CS 604: Artificial Intelligence: Spring 2016

Suggested Projects

General Notes

- Any code written should meet GNU's coding and documentation standards, and be released under the GNU GPL Version 3. Proper source code management using SVN should be practiced from the beginning.
- It is recommended but not required that any software development be interfaced and consistent with OpenAI: `http://openai.sourceforge.net`.

Problems

1. Multi-Purpose Water Resource Management

   Water is perhaps the most critical natural resource for human society; it is not a coincidence that all major civilizations have evolved close to major water bodies. Water is also a resource whose supply is severely limited, and which is needed for different, incompatible purposes (e.g., drinking and domestic consumption, crop irrigation, hydropower generation and industrial uses). Water resource management thus implies the careful use of water for diverse purposes, considering different parameters such as past, current, and predicted rainfall; reservoir capacities and levels; demands from different user communities, etc.

   Data about water and related matters (rainfall, weather patterns, reservoir levels, etc.) is available at `https://www.ksndmc.org/Default.aspx`.

   Build a model using appropriate machine learning techniques, for dynamic water resource management using historical as well as forecast data.

2. Commonsense Reasoning on Streaming & Static Data

   Using the Data Sketches stream processing algorithm made available by Yahoo (`http://venturebeat.com/2015/12/17/yahoo-data-sketches/`), create a system for commonsense reasoning with streaming data.

Such commonsense reasoning may be applied for economic or financial forecasting, or sentiment analysis (just examples, not a complete list).

Sample conclusions (just for illustration; not claimed to be true) include: *the production of rice in Thailand is inversely correlated, with a six-month delay, with the price of Brent North Sea crude oil*; *married women are more likely to purchase furniture online than single women, or either married or single men*; *the demand for luxury apartments is seasonal in London and Brussels but not in San Francisco or Paris*; etc.

Some data sources of interest are:

- http://jmcauley.ucsd.edu/data/amazon/
- http://archive.ics.uci.edu/ml/datasets.html
- https://www.cia.gov/library/publications/the-world-factbook/

3. Cake-cutting on a Resource-Starved Cloud

   "You cut, I choose" is a classical sharing algorithm for fair division of a resource (e.g., a cake) among two agents; it has desirable properties like *envy-freeness*. More sophisticated algorithms exist for multiple agents. In certain multi-user computing systems, e.g., those running high performance computing (HPC) applications with significant computing loads, it is necessary to use resources fairly but to the utmost.

   Create a cloud system with fair division where the cloud-based resources are shared fairly among any number of users, given a variable availability of resources (e.g., hardware, or energy to run it).

4. Agent-Based Modeling and Simulation of Zoonotic Diseases

   Some rare but deadly diseases such as Ebola, Rabies, and Q Fever are "zoonotic," and can be transmitted from animals (which may harbor pathogens without being severely affected) to humans. It is of much contemporary interest to understand such diseases using computational tools.

   (a) Multi-agent modeling of pathogens and hosts under different conditions (which would reduce the need for extensive field work).
   (b) Using game theory and mechanism design to formulate effective strategies to combat disease outbreaks, including ones that are yet to occur.

   Some data sources of interest are:

   - `http://www.doh.wa.gov/DataandStatisticalReports/EnvironmentalHealth/ZoonoticDiseases`
   - `http://www.absa.org/pdf/ZoonoticFactSheet.pdf`

- https://www.gov.uk/government/collections/notifications-of-infectious-diseases-noids

5. Using Artificial Immune Systems Against Real Diseases

   The core theory to be understood is *machine learning using clonal selection*; see L.N. De Castro & F.J. von Zuben, *Learning and Optimization Using the Clonal Selection Principle*, IEEE Trans. Evol. Comput. 6 (3), June 2002, doi:10.1109/TEVC.2002.1011539, and succeeding papers.

   This can be applied to age-old diseases such as Alzheimer's, TB, Diabetes, and Cancer. Issues of interest include models depicting the behavior of the diseases under different conditions, and the effects of different treatments on the diseases.

   Some data sources of interest are:

   - http://yoda.yale.edu/
   - http://www.who.int/tb/country/data/download/en/
   - https://dtp.cancer.gov/databases_tools/default.htm