Artificial Intelligence (H) 2018-2019 - Assessed Exercise

Student first name:	*Chongbin*
Student surname:	*Ren*
Student id:	*2370060R*
Marker (first):	*FP*
Marker (second):	*BSJ*
Final score (out of 100):	*76*

Max possible marks [fixed]	Achieved marks [computed automatically]		Element		Awarded score for specific (sub)element [0-100]	Comments (if any):	Explanation/criterion
15	2,30	0,10	**Analysis**	Introduction and motivation	*10*	*Not provided but elements throughout.*	Does the report contain a suitable introduction and motivation relevant to the AE?
		1,60		PEAS analysis	*80*	*The analysis covers mainly the very basics without dealing with all the details that we'd normally expect in a complete analysis of the task env allthough some are considered later on.*	Basic and well-presented PEAS analysis, marks equally distributed among four aspects. P: Reward (/successrate)=25%, E=25%, A=25%, S=25%.
		0,60		Task Environment analysis	*5*		Is the analysis correct, complete and well-justified? Marks equally distributed among AIMA defined aspects (static., obs., known, seq., determ.). Agent dependent analysis weighted 5/10/85.
2	2,00	0,75	**Random Agent**	Design/method	*100*		Is the method well-justified (80) with reference to the analysis/requirements (20). Normally: We would expect only a brief comment saying that it is senseless and takes (uniform) random actions
				Description/theory	*100*		Is the description correct, complete and concise (=can you with reasonable means reproduce the results from the description and references)
		1,25		Agent Implementation (25% overall)	*100*		Is the implementation present, correct & executable (50), reflects design (20), well-structed (20), well-documented (10)
4	3,80	1,35	**Simple Agent**	Design/method	*90*	*Would benefit from a stronger justification of the method and explicit reference to an authoritative source of information.*	Is the method well-justified and suitable given the requirements and assumptions: optimal, complete, reaches goal ? Normally: we would expect to see A-star (or a std tree/graph search) with a reference (e.g. to AIMA book) and a mention of properties (optimal, consistent and applied heuristic)
				Description/theory	*90*		Is the description correct, complete and concise (=can you with reasonable means reproduce the results from the description and references). The report would need to mention consistency and optimality for full marks.
		2,45		Agent implementation (25% overall)	*98*		Is the implementation present, correct & executable (50), reflects design (20), well-structured (20), well-documented (10)
34	32,30	11,48	**Realistic Agent**	Design/method	*90*	*Would benefit from a stronger justification of the method and explicit reference to an authoritative source of information.*	Is the method well-justified and suitable given the requirements and assumptions: online, noisy transition model, learning, avoid dead-ends / cyclic behavior, has a chance of reaching goal (consistent), optimal. Normally we would expect to see RL learning method (Q-learning, ADP or other RL technique from the book); is a custom method is proposed we would need details about the specific method.
				Description/theory	*90*		Is the description correct, complete and concise (=can you with reasonable means reproduce the results from the description and references). Normally, we would expect a Q-learning variant, but if a custom method is suggested specific details would be required to reproduce it , whereas for a std textbook method a reference might be enough.
		20,83		Agent implementation (25% overall)	*98*		Is the implementation present, correct & executable (50), reflects design (20), well-structed (20), well-documented (10)
40	31,30	16,30	**Evaluation (note evaluation is subject to the 5/15/85 weighting)**	Strategy	*90*	*The figures are generally relevant. It is unclear if the last figure really displays percentages or rations (from the code, I'd suggest ratios). The presentation does not include the simple agent. Would have benefited from providing more insight/discussion into the results.*	Is the evaluation procedure well-defined, clearly explained (=reproducible) and suitable. Normally: We would expect to see a definition of the overall goal of the evaluation, the metric used and a discussion of the need to consider multiple problem_ids.
				Execution & Presentation	*80*		Is the evaluation well-executed and are the results presented in a suitable manner to demonstrate the performance of the agents also allowing comparison: table/figure (or clearly outlined in the text), relevant metrics (rewards or success rate must be considered; time/steps is also relevant) - weighted by the 5/10/85 rule. We would expect to see at least an evaluation and relevant performance plots (reward vs episode) on two problem_ids (=generalization). Potentially, added value in presenting other convergence plots (e.g. number of actions, time, etc). Graphs must follow best practices (e.g. labels, legends, etc.).
				Comparison	*85*		Does the report clearly compare the results from different agents in a sensible manner.
		15,00		Evaluation script	*75*	*The script contains the implementation of all the agents which is not ideal; you should really just be calling the agents from here instead of having duplicate code.*	Is the run_eval script included, executable and does it call the agents in a sensible manner. Does it reproduce (mostly) the results (tables/figures/text)/ presented in the report. Is it well-structured and well-documented.
5	4,00	4,00	**Discussion & Conclusion**		*80*		Does the report contain a suitable discussion of the results and limitations and a critical reflection on own work or the assignment as a whole.
Total:	*76*						

需要对每个agent的static,known,分析

Overall Comments (if any):