Experiment Design

Lecture 1: Introduction[1]

Claus Aranha
caranha@cs.tsukuba.ac.jp

Department of Computer Science

April 22, 2014

[1] version 1.2

Some notes

- Next week: No class on the 29th;
- The other week: Class on Friday (9th);

Introduce Your Research!

Important Principles for experimental research on algorithms

- Relevant Experiments;
- Relation to the literature;
- Significant test problems;
- Experimental Design;
- Efficient Implementation;
- Reproducibility;
- Comparability;
- Complete Description;
- Claims supported by results;
- Data properly presented.

Pre Experimental Design (1)

Before we begin

- Do you have all knowledge needed?
  - Very common in information sciences
  - Specific factors/problems in the specific application
  - Teamwork

Pre Experimental Design (1)

Before we begin

- Do you have all knowledge needed?
  - Very common in information sciences
  - Specific factors/problems in the specific application
  - Teamwork
- Relevance of the experiment
  - Think about the target audience;
  - Check the literature throughly/avoid reinventing the wheel;
  - Is the proposed question relevant?
  - Break down the problem if necessary;

Pre Experimental Design (2)

Decide what is the question you want to answer:

Objetives of an experiment

- Determine the most influential variables in a system;
- Determine desired parameter values to reach an output;
- Describe the behavior of a system;
- Etc.

Pre Experimental Design (2)

Decide what is the question you want to answer:

Objetives of an experiment

- Determine the most influential variables in a system;
- Determine desired parameter values to reach an output;
- Describe the behavior of a system;
- Etc.

(some) Types of Experiments

- Time-series analysis;
- Observation of the system;
- Planned interference;

Experimental Design (1)

Parameter/Factor Selection

- Educated Guesses
- One Parameter at a time
- Factorial Planning

- Arbitrary combination of parameter/values for study;
- Observe the behavior; Change some parameters; Repeat;
- Risks premature convergence ("good enough");

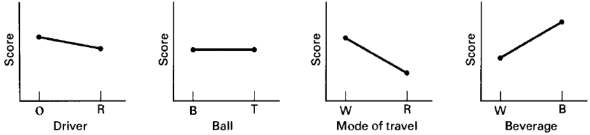
Requires solid knowledge about the underlying problem.

Experimental Design (1)

Parameter/Factor Selection

- Educated Guesses
- One Parameter at a time
- Factorial Planning

- Selects and changes one parameter at a time;
- Often used, safe;
- Unable to determine parameter integration;



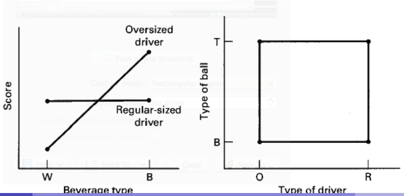
(Baseline: (O)versized driver, (B)alata ball, (W)alking, drinking (W)ater)

Experimental Design (1)

Parameter/Factor Selection

- Educated Guesses
- One Parameter at a time
- Factorial Planning

- Testing the influence of different factors on the system;
- Values of each factor are varied in an organized fashion;
- Determines individual effects and iteractions;
- Experiment size increases with number of factors;



Experimental Design (2)

Data Selection

- Benchmarks
- Natural Data
- Random Data

Experimental Design (3)

Design Principles

- Repetition
- Randomization
- Blocking

Repetition:

- Repetition of the experiment allows for estimation of the error
- Used for calculating statistical significance
- Allows for a more precise measure
- Decide beforehand the number of repetitions!

Experimental Design (3)

Design Principles

- Repetition
- Randomization
- Blocking

Randomization:

- Random order of the experiments, Random allocation of resources;
- Reduce the bias from data, reduces influence from unrelated factors;
- Where randomization is not possible: partial randomization, blocking, etc;

Experimental Design (3)

Design Principles

- Repetition
- Randomization
- Blocking

Blocking:

- Removes the influence of unrelated factors
- Breakdown experiments into blocks based on these factors
- Reduces the number of available observations

Experimental Design (4)

Comparisons

- Selecting Comparison methods
  - Recent Methods;
  - “Traditional” methods;
  - Methods Outside your discipline
- Tweaking of Code and Parameter

Experimental Design (4)

Comparisons

- Selecting Comparison methods
  - Recent Methods;
  - “Traditional” methods;
  - Methods Outside your discipline
- Tweaking of Code and Parameter

Statistical Analysis (1)

- A good experimental design should give you enough information to select the correct statistical model.

  (Or close enough)

- Export all possible data from your experiment, use known statistical tools to deal with the data

  (Programmers have a tendency to reinvent the wheel)

- R, Matlab, Octave, etc...

Statistical Analysis (2)

General Procedure for Hypothesis Testing

1. Define the Null Model and the desired significance level;

Statistical Analysis (2)

General Procedure for Hypothesis Testing

1. Define the Null Model and the desired significance level;
2. Determine $P(\text{data}|\text{null model})$

Statistical Analysis (2)

General Procedure for Hypothesis Testing

1. Define the Null Model and the desired significance level;
2. Determine $P(\text{data}|\text{null model})$
3. Decide Whether or not to reject the Null Hypothesis

Statistical Analysis (2)

General Procedure for Hypothesis Testing

1. Define the Null Model and the desired significance level;
2. Determine $P(\text{data}|\text{null model})$
3. Decide Whether or not to reject the Null Hypothesis

Many people stop here ...

Statistical Analysis (2)

General Procedure for Hypothesis Testing

1. Define the Null Model and the desired significance level;
2. Determine $P(\text{data}|\text{null model})$
3. Decide Whether or not to reject the Null Hypothesis
4. Validate the premises of the Statistical Model

Statistical Analysis (2)

General Procedure for Hypothesis Testing

1. Define the Null Model and the desired significance level;
2. Determine $P(\text{data}|\text{null model})$
3. Decide Whether or not to reject the Null Hypothesis

4. Validate the premises of the Statistical Model
5. Estimate the magnititude of the differences

Analyzing the results

Stick to the Facts!

Beware of overgeneralizing the results and making claims not supported by the data.

Stick to ALL the facts

Avoid cherrypicking the results to make the system look better than it is. Complete descriptions are actually more useful.

Dealing with anomalies

Any anomalies must be reported, but beware of “anomaly hunting”.

Writing that paper (1)

Be Through

Report everything that can help the reader understand the extents and limitations of your experiments:

- Stop criterions;
- Computational costs (if time, how it was measured);
- Parameters, parameter selection methods;
- Experimental setup (and its reasoning);
- Etc;

Writing that paper (2)

Strive for replicability!

- Be very explicit describing your data
  - Cite literature benchmarks;
  - Algorithms/parameters for artificial data sets;
  - If hard to obtain, consider distributing your data online;
- If at all possible, distribute the source code for your experiment!

  make sure you add all information needed for compilation...
- Distribute online experimental data that did not fit the paper;
- Consider open access models of publication;

  remember you are usually allowed to distribute pre-prints