July 9-13, 2012,
University of Sydney,
Sydney, NSW, Australia

Experiments in Robotics Research Practical Issues

Fabio Bonsignorio[1,2]

RoboticsLab
University Carlos III of Madrid

Heron Robots

Euron GEM Sig[1] IEEE RAS TC-PEBRAS[2]

- 'Look Ma, No Hands' syndrome?
- Replication of experiments
- Performance benchmarks

  to allow comparisons of results
- Needed to foster research advancement and enable practical application of research achievements

ROBOTICS
SCIENCE AND SYSTEMS

July 9-13, 2012,
University of Sydney,
Sydney, NSW, Australia

From theory to practice of performance comparison and result replications in Robotics Research

(Flexible!) schedule:

July 12th

9:00-10:00 A common ground between Visual Servoing and SLAM, from SLAM side (main speakers: J.Tardos, University of Zaragoza, W.Burgard, University of Freiburg)

10:00-10:30 *Coffee Break*

10:30-11:30 A common ground between Visual Servoing and SLAM, from manipulation, grasping and other sides (main speakers: Sidd Srinavasa, Carnegie Mellon University, Fabio Bonsignorio, University Carlos III of Madrid)

11:30-12:00 *Discussion*

Experiments in Robotics Research: Practical issues

July 9-13, 2012,
University of Sydney,
Sydney, NSW, Australia

****

The WEBCAST will be accessible from here:

http://webconf.ucc.usyd.edu.au/r11knrdtwe7/

The facebook page for virtual participation is :

https://www.facebook.com/groups/419682008082861/

****

Experiments in Robotics Research: Practical issues

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

What is an 'experiment' in robotics?

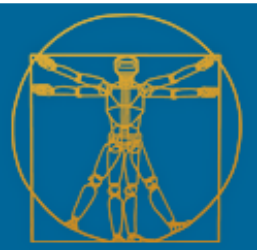
What about similar issues in Biology?

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.

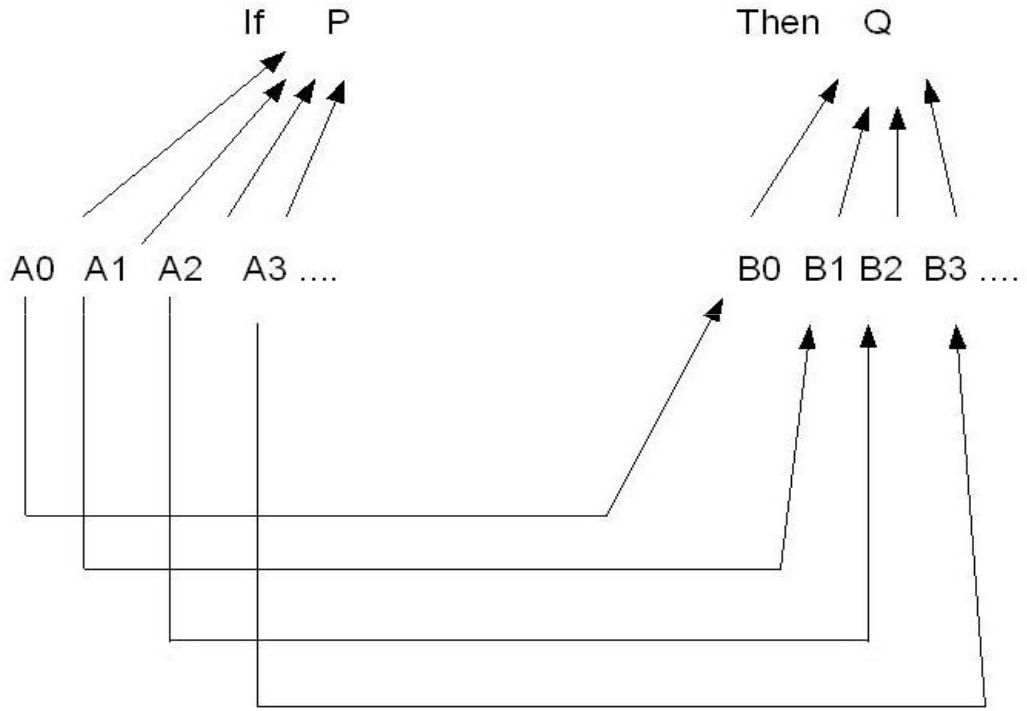
An interesting point is that the laws regarding a specific function in a species become true at a specific time, as a new function evolve, as depicted afterwards., and only if some initial conditions occur.



I	→	If P	then Q
t_0		t_1	t_2

Time dependence of biological 'laws'

Experiments in Robotics Research: Practical issues



'Causality at different levels'.

Experiments in Robotics Research: Practical issues

L1, L2, L3, ..., Ln covering laws

explanans

C1, C2, ..., Cn initial conditions

E

explanandum

Hempel-Oppenheim Schema

Experiments in Robotics Research: Practical issues

Discussion

Why we need both replication AND benchmarking?

FACT: Benchmarking is more studied than Replication (with caution ☺)

- SLAM
- Mobile Robots' Motion Control
- Robot Obstacle Avoidance
- Grasping
- Visual Servoing
- Autonomy/Cognitive tasks

Discussion

A new kind of ‘papers’?

Replication of robotics experiments

- Research Reporting in Biology and Medicine
- Evidence Based Medicine
- Early activities in robotics research replication

Discussion

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)

…

EXAMPLE:

throughput.sourceforge.net

A BETA repository for replicable experiments with human-robot systems

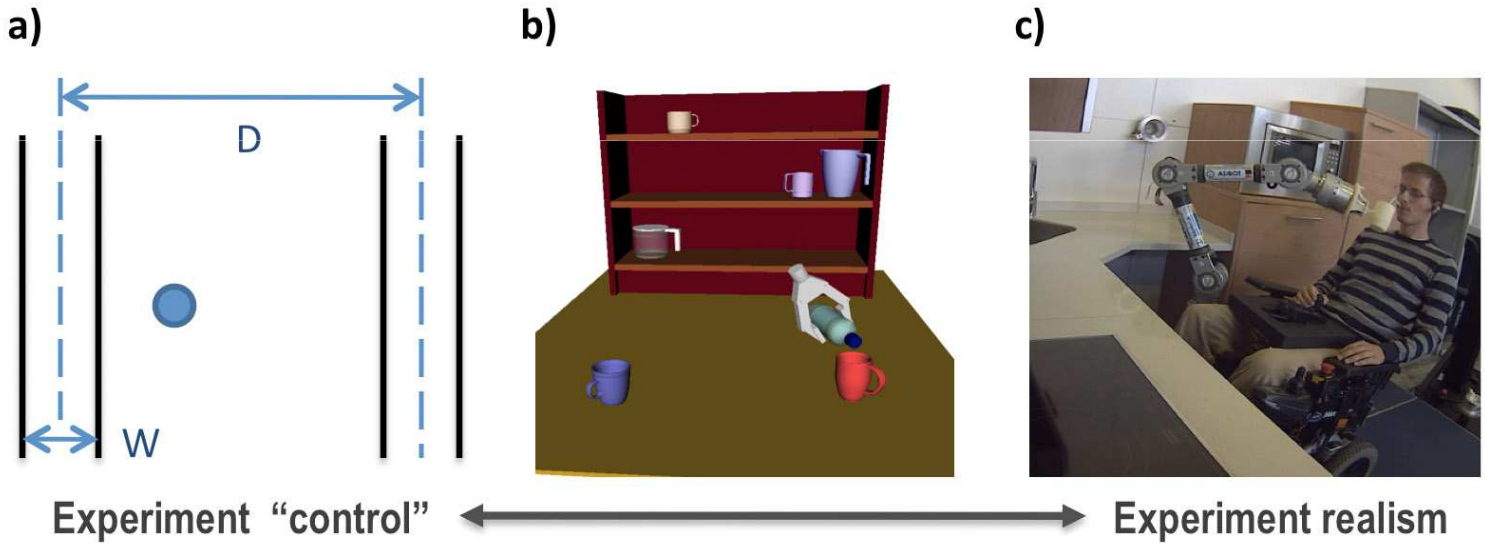
The Euron Good Experimental Methodology Guidelines

1. Is it an experimental paper?
2. Are the system assumptions/hypotheses clear?
3. Are the evaluation criteria spelled out explicitly?
4. What is being measured and how?
5. Do the methods and measurements match the criteria?
6. Is there enough information to reproduce the work?
7. Do the results obtained give a fair and realistic picture of the system being studied?
8. Are the drawn conclusions precise and valid?

The Euron Good Experimental Methodology Guidelines

1. Is it an experimental paper?
2. Are the system assumptions/hypotheses clear?
3. Are the evaluation criteria spelled out explicitly?
4. What is being measured and how?
5. Do the methods and measurements match the criteria?
6. Is there enough information to reproduce the work?
7. Do the results obtained give a fair and realistic picture of the system being studied?
8. Are the drawn conclusions precise and valid?

Replicable experiments for human-robot systems?



(a couple of) Critical Points

Robot requirements:

- *simulated* robot on commercially available simulator
- available (commercially or open source) *real* robot platform

Human experiment requirements :

•Sufficient data: Participant statistics, participant instructions, experiment protocol

•Sufficient “control”: Sufficient repetitions, clear success/failure criteria, confidentiality ++

•Informed consent, ethical issues (including forms ☺)
(boundary with clinical protocols)

What should be 'Opensource'?

Should ALL experiment software be open?
- Not necessarily: for hypothesis confirmation/refutation
- Sourceforge: Open source "experimentware" ok?

Minimum software requirement:
- All executables and configuration files
- All shared libraries or instructions for install

Licensing:
- Code: LGPL-3.0
- All other content: Creative Commons Attribution 3.0

Example (from next IROS)

What metrics can be applied to evaluating shared control?

- Mean Time (MT)
- Errors
- MT and errors (one metric)?
- Information metrics?

Case study:

- Online adaptation of collision limitation behavior
- 5 participants @ 1 hour each

What should be 'Opensource'?

Should ALL experiment software be open?
- Not necessarily: for hypothesis confirmation/refutation
- Sourceforge: Open source "experimentware" ok?

Minimum software requirement:
- All executables and configuration files
- All shared libraries or instructions for install

Licensing:
- Code: LGPL-3.0
- All other content: Creative Commons Attribution 3.0

Description

Abstract

Assistive robots are increasingly being envisioned as an aid to the elderly and disabled. However controlling a robotic system with a potentially large amount of Degrees of Freedom (DOF) in a safe and reliable way is not an easy task, even without limitations in the mobility of the upper extremities. Shared control has been proposed as a way of aiding disabled users in controlling mobility aids such as assistive wheelchairs, by using the sensors of the robotic platform to predict the user's intent and assist in navigation. Assistive manipulators, that aim to perform physical Daily Life Activities (DLA), is a more complex problem however. The problem arise from the exponential increase in the size of the state-space with DOF and the increased level of accuracy required for manipulation. Another complication is the potential need for adapting the system to each user's abilities and disabilities. This calls for good experimental practices to ensure repeatability, reproducibility, and steady progress. The work presented here attempts to model the complete system for assistive manipulators, and in the context of this model define metrics and good practices for benchmarking shared control for such robots. An adaptive shared control approach for limiting collisions during teleoperation is used as a case study.

Link

Coming soon...

Citation

@article{Maintainer,
author = {Stoelen, M.F. and Tejada, V.F. and Jardon, A.H. and Bonsignorio, F. and Balaguer, C.},
journal = {to be presented at the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems},
title = {Benchmarking Shared Control for Assistive Manipulators: From Controllability to the Speed-Accuracy Trade-Off},
year = {2012}
}

Data sets

All Cartesian trajectories. All trial-level results. Development of neural network weight matrices over time. Matlab examples for loading trajectories, plotting the trial-level results and visualizing the development of the neural network weights. See [AdaptiveCollisionHand-DataSets].

Code

Complete set of executables, Linux 32 and 64 bit. Tested on Ubuntu 10.04 for 32 bit and 11.04 for 64 bit. See

Experiments in Robotics Research: Practical issues

Archive content

Code

- 7 executables, Linux 32 and 64 bit
- Shared libraries
- Models and configuration files

data_sets

- 184 MB of raw data (csv text files)
- Trajectories, MT, errors, NN weights
- method
- Example video
- Questionnaires
- Consent form w/ participant instructions

Replication

Hardware requirements:

- Modern multicore desktop computer
- 3DConexion SpaceNavigator: 99 $ (US)

Full instructions for installation and setup

- All experiment executables provided
- Freely available shared libraries used: (OpenRAVE, YARP, ++)

Reference results

Trial-level data

- MT (medium time), errors and number of attempts

Cartesian trajectories

- User velocity input
- Noise input
- Shared control output

Neural network data

- Weight matrix stored every 3 seconds

Human experiment specifics

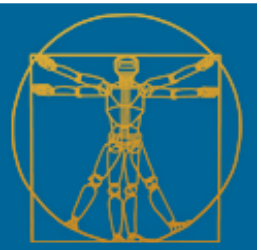
- "Human" number of trials/repetitions
- Detailed method for experiment required
- Informed consent for publishing data(?)

Replication

Total effort to replicate: 1-2 hours (on clean Ubuntu machine) but… same person ☺…we need naive testers

Time needed for the setup: a few days (under pressure, but first time)→ a very small fraction of the overall effort to set up the experiments and get the results….

→ Conclusions?☺



Thank you!