

COMP3004/COMP4105

Designing Intelligent Agents

Coursework

Colin Johnson

Overview

The coursework for this module is based around (1) designing intelligent autonomous agents and an environment with which they interact, (2) setting those agents a task, (3) asking one or more questions about that task, and (4) evaluating it using experimental methods. You will then present the results from this in a report, which will also explain the context for the work.

Details

An autonomous intelligent agent is a program that operates in a particular environment, perceives aspects of that environment, and then carries out actions that change that environment to carry out some task. Typically, these actions are a mixture of responses to its perception and proactive actions such as exploration.

Your task for this coursework is to design an agent-based system containing the following four aspects:

An Environment. This is the (virtual) place where the agents will operate. It could be one of:

- A simulation of a physical environment in which mobile robotic agents move. This could be the simulation used in the classes earlier in the semester (perhaps extended), a robot environment such as *The Player Project* (<http://playerstage.sourceforge.net>), or a project in Unity or a similar game environment if you are familiar with one from elsewhere.
- A chatbot environment such as the ones used in the classes.
- The blackboard system used in the class where we discussed language agents writing poetry
- The *Bristol Stock Exchange* system (<https://github.com/davecliff/BristolStockExchange>) or a similar simulation of some aspect of the economy or society
- A game environment such as *Ms. PacMan* (<https://gym.openai.com/envs/MsPacman-v0/>), the *Open Racing Car Simulator* (<http://torcs.sourceforge.net>), RoboCup (<https://www.robocup.org/leagues/23>) or similar (see e.g. <http://www.gvgai.net>)
- One of the more complex task environments from the *OpenAI Gym* (<https://gym.openai.com>)

There is no need to develop the environment yourself—the focus of the project will be on the agents in the environment (chatbots, robots, trading agents, game-playing agents, autonomous drivers, etc.) – but it is likely that you will set up the details of the environment to address your specific question. You are allowed to use the code from the classes, but please try to make it clear broadly which parts of the code are taken from the class examples, and which is your own work (we appreciate that this is sometimes complicated to do at a line-by-line level, but you should indicate this in broad terms).

Autonomous Agents. You should introduce one or more autonomous intelligent agents into the environment, which use some kind of AI to solve a task.

- Examples of AI could be an AI planning system such as Goal Oriented Action Planning (<http://alumni.media.mit.edu/~jorkin/goap.html>), a search algorithm such as

A* search, a genetic or swarm search, a reinforcement learning algorithm, fuzzy logic, or a hard-coded reactive or state-machine AI.

- The task will be one relevant to the environment: e.g. a robot vacuum cleaner clearing up dirt, a chatbot taking an order from a customer, a trader trying to optimise its returns, a game player trying to get a high score in a game, etc.

Within reason, you can use any language to do this. If you are planning to use anything other than Python, Java, MATLAB/Octave, JavaScript, and standard web technologies such as HTML/CSS, then please mention this in your topic approval.

A Question. You should ask a specific question (or a set of related questions) about your system. For example:

- How do different approaches (a genetic algorithm, an A* search algorithm, a hard-coded heuristic) compare in terms of task performance?
- How does the performance of the system change as we vary the number of agents in it?
- If the system is trained on one version of the environment, does that learning transfer over to a new version of the environment?
- How do different kinds of communication/coordination between agents effect the efficiency of those agents on the task?
- How much improvement does storing some information (e.g. a map of the environment) make compared to carrying out the task in a purely reactive way?
- How do different kinds of sensing/perception systems affect the capacity of the agent to carry out its task?
- How sensitive is the agent to error/noise?

A Set of Experiments. You should answer your question by carrying out a set of experiments. Remember the structure that we talked about in one of the lectures:

- implement code that carries out a run of the agent's behaviour and measures performance
- then, run that code multiple times to get a measure of average performance
- then, repeat that process for the different conditions in your question, and use descriptive statistics, charts/visualisation, and/or inferential statistics (e.g. significance tests) to test your question

Then, you are in a position to discuss the question using these experimental results as your evidence.

If this evaluation involves asking people to interact with your code, then please read the notes on research ethics that are in the *Coursework* section of the module Moodle page.

Constraint. You should not just run a standard reinforcement learning or convolutional neural networks tutorial on a standard environment from AI Gym or similar, though you can use these as part of your work.

Examples

Here are a few examples of things that you could do. You don't have to do one of these—indeed, we would prefer you to come up with your own idea—but, these would all be acceptable project ideas if you want to do them:

- To take the “robot vacuum cleaner” from the early classes, and experiment with different numbers of robots, and different coordination strategies (e.g. robots try to stay a fixed distance from each other, compared to sharing a map that they build up)
- Contrast random, fixed and planned orders of asking questions in a chatbot, and see (perhaps by doing a brief user test) which one is better.

- Take a number of different trading strategies and run them in the *Bristol Stock Exchange* system with varying amounts of noise/uncertainty, to see how robust each strategy is.
- Take the “avoid the cats” problem from the class, and compare a number of strategies for the problem: warning the cats vs. moving out of the way, and learning when to act based on a simple statistical approach vs. a decision-tree approach.
- Consider the problem of planning a robot’s movement around a mapped environment (e.g. the map generated from WiFi triangulation introduced in one of the classes). Contrast A* search and genetic algorithms on this problem, and compare them both against random wandering.

Topic Approval

You should submit a short description of your project idea (around a paragraph, 100 word) on the Moodle page by 15:00 on the 1st April 2022. We will then give you feedback on whether the project is an acceptable one, and how it might be modified or improved. You do not have to wait until then before submitting your idea; we will start looking at them from mid-April onwards.

Submission

By 3pm on 6th May 2022 you should submit the following. This may be extended if you have a support plan or extenuating circumstances. Late submissions will incur a [penalty of 5% per working day](#), up to 13th June 2022, after which you will receive a mark of zero.

COMP3004

A report, around 2500 words (up to 3500 little longer if you need it), where you describe:

- The core ideas of your project; clearly state the question that you are trying to answer
- A review of relevant ideas, technologies and research papers
- How you designed the environment and agents in order to address that question
- Technologies used, and challenges that you met in doing the implementation
- How you set up and ran your experiments
- The results from your experiments
- A discussion of the question in light of the experimental results
- A conclusion, where you summarise the work, reflect on its successes and limitations, and briefly mention some ideas for how you would take the work forward if you had more time

A copy of your code, either as an upload or a link to a repository

Anything else that you think would be helpful for the markers, e.g. sample outputs from your system, a link to a brief video demonstrating it working, etc.

COMP4105

A report, around 2000 words (up to 3000 if you need it), where you describe:

- The core ideas of your project; clearly state the question that you are trying to answer
- How you designed the environment and agents in order to address that question
- Technologies used, and challenges that you met in doing the implementation
- How you set up and ran your experiments
- The results from your experiments

- A discussion of the question in light of the experimental results
- A conclusion, where you summarise the work, reflect on its successes and limitations, and briefly mention some ideas for how you would take the work forward if you had more time

A report, about 1500-2000 words, where you give an extended review of relevant ideas, technologies and research papers

A 10-minute presentation about your work (dates/times will be arranged)

A copy of your code, either as an upload or a link to a repository

Anything else that you think would be helpful for the markers, e.g. sample outputs from your system, a link to a brief video demonstrating it working, etc.

Academic Integrity

This is an individual assessment that should consist of your own unaided work. You should make any direct quotations clear both by using quotation marks and by providing a clear reference to the paper immediately after the quotation. If you are building on someone else's code (e.g. our code from the classes, open-source projects, etc.), please make it clear which aspects of the code are your work through the use of comments. The University has [detailed advice about academic integrity](#), and submissions that demonstrate a lack of that integrity will be treated under appropriate disciplinary procedures.

How the Work will be Marked

Marking will take into account:

- background research and how you have used it to contextualise your work
- the choice of task environment and how you have used it/adapted it for your specific project
- the effective use of artificial intelligence and agent-based systems ideas from the course and your wider studies in designing your autonomous agents
- how clear your question(s) are, how well the experiments have been designed to answer them, and your level of rigour in planning and analysing the experiments
- how well the report answers the question by using the evidence from the experiments
- the overall clarity and structure of the report, appropriate use of scientific and technical English, and the quality of charts, diagrams, pseudocode where relevant
- the quality of reflection on the successes and limitations of the work
- (for students doing a presentation) the structure of the presentation, the clarity of explanations, and good use of slides or other visual aids

COMP3004 Marking Scheme

Each of the following descriptors gives a broad idea of the achievement expected for a mark in that range. Clearly, individual projects may fall short in some areas and show excellence in others. The marking should also be adjusted to reflect the intrinsic difficulty of the project.

Band	Guidelines
90-100	Marks in this range are reserved for a superb all-round performance. Work done in all aspects of the project go beyond even high expectations. The student has shown a thorough understanding of the problem. All expected tasks have been successfully completed, the project shows depth and engagement with research ideas, and everything has been completed to a high standard. The report could form the basis of a publishable conference/workshop paper.
80-89	Excellent contributions to all areas of the project. Exceeded expectations in some areas. Demonstrates knowledge and understanding of the project that is beyond standard resources covered in the module. Clear appreciation of the project as a whole, its adequacies, limitations and possibilities for future development. The project demonstrates insight and depth beyond that usually expected in undergraduate work.
70-79	Very good contributions to all areas of the project. Successful completion of the project tasks. Demonstrated initiative and creative problem-solving ability. Able to undertake the work in a competent and independent manner. Able to reflect accurately on adequacy and limitations of the project's achievements.
60-69	Good appreciation of background. A good attempt at applying this to the task, with demonstrated ability to cope with difficulties. Good technical skills in several areas. Whilst most of the core aims of the project have been achieved, it might come a little short in some areas. Good reflective understanding of the project.
50-59	Satisfactory background reading and a competent attempt at their tasks. Reasonable technical competence demonstrated. The core task completed satisfactorily, but little achieved beyond that. Able to reflect satisfactorily on the project.
40-49	Pass level. Competent background reading and appreciation of the project area. Basic technological competence. Some areas of the core tasks may be incomplete, but a decent attempt has been made at them. Able to reflect in a limited way on the project.
30-39	Unsatisfactory. Some attempt has been made at the background reading but clearly only partial understanding of project topic. Incomplete attempt at the core tasks. Weak technical competence. Little ability to reflect adequately on the project.
20-29	Inadequate background reading, but shows some limited understanding of how ideas can be linked to the task. Minimal attempt at the core tasks, showing poor understanding. A substantial amount of work is still needed to achieve the core tasks. Minimal reflection on the project.
10-19	Minimal attempt at background reading, inappropriate use of material, almost no attempt at core tasks. Very poor understanding of the problem. Minimal or no reflection on the project.
0-9	No or almost no significant attempt.

COMP4105 Marking Schemes

Each of the following descriptors gives a broad idea of the achievement expected for a mark in that range. Clearly, individual projects may fall short in some areas and show excellence in others. The marking should also be adjusted to reflect the intrinsic difficulty of the project.

Presentation COMP4105

Band	Guidelines
9-10	A professional-level presentation of exceptional clarity and very clear structure, with very high-quality slides or other visual aids that flow seamlessly with the spoken presentation
7-8	A clearly structured presentation, which explains all aspects of the project well, and which has high-quality slides or other visual aids that are tied in strongly with the spoken presentation
5-6	Pass level. A competent presentation that has a decent structure, gives a competent description of most aspects of the project, and where the slides and other visual aids are largely clear and related to the spoken presentation
3-4	A presentation that has some level of organisation but where the topics are not presented in a clear order or where the presentation jumps from topic-to-topic, some explanations not clear, visual aids provided but not very clear and/or not very related to the spoken presentation
1-2	A presentation that mentions some aspects of the project work but is largely disorganised to the point where it cannot be followed and where most explanations are unclear, and where visual aids are unclear and/or not related to the spoken presentation
0	No significant attempt at presentation

Review report COMP4105

Band	Guidelines
90-100	A professional-level review of the literature/technology, demonstrating a clear and deeply analytical/critical understanding of relevant work, an original structure to the review, insightful links between the various papers and technologies reviewed leading to an innovative thematic analysis, and a very clear link to the project work.
80-89	A clear and deeply analytical/critical understanding of relevant work, an insightful structure to the review, a clear thematic understanding making links between the various papers and technologies clear, and a very clear link to the project work.
70-79	A clear and analytical/critical understanding of relevant work, a structure to the review, links between the various papers and technologies brought out, and a clear link to the project work.
60-69	A clear understanding of relevant work with some analysis/critique, some structure to the review, some thematic links between papers and technologies identified, and connected in parts to the project work.
50-59	Pass level. A competent understanding of relevant work with some attempt at analysis/critique in parts, some structure to the review, occasional attempts to make thematic links between papers and technologies, and broadly related to the project work.
40-49	Understanding of some relevant work but incomplete/misunderstood in parts, with little analysis/critique, largely unstructured, not many links between papers/technologies, link with project not very clear
30-39	Some papers/technologies have been studied and there is some understanding, but much is incomplete/misunderstood/irrelevant, no meaningful analysis/critique, papers/technologies presented independently, link with project not very clear
20-29	A few papers/technologies have been looked at, but understanding is low and there is no analysis/critique, no links between items studied, minimal structure, little link with project.
10-19	A couple of paper/technologies have been mentioned, but with very little understanding, no links between them or structure, and not clear how they relate to the project
0-9	No or almost no significant attempt.

Project work COMP4105

Band	Guidelines
90-100	Marks in this range are reserved for a superb all-round performance. Work done in all aspects of the project go beyond even high expectations. The student has shown a thorough understanding of the problem. All expected tasks have been successfully completed, the project shows depth and engagement with research ideas, and everything has been completed to a high standard. The report could form the basis of a publishable conference/workshop paper.
80-89	Excellent contributions to all areas of the project. Exceeded expectations in some areas. Demonstrates knowledge and understanding of the project that is beyond standard resources covered in the module. Clear appreciation of the project as a whole, its adequacies, limitations and possibilities for future development. The project demonstrates insight and depth beyond that usually expected in undergraduate work.
70-79	Very good contributions to all areas of the project. Successful completion of the project tasks. Demonstrated initiative and creative problem-solving ability. Able to undertake the work in a competent and independent manner. Able to reflect accurately on adequacy and limitations of the project's achievements.
60-69	Good appreciation of background. A good attempt at applying this to the task, with demonstrated ability to cope with difficulties. Good technical skills in several areas. Whilst most of the core aims of the project have been achieved, it might come a little short in some areas. Satisfactory or good reflective understanding of the project
50-59	Pass level. Competent background reading and appreciation of the project area. Basic technological competence. Some areas of the core tasks may be incomplete, but a decent attempt has been made at them. Able to reflect in a limited way on the project.
40-49	Unsatisfactory. Some attempt has been made at the background reading but clearly only partial understanding of project topic. Incomplete attempt at the core tasks. Weak technical competence. Little ability to reflect adequately on the project.
30-39	An attempt has been made at the background reading but clearly only partial understanding of project topic. Attempt at the core tasks, but not much achieved overall. Weak technical competence. Minimal reflection on the project.
20-29	Inadequate background reading, but shows some limited understanding of how ideas can be linked to the task. Minimal attempt at the core tasks, showing poor understanding. A substantial amount of work is still needed to achieve the core tasks. Minimal reflection on the project.
10-19	Minimal attempt at background reading, inappropriate use of material, almost no attempt at core tasks. Very poor understanding of the problem. Minimal or no reflection on the project.
0-9	No or almost no significant attempt.