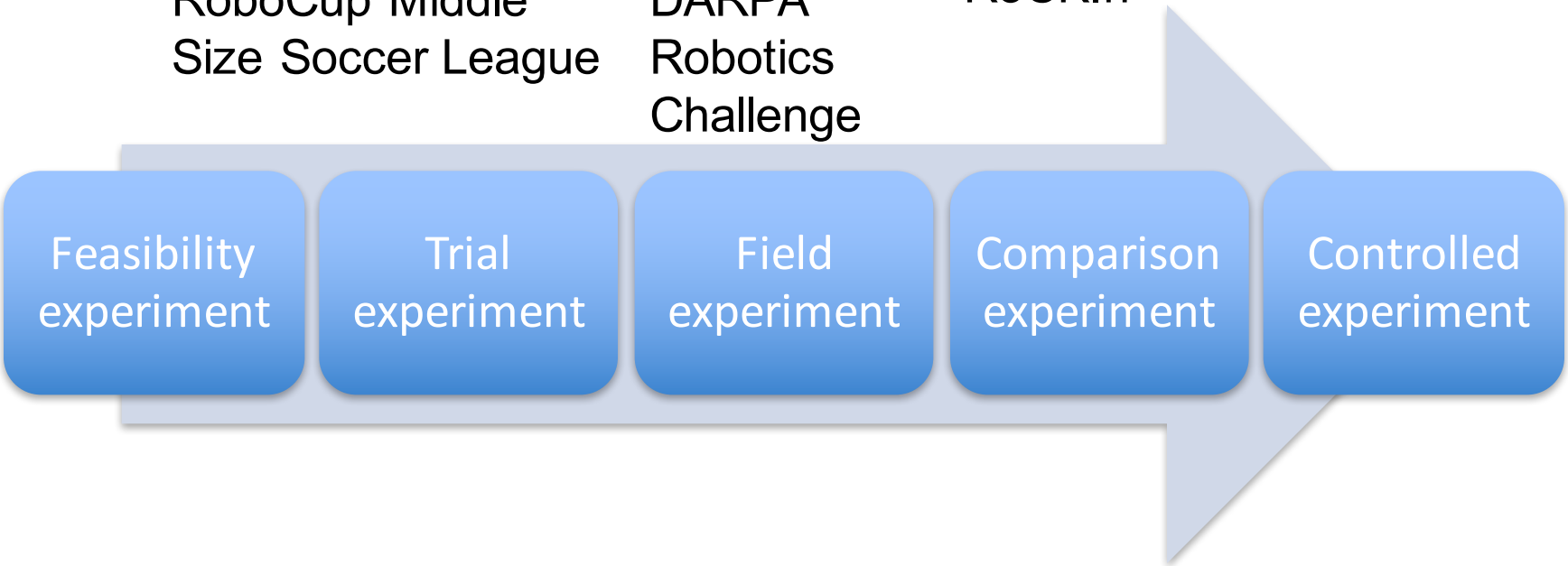
ICRA Humanitarian Robotics and Automation Technology Challenge

euRathlon

RoboCup Middle Size Soccer League

DARPA Robotics Challenge

RoCKIn



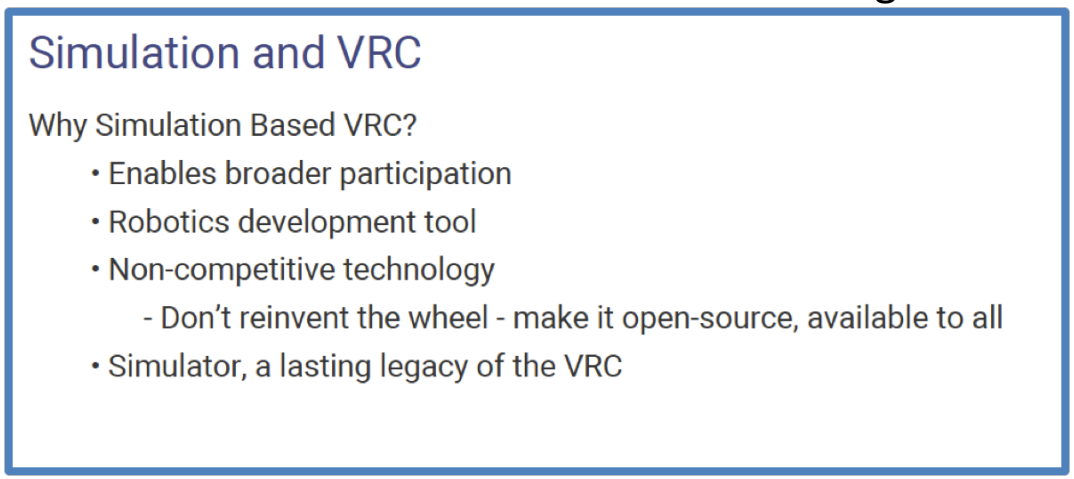
Simulation experiments

Simulations are increasingly used in robotics!

Simulations serve several (sometimes overlapped) purposes in robotics:

- preliminary test robot abilities (before going to the real world)
- perform large number of runs
- create situations that are impossible in reality
- …

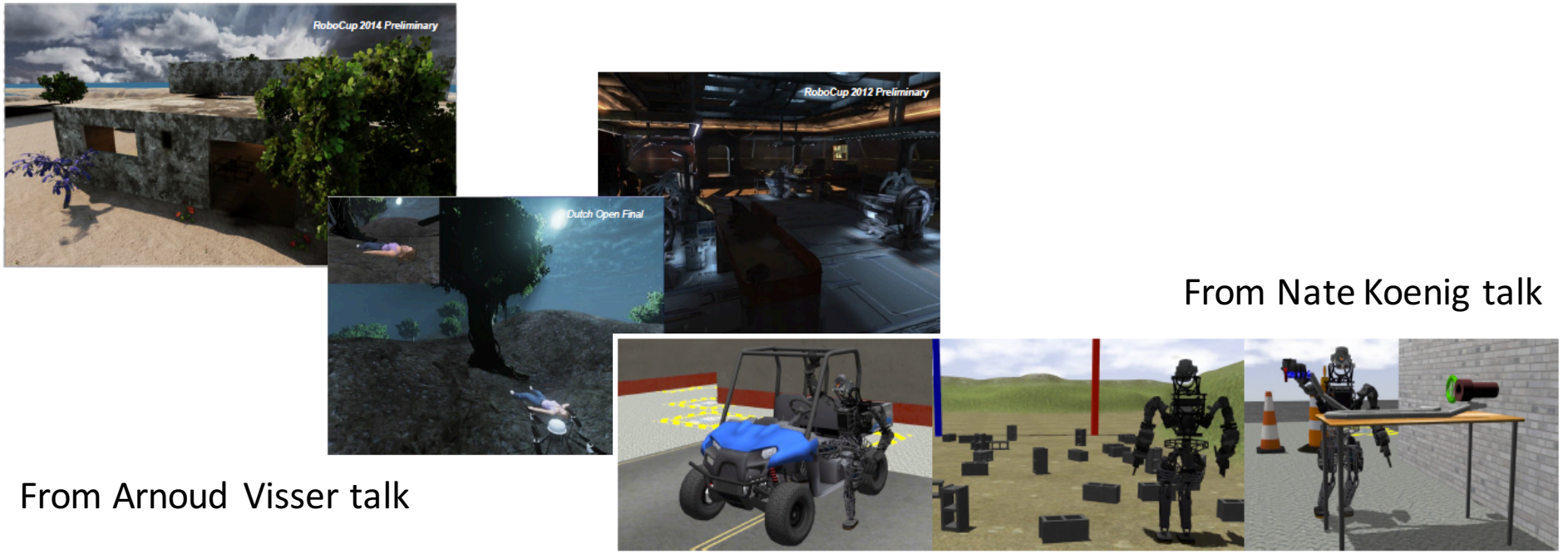
From Nate Koenig talk



International Conference on Simulation, Modeling, and Programming for Autonomous Robots (SIMPAR)

Simulations in robotics (1)

- Simulation = model + execution
- System to be reproduced: robots that interact with an environment
- Model: representation of the robots and of their interaction with the environment



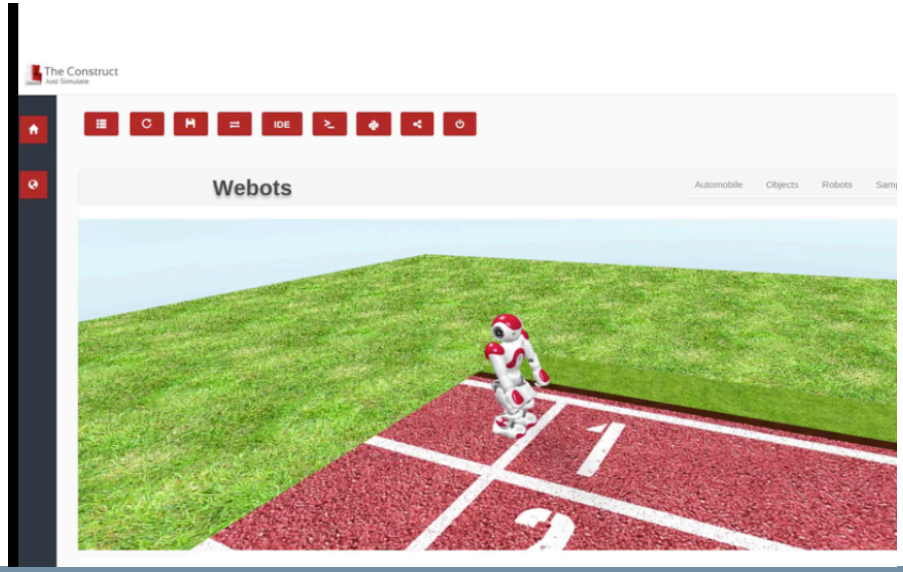
From Nate Koenig talk

From Arnoud Visser talk

Simulations in robotics (2)

- Good models of the robots are available: they are artifacts and we (should!) know how they work
- Models of interaction with the environment are more complex
  - Real world is complex
  - Errors in movements: slippage, …
  - Errors in perception: lights, reflections, …
  - Unpredictable interactions

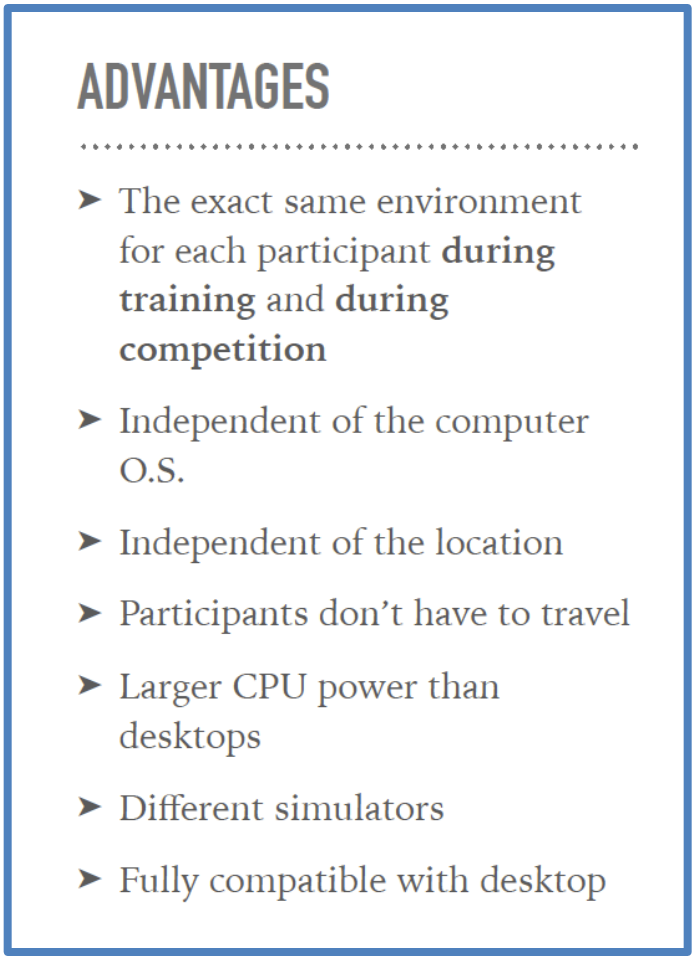
From Ricardo Tellez talk



(Some of the) Advantages of simulations in experiments

- Comparison
  - Common ground for comparing performance
  - Controlled settings
  - Uniformly measured parameters
- Reproducibility and repeatability
  - Ease of setting up experiments
- Justification/explanation
  - Availability of ground truth
- Generalization

From Ricardo Tellez talk



Simulations and experiments in robotics (1)

Simulations can be used as parts of experiments

Experiments and simulations are controlled experiences…

…but experiments are something more: purpose (e.g., confirm/refuse a theory), interpretation of results, …

A simulation *per se* is not an experiment!
Example: flight simulator

Simulations and experiments in robotics (2)

Why

To an d?”
by te
situat

“I wa est it
in sit

- rld
- o the
-

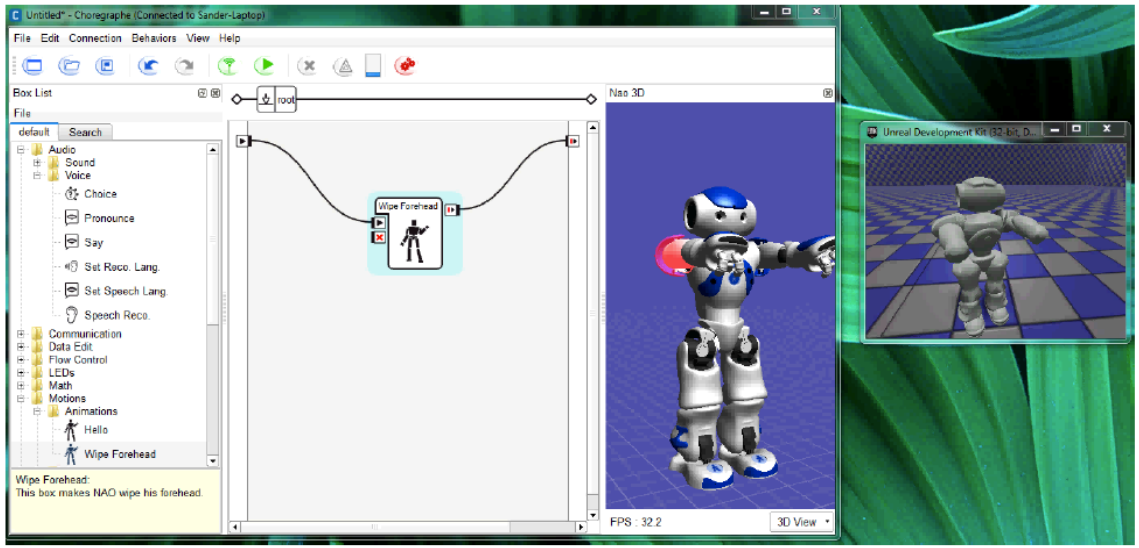
Simulations and experiments in robotics (3)

Why are simulations commonly used in experiments in some areas of robotics and not in others?

Example: locomotion of legged robots and manipulation
vs. detecting people for service robots

Are some models
"better" than others?

From Arnoud Visser talk

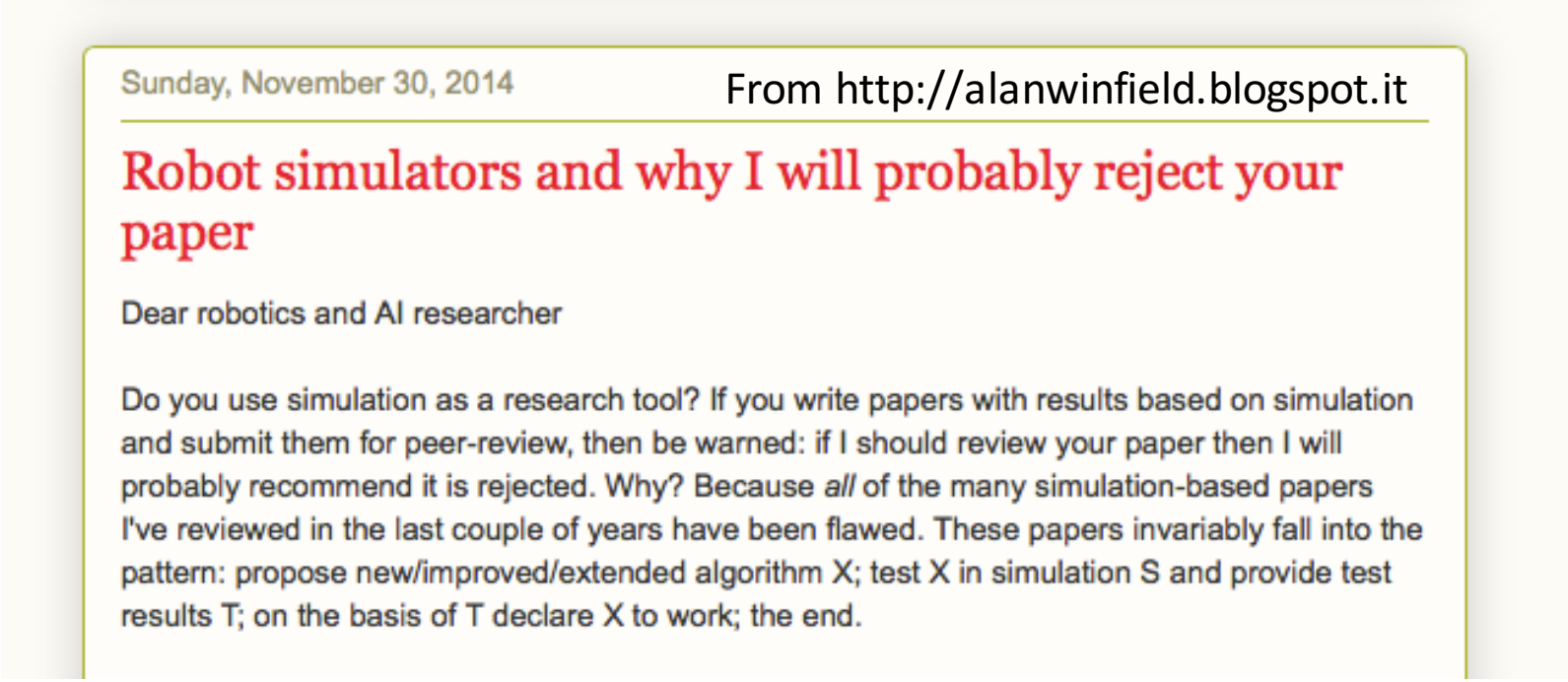


Simulations and experiments in robotics (4)

Why do we trust a simulated experiment we read about in a paper?

Pragmatic answer: because it appears in an accepted paper, meaning that the reviewers (or the community) considered the experiment convincing

But we should provide a more theoretical answer…



Sunday, November 30, 2014

From http://alanwinfield.blogspot.it

Robot simulators and why I will probably reject your paper

Dear robotics and AI researcher

Do you use simulation as a research tool? If you write papers with results based on simulation and submit them for peer-review, then be warned: if I should review your paper then I will probably recommend it is rejected. Why? Because *all* of the many simulation-based papers I've reviewed in the last couple of years have been flawed. These papers invariably fall into the pattern: propose new/improved/extended algorithm X; test X in simulation S and provide test results T; on the basis of T declare X to work; the end.

Simulations and experiments in robotics (5)

Pool of strategies for reliability of robot simulations:

- prior successes of the model building techniques adopted
- production of outcomes fitting well with previously accepted data, observations, and intuitions
- capability of making successful predictions
- capability of producing practical accomplishments

Simulations in robotics: Data sets

- Many data sets collected in real world experiments are publicly available (Radish, Rawseeds, …)
- From the one hand, using these data sets amounts to set up a simulation
  - Data sets are models of the interaction of the robots with environments
  - "World is the best model of itself"
- From the other hand, using these data sets can be seen as part of real experiments
  - Collecting and processing data are performed in different places at different times
- Not useful for closed-loop systems!

Competitions, simulations, and experiments in robotics

Many open issues, no definite answer

Things not discussed here:

- related topics: benchmarking, standardization, …
- quantitative evaluation of how many papers in some venues use simulations or real experiments
- custom-made vs. “standard” simulators
- …