Homework: $n-$armed bandit and Matlab

First, be sure that you have downloaded the code for the three techniques we have discussed in class:

- $\epsilon-$greedy
- "Softmax"
- "Win-Stay, Lose-Shift" (or Pursuit). For this one, you'll need to copy it or type it in from the course notes.

Run the two script files for $\epsilon-$greedy, `banditScript01.m` and `banditEplot.m`.

Run the two script files for softmax: `softmaxScript01.m` and `softmaxScript02.m`.

Be sure to put the files together for the last technique and check by running the code in Matlab (if you have no errors, then you've done it correctly!).

For the following, write down the changes you would make to the script files to do the following:

1. Change the actual payouts to the following vector of payouts: $[-2, -1, 1, 2, 3]$.
2. How might we test the effectiveness of each of the algorithms, in order to decide which one is "best"?
3. Explain how we go about "choosing machine a with probability $P(a)$". That is, given a vector of probabilities in P, what do we do?
4. Explain what the role of τ is in the algorithm for softmax.
5. If we want to go from some intial value of τ to some final value of τ in N steps, what formula might be best to use?
6. Would it be difficult to change the code for $\epsilon-$greedy so that ϵ starts high and goes low? Think about what would need to change (and why we might like to do that).
7. What is the role of the parameter β in the pursuit strategy?