Current state: X Y Z

Problem: The agent gets in a position where there are several possible moving actions. After a sequence of actions, one of the previous possible actions becomes unavailable ( a free space is now occupied by a block placed by the player ). The agent learns this and in all future episodes he thinks there is a block there, forcing him to choose from other actions without regarding the possible action of moving into that position.

Proposed new state: X Y Z + 3x3x3 observation(optimise by removing unnecessary blocks)

Reason: When the agent is in the same X Y Z, the agent is also aware of the blocks surrounding him and will act based on an entry in the Q table that corresponds to the blocks surrounding him. This means that the same position with a block in front is different state than the same position without a block in front which allows different actions and therefore different behaviour.