EUCogII-Euron Workshop
Challenges, Good Experimental Methodology & Benchmarking
http://www.eucognition.org/eucogii-euron-workshop

8.04.2011, 13:00-18:00

Organisers: Fabio Bonsignorio & Vincent C. Müller
with the backing of Angel P Del Pobil and John Hallam

Speakers:
Anne Bajart (EU Commission, DG INFSO, Luxembourg)
Fabio Bonsignorio (U Carlos III, Madrid & Heron Robots)
Massimo Caccia (CNR, Italy)
Manja Lohse (Bielefeld University)
Matteo Matteucci (Polimi, Milano)
Giorgio Metta (IIT, Genoa)
Vincent C. Müller (Anatolia College Thessaloniki, Thessaloniki)

Why should we think about *Challenges, Good Experimental Methodology & Benchmarking*?

- **to *shed light on cognitive systems and robotics research* in term of challenges, experiments and performance measurement**

- **to pass *from a pre-paradigmatic to a mature scientific phase* by means of a more rigorous scientific**

- **to foster a cumulative advancement of our knowledge of natural and artificial cognitive agents**

- **to correctly appreciate disruptive innovation and to give solid ground to theoretical questions (e.g.: "Embodiment fad or future?", the subject of the next EUCogII meeting)**

EUCOGII

European ROBOTICS Research Network EURON

PROGRAM

13:00-15:00 Good Experimental Methodology and Benchmarking

13:00 A. Bajart: Benchmarking activities in FP7 Challenge 2 - Cognitive Systems and Robotics

13:20 F. Bonsignorio: Ongoing Activities on Replicable Robotics Research and Benchmarking

13:40 G Metta: Common Platforms and Benchmarks for Cognitive systems

14:00 M Matteucci: Rawseeds and SLAM benchmarking

14:20 M. Lohse: Challenges in the Evaluation of Interactive Cognitive Systems

14:40 M. Caccia: Benchmarking in Marine Robotics

15:00-15:30 Coffee Break

15:30-18:00 Experimental method, Benchmarking and Systematic Challenges

15:30 V. Müller: Presentation of the EUCogII Rapperswil workshop on 'challenges for artificial cognitive systems'

15:50 Open Discussion

17:30 A manifesto early draft

www.heronrobots.com/GEMSIGEvents.html

PerMIS'08 Performance Metrics for Intelligent Systems Workshop (NIST Gaithersburg, USA, 19-21 August, 2008)

Workshop onExperimental Methodology and Benchmarking in Robotics Research

at RSS'08 (Zurich, Switzerland, June 28, 2008).

CogGemBench 2008 (Karlsruhe, Germany, April 1,2008)

GEMBenchForum 2008 (Prague, Czech Republic, March 25-26, 2008)

IEEE/RSJ IROS 2007 Workshop on Performance Evaluation and Benchmarking for Intelligent Robots and Systems (San Diego, USA, November 2, 2007).

IEEE/RSJ IROS 2006 Workshop on Benchmarks in Robotics Research (Beijing, China, October 10, 2006).

www.heronrobots.com/GEMSIGEvents.html

IEEE/RSJ IROS 2009 Workshop on Performance Evaluation and Benchmarking for NEXT Intelligent Robots and Systems (St. Louis, MO, USA, October 15, 2009)

PerMIS'09 Performance Metrics for Intelligent Systems Workshop (NIST Gaithersburg, USA, 21-23 September, 2009)

Workshop on Good Experimental Methodology in Robotics at RSS'09 (Seattle,WA, USA, June 28, 2009).

Workshop on Good Experimental Methodology and Benchmarking in Robotics Research and Applications - GEMBENCH09 at the Euron Annual General Meeting (Leuven, Belgium, 6-8 April, 2009)

IEEE/RSJ IROS 2008 Workshop on Performance Evaluation and Benchmarking for Intelligent Robots and Systems (Nice, France, 26 September, 2008), check the program

Clawar 2008Session on Benchmarking and standardization for clawar and mobile robots (Coimbra, Portugal, 8-10 September, 2008)

www.heronrobots.com/GEMSIGEvents.html

Upcoming events

Workshop on the path Towards Replicable Experiments in Robotics Research at IEEE ICRA 2011 (Shanghai, China, May 13, 2011)

EuCogII-Euron Workshop: Challenges, GEM & Benchmarking in Cognitive Systems and Robotics during the European Robotics Forum (Vasteras, Sweden, April 8, 2011)

Recent past events

Excellence Cluster Cognitive Interaction Technology (CITEC) at Bielefeld University Workshop on Benchmarking and Evaluation of Interactive Cognitive Systems (Bielefeld, Germany, February 7-8, 2011)

EuCogII Workshop on Challenges for Cognitive Systems (Rapperswil, Switzerland, January 28-30, 2011)

IEEE/RSJ IROS 2010 Workshop on Performance Evaluation and Benchmarking for Intelligent Robots and Systems with Cognitive and Autonomy Capabilities (Taipei, Taiwan, October 22, 2010)

www.heronrobots.com/GEMSIGEvents.html

Upcoming events

Workshop on the path Towards Replicable Experiments in Robotics Research at IEEE ICRA 2011 (Shanghai, China, May 13, 2011)

EuCogII-Euron Workshop: Challenges, GEM & Benchmarking in Cognitive Systems and Robotics during the European Robotics Forum (Vasteras, Sweden, April 8, 2011)

- 'Look Ma, No Hands' sindrome?
- Replication of experiments
- Performance measure benchmarks to allow results comparison
- Needed to foster research advancement and enable practical application of research achievements

if robotics aims to be serious science replication of experiments deserves serious attention.
Are we really able to verify if and by which measure new procedures and algorithms proposed in research papers constitute a real advancement and can be used in new applications?
New more successful implementations of concepts already presented in literature, but not implemented with exhaustive experimental methodology, risk to be ignored, if appropriate benchmarking procedures, allowing to compare the actual practical results with reference to standard accepted procedures, are not in place.

The analysis of the state of the art of experimental methodology (Amigoni et al. 2009) evidences that a stable experimental methodology is still lacking

- Even in the engineering sense of a set of strategies for good experimental design practices

- Do-it-your self approach

Table 1 Papers selected to provide a representative picture of experimental activities in mobile robot localization and mapping

Authors (Year)	Title	Published in
Amigoni (2008)	*Experimental evaluation of some exploration strategies for mobile robots*	IEEE Int. Conf. on Robotics and Automation
Comport et al. (2007)	*Accurate quadrifocal tracking for robust 3D visual odometry*	IEEE Int. Conf. on Robotics and Automation
Davison (2003)	*Real-time simultaneous localisation and mapping with a single camera*	IEEE Int. Conf. on Computer Vision
Davison et al. (2007)	*MonoSLAM: Real-time single camera SLAM*	IEEE Trans. on Pattern Analysis and Machine Intelligence
Frese et al. (2005)	*A multilevel relaxation algorithm for simultaneous localization and mapping*	IEEE Trans. on Robotics
Grisetti et al. (2007)	*Improved techniques for grid mapping with rao-blackwellized particle filters*	IEEE Trans. on Robotics
Gutmann and Konolige (1999)	*Incremental mapping of large cyclic environments*	IEEE Int. Symp. on Comp. Intelligence in Robotics and Automation
Hahnel et al. (2003)	*An efficient fastSLAM algorithm for generating maps of large-scale cyclic environments from raw laser range measurements*	IEEE/RSJ Int. Conf. on Intelligent Robots and Systems
Leonard and Durrant-Whyte (1991)	*Simultaneous map building and localization for an autonomous mobile robot*	IEEE/RSJ Int. Workshop on Intelligent Robots and Systems
Liu and Thrun (2003)	*Results for outdoor-SLAM using sparse extended information filters*	IEEE Int. Conf. on Robotics and Automation
Lu and Milios (1997)	*Globally consistent range scan alignment for environment mapping*	Autonomous Robots
Minguez et al. (2006)	*Metric-based iterative closest point scan matching for sensor displacement estimation*	IEEE Trans. on Robotics
Montemerlo et al. (2003)	*FastSLAM 2.0: An improved particle filtering algorithm for simultaneous localization and mapping that provably converges*	Int. Joint Conf. on Artificial Intelligence
Neira and Tardos (2001)	*Data association in stochastic mapping using the joint compatibility test*	IEEE Trans. on Robotics and Automation
Newman et al. (2002)	*Explore and return: Experimental validation of real-time concurrent mapping and localization*	IEEE Int. Conf. on Robotics and Automation
Newman et al. (2006)	*Outdoor SLAM using visual appearance and laser ranging*	IEEE Int. Conf. on Robotics and Automation
Paz et al. (2008)	*Large-Scale 6-DOF SLAM With Stereo-in-Hand*	IEEE Trans. on Robotics

Table 2 Main relations between trends in experimental activities in mobile robot localization and mapping (Sect. 2) and principles of an experimental methodology (Sect. 3.3)

	Principles of an experimental methodology		
	Comparison	Reproducibility and repeatability	Justification/ explanation
Purpose of experiments			
Demonstrating that system works			X
Getting insights on the behavior of the system and assessing limits of applicability		X	X
Comparing the system with competing ones	X		
Data sets			
Publicly available instances and code	X	X	
Use of different environments		X	X
Measured quantities			
Computational complexity	X		X
Computational time/Memory usage	X	X	
Profiling of the total time	X	X	X
Precision	X	X	X
Accuracy	X	X	X
Robustness	X	X	X
Report anomalies in performance		X	

Conference tracks

The Permis workshop started in 2000 and in 2010 it reached the tenth edition. This workshops aim to define measures and methodologies for the evaluation of performance of intelligent systems.

Conference tracks

There are some connections between these workshop topics, those addressed in this talk and the scope of the Workshops on Technical Challenges for Dependable Robots in Human Environments co-sponsored by IARP and the IEEE Robotics and Automation Society. The 2005 workshop was co-sponsored by Euron, the EU network of excellence in robotics.

Conference tracks

From 2006 to 2010 at IROS conference was organized a workshop on performance metrics of robots. And the same at RSS 2008. At RSS2009 and 2010 and ICRA 2010 there were workshops more focused on the experimental methodology. There will be another one at ICRA2011…
Others at ECAI…this one ☺, Rapperswil…

Euron GEM SIG, IEEE TC-PEBRAS

Tools for Performance Metrics

Performance metrics have been developed in various area of robotics for specific purposes.
There are a number of initiatives devoted to define adequate performance metrics in specific subfields.
Here below follows a non exhaustive list, whose main purpose is to exemplify the community attempts to cope with the benchmarking problems.

Tools for Performance Metrics: Data sets

Radish, started in 2003, by Andrew Howard and Nick Roy is a repository of standard data sets with, currently, a main focus on localisation and mapping.
The more common format CARMEN, the open source Carnegie Mellon Robot Navigation Toolkit, for mobile robots. Control environment that provides basic navigation primitives.

Tools for Performance Metrics: Data sets

At present it contains mostly logs of odometry, laser, sonar and other sensor data taken from real and simulated robots and environment maps created by robots or manually.

Tools for Performance Metrics: Data sets

The RAWSEEDS project was an SSA (Specific Support Action) in the EU 6th Frame Program, providing a comprehensive, benchmarking toolkit for SLAM (Simultaneous Localization And Mapping).

EUCOG II

European ROBOTICS Research Network EURON

Performance Metrics: Data sets

It (will) provides a web accessible repository storing standard data sets, based on different sensor sets, and related benchmarks , state-of-the-art solutions to SLAM problems in the form of algorithms and software, and methodologies for the validation of algorithms.

Tools for Performance Metrics

The NIST USAR (Urban Search And Rescue) 'after disaster' scenarios, ranked as yellow, orange and red, are used in RoboCup USAR. They provide an useful conventional reference scenario for USAR applications together with USARsim the open source simulation environment based on the Unreal Tournament gaming engine. The VMAC competition in Virtual Manufacturing.

Competitions and Challenges

RoboCup is probably the most famous competition in robotics. RoboCup is mostly focused on soccer game as a primary domain, and organizes the Robot World Cup Soccer Games and Conferences. Soccer is a very good testbed for multi (robot) agent technologies.

New competitions in search and rescue, based on NIST scenarios, and home assistance have been added, @Home, @Work…

EUCOG II

European ROBOTICS Research Network EURON

Competitions and Challenges

The DARPA Grand Challenge is a famous competion for outdoor robot race on an about 200km circuit in the desert. The 2005 edition was won by Stanford team, with a modified version of a VW Tuaregh, and five teams were able to complete the race. It was the first time
Later DARPA organized the Urban Challenge where the robot have to cope with an urban traffic scenario.

Competitions and Challenges

The European Land-Robot Trial (ELROB), organized by the German Federal Armed Forces (Bundeswehr), is an outdoor robot demonstration with no real competitions or prizes, but otherwise similar to the DARPA Grand Challenge.
It focuses on mobility and RSTA (Reconnaissance, Surveillance, and Target acquisition). It took place in 2006 for the first time, in 2007 a civilian version was organized in Switzerland.

Discussion

It looks apparent that the bare replication of experimemts and the quantitative comparison of research results in robotics raise many challenging issues.
This is due to the variety of applications, tasks, mechanical structures, sensor sets, actuators, control system, software architectures, required levels of flexibility and autonomy, and so on.

Discussion

When we are dealing with Human Robot Interaction in everyday settings also human psychology is involved.
On the other end, there are many initiative trying to define proper standards.
There are benchmarks in some specific areas like visual servoing, SLAM, motion planning, but there is still a lot of work to do.

Discussion

In some experimental works 'entropy measures' on the 'sensory-motor' coordination of different 'robotics' equipment have shown that information metrics can be used to classify, at least, and to get an insight on (semi) autonomous robotics devices, which show an 'emergent behavior', while, in [Chatila,2006], entropy measures are used to rank environment complexity, with reference to the navigation task.

The (Shannon) entropy:

$$H(x) = -\sum_{x \in X} p_x(x) \log p_x(x)$$

Mutual Information

It is given by a function of the mutual information, between the sensors and the actuators connected to that node. The mutual information between two given variables is given by equation (4), where X and Y two random variables:

$$I(X,Y) = -\sum_i \sum_j P_{XY}(i,j) \log \frac{P_X(i)P_Y(j)}{P_{XY}(i,j)}$$

If X and Y are statistically independent eq above gives I(X,Y)=0

EUCOGII

European ROBOTICS Research Network EURON

Lungarella, Sporns (2006)

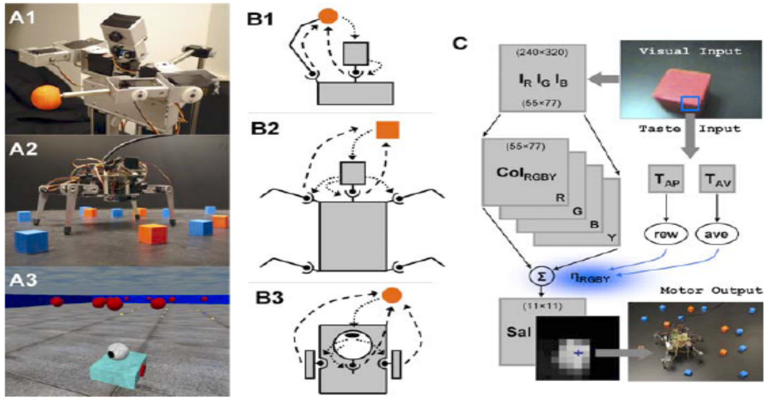


Figure 1. Robots, Sensorimotor Interactions, and Neural Control Architecture

(A1) *Roboto* has a total of 14 DOF, five of which are used in the current set of experiments. Note the system (2 DOF), and the moveable left arm with shoulder, elbow, and wrist joints (3 DOF). The object is tip of the last joint.

(A2) *Strider* has a total of 14 DOF, with four legs of 3 DOF each and 2 DOF in the pan-tilt head system. *Strider* is situated in an environmental enclosure with black walls.

(A3) *Madame* has 4 DOF, with 2 DOF in the pan-tilt system and 2 DOF for the wheels, which are both lc The environment is a square arena bounded by blue walls containing 20 red-colored floating spheres

(B1) *Roboto* engages in sensorimotor interactions via the head system and arm movements; sensor (dashed arrows).

(B2) *Strider* engages in sensorimotor interactions via the head system, as well as via steering signals ger legs.

(B3) *Madame*'s behavior consists of a series of approaches to colored objects and ovations. Fixations action of head and body.

(C) Neural control architecture. The architecture common to all robots is composed of color image ar saliency map *Sal* (see text for details). The peak of the saliency map (blue cross) determines the pan-tilt *Strider*'s neural system contains a value system with taste sensory inputs relayed via a virtual taste sensc ($T_{AP,AV}$), which in turn generates reward and aversiveness signals (rew, ave). These signals are used to η_{RGBY} (see text for details).

DOI: 10.1371/journal.pcbi.0020144.g001

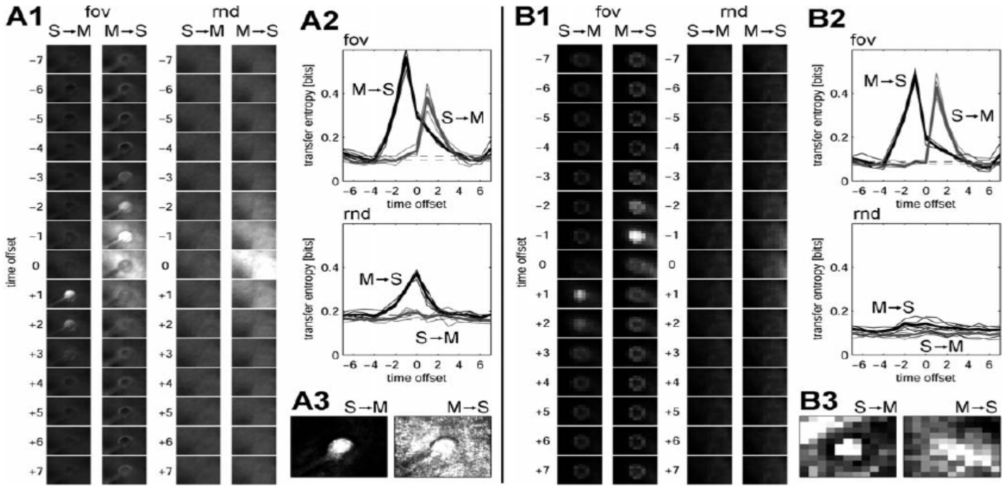


Figure 3. Information Flow (Transfer Entropy) between Sensory Input, Neural Representation of Saliency, and Motor Variables in *Roboto*

(A1) Transfer entropy between array I_R (variable S) and pan-tilt amplitude (variable M). Series of plots show maps of transfer entropy from S to M (S → M) and from M to S (M → S) over visual space (55 × 77 pixels), calculated for offsets between −7 ("M leading S") and +7 ("S leading M") time steps. Plots show data for conditions "fov" and "rnd." The gray scale ranges from 0.0 to 0.5 bits (for all plots in panels A1 and B1).

(A2) Curves show transfer entropy for five individual runs (thin lines) as well as the average over five runs (thick lines) between the single central pixel of array I_R (S) and pan-tilt amplitude (M), for directions M → S (black) and S → M (gray).

(A3) z-Score maps of significant image regions (plotted between $z = 0$ and $z = 6$). The z-scores are expressed as number of standard deviations above background at time offset +1 (S → M) and −1 (M → S). Mean and standard deviation of background is calculated from transfer entropy values at maximal time delays (−7,+7 time steps).

(B) All three panels have the same format as (A), but the neural activations of the saliency map *Sal* are substituted as variable S (11 × 11 neural units).

DOI: 10.1371/journal.pcbi.0020144.g003

Lampe, Chatila (2006): Environment complexity

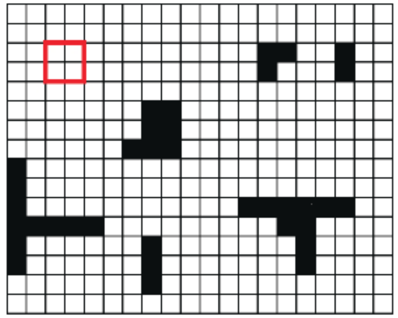
•H is defined as the entropy related to density of obstacles:

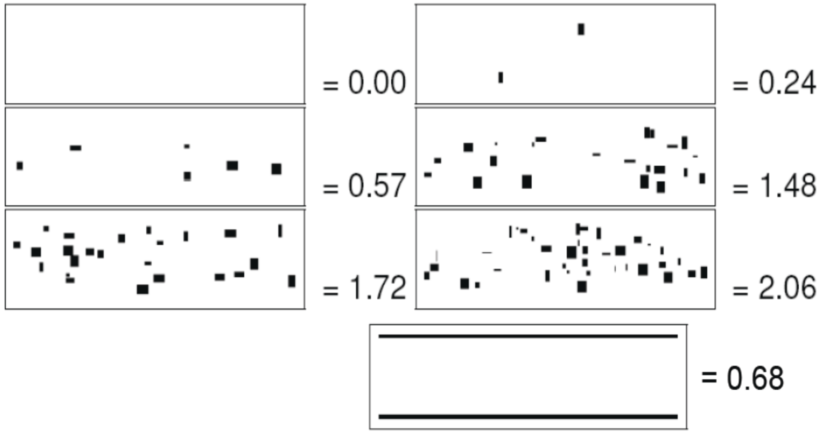
$$H = \sum_i -p(d_i) \log p(d_i)$$

in the occupancy grid,

with:

$$\sum_i p(d_i) = 1$$





Robotics 'OSI levels'?

It might help dividing the robot functionalities into level with an approach similar to the communication OSI level, starting, for instance, from the phisical level, to the control, perception, planning and 'cognitive' levels?

If robotics aims to be serious science, serious attention must be paid to the experimental method.

What is an 'experiment' in robotics?

Both replication and benchmarking are needed to foster a cumulative advancement of our knowledge of intelligent physical agents and even to correctly appreciate disruptive innovation in the science (?) and technology of robots.

Should we take inspiration from biology and medicine ?

Replication&Falsifiation

A clinical trial protocol is the detailed written plan of a clinical experiment.
It may be inspiring looking at the US NCI guidelines for drafting a clinical trial protocol: the enphasis on signaling 'adverse events' , the definition of 'criteria for response assessment', the necessity of defining clearly principal and secondary hypotheses to be validated.
The statistical section of the protocol is asked to define how the data will be analyzed in relation to each of the objectives.

In particular it expects that an acceptable trial specify, with reference to the study objectives:

* Method of randomization and stratification

Total sample size justified for adequate testing of primary and secondary hypotheses

* Error levels (alpha and beta)
* Differences to be detected for comparative studies
* Size of the confidence interval of the estimates.

Clinical Research example

From 'Bayesian Statistical Analysis in Medical Research'

David Draper

Department of Applied Mathematics and Statistics University of California, Santa Cruz

draper@ams.ucsc.edu www.ams.ucsc.edu/~draper

ROLE Steering Committee Meeting New York NY

25 April 2007

Clinical Research example

'The Big Picture

Statistics is the study of uncertainty: how to measure it, and what to do about it.

How to measure uncertainty: probability; two main probability paradigms: frequentist and Bayesian.

What to do about uncertainty: two main activities —

• Inference: Generalizing outward from a given data set (sample) to a larger universe (population), and attaching well-calibrated measures of uncertainty to the generalizations (e.g., "Nonwhites in the population of people at substantial risk of HIV–1 infection are 88% more likely to get infected if they don't receive this rgp120 vaccine than if they do receive it (relative risk of infection 1.88, 95% interval estimate 1.14–3.13)").

• Decision-Making: Taking or recommending an action on the basis of available data, in spite of remaining uncertainties (e.g., "Based on this trial, for whom nonwhites were a secondary subgroup, it's recommended that the vaccine be studied further with nonwhites as the primary study group"). '

Predictability

Epistemological issues

An information theoretic standpoint

Predictability in Mechanics: determinism, undeterminism, Chaos

Predictability

Schlick, Popper and the ‘demarcation problem’

Kuhn, Lakatos and Feyerabend

The ‘operational’ view

Biology and Robotics ☺ issues

Euron GEM Guidelines

Robotics papers come in many varieties.

For example, a paper may present a new theoretical advance; it may describe a new system concept; it may advance an argument based on discussion; it may present comparisons between a set of known techniques; it may do more than one of the foregoing...

.

1. Is it an experimental paper?

An experimental paper is one for which results, discussion and/or conclusions
depend crucially on experimental work. It uses experimental methods to answer
a significant engineering or scientific question about a robotic (or robotics-related) system. To test whether a paper is experimental, consider whether the paper would be acceptable without the experimental work: if the answer is no, the paper is experimental in the context of this discussion.

2. Are the system assumptions/hypotheses clear?

The assumptions or hypotheses necessary to the function of the system must be clearly stated. System limits must be identified.

3. Are the performance criteria spelled out explicitly?

An experimental paper should address an interesting engineering (or scientific)
question. Such questions will generally concern the relationship between system or environment parameters and system performance metrics. The performance criteria being studied must be clearly and explicitly motivated, and the parameters or factors on which they depend must be identified.

4. What is being measured and how?

The performance criteria being studied must be measurable; the paper must identify measurements corresponding to each criterion and motivate the choice of measurements employed. The data types of measurements should be clearly given or obvious — categorial (e.g. yes/no), ordinal (e.g. rankings), or numerical.

5. Do the methods and measurements match the criteria?

Measurement methods and choices must be clearly and explicitly described and,
where appropriate, explained and justified. The paper must demonstrate (unless it
is self-evident) that the chosen measurements actually measure the desired criteria and that the chosen measurement procedures generate correct data (for example, that implementations are plausibly correct).

6. Is there enough information to reproduce the work?

It is fundamental to scientific experimentation that someone else can in principle repeat the work. The paper must contain a complete description of all methods and parameter settings, or point clearly to an accessible copy of that information (which should be supplied to the paper's reviewers). Known standard methods need not be described, but any variations in their application must be noted. If benchmark procedures are used, they must be referenced, and any variations from the standard benchmark must be documented and justified.

7. Do the results obtained give a fair and realistic picture of the system being studied?

Care must be taken to ensure that experiments are properly executed: factors affecting measured performance that are not the subject of study must be identified and controlled for. In particular, uncontrolled variations in the system or the environment must be identified and dealt with by elimination, grouping techniques or appropriate statistical methods. The task tackled by the system must neither be too easy or too hard for the system being studied. Outlying measurement data may not be eliminated from analysis without justification and discussion.

8. Are the drawn conclusions precise and valid?

The experimental conclusions must be consistent with the experimental question(s) the paper poses, the criteria employed and the results obtained. System limits must be presented or discussed as well as conditions of successful operation. Conclusions should be stated precisely. Those drawn from statistical analysis must be consistent with the statistical information presented with the results.

Replication&Falsification

There are different modulation of this concept, but whether we think we are in a cumulative phase in the development of a scientific field or in presence of a 'disruptive' creative paradigm shift, as somebody is claiming in nowadays robotics, AI and Cognitive Sciences, a kind of widely accepted experimental methodology is needed in order to be able to ground the advancement of research on a shared quantitative language.

Replication&Falsification

It seems clear that in robotics the experimental methodology standards are currently in many cases weaker, and the syndrome 'it worked once, in my lab' could be more widespread than we may think.

Replication&Falsification

A limit to replication is given by the huge variability of robot machines.
Perhaps, following the biomedical analogy, we have to compare behaviors and performances of different 'animals'.

What about similar issues in Biology?

The definition of what should be considered a 'law of nature' in biology raises a number of issues. For reasons not very different from those raised from robotics research. The laws are usually not universal but apply to specific species: the Mendel laws apply to species with sexual reproduction, but not to all living species.
Almost every theoretical enunciate refer to a species or a set of species and has stochastic characteristics.

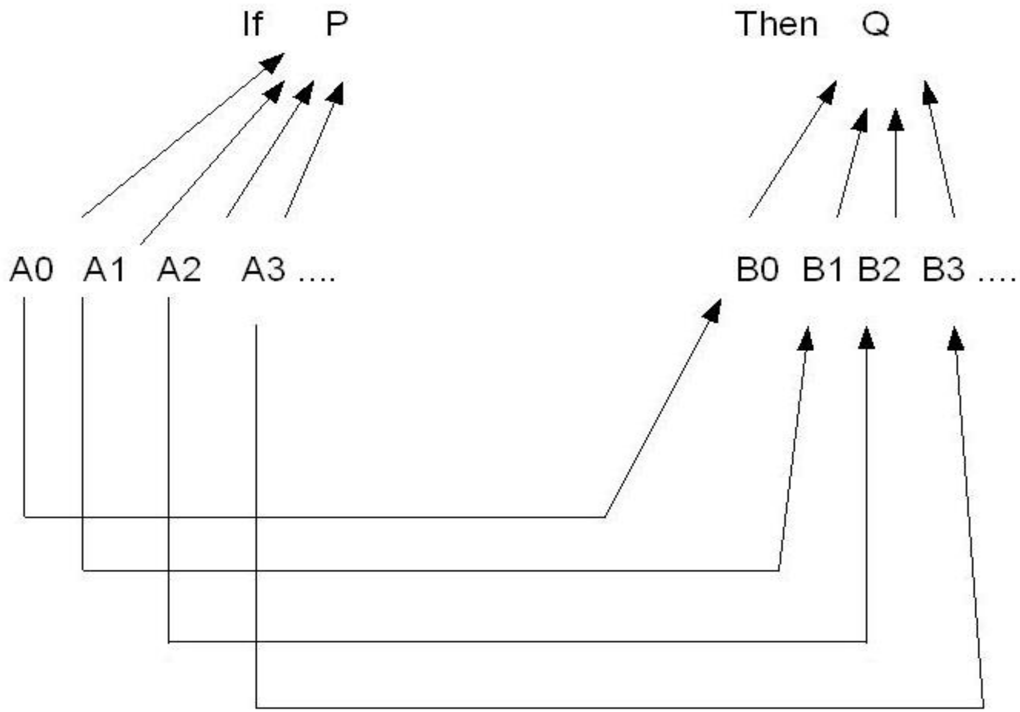
What about similar issues in Biology?

Systems are usually very complex, involve a huge numbers of variable and work in open ended stochastic environments. The same function, for example flight, can be performed in many different ways. The wing morphology and dynamics of a fly are quite different from those of a bird. On an other end, the wing of a penguin are used to stabilize swimming.

An interesting point is that the laws regarding a specific function in a species become true at a specific time, as a new function evolve, as depicted afterwards., and only if some initial conditions occur.

I	→	If P	then Q
t_0		t_1	t_2

Time dependence of biological 'laws'



'Causality at different levels'.

L1, L2, L3, ..., Ln	covering laws	
		explanans
C1, C2, ..., Cn	initial conditions	
E		explanandum

Hempel-Oppenheim Schema

Discussion

Why we need both replication AND benchmarking?

FACT: Benchmarking is more studied than Replication

- SLAM
- Mobile Robots' Motion Control
- Robot Obstacle Avoidance
- Grasping
- Visual Servoing
- Autonomy/Cognitive tasks: well, if scenarios are ok, Turing's test etc etc…from the very beginning, otherwise… very little

Discussion

The bare replication of experiments and the quantitative comparison of research results in robotics and cognitive sciences raise many challenging issues.
This is due to the variety of applications, tasks, mechanical structures, sensor sets, actuators, control system, software architectures, required levels of flexibility and autonomy, and so on.

A new kind of papers?

We may think of theoretical/concept papers, proof of concept papers, and experimental papers , as we have started to define here, as steps in a research idea 'life-cycle'. We believe that more paper of the 'experimental' kind would greatly help the research activities in robotics and the industrial exploitation of the results.

A new kind of papers?

- ‘description’ : a journal paper text+figures+ multimedia ….according to GEM Guidelines (or similar)

- Data sets (similar to IJRR ‘Data paper’

- Complete ‘code’ identifiers and or downloadable code (executables may be enough)

- ‘HW’ description or HW identifier (if it is identifiable)

…

EUCOG II

European ROBOTICS Research Network EURON

Resources

http://www.heronrobots.com/EuronGEMSig/

http://www.robot.uji.es/EURON/en/index.htm

http://www.nist.gov/mel/isd/permis2010.cfm

EUCOG II

European ROBOTICS Research Network EURON

Discussion

Should we take inspiration from biology and medicine?

Are replication and benchmarking really possible in robotics?

Are current limits in reporting a bottleneck to industrial exploitation?

Is there an impact on robotics research funding policy evaluation?

Should we adopt the proposed ‘extended paper’ reporting model?