KBAI Project 2 (classification)

(10/27/2018)

	Project 1 Understanding	Project 2 Classification	Project 3 Learning
KBAI Algorithms	• Thematic Roles • Frames	• Semantic Networks • Generate and Test • Production Systems (rules based systems) (decision trees) • Means-Ends Analysis • Problem Reduction • Case Based Reasoning	• Incremental Concept Learning • Version Spaces • Analogical Reasoning • Logic • Planning • Scripts • Learning by Correcting Mistakes • Learning by Recording Cases

Classification

"*Classification is mapping sets of percepts in the world into equals classes, so that we can take actions in the world in an efficient manner.*" [1]
https://gatech.instructure.com/courses/40559/files/folder/Video%20Slides?preview=1040865

Semantic Networks

"*A semantic network purports to represent concepts expressed by natural-language words and phrases as nodes connected to other such concepts by a particular set of arcs called semantic relations.*" [2]
https://gatech.instructure.com/courses/40559/files/folder/Video%20Slides?preview=1040858

Generate and Test

"*Generate and test is a problem solving method.*" [1] "*Given a problem, generate potential solutions to it, and then test the solutions for the efficiency for addressing the problem.*" [1]
https://gatech.instructure.com/courses/40559/files/folder/Video%20Slides?preview=1040859

Production Systems (rules based systems) (decision trees)

"*A production system consists of three components: a long-term memory (in the form of a rule base), a working memory, and an inference engine. The rule base contains rules, the conditions of which must be matched to elements in working memory. The inference engine determines which of the rules in the rule base have all their conditions matched to objects in working memory, and then decides which rule to apply. The application of a rule will usually cause elements in working memory to be added, removed, or altered. Further*
rules can then be matched by the inference engine." [7]
https://gatech.instructure.com/courses/40559/files/folder/Video%20Slides?preview=1040860

Means-Ends Analysis / Problem Reduction

"*The general attributed his success to an important principle: If you need a large force to accomplish some purpose, but are prevented from applying such a force directly, many smaller forces applied simultaneously from different directions may work just as well.*" [8]
https://gatech.instructure.com/courses/40559/files/folder/Video%20Slides?preview=1294354

Case Based Reasoning

"*Case-based reasoning means using old experiences to understand and solve new problems.* " [9]
https://gatech.instructure.com/courses/40559/files/folder/Video%20Slides?preview=1040863

Project Details

Design an agent, using KBAI techniques, to answer **simple** questions about the class syllabus.

Domain

Intent	Object	Category	Answer (vocabulary)
0			I do not know
1	Assignment 1	RELEASEDATE	Assignment 1 will be available in the second week.
2	Project 1	RELEASEDATE	You can download project 1 during week 3.
3	Assignment 2	RELEASEDATE	Assignment 2 will be released on week 6.
4	Midterm	RELEASEDATE	The midterm occurs in week 7.
5	Project 2	RELEASEDATE	Project 2 will be distributed in week 8.
6	Assignment 3	RELEASEDATE	Assignment 3 starts in week 11.
7	Project 3	RELEASEDATE	Project 3 will be available in week 13.
8	Final	RELEASEDATE	The final occurs in week 16.
9	Assignment 1	DUEDATE	Assignment 1 is due at the start of week 3.
10	Project 1	DUEDATE	Project 1 must be turned in at the beginning of week 6.
11	Assignment 2	DUEDATE	Assignment 2 must be submitted at the start of week 7.
12	Midterm	DUEDATE	The midterm must be completed by the end of week 7.
13	Project 2	DUEDATE	Project 2 must be submitted by the end of week 11.
14	Assignment 3	DUEDATE	Assignment 3 is due by the end of week 12.
15	Project 3	DUEDATE	Project 3 must be submitted by the end of week 16.
16	Final	DUEDATE	The final is due at the end of week 16.
17	Assignment 1	DURATION	1 week
18	Project 1	DURATION	3 weeks
19	Assignment 2	DURATION	1 week
20	Midterm	DURATION	1 week
21	Project 2	DURATION	3 weeks
22	Assignment 3	DURATION	1 week
23	Project 3	DURATION	3 weeks
24	Final	DURATION	1 week
25	Assignment 1	WEIGHT	4% of final grade
26	Project 1	WEIGHT	15% of final grade
27	Assignment 2	WEIGHT	4% of final grade
28	Midterm	WEIGHT	15% of grade
29	Project 2	WEIGHT	15% of final grade
30	Assignment 3	WEIGHT	4% of final grade
31	Project 3	WEIGHT	15% of final grade
32	Final	WEIGHT	20% of grade
33	Assignment 1	PROCESS	Turn in to Canvas as PDF.
34	Project 1	PROCESS	Turn in code as zip file, and report as pdf into Canvas.
35	Assignment 2	PROCESS	Turn in to Canvas as PDF.
36	Midterm	PROCESS	Turn in to Canvas as PDF.
37	Project 2	PROCESS	Turn in code as zip file, and report as pdf into Canvas.
38	Assignment 3	PROCESS	Turn in to Canvas as PDF.
39	Project 3	PROCESS	Turn in code as zip file, and report as pdf into Canvas.
40	Final	PROCESS	Turn in to Canvas as PDF.
41		GENERAL	Announcements, question answering and discussions: We will use the Piazza forum for announcements, question answering, discussions, and collaboration. It is important that you log into Piazza regularly and frequently (at least two or three times a week, daily if possible).
42		GENERAL	The class is organized around three primary learning goals. First, this class teaches the concepts, methods, and prominent issues in knowledge-based artificial intelligence. Second, it teaches the specific skills and abilities needed to apply those concepts to the design of knowledge-based AI agents. Third, it teaches the relationship between knowledge-based artificial intelligence and the study of human cognition.

Table 1. Responses and Vocabulary

Example Questions (vocabulary) (9 word limit)	Words	Object	Category	Intent
when will **assignment 4** be released				**0**
when will **assignment 1** be released	6	Assignment 1	RELEASEDATE	**1**
when can we begin working on **project 1**	7	Project 1	RELEASEDATE	**2**
when can we start on **assignment 2**	7	Assignment 2	RELEASEDATE	**3**
what week is the **midterm**	5	Midterm	RELEASEDATE	**4**
when can i download **project 2**	6	Project 2	RELEASEDATE	**5**
What week does **assignment 3** start	6	Assignment 3	RELEASEDATE	**6**
when can we begin working on **project 3**	7	Project 3	RELEASEDATE	**7**
when will the **final** open	5	Final	RELEASEDATE	**8**
when is **assignment 1** due	7	Assignment 1	DUEDATE	**9**
when will **project 1** need to be submitted	8	Project 1	DUEDATE	**10**
when will i need to submit **assignment 2**	8	Assignment 2	DUEDATE	**11**
when is the **midterm** due	5	Midterm	DUEDATE	**12**
when should we have **project 2** completed by	8	Project 2	DUEDATE	**13**
when do I need to turn in **assignment 3**	9	Assignment 3	DUEDATE	**14**
when will submissions close for **project 3**	7	Project 3	DUEDATE	**15**
when do I need to turn in the **final**	9	Final	DUEDATE	**16**
how much time is there for submitting **assignment 1**	9	Assignment 1	DURATION	**17**
how long do we have to complete **project 1**	9	Project 1	DURATION	**18**
how long do we have to finish **assignment 2**	9	Assignment 2	DURATION	**19**
how long do we have to complete the **midterm**	9	Midterm	DURATION	**20**
how long do we have to do **project 2**	9	Project 2	DURATION	**21**
how many weeks to write **assignment 3**	7	Assignment 3	DURATION	**22**
How many weeks to code **project 3**	7	Project 3	DURATION	**23**
How many weeks to complete the **final**	7	Final	DURATION	**24**
what percentage of my grade is **assignment 1** worth	9	Assignment 1	WEIGHT	**25**
how much is **project 1** worth	6	Project 1	WEIGHT	**26**
what is the weight of **assignment 2**	7	Assignment 2	WEIGHT	**27**
how much will the **midterm** be worth	7	Midterm	WEIGHT	**28**
what percentage of the total grade is **project 2**	9	Project 2	WEIGHT	**29**
how much does **assignment 3** contribute to my grade	9	Assignment 3	WEIGHT	**30**
how much is **project 3** contributing to my grade	9	Project 3	WEIGHT	**31**
how much is the **final** worth	6	Final	WEIGHT	**32**
what is the process for submitting **assignment 1**	8	Assignment 1	PROCESS	**33**
what is the process of submitting **project 1**	8	Project 1	PROCESS	**34**
where do i turn in **assignment 2**	7	Assignment 2	PROCESS	**35**
where do i turn in my **midterm**	7	Midterm	PROCESS	**36**
How do I turn in **project 2**	6	Project 2	PROCESS	**37**
where do i submit **assignment 3**	6	Assignment 3	PROCESS	**38**
where do i submit **project 3**	6	Project 3	PROCESS	**39**
what is the procedure for submitting the **final**	8	Final	PROCESS	**40**
where do i go to get class announcements	8		GENERAL	**41**
what are the primary learning goals for the class	9		GENERAL	**42**

Table 2. Example Questions and Vocabulary

Question Limitations

- Questions start with: ‘when’, ‘what’, ‘where’, ‘how’.
- The questions are limited to 9 words including digits.
- The vocabulary is limited to the words (all the words) in tables 1 and 2.
- Your agent will be given a question and should return an intent number.
- The test questions will be worded differently from the example questions (table 2).
- The questions will be all lowercase, no question mark.

Vocabulary

when	on	9	turn	worth	procedure	least	prominent
will	project	10	agents	4%	go	two	discussions
ai	you	final	close	15%	get	or	knowledge-based
4	design	open	for	weight	class	three	~~discussions~~
be	during	13	how	total	abilities	times	intelligence
released	3	due	much	contribute	question	a	collaboration
i	start	at	time	contributing	answering	daily	submissions
do	2	of	there	20%	specific	if	announcements
not	6	need	skills	process	use	possible	assignment
know	~~what~~	to	long	canvas	piazza	are	distributed
1	is	needed	complete	as	forum	primary	download
available	midterm	must	weeks	pdf	artificial	learning	submitted
in	occurs	turned	finish	getting	issues	goals	beginning
the	7	apply	many	zip	it	organized	completed
second	third	submit	write	file	important	around	submitting
week	8	those	code	and	that	first	relationship
can	what	by	days	report	log	this	between
we	day	end	study	into	regularly	teaches	percentage
begin	does	should	my	where	frequently	concepts	human
working	starts	have	grade	specification	at	methods	cognition

What to Do

1. Create 60 question based on the question limitations and the vocabulary
2. Parse the sentences
3. Classify the parsed result into an intent (a intent is just a label)
4. Note how you classified the sentences
5. Write an agent to classify the sentences using what you learned when you did it manually in step 3
6. Your agent should take in a sentence and return an intent (integer between 0 and 42)

Not as Simple as it Seems

- Your agent must return intent 0 if it does not have enough data to answer the question
    - When is assignment 4 due?
    - Where is the class?
    - What time is the class?
    - When will the project grade be released? (your agent does not know this answer)
- The test sentences for project 2 will be grammatically difficult compared to project 1
    - What are the goals for this class?
    - When is the beginning of project 1?
    - What is the final date for assignment 3 submissions?
- This is not a simple word search. You must build and code a model based on the available information.
    - **What project can we start during week 4 - intent 2**
    - **What can I download in week 4 - intent 0 (look closely at the tables)**
    - **What project do I turn in by week 6 - intent 10**

Differences Between Project 1 and Project 2

- Project 1 was about inferring: "*Deduce or conclude (information) from **evidence and reasoning** rather than from **explicit statements**.*"
- Project 2 is about classification: "*The action or process of classifying something according to **shared qualities or characteristics**.*"
- The prior knowledge for project 1 was the English language rules (and a few keywords).
- The prior knowledge for project 2 is a complete vocabulary in a well defined domain (more prior information).
- The goal for project 1 was to **extract** information from a sentence.
- The goal for project 2 is to **compare** information in sentences to look for ***shared qualities or characteristics***.
- Project 1 required the use of thematic roles.
- Project 2 **does not** require the use of thematic roles.
- The test sentences for project 2 will be grammatically difficult compared to project 1.
- For project 2, not all test questions will be in domain (answerable). Must reply 'I do not know'.
    - This requires that your agent understand the boundaries of the domain.

Getting the code

```
git clone https://github.gatech.edu/Dilab/CS7637AOProjects.git
```

Look in project 2 directory

Executing the code (you must use Python 3)

```
python AgentGrader.py -v

==========================================
Autograder - Version Project 2 student092518


usage: -q <json containing questions/intents>
       -l <path/filename to log file>
       -v verbose output to console
       -h this message to console
==========================================
Opening questions: ExampleQuestions.json
Redirecting to file: results.out
Logging to file: results.log
Instantiating student agent

Starting test
question,intent,studentIntent,count,# correct,
when will assignment 4 be released,0,0,1,1,

……...

count,match,
43,1,
Done
```

Code (in project 2 directory)

File	Change?	Description
ExampleQuestions.json	YES	Add your test questions in this file. There should be multiple questions for each intent.
StudentAgent.py	YES	Add your code here
Vocabulary.txt	NO	Project vocabulary
AgentInterface.py	NO	Autograder to agent interface
AgentGrader.py	NO	The autograder will test your agent and output a result

You can ADD more files for your project. The autograder simply overwrites the files in red above.

Grading

Your agent will be asked 100 questions. Each question is worth 0.5 points. Any answer outside of 0 to 42 is marked as incorrect. The results from your agent represent 50% of your grade. The report represents the other 50% of your grade.

Reflection Rubric (50% of grade)

Item #	P	Address in your report
1	10	**Human thought process** Explain your mental process for classifying the sentences. How did you, as a human, classify the sentences? What prior knowledge did you as a human require to mentally classify the sentences? As a human, did you fail to classify any sentences? If so, **why**? As a human what did you do when you encountered a sentence you could not classify?

2	10	**Map human thought processes to KBAI techniques** How do the KBAI techniques taught in class fit into the mental processes you used to classify the sentences. Create a diagram, that describes your mental process for classifying sentences using KBAI techniques.
3	10	**Create an agent based on your mental process** **Compare your agent's processes to your mental processes** Create a diagram of your agent. Compare this diagram to the diagram you created in item 2. Explain the comparison with an emphasis on why. **Why** are there differences? **Why** did you make the decisions you did with your agent design?
4	10	How did you create test questions? Please include 10 test questions in your report. 5 questions your agent classifies correctly, and 5 questions your agent classifies incorrectly. **Why** did your agent answer the questions incorrectly?
5	10	**Metacognition** What did you learn about human cognition? Did the act of designing your agent highlight anything about your thought process and how humans approach problems?

Please refer to the class syllabus for guidance on report writing and word limits.

Libraries

Only libraries listed below are allowed. Your code can only import modules you have created or modules listed in the table below.

Legal Libraries (ALL OTHER LIBRARIES ARE NOT PERMITTED)

Python 3 standard library (https://docs.python.org/3/library/)

The autograding machine will not have any additional libraries installed. Please verify your imports do not require any libraries not listed above. **If you use any libraries not listed above, your code will not run on the autograding machine and you will get a 0.**

Some Python IDE's will automatically import debug libraries. Please test your final code outside of the IDE directly from the command line in a clean Python environment.

Project Goals

The project goals center around metacognition. The goal is to get you thinking about thinking. Your first step for this project should be to classify the sentences manually. Think about your internal thought process while you do it. Maybe ask friend or family to do it and explain their mental process outloud.

References

1. KBAI Ebook
2. http://www.cs.cmu.edu/~dgovinda/pdf/semantics/Semantic%20Networks.pdf
3. http://www.aclweb.org/anthology/W16-6609
4. https://link.springer.com/content/pdf/10.1023/A:1022642026684.pdf
5. https://pdfs.semanticscholar.org/657d/9008cc645be0521055f7bdd19821d42c7d91.pdf
6. https://www.cs.bham.ac.uk/~jxb/IAI/w7.pdf
7. https://pdfs.semanticscholar.org/a5e0/335e0597cfd489c2373a67d0e376d6668c1e.pdf
8. http://www.psychology.sunysb.edu/hwaters-/psy325/slidesets/325slideset10b.pdf
9. http://alumni.media.mit.edu/~jorkin/generals/papers/Kolodner_case_based_reasoning.pdf
10. https://www.researchgate.net/profile/Rosina_Weber/publication/225070215_Textual_Case-Based_Reasoning/links/5759ad5308aed884620b2481.pdf
11. Dialogue Management for Conversational Case-Based Reasoning
12. A Tutorial on Case-Based Reasoning
13. Reasoning with Textual Cases
14.