Experiment Design for Computer Sciences (01CH740)

Topic 01 - What is an experiment?

Claus Aranha
caranha@cs.tsukuba.ac.jp

University of Tsukuba, Department of Computer Sciences

April 30, 2020

Version 2020.1

Outline

- What is Science?
- What is an Experiment?
- The characteristics of a good Experiment;
- Good experimental practices;
- Report 1 outline

What is Science?

As master degree students, what is science for you?

What is Science?

As master degree students, what is science for you?

Answers from past years

- A method to learn about the world;
- A method to reach the truth;
- Science is useful when it contributes to society;
- How we develop new technologies;

Marie Curie

Marie Curie
1867 – 1934

One way to learn about science, is to learn a little bit more about proeminent scientists.

Marie Curie

- Physicist and Chemist
- Pioneer of radioactivity
- First woman to win the Nobel Prize
  - First person to win the Nobel twice

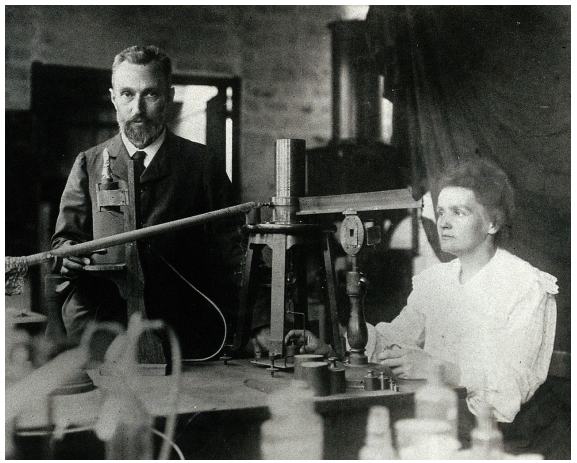
Marie Curie

Humble Beginnings

- Born in Poland
- Could not enroll in a regular university because she was a woman, so she got educated at the clandestine "Flying University"
- Sustained herself working as a tutor and as a teacher for families;

Marie Curie

Moving to Paris



- Earned her Physics degree at the University of Paris
- Worked at a small shed and had difficulty acquiring funding;
- Found out that the emission of radiation from uranium depended only on the size of the sample;
- Did not patent her techniques, so science could proceed unimpeded;

Marie Curie

Applications

- Observed that tumour cells died more quickly to radiation than healthy cells;
- Developed mobile X-Ray units to be used for surgery during World War I ("little curies")
- Developed "Radium Needles" for sterilizing tissue;
- Died of radiation related diseases; Some of her research notebooks are still radioactive!

For you to think at home

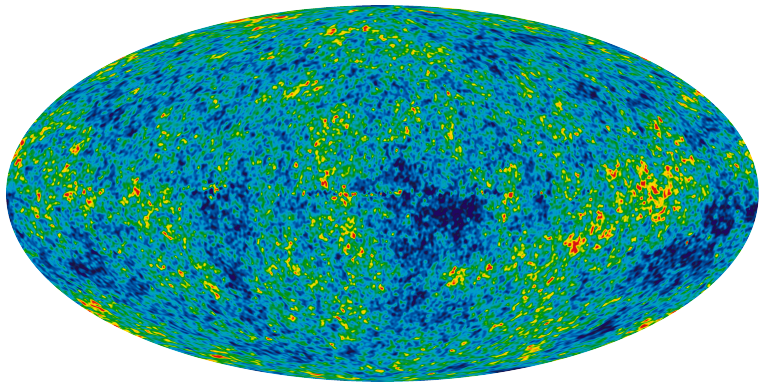
Who is a scientist that inspire you? Do you know their story?

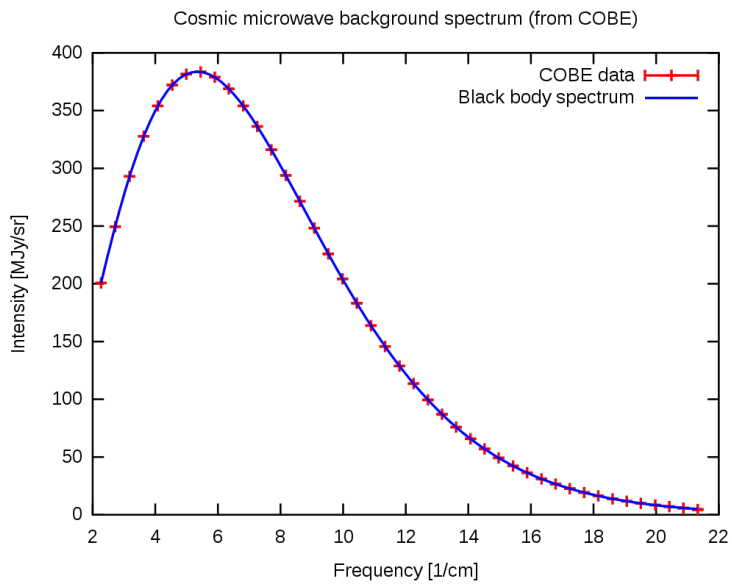
Let's talk a little about scientific discoveries;

Examples of Scientific Discoveries

The Big Bang Theory

- The big bang theory describes how the universe behaved in the first moments of its existence;
- One of the predictions made by the big bang theory is the distribution of the Cosmic Background Radiation, energy remaining from the early universe;
- The NASA COBE mission measured the CBR in space, and found its distribution to match near perfectly the predicted values;





Images from the NASA Cobe project

Examples of Scientific Discoveries

Vitamin C prevents scurvy

James Lind (1747):

- Observation: scurvy in sailors;
- Conjecture: Caused by the body rottenning;
- Idea: attempt to avoid/reverse effects with acidic substances;

Separation of a group of 12 affected sailors in six groups with identical diets, except for the addition of a supplement:

Group 1
Cider.

Group 2
Vitriol.

Group 3
Vinegar.

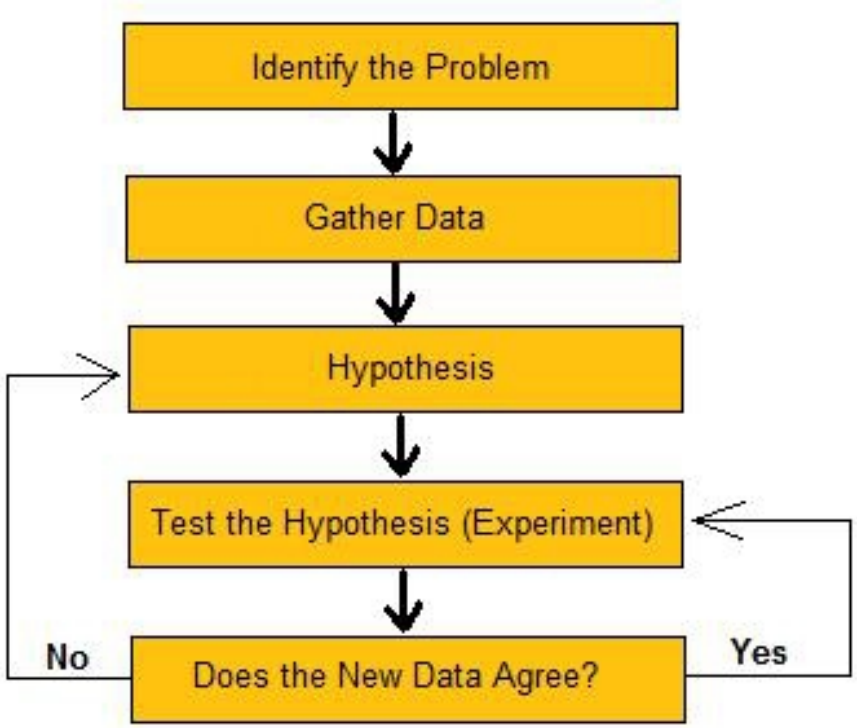
Group 4
Sea water.

Group 5
Oranges and lemons.

Group 6
Tea.

The Scientific Method

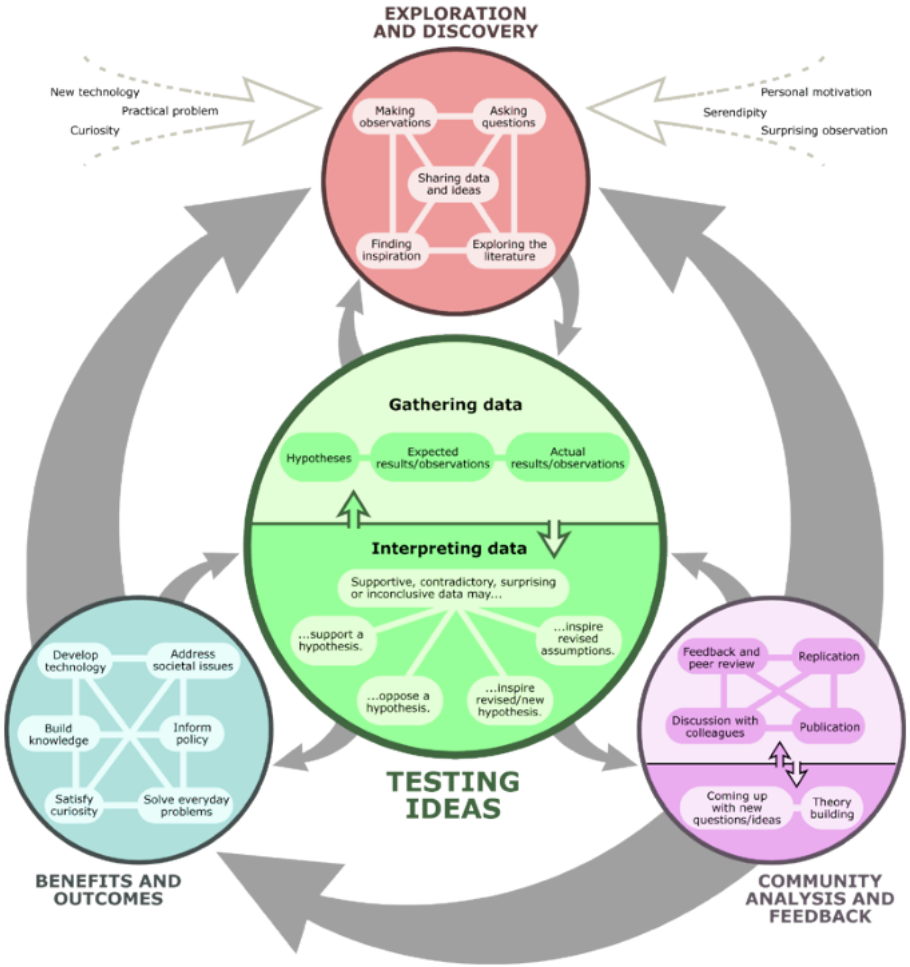
- The examples we saw demonstrate the familiar idea of the scientific method;
- Hypothesis, **Experiment**, Analysis;
- But is this really all that there is to the scientific method?



The Scientific Method

Science as an interactive process

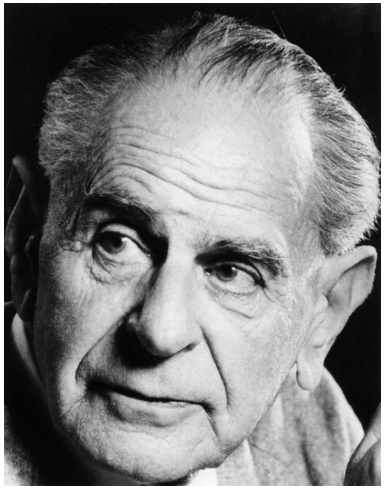
The scientific process can be more complex than a simple recipe.



The Role of Experimentation

- Both in the simple definition of the scientific method, and on the more complete one, the experiment takes a central role;
- An experiment is how we test hypothesis, how we learn more about the world, how we examine our ideas;
- But what is an experiment? It is more than just collecting data!

What is an experiment?



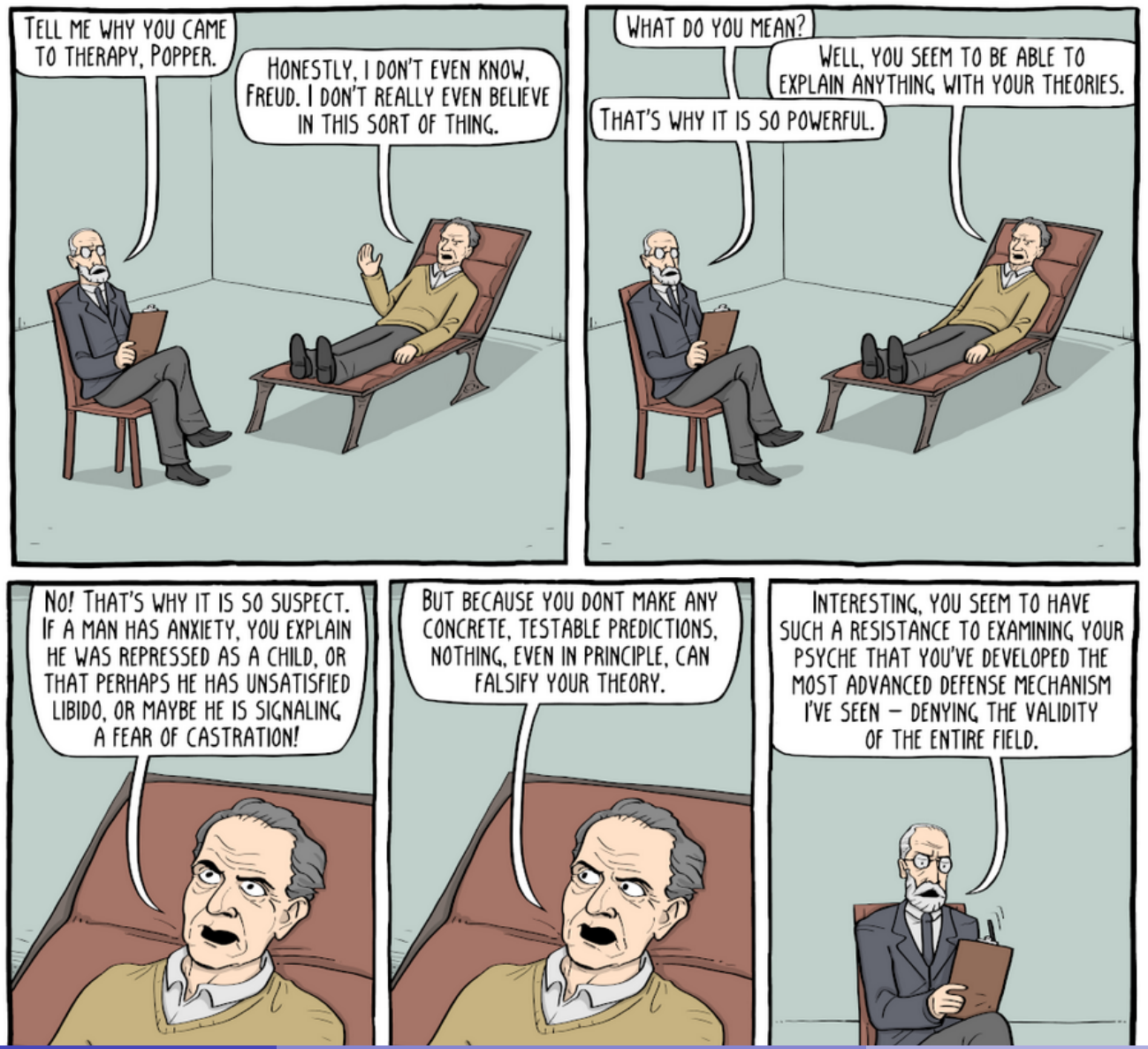
Karl Popper
(1902–1994)

- Philosophy of Science: How do we obtain knowledge about the world?
- Scientific theories can only be tested by observing their implications;
- Reject theories that cannot be confirmed by experiment;

What are the characteristics of a good experiment?

- Falsifiable Hypothesis;
- Useful predictions;
- Data Collection;
- Reproducibility;

Falsifiability



Falsifiability

A scientific hypothesis is **falsifiable** if there is some observation that would render it false.

Specific Predictions

Falsifiable hypothesis make specific predictions about how the world behaves, not only if the hypothesis is true, but also if the hypothesis were false.

Useful/Strong Predictions

It is not very hard to make many trivial predictions about the world. Scientific hypothesis should not only be falsifiable, but also strong and/or useful.

Types of Experiments

There are many different types of experiments, depending on what kind of data you want to obtain. Based on the data collection method, for example, we can classify an experiment in three types:

- Observational Experiments;
- Retrospective Experiments;
- Controlled Experiments;

Types of Experiments

Observational Experiments

In an Observational Experiment, you obtain data by observing a phenomena without interacting with it directly.

Example: you count the number of people who use the train with and without masks every day.

- Requires care to observe representative situations;
- Allows the researcher to choose general conditions for observation;
- The situation of interest may be too rare to observe naturally;

Types of Experiments

Retrospective Experiments

In a Retrospective Experiment, the researcher obtains data from historical records (newspaper, reports, other scientific papers).

Example: you search from the relationship between announcements of celebrity marriages, and total number of registered marriages;

- Generally cheaper, and may be the only way to gather data over a very long period of time;
- Susceptible to missing records or bias in recording;

Types of Experiments

Controlled Experiments

In a Controlled Experiment, the researcher is able to define several variables in the experiment, and perform it in the conditions desired.

Example: You develop a new algorithm, and test it on some selected data sets, on a collection of different computational architectures;

- Gives a lot of control for the researcher;
- If not designed carefully, allows for the introduction of biases into the experiment;
- Can be the most expensive kind of experiment (although not always in CS);

What is Experiment Design?

To perform any experiment, we have to make several technical and scientific decisions:

- Which methods we compare in the experiment?
- Which data sets are used?
- How many times do we interview each participant?
- In what order do we perform the experiments?
- Which data is reported, and how is the data summarized?
- What criteria determines that the hypothesis was accepted or rejected?
- What hyper-parameters do we use?
- How many times is the experiment repeated? How are these repetitions sumarized?

Experiment Design is how we answer each of these questions.

Experiment Design

Example: Controlling for Variation

Let's say you are comparing two computer programs by measuring their running time (wallclock time).

You know that the running time of a program is affected by other programs that are running in the background of the operational system. For example, if a software update happens in the background, it could make a run much slower.

To control for this variation, you make sure to run your experiment in a system with a minimum number of running processes, and you also repeat the experiment many times and take the average running time;

Experiment Design

Example: Controlling for Independence

Imagine that you are comparing two website designs with the following experiment: You measure the time for a user to find some information on website A, then you measure the time for website B.

If you make this comparison always in the same order for all users, you discover that the users are a bit faster for website B, because they get used to the testing environment and are more relaxed.

To remove this influence, you make sure that the test order is always random, or you make sure that each user tests only one website.

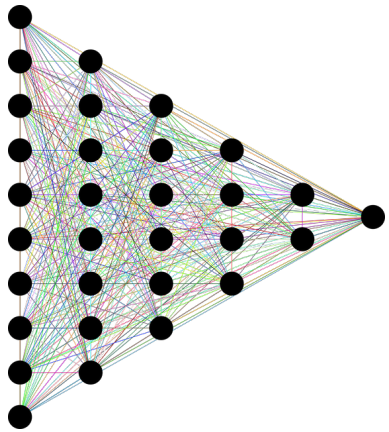
Experiment Design

Example: Controlling for Fairness

You propose a neural network architecture for a new vision problem, and you compare it against traditional architectures.

Because of the special characteristics of the problem, you fine-tune the hyper-parameters of your architecture to achieve the best performance.

To make sure that the comparison is fair, you use the same fine-tune techniques to the traditional architecture that you are comparing against, not using its old hyper-parameters from the literature.



Pre-registered Experiments

Pre-registration is the act of fully defining your research protocol **before you begin to collect or analyse data**.

By pre-registering your research, you avoid modifying your methods to fit your hypothesis (or modifying your hypothesis to fit your data)

Public pre-registration can prevent the loss of negative results. Private pre-registration can help you keep yourself in check.

Learn more: Center for Open Science
`https://cos.io/prereg/`

Reproducible Experiments

Reproducibility is an important property of good research:

- Others can confirm your results;
- Others can build on your results;
- Others can improve your results;
- Society can use your results;

Reproducible Experiments

How can we make experiments more reproducible?

Clear Experiment Design

Detailed steps taken to perform the experiment; Values of relevant parameters; How the results are processed and evaluated;

Open Data and Open Source

Data acquisition protocol is clearly defined; Raw data and pre-processing scripts are available; Data is well documented;

For CS, open source of proposed algorithms is essential;

Open Documentation

Code used for statistical analysis and data visualization;

Summary of the Lecture

- Experimentation is a key part of Science;
  - Experiments acquires data that can be used to validate or falsify scientific ideas, and to answer scientific questions;
- An experiment has to be performed carefully to guarantee its usefulness;
  - **Experimental design** defines the type of experiment, and how data is gathered;
  - Several factors can affect the **fairness and meaningfulness** of experiments;
  - **Reproducibility** is essential to guarantee the usefulness of an experiment;

Report 1

Design and execute a scientific experiment, and report your results

For this report, you must choose a simple experiment to design, perform, and analyse the results. Your report should consist of:

- **Introduction**: Describe your scientific question, its relevance, and why do you need an experiment for it;
- **Experiment Design**: Describe how you will collect data to answer your scientific question; Make sure to mention any parameters or factors that must be controlled;
- **Data Collection:** Report on your data collection, if anything happened outside of expected from the experimental design;
- **Analysis:** Describe your results in detail, and what answer they provide to your scientific question;

Remember to follow practices of **reproducible science**

Report 1

How to choose an experiment for your report

- If possible, choose something from your own research;
- Experiments from your day to day life are also good;
  - Comparing cooking techniques is always fun;
  - When collecting data, be careful of measuring errors;
- When in doubt, comparing algorithms is an easy choice;
  - Make sure to choose an appropriate metric to report!
- Make sure you choose an experiment that you can perform!
- Next lecture, we will talk about a bit about how to analyse and report experimental data;

Recommended Reading

- Understanding Science
  `https://undsci.berkeley.edu/article/intro_01`
- Existential Comics `http://existentialcomics.com`;
- Crash Course Psychology (Youtube);

About these Slides

These slides were made by Claus Aranha, 2020. You are welcome to copy, re-use and modify this material.

These slides are a modification of "Design and Analysis of Experiments (2018)" by Felipe Campelo, used with permission.

Individual images in some slides might have been made by other authors. Please see the references in each slide for those cases.

Image Credits I

[Page 4] Marie Curie sketch from https://www.irasutoya.com
[Page 5] Curie sisters image from Wikipedia (public domain)
[Page 6] Marie and Pierre Curie image from Wikipedia (Public Domain)
[Page 9] CMB images from NASA (public domain)
[Page 10] James Lind image from Wikipedia, Public Domain
[Page 12] Image from "understanding science, University of berkerley"
[Page 14] Karl Popper picture from Wikipedia
[Page 16] "Froid and Popper" from http://existentialcomics.com/
[Page 17] White swan image from pixabay
[Page 23] Clock image from https://pixabay.com
[Page 24] Website images from https://www.irasutoya.com
[Page 25] Neural Network image by Gordon Johnson from https://pixabay.com