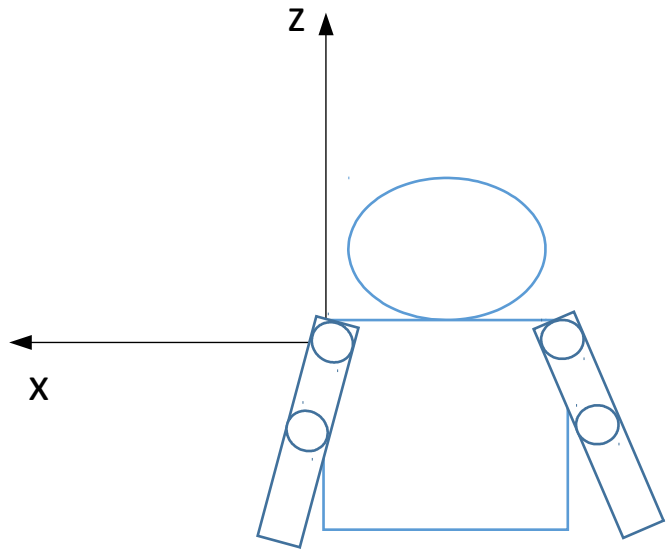
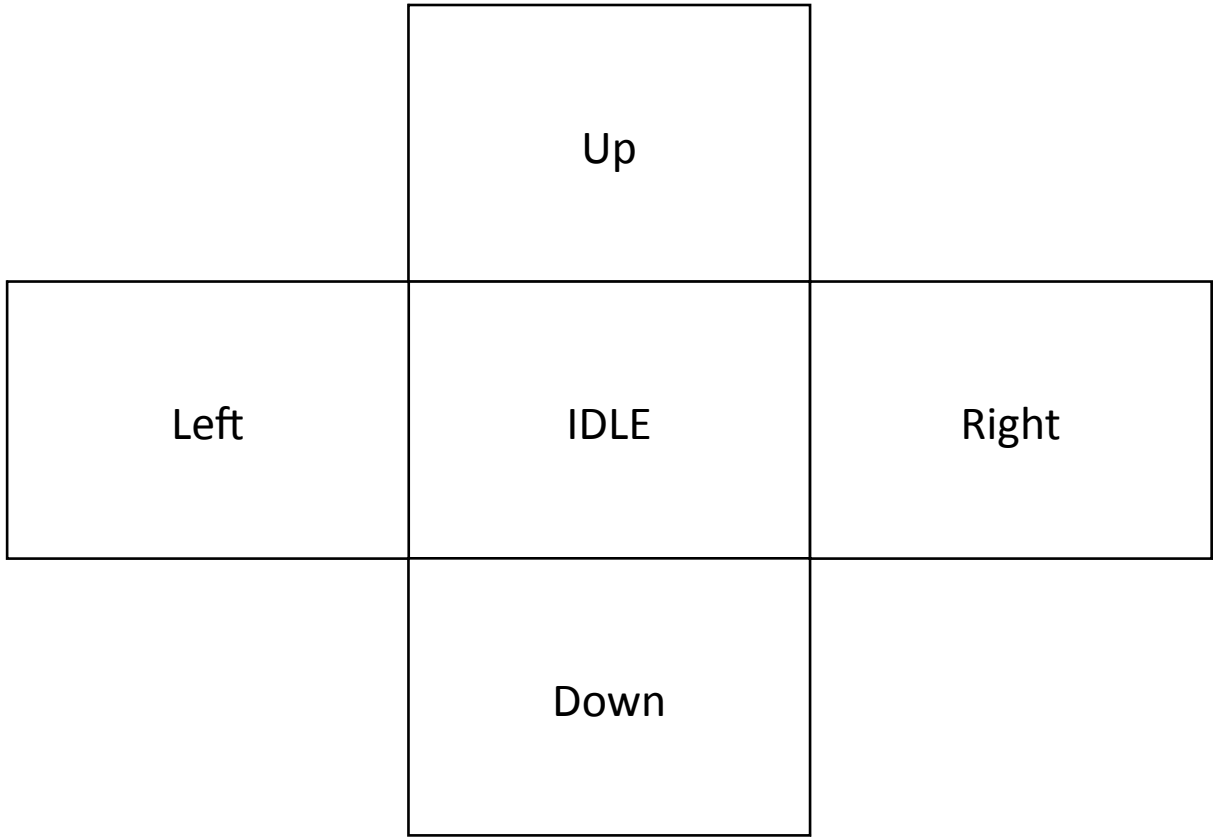
# Nao Q-Learning Project

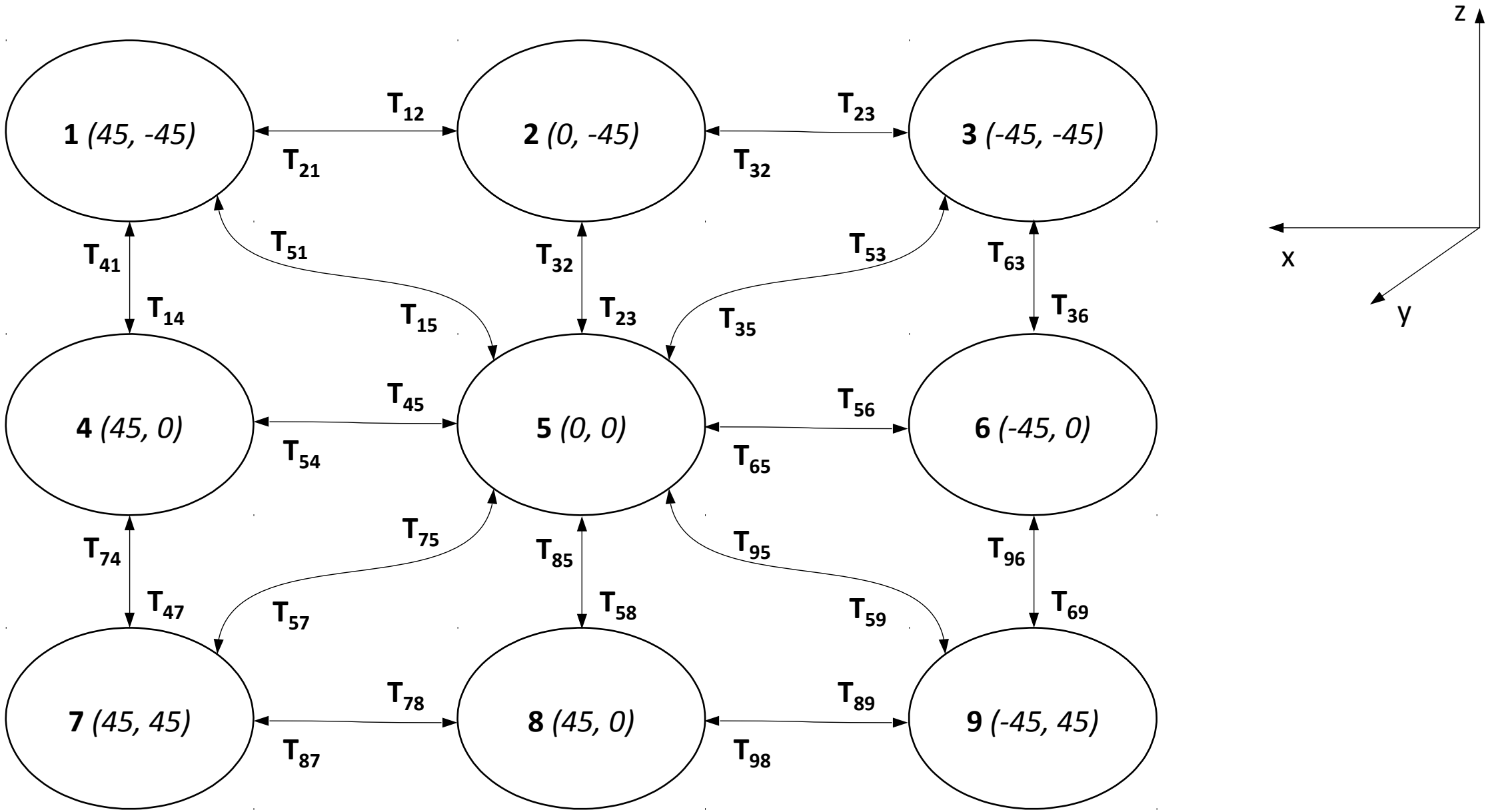
## States and Actions




- The space is split into 9 states determined by the robot's arm's elevation and azimuth.

**Five actions :**




1 (45, -45)
2 (0, -45)
3 (-45, -45)
4 (45, 0)
5 (0, 0)
6 (-45, 0)
7 (45, 45)
8 (45, 0)
9 (-45, 45)
T12
T21
T23
T32
T41
T14
T51
T15
T32
T23
T53
T35
T63
T36
T45
T54
T56
T65
T74
T47
T75
T57
T85
T58
T95
T59
T96
T69
T78
T87
T89
T98
z
x
y

**Policies** :

- For picking actions, several policies are to be implemented. The first Matlab script picks a possible action randomly, but the choice can be based on values of the Q matrix.

**Rewards** :

- Rewards are to be indicated at each step by the human partner, through signs which are to be identified by the robot using for instance image recognition or through the voice.
If the robot brings his arm closer to the «goal state », the user rewards the robot, and if it brings it farther, he penalizes it.