CS 6300 **HW06: Q Learning with Functional Approximation** **Due Feb 26**

Please use the LaTeX template to produce your writeups. See the Homework Assignments page on the class website for details. Hand in through gradescope.

1 Functional Approximation

We revisit the simplied version of blackjack from Homework 5. The deck is infinite and the dealer always has a fixed count of 15. The deck contains cards 2 through 10, J, Q, K, and A, each of which is equally likely to appear when a card is drawn. Each number card is worth the number of points shown on it, the cards J, Q, and K are worth 10 points, and A is worth 11. At each turn, you have two possible actions: either *hit* or *stay*.

Unhappy with your experience with basic Q-learning, you decide to featurize your Q-values. Consider the two feature functions:

$$f_1(s,a) = \begin{cases} 0 & a = stay \\ +1 & a = hit, s \geq 15 \\ -1 & a = hit, s < 15 \end{cases} \quad \text{and} \quad f_2(s,a) = \begin{cases} 0 & a = stay \\ +1 & a = hit, s \geq 18 \\ -1 & a = hit, s < 18 \end{cases}$$

Which of the following partial policy tables may be represented by the featurized Q-values unambiguously (without ties)? Derive your answers for each policy table.

s	$\pi(s)$
14	hit
15	hit
16	hit
17	hit
18	hit
19	hit

(a)

s	$\pi(s)$
14	stay
15	hit
16	hit
17	hit
18	stay
19	stay

(b)

s	$\pi(s)$
14	hit
15	hit
16	hit
17	hit
18	stay
19	stay

(c)

s	$\pi(s)$
14	hit
15	hit
16	hit
17	hit
18	hit
19	stay

(d)

s	$\pi(s)$
14	hit
15	hit
16	hit
17	stay
18	hit
19	stay

(e)