MiniMax and Alpha-Beta Pruning
----

1. Alpha beta pruning can greatly reduce the running time of the min-max algorithm. A best case analysis shows a time complexity of $O(b^{d/2})$ where *b* is the branching factor and *d* the tree depth. Under what conditions is this best case situation obtained.

2. Consider adversarial search in a *three-player game*. Assume that no alliances are allowed. We will call the players A, B, and C for convenience. The first change is that the evaluation function will return a list of three values indicating (say) the likelihood of winning for players A, B, and C, respectively. Complete the following game tree by filling out the backed-up value triples (i.e. find correct values for the "??") for all remaining nodes, including the root. This question is worth 8 points.

```
to move

A                                     (?? ?? ??)

                                /                    \
                               /                      \

B                  (?? ?? ??)                                 (?? ?? ??)

                  /          \                               /          \
                 /            \                             /            \

C       (?? ?? ??)            (?? ?? ??)            (?? ?? ??)            (?? ?? ??)

        /        \            /        \            /        \            /        \
       /          \          /          \          /          \          /          \

A (+1 +5 +3) (+7 +2 +1) (+2 +2 +2) (+7 +6 -1) (+5 -2 -1) (-5 +3 +2) (+8 +7 -1) (+5 +7 +5)
```

Q- Learning
----

1. In Q learning a reward maybe obtained after each action. What utility is the Q learning agent attempting to maximize. Write the form of this utility function.

2. The standard Q-learning algorithm is shown below:

$$
\begin{array}{l}
\forall s, \forall a : \hat{Q}(s,a) \leftarrow 0 \\
\text{Loop Forever} \\
\quad s \leftarrow \mathit{PerceiveState}() \\
\quad a \leftarrow \mathit{SelectAction}(Q) \\
\quad r \leftarrow \text{Reward}() \\
\quad s' \leftarrow \mathit{PerceiveState}() \\
\quad \hat{Q}(s,a) \leftarrow r + \gamma \max_{a'} \hat{Q}(s',a') \\
\text{EndLoop}
\end{array}
$$

Figure 1: Q-Learning Algorithm

a) The Q-learning algorithm guarantees two formal properties regarding $\hat{Q}_n(s,a)$ where the subscript n refers to the n^{th} "sweep" of all state action pairs. State these two properties.

3. Consider using Q-learning to allow a robot to play black-jack aka 21.

The object of blackjack is to get a total value of cards as close to 21 as possible without going over and beating your opposition's total. Numbered cards are worth their face value. Face cards -- Kings, Queens and Jacks -- are each worth 10. Aces are worth either 1 or 11, whichever is better in the circumstances.
If the total value of your cards exceeds 21, you "bust", which means you lose. If your total is less than or equal to 21 and the dealer's total is over 21, then the dealer busts and you win. If neither your total or the dealer's total exceeds 21, then the higher total wins. In the event of a tie, no one wins and no one loses. A tie is called a "push". The dealer deals two cards to you and herself/himself. These cards are dealt face-down. Now you have to make a decision: do you take another card ("hit") or pass to the dealer ("stand" or "stay"). If you hit, the dealer gives you another card and again asks if you want to hit or stand. You keep hitting until you are satisfied with your total or you bust. The dealer has the same options as you.

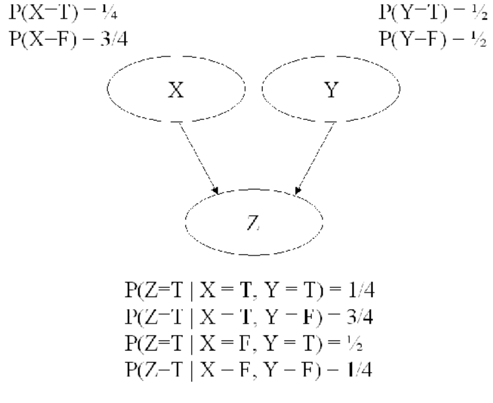
a) Formulate playing black-jack as a Q-learning problem by listing the states, actions, rewards and discount factor?
b) One limitation of Q-learning that this proof obscures is that the agent must somehow stumble across a goal state. Suggest a principle method to overcome this limitation for blackjack?
c) What is another requirement of the Q-learning proof of convergence?

Belief Networks
----

1. Consider a belief network over m nodes $X_1 \ldots X_m$. We can apply the global semantic to obtain an expression for the joint distribution $P(X_1 \ldots X_m)$.

a) Write the expression to compute $P(X_1 = T \ldots X_i = T,\ X_{i+2} = T \ldots X_m = T)$.

2. Consider the belief graph:



a) What relationship do X and Y have given Z ?

b) For the graph on the left, calculate P(Z = F)?

c) If the graph was cyclic, what form would the answer to question b) take?

d) Name two type of reasoning that can occur in a belief network.

e) If Z occurs and we know X occurs what should happen to our belief in Y.