Report 1: Research Description

Due: Friday of week 2

Purpose: To think clearly about and reflect on the domain for your research, as well as begin/continue the necessary *writing* component of scientific endeavors

Overview

As we have discussed, the behavior of an algorithm (or larger system) is a combination of the algorithm itself, a task, and an environment. For example, the behavior of a database system depends on aspects of the system (e.g., methods for query optimization), aspects of the task (e.g., the specific queries the system is asked to execute), and aspects of the environment (e.g., disk speed and available RAM).

In this report, describe key aspects of the algorithm/system, task, and environment that you will be studying for your research project. Recall that the algorithm or system is what a computer scientist controls, the task is what a user controls, and the environment is what neither controls.

In the case of the algorithm or system, describe it in sufficient detail that the reader could design an implementation without further information. Note the use of the term *design*; actual implementation would almost certainly take additional information, but the reader should be able to identify the key components (e.g., classes and methods) necessary for an implementation.

Also, describe the task and environment that you are expecting to study. Identify key variables of the task and environment (e.g., for the database example: query complexity, structure of the database, number of records, available RAM). Note which variables you can directly control and which variables you can only measure but not directly control.

Finally, describe the behaviors you can measure, identifying the key variables (e.g., summary statistics, benchmark metrics, etc.).

Identify and briefly describe key resources that you expect to use in future phases of the project, including benchmark datasets, simulation environments, text corpora, query sets, etc. Note the extent to which these resources provide coverage of the range of the key variables mentioned above.

The length of this report should be approximately 2-4 pages of single-spaced text with additional space for diagrams and citations. Diagrams, tables, and (complete) citations are encouraged.

Details

You should produce single, collaboratively-generated report for each team. Parts of this report will be reflected in future writings and perhaps your final product. You should find that thinking clearly now will be extraordarily helpful to your productivity and future work.

In addition, I will *very strongly* encourage you to use the standard document typesetting system for producing scientific documents, *LaTeX*. It is a markup language that produces functional and attractive documents relatively easily. Rather than beginning with *LaTeX*, I will equally strongly encourage you to become familiar with an open-source graphical front-end called *LyX*. It is a "WYSIWYM" editor (what you see is what you mean). This allows you to focus on content, rather than on layout, letting *LaTeX* do the hard work. *LyX* (as well as the necessary *LaTeX*) is available on the MathLAN and our compute servers. When we (eventually) produce a report for submission to peer-reviewed conferences, we will use these tools. Having the document pieces (esp. figures, tables, and equations) already in that format will simply life beyond imagination. Trust me.

Should you need to, an easy way to generate diagrams for incorporation into a *LyX*/*LaTeX* document is with the program `xfig`. Though its interface design is quite old, it is quite simple, stable, and robust, and it creates vectorized files that reproduce well under almost any circumstances.
