CSC2542
Class Project

Sheila McIlraith and Rick Valenzano
Department of Computer Science
University of Toronto
Fall 2016

Class Project

The purpose of the project is for you to deepen your knowledge of some aspect of sequential decision making by exploring a research problem in further detail.

The course project must be on the general topic of sequential decision making or subtopics (such as search) that contribute to sequential decision making. A set of potential topics will be provided, but we encourage students to choose their own topic and to use this as a vehicle to learn something that will be of value to them, to jumpstart a new research project or to investigate a new aspect of ongoing research. (Marks: 55% of overall mark)

You can do the project on your own, or in groups of 2 or 3, but the expectation for the project will commensurate with the size of the team.

...

Class Project

Important Dates

Now: Start thinking about your project and talking to Sheila/Rick

Oct 31: Last day to hand in project proposal
(earlier is better for you, so you start on the work asap)

Dec 1 and/or Dec 8 (TBD):
Short in-class project presentations (status report)

Dec 15: Projects due

...

Class Project -- Evaluation

Evaluation of the project (55 marks) will be as follows:

- (5 marks) Your project proposal. **(Oct 31: last day to hand in)**
- (10 marks) Your project presentation. Your presentation will be given in a class towards the end of term. As such, your presentation may have to be given before your project is completed. **(Dec 1: start)**
- (40 marks) For the overall quality of your project, based in part on its level of difficulty, on the insights you exposed, and any novel ideas of your own that you are able to explore, and your final analysis of your project. A major proportion of this mark will depend on the students' presentation of their final results. This should usually be in the form of a formal written paper, perhaps with a well-structured web site to show results, if relevant.

...

Class Project – Evaluation (cont.)

- (40% cont.) A major component of the report will be a review and analysis of the related literature, along with your assessment of the effectiveness and relative merits of each approach. This will focus mainly on the 2-4 papers you chose, but will also likely require several further sources in order to provide sufficient groundwork. The written report and/or website should also include a detailed description of any algorithms you implemented. This should include problems you faced, the mathematical details of what was implemented, and an assessment of any empirical results. **(Dec 15: report due)**

Class Project – Project Proposal

Your 2-page **project proposal** is worth 5% of your course mark. The proposal should comprise:

- a careful description of the problem your project will address;
- a set of approx 2-4 research papers from which the projects will be drawn;
- a description of the approach you will take to addressing the project;
- a description of how you will evaluate the success of the project;
- a rough schedule for when you'll accomplish the work

Start thinking about your project early. **Come and talk to us now and before submitting your proposal!**

...

Class Project – Project Presentation

Your **project presentation** is worth 10% of your mark. Your presentation should be n minutes in length (n minutes for actual presentation and m minutes for questions). n minutes = approx n/2 slides. (Apologies. Since we are allowing group projects, n and m will be determined based on the number of project proposals we receive.) You should try cover the following points:

- Statement of the problem you are trying to address
- Why the problem is interesting or important
- What has been done in the area until now (can be short)
- Your approach
- How you expect to evaluate your work
- Progress you've made to date

➔ Use this as a guide. Don't use it if it doesn't seem to fit your project.
➔ We expect each of you to talk to Sheila/Rick at least 4 days before your presentation to go over the outline of your slides.

Examples

- The next few pages contain examples of what you might need to do for various kinds of projects. These are just samples to get you started thinking …

Example 1

- Suppose you want to compare algorithm X with algorithm Y
  - E.g., compare model-based RL to model-free RL
- Has someone already done it? (check the literature)
- Theoretical comparison?
  - Expressivity: are there domains that one algorithm can represent/ solve but the other can't (or can't represent as well)?
  - Worst-case complexity: does it mean much in practice?
  - Average-case analysis: probably pretty difficult
- Experimental study?
  - How can you be sure the experiments mean anything?
    - Randomly generated problems? What about built-in bias?
    - Problems from the planning competition?
    - Statistical significance? Confounding factors?
- Overall significance of the results?
  - Does it say something general, or just about one kind of problem?

Example 2

- Suppose you want to apply search, automated planning, MDPs, RL to some application *X*
  - Do you know enough about application *X* to do a credible job?
  - Do you need data and do you have access to the data and is it clean and in a usable form?
  - Has someone already done it? (check the literature)
  - Analyze real-world requirements
  - Formalize the problem
  - Which algorithm? What kind of algorithm?
    - Domain-independent, domain- customized, domain-specific?
  - What makes your algorithm better than others for this problem?
    - Maybe compare two algorithms? (see example 1)
  - What will you do if the algorithm doesn't work well?

Example 3

- Suppose you want to extend algorithm X to include feature Y
- E.g., modify Satplan, FastForward or some other planner
  - To include preferences, state variables, temporal reasoning, resource management, control rules, or dependency-directed backtracking
  - To produce conditional or conformant plans
  - To work in multi-agent environments (cooperative? adversarial?)
- Has someone already done it? (check the literature)
- What is the motivation for the feature? (sometimes trivial to address)
- What characteristics does your algorithm have?
  - Soundness, completeness, efficiency?
  - General idea or just a hack?
- Experimental evaluation (see example 2)
- Overall significance?
  - How general is your approach? Could it be made to work for other algorithms as well?
  - Does your approach extend to other algorithms?
  - Good for just one kind of problem, or for many? Which ones?

Further Options/Dimensions

- Literature survey
- Team project
- ….

Some things to think about…

- If you succeed in carrying out your idea, will the result be interesting?
  - Interesting to you? Interesting to others?
- What is needed to carry out your idea
  - Is it too hard to accomplish in the amount of time that you have? Is it too easy to count as a "real" project?
- How do you ensure success regardless of how the result turns out?
  - If the only interesting/significant result is "yes it works," then you'll be in trouble if you can't actually get it to work. You should either
  - (1) think enough about it to be pretty sure you can get it to work, or
  - (2) reformulate it in such that any of the likely outcomes will be interesting as the main result of your project

Assessment Criteria

Your project doesn't have to necessarily make a new research contribution, though this would be nice! The primary objective is pedagogical. Nevertheless, as you train to be good researchers, it is good to evaluate whether your work is a research contribution. To that end, here are some of the questions that are often asked of reviewers assessing your work for publication (ICAPS-07 review form):

Rank (1=bad, 5=middle, 10=good)

- Reviewer familiarity *
- Relevance to the call of papers *
- Technical quality and soundness *
- Originality and novelty *
- Significance to theory and practice *
- Readability and organization *
- Overall recommendation *

Assessment Criteria (cont.)

- ORIGINALITY
  - Is the work described in the paper novel?
- SIGNIFICANCE
  - Is the work important?
  - [Research Track] Does the work present theoretical or experimental results that advance the current state of the art in planning and scheduling?
  - [Application Track] Does the work describe a high-impact application or use of planning and scheduling technologies in anoperational setting?
  - [System Track] Does the paper describe a significant integration of diverse component technologies into complex planning, scheduling and execution systems?

Assessment Criteria (cont. 2)

- TECHNICAL QUALITY
  - Is the work technically sound?
  - Are the paper's arguments compelling?
  - Is there a compelling empirical evaluation?
  - [Application Track] Does the work clearly rationalize the use of planning/scheduling technologies in the target application and quantify the benefit over current practice?
  - [System Track] Does the work present a successful integration of planning, scheduling, learning, constraint satisfaction or other technologies into a system that solves a well-defined problem? Does the system represent a well-engineered solution to the problem?
- QUALITY OF PRESENTATION
  - Is the paper clearly written?
  - Does the paper motivate the research or application?
  - Is the paper well organized?

Specific Project Ideas

<we can discuss one on one>

General ideas are contained in a living document to be posted via Piazza