

Vegas (WIP) Presentation

Elliott Brossard
Kellen Donohue
Zach Stein

Aspirations & reality

Aspirations



Reality



Vegas

- Blackjack simulation & AI in *python*
- Support for all blackjack rules & betting
- Human, reflex (from Wikipedia), q-learning, value-iteration, and always standing agents

State

- Comprises three different values
 - Hard count of hand
 - Whether player's hand contains an ace
 - Value of dealer's visible card
- This makes for a relatively small state space
- Nothing yet on counting, but that will come later

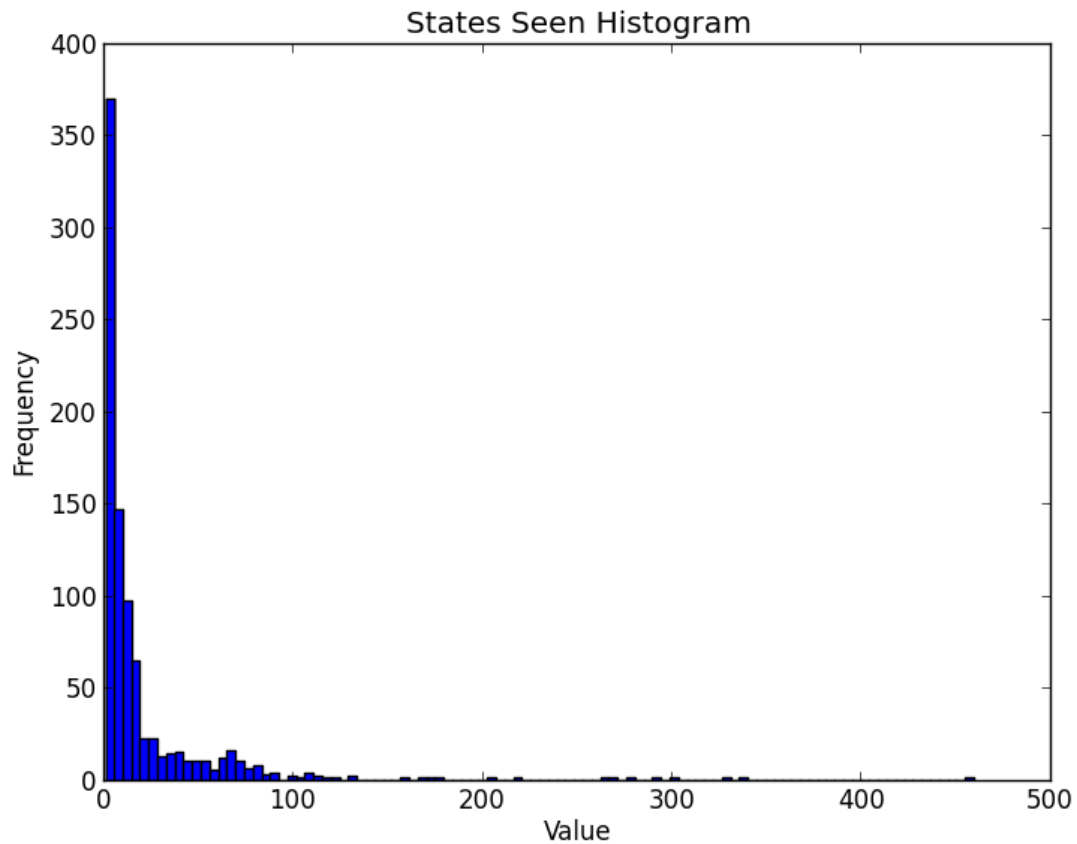
Demo

What did you learn

- Lots of states don't get seen very often
- Always standing is not as bad a strategy as you'd think
- State modeling is important
- With value iteration modeling using infinitely many decks you could get more accurate results in some cases, but can't do counting

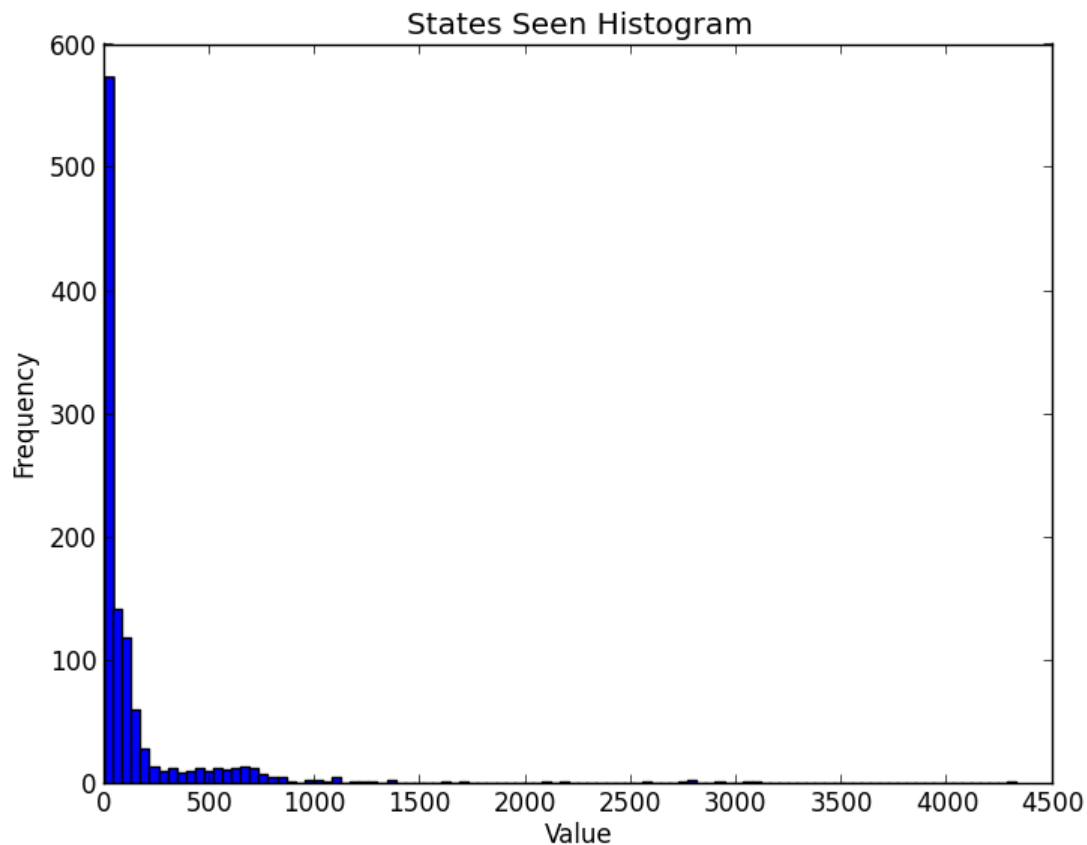
States seen histogram

states seen: 886 (10,000 training rounds)



States seen histogram

states seen: 1092 (100,000 training rounds)



239 states visited during the 1000 testing rounds never seen during training!

Experiments

Comparing our AI agents to

- Agent that always stands
- Blackjack strategy tables
- QLearning with counting vs. value iteration

Agent policies

ReflexAgent (strategy from Wikipedia):

Player hand	Dealer's face-up card										
	2	3	4	5	6	7	8	9	10	A	
Hard totals (excluding pairs)											
17-20	S	S	S	S	S	S	S	S	S	S	S
16	S	S	S	S	S	H	H	SU	SU	SU	SU
15	S	S	S	S	S	H	H	H	SU	H	H
13-14	S	S	S	S	S	H	H	H	H	H	H
12	H	H	S	S	S	H	H	H	H	H	H
11	Dh	Dh	Dh	Dh	Dh	Dh	Dh	Dh	Dh	H	H
10	Dh	Dh	Dh	Dh	Dh	Dh	Dh	Dh	H	H	H
9	H	Dh	Dh	Dh	Dh	H	H	H	H	H	H
5-8	H	H	H	H	H	H	H	H	H	H	H

Soft totals											
	2	3	4	5	6	7	8	9	10	A	
A,8-A,9	S	S	S	S	S	S	S	S	S	S	
A,7	S	Ds	Ds	Ds	Ds	S	S	H	H	H	
A,6	H	Dh	Dh	Dh	Dh	H	H	H	H	H	
A,4-A,5	H	H	Dh	Dh	Dh	H	H	H	H	H	
A,2-A,3	H	H	H	Dh	Dh	H	H	H	H	H	

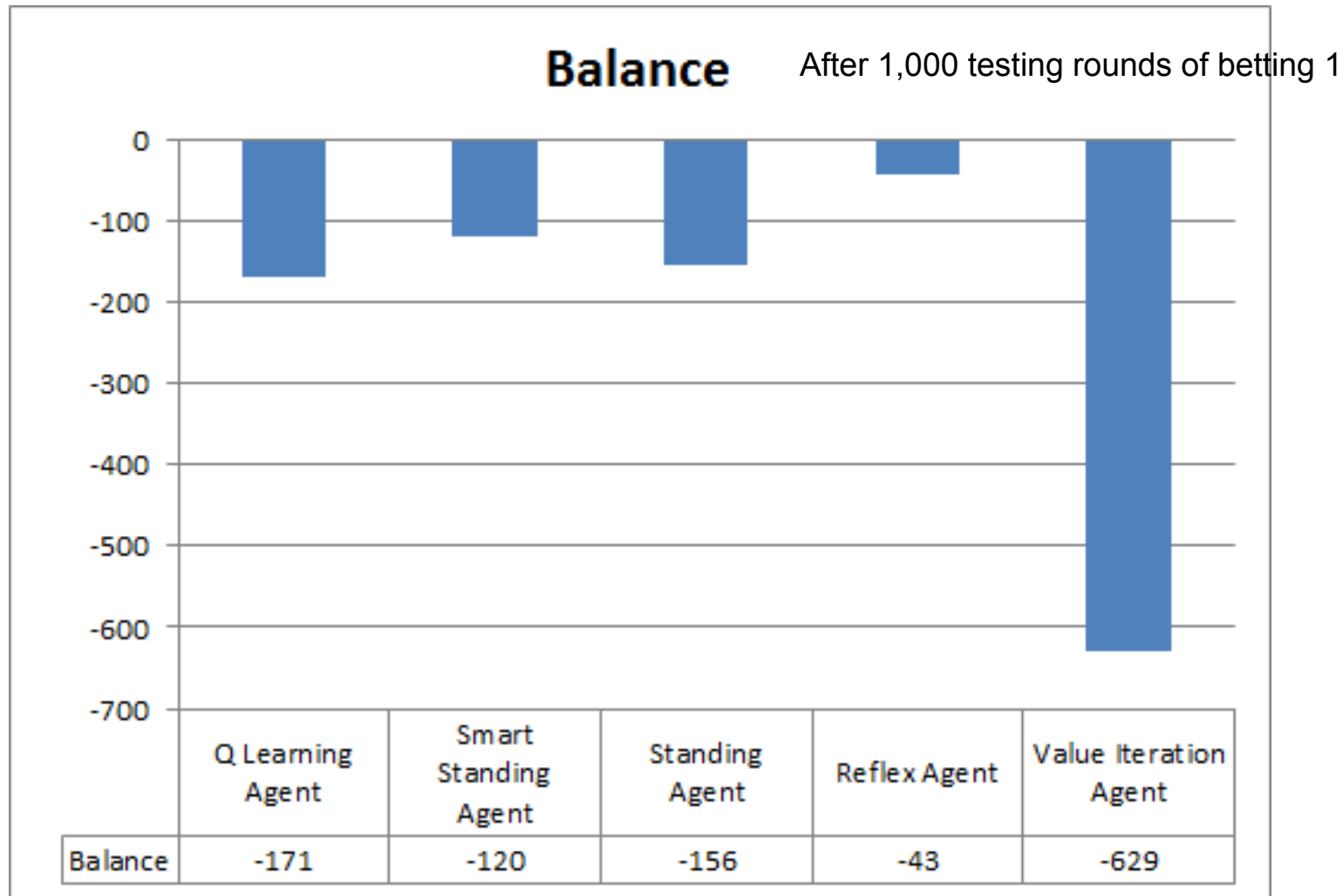
Pairs											
	2	3	4	5	6	7	8	9	10	A	
A,A	SP	SP	SP	SP	SP	SP	SP	SP	SP	SP	
10,10	S	S	S	S	S	S	S	S	S	S	
9,9	SP	SP	SP	SP	SP	S	SP	SP	S	S	
8,8	SP	SP	SP	SP	SP	SP	SP	SP	SP	SP	
7,7	SP	SP	SP	SP	SP	SP	H	H	H	H	
6,6	SP	SP	SP	SP	SP	H	H	H	H	H	
5,5	Dh	Dh	Dh	Dh	Dh	Dh	Dh	Dh	H	H	
4,4	H	H	H	SP	SP	H	H	H	H	H	
2,2-3,3	SP	SP	SP	SP	SP	SP	H	H	H	H	

Agent policies

Value iteration agent

- Bad values right now
- We're going to change rewards

Agent Comparison



Plans for remaining week

- Value iteration improvement (new state representation)
- Improving q learning exploration coverage
- Counting
 - Traditional: +1 for face cards, -1 for low cards
 - More complicated state knowledge?
- Additional knowledge?
 - Humans sometimes get this -- seeing dealer's other card, marked cards

Who did what

- **Elliott**
 - Deck, cards, game logic
 - Modeling as MDP, value iteration
 - Betting, doubling, splitting
 - Human agent, reflex agent
- **Kellen**
 - Hand, game state, game logic
 - Policy analysis
 - Reflex agent
 - Presentation slides
- **Zach**
 - Win-loss-tie tracking, standing agent
 - Q-learning, policy analysis

Questions