Homework01

Task 01

Sate description
8x8 fields where each of the game characters can be placed on

Actions
all actions that each character can make

Reward
positive or negative reward for winning or loosing (has to be higher than beating all other characters), +1 for defeating a character, -1 for losing own character

Policy
maximizing reward

Task 02

State description
There are 8 states: the coordinates of the lander in x & y, its linear velocities in x & y, its angle, its angular velocity, and two booleans that represent whether each leg is in contact with the ground or not.

Actions
Boosting each of the three engines

Policy description
chooses the action based on the current state

Task03

The state transition function contains all moves that can and are allowed to be made to get from a state s to s'.
For example if we imagine a 10*10 pixel board and have a character with position x, y that wants to walk towards a goal the state. The character is only allowed to move one step at a time so all the allowed state transitions would be x+1, y| x-1 ,y | x ,y +1 | x, y-1
An other example for a state transition function would be the rules of chess. Every piece can only be moved according to its movement rules rules, aka the pawn can only move one step in front unless it is at its baseline in which case it can move 2 forward etc.

The reward function describes how the transition from s to s' changed the overall evaluation of the state. Examples for a simple reward function could be in a labyrinth the distance from the goal that is reevaluated in every step.
A more complex example could be the moves left for mate in a chess game. This of course is not really practicable since calculating this takes an absurd amount of computational power

Environment dynamics:
In most cases we could know the environment dynamics, but if we have complex problems this is not practical. For example a self driving car would need everything that affects its functionality to be able to make a state transition function. The amount the car steers to the left given a move of the steering wheel could be impacted by wind, humidity, the temperature or the road etc. As you can clearly see there are too

many parameters to create a perfectly accurate state transition function, although it technically exists. However there are also some situations where we don't think that it is possible to create a state transition function. Examples for this would be the quantum mechanics or radioactive decay. We just simply cant predict what will happen next. We can give a probability but we can not be sure that the next state will occur