CS 294-115 Algorithmic Human-Robot Interaction **Fall 2016**

Lecture 8: Experimental Design I

Scribes: Thanard Kurutach and Jonathan Shum

8.1 Overview

In this series of lectures, we will talk about the following topics:

- What is an experiment?
- What is good experiment?
- Analysis of experiments

8.2 What is an experiment?

8.2.1 Origins

Experimental design is originated from the following fields:

- Medical (drug/treatment effectiveness)
- Agriculture (pesticide effectiveness) - Fisher (1920's)
- Psychology (understanding human behavior)

In this class, we are interested in experimental design for:

- HRI (studying the effectiveness of design or interaction strategies)
- Robotics (studying algorithm effectiveness)

Note: A common paradigm for experiments in HRI is the Wizard of Oz study, in which participants believe robots are autonomous with artificial intelligence but there is actually a hidden person manually controlling the robot (motions/speech).

8.2.2 Components

An experiment traditionally consists of four components:

- *Treatments* (conditions)

- *Responses* (measures)
- *Experimental Units* (subjects)
- *Assignment Method* (subject allocation method)

In medical domain, we are interested in drug effectiveness. In their experiments, the treatment is drug versus placebo because we want to know if the drug really helps patients, the subjects in the experiment. The response is the measured health outcome after the treatment. The assignment method for the treatment may be randomly selected between the drug and the placebo.

In motion planning, we are interested in comparing the results of different motion planning algorithms. In an example experiment, the treatment is random versus optimal motion for the robot. The measured responses may be more specific such as conflict/comfort. The subject of the experiment are people, participants, or users, and the assignment method is dependent on the goals of the experiment but often the participant sees both, comparing for less variance overall.

In trajectory optimization, we may propose an experiment to analyze CHOMP versus TrajOpt. The treatment would be the two algorithms. The experiment may measure computation time, success rate, or path length but ideally there would be a single measurement. The experimental units are different motion planning problems, defined by a known robot in a particular environment with a start and goal configuration. The assignment method would present each motion planning problem with both algorithms because the robot does not remember previous treatments and we are interested in head-to-head comparisons.

A study with few people is often considered a user study while a study with many people is considered an experiment.

8.2.3 Assignment Method

There are a few assignment designs to characterize the distribution of conditions to participants in an experiment:

- *Between-Subjects*: One condition per user $\rightarrow$ No biases
- *Within-Subjects*: All conditions per participant $\rightarrow$ Direct comparisons $\rightarrow$ Equal populations.
- *Mixed Designs*

There are many cases where within-subjects designs are not possible or should be avoided. For example, it may not be a good idea to give a patient both the drug and the placebo. In general, within-subjects designs should be avoided when there is a study with learning effect or some element of surprise which affects future results. For many studies, only the first exposure to a situation is valuable.

In mixed designs, conditions may vary in multiple dimensions. In some of these conditions a within-subjects design is used while in other conditions a between-subjects design is used.

8.2.4 Operationalization of Variables

Two main categories of variables in experiments are the following:

- *Independent Variables* (conditions/treatments) Independent variables are quantities that can be manipulated in the experiment. The conditions depend on how many values the independent variable can have. These are known as levels and each level is a condition. An example variable is drug type. The levels, possible conditions, for the independent variable may be drug and placebo.

  Experiment designs should consider on all possible conditions generated by independent variables. For example, with one independent variable with n levels, we would need to run n experiment settings. With two independent variables with n and m levels, we would need to run mn experiment settings.

- *Dependent Variables* (measures/responses) Dependent variables are quantities or responses that are measured. For example, health, comfort, or length may be dependent variables in an experiment.

8.2.5 Hypothesis

A hypothesis is a claim that an independent variable x affects dependent variable y. An even better hypothesis includes directionality in the form: an independent variable x positively or negatively affects dependent variable y. An example is the hypothesis that optimizing motion increases comfort. The experiment compares optimized and non-optimized motion while measuring comfort to determine if the data supports or rejects the hypothesis.

A mechanism hypothesis is one that is backed up by an argument or an analogous experiment. These hypotheses are often written in the form: because a, we hypothesize b. An example of a mechanism hypothesis may be: because more efficient motion is more predictable for people, we believe that efficient motion is more comfortable for the user. For this hypothesis, the experiment must show the effect and show that the effect is caused by the proposed mechanism.

In the typical robotics scenario, researchers often write papers that focus on showing that the performance of their algorithms are better than the performance of the state-of-the-art algorithms. Readers are left to infer the independent variables in the experiment from the differences between the algorithms. Coming up with a hypothesis requires extracting the manipulated independent variables from the experiment. The differences between the algorithms may be applicable improvements to algorithms in other domains. In general, a hypothesis is important because it helps extract the key insight or main contribution of a paper.