The Role and Relevance of Experimentation in Informatics

Viola Schiaffonati
Artificial Intelligence and Robotics Laboratory
Politecnico di Milano

Experimentation: role and relevance

- Starting point: philosophy of science perspective (philosophy of experimentation)
- Ending point? Philosophy and engineering
- In the middle: good experimental methodologies in computer science and engineering
  - Grounded philosophy!

Relevance

- Sure, experiments are relevant
  - Experimental scientific method taking center stage in computer science and engineering (Freeman 2008, Morrison and Snodgrass 2011)
- Why are they relevant?
  - Help in building a reliable base of knowledge, in leading to useful and unexpected insights, in accelerating progress (Tichy 1998)

Role

- But what about their role?
  - What is an experiment in general and in informatics in particular
  - Experiments in informatics between science and engineering (research)

Taking inspiration from science

- Experimental methodologies in informatics have not yet reached the level of maturity of other scientific disciplines
  - Idea: look at how experiments are performed in traditional scientific disciplines
  - Principles: comparison, reproducibility and repeatability, justification and explanation

Consequences

- Terminological and conceptual clarification
  - Definition of experiment (experimental methods not to be confused with empirical methods)
  - Replication not enough!
- Application of traditional scientific method to computer science and engineering
  - Comparison, reproducibility/repeatability, justification/explanation declined
- Consideration of peculiar aspects of experiments in engineering

What is an experiment?

- Experiment is controlled experience (from Galileo's '*sensate esperienze*')
- Set of observations and actions, performed in a controlled context, to test a given hypothesis
  - The phenomenon under investigation must be treated as an isolated object
  - It is assumed that other factors not under investigation do not influence the investigated object

Observing vs. experimenting

- Observing a drop of water through a microscope is not an experiment
- Observing the same drop through a microscope, after having colored it with a chemical reagent in order to evidence some microorganisms, is an experimental procedure
  - Ability to control some of the features of a phenomenon under investigation
  - Purpose of testing the behavior of the drop under some controlled circumstances

Experimental principles declined

- Comparison, repeatability/reproducibility, justification/explanation in a computer engineering field (Amigoni *et al.* 2009, Amigoni and Schiaffonati 2010)
- Autonomous mobile robotics
  - Robots with the ability to maintain a sense of position and to navigate without human intervention

Comparison

- Comparison presupposes to know what has been already done in the past to evaluate new results with the old ones
- Comparison in autonomous mobile robotics
  - Increasing use of publicly available data sets (Victoria Park, RADISH, and Rawseeds) to set a common ground for comparing different systems
  - Development of comparable implementations, starting from the description provided in papers and reports and also from the use of the same code

Reproducibility and repeatability

- Reproducibility is the possibility to independently verify the results of a given experiment
- Repeatability concerns the fact that a single result is not sufficient to ensure the success of an experiment
- Reproducibility and repeatability in autonomous mobile robotics
  - Implementation of similar experiments to understand the parameters influencing the system
  - Public distribution of code and/or problem instances
  - Adoption of standard data sets as benchmarks
  - Report of anomalies in performance

Justification and explanation

- Justification deals with drawing justified conclusions on the basis of the information collected during an experiment
- Explanation requires a deep analysis of data to derive correct implications
- Justification and explanation in autonomous mobile robotics
  - Use of several data sets to derive well justified conclusions
  - Correct behavior of systems verified according to ground truth or visual inspection
  - Difficulty in generalizing when ground truth is not available

Experiments from science to engineering

- Not just different objects
  - Natural objects (science)
  - Technical artifacts (engineering)
- But different purposes
  - To understand a natural phenomenon (science)
  - To test an artifact (engineering)

Experiments and technical artifacts

- The notion of technical artifact is central to reflect on experiments in computer science and engineering
- Why?
  - Engineering is an activity producing technology
  - Technology is a practice focused on the creation of artifacts and artifact-based services (Franssen *et al.*, 2010)

Technical artifacts

- Material objects deliberately produced by humans in order to fulfill some practical functions
  - Technical function: what is the technical artifact for?
  - Physical composition: what does it consist of?
  - Instruction for use: how must it be used?
- Mutual dependency
  - Technical artifact as a physical object with a technical function and use plan designed and made by human beings

Informatics and technical artifacts

- Why informatics products are technical artifacts?
- They are physical objects deliberately produced by humans with a technical function and use plan designed and made by human beings (vermaas *et al*. 2011)

Experiments and technical function

- Experiments in engineering evaluate technical artifacts according to whether and what amount the function for which they have been built is fulfilled
- Normative claims are introduced depending on a given reference function or set of functions
  - The artifact as ‘good’ or ‘bad’
- Is this enough?

Between science and engineering

- Informatics between engineering and science (even with respect to experiments)
  - Experiments performed to test how well an artifact works with respect to a reference model and a metric
  - Experiments performed to understand how complex artifacts (whose behavior is hardly predictable) work and interact with the environment (at different degrees)

Again on the role of experimentation

- More rigor, better progress?
- Internal and external role of experimentation
  - Internal: reflection on the disciplinary status of computer science and engineering from a methodological point of view (not just the object, but also the method)
  - External: toward the philosophy of engineering (with the contribute of philosophy of science and technology)

References

- Amigoni, F., Reggiani, M., Schiaffonati, V. (2009) "An Insightful Comparison between Experiments in Mobile Robotics and in Science". *Autonomous Robots* 27(4), 313-325.
- Amigoni, F., Schiaffonati, V. (2010). "Good Experimental Methodologies and Simulation in Autonomous Mobile Robotics" in Magnani, L., Carnielli, W., Pizzi, C. (eds.), *Model-Based Reasoning in Science and Technology*, Springer, 315-332.
- Freeman, P. (2008) "Back to Experimentation" in *Communications of the ACM* 51 (1), 21-22.
- Franssen, M., Lokhorst, G., van de Poel, I. (2010) "Philosophy of Technology " in The *Stanford Encyclopedia of Philosophy* (Spring 2010 Edition), Edward N. Zalta (ed.), http://plato.stanford.edu/archives/spr2010/entries/technology/
- Morrison, C., Snodgrass, R. (2011) "Computer Science Can Use More Science" in *Communications of the ACM* 54 (6), 36-38.
- Tichy, W. (1998) "Should Computer Scientists Experiment More?" *IEEE Computer* 31 (5), 32-40.
- Vermaas, P., Kroes, P., van de Poel, I., Franssen, M., Houkes, W. (2011) *A Philosophy of Technology. From Technical Artefacts to Sociotechnical Systems*. Morgan and Claypool.