

Robot remembers path of current episode

Sensor failures checked on each new "step"

- Each (i.e. false positive/false negative) have a 10% chance of failure

*The agent should use the following rules to navigate:*

1. *If any of the grid positions in its neighborhood refers to an item ordered, then it moves to that grid position. If two or more neighboring positions contain ordered items, then the tie is broken by making a random choice between the positions involved.*
2. *If none of the neighboring positions contain ordered items, then a random choice is made to move to the next node.*

*Each visit to a position in a grid that does not contain an ordered item incurs a score of -1, while each visit to a position containing an ordered item scores +3. Thus a brute force algorithm that sweeps through the grid from the starting position P will score a total score of  $(35-n)*(-1)+n*3=-35+4*n$  where n is the number of items ordered.*

Stock never runs out

Number of any given item type is irrelevant (e.g. it gets however many it needs of "A" while it is stopped at the shelf)

Assume robot restarts from [0, 0]

- Although this could be changed easily by not resetting its position at the beginning of a new episode

Shortest and longest path -> highest scoring and lowest scoring path