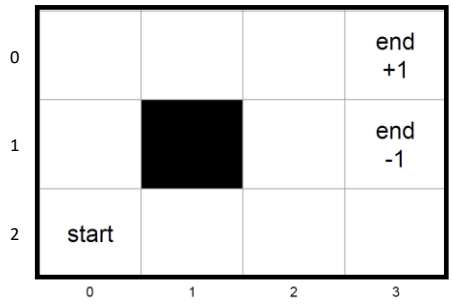
Workshop Lecture 3

Part A: Continuing with the simple maze

In the Blackboard folder for this week, you will see the python script "`lecture3-simulation.py`" under the workshop materials. Save locally, and run with (in terminal, `cd` to location of script):

```
python lecture3-simulation.py
```

This script provides the same simple maze as in week 2 (class `State`): you are encouraged to tweak in whatever way you see fit to improve visualisation, usability, etc. There is one 'agent' implemented: an `IRLAgent` class that provides a basic mechanism for a human to provide a reward after every action that the agent performs in the maze environment.



Part B: Investigating the IRL Agent

The `IRLAgent` requires the human supervisor to provide a reward signal/feedback after every action has been executed and the state subsequently updated – what is this reward signal? Try out this mechanism, and see what the behaviour looks like – can you explain this behaviour? How does this change over the course of multiple episodes? What *load* is there on you as the human supervisor?

Compare the performance of this new interactive learning agent with the performance of the basic learning agent from week 2. Consider the following potential extensions `IRLAgent`:

- As in week 2, consider whether the learning algorithm currently implemented is suitable for the task/environment at hand.
- Instead of merely providing positive/negative reward for the present state, how could you explicitly guide which action the agent should take next? (*Note that this may require a change of learning algorithm*). This is a starting point for the lecture/workshop content in week 4.