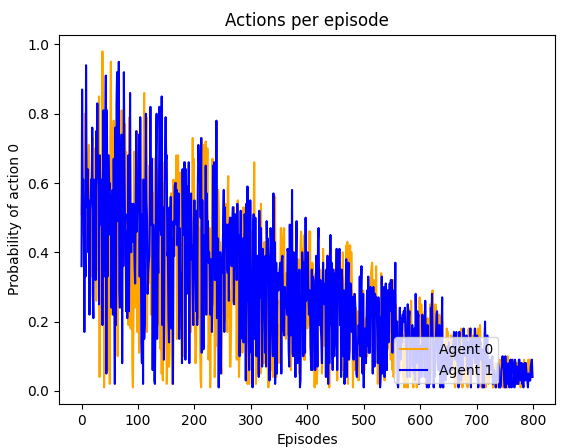
**Prisoner**, with auto set for nash target on action 1(cooperate) of 0:



**Prisoner**, with manual set target on action 1(cooperate) of 1:

A graph of a chart

Description automatically generated with medium confidence

**Prisoner**. Using Softmax, no recourse:

A graph of a chart

Description automatically generated with medium confidence

**Matching Pennies**, with recourse:

A graph of a blue and orange line

Description automatically generated

Agent 0 final actions: [0.49800579 0.50199421]

Agent 1 final actions: [0.53024114 0.46975886]

**Matching Pennies**, with Softmax:

A blue and orange barcode

Description automatically generated

Agent 0 final actions: [0.08349227 0.91650773]

Agent 1 final actions: [0.04797961 0.95202039]

**NOTE:** Leans towards tails.

**Conclusion:**

Softmax is great for NFG like stag and prisoner, when you want them to learn a strategy which is Nash. On prisoner for example, if you want agents to learn a strategy which is Nash, both agents defect, then go with Softmax. Convergences to desired policy will occur with much less episodes. Training for agents will be shorter. However when you may want agents for learn a different strategy or a Nash which involves a mixed strategy then Softmax is not much use. This is where recourse can help.