**Question A:**

1. (15 pts) You will be designing the control system for a robot that will be sent into an unknown static environment in which there are features that are likely to result in the robot's destruction if a poor sequence of actions are taken (e.g., very steep inclines that are difficult for the robot to recover from).  Using a table, explain the advantages and disadvantages of the schemes discussed in this course for this problem.
2. (5 pts) Explain which of the above is most appropriate to the problem, and why.
3. (15 pts) Redo the analysis of (1) for a dynamic environment.
4. (5 pts) Explain which of the above is most appropriate to the dynamic variant problem, and why.

**Question B:**

1. (15 pts) You will be designing the control system for a robot that will be sent into an unknown static environment, in which there are features that will not result in the robot's destruction if a poor action is taken (e.g., no steep inclines that the robot can not correct for).  Using a table, explain the advantages and disadvantages of the schemes discussed in this course for this problem.
2. (5 pts) Explain which of the above is most appropriate to the problem, and why.
3. (15 pts) Redo the analysis of (1) for a dynamic environment.
4. (5 pts) Explain which of the above is most appropriate to the dynamic variant problem, and why.

**Question C:**

1. (15 pts) You will be designing the control system for an agent that needs to successfully play a video game; the scheme will be used in a competition that will be held sometime in the future (you have access to the video game prior). Using a table, explain the advantages and disadvantages of the schemes discussed in this course for this problem.
2. (5 pts) Explain which of the above is most appropriate to the problem, and why.
3. (15 pts) Redo the analysis of (1) for the case where you do not have access to the video game prior to the competition.
4. (5 pts) Explain which of the above is most appropriate to the problem, and why.

**Question D:**

1. (5 points) Using python, write a function that shows how to compute the argmax for an action-value function when using tabular methods.
2. (5 points) Using python, write a function that shows how to compute the argmax for an action-value function when using function approximation methods.
3. (10 points) Why is it difficult to use function approximation methods when the action space is continuous, and why does this problem does not appear when you have a continuous state space?

**Question E**:

(10 points) Clearly explain why Q learning is considered to be an off-policy scheme when there is no reference to distinct pi and b policies?

**Question F:**

1. (10 points) Clearly explain how exploration arises in the various schemes that we have covered in this course. Use a table.
2. (5 points) Assuming that human beings learn by reinforcement, use RL concepts to explain why practical experience via challenging assignments is essential in effectively learning a new topic (... hopefully this will make clear why the heavy assignments and exercises were required in APS1080, regardless of weighting ...).