AI Poker Overview/ Structure Map

ENVIRONMENT

Observation Space:

player hand strength
$ on table
$ remaining
round type
...

Action Space

fold
raise
...

state: (hand strength, $ on table, $ remaining)

step from initialization
based on rules of game

MAIN

create episodes

episode: (state, action, reward)

(s1, a1, r)
(s1, a2, r)
(s2, a1, r)
(s1, a1, r)
(s1, a1, r)
...

100 instances by default
...break if....
everyone folds,
everyone checks

episode: (state, action, reward)

episode: (state, action, reward)

episodes
(static user input)

choose action
by probability
(semi-exploratory)

train

for each episode....

for each state-action pair....

(s1, a1)
(s1, a2)
(s2, a1)
(s1, a1)
(s1, a1)
...

average r's

(running
average
from all
episodes)

look up reward for
state and action

update / train!

POLICY

Q nested dictionary
state --> (action --> value)

Observation
(state)

look up

s1 --> a1 --> reward
a2 --> reward
....
s2 --> a1 --> reward
a2 --> reward
...

increase probability of action
with highest reward

Action Probabilities

E +
E
...

epsilon is a
static user input