Play Selection in American Football

Daniel Bestard, Michael Cameron, Hans-Peter Höllwirth, Akhil Lohia

April 4, 2017

1 Motivation

In the following sections we assume some basic understanding of the American football rules.

2 The Football Model

In our American football model we consider the score of one offense drive and add to it the expected points gained by the opposing team from our final field position. More concretely, we want to maximize the expected difference between our drive-score and the opposing team's responding drive-score (which is simply a function of our final field position).

The state of the system $i \in S$ is described by a vector of 3 quantities: $i = [x, y, d]$:

- x = the number of yards to the goal (discrete value between 1 and 100)

- y = the number of yards to go until the next first down (discrete value between 1 and 100)

- d = down number ($\in \{1, 2, 3, 4\}$)

At each state, the team can choose from one of 4 play options (actions) $u \in U$ with $U = \{R, P, U, K\}$:

- (R)un: moves $D_p - 2$ yards forward with $D_p \sim \text{Poisson}(6)$

- (P)ass: moves $D_p - 2$ yards forward with $D_p \sim \text{Poisson}(12)$

- P(u)nt: moves $6D_p + 6$ yards forward with $D_p \sim \text{Poisson}(10)$

- (K)ick: successful with probability $\max(0, .95 - .95x/60)$

The set of state-action pairs determine the stationary policy μ. The reward of the drive is determined by the final state transition:

- Touchdown: 6.8 points (from run or pass)

- Opponent's touchdown: -6.8 points (from run or pass)

- Safety: -2.0 points (from pass or run)

- Field goal: 3.0 points (from kick)

- No score due to fumble (from run), interception (from pass), missed 4th down (from pass or run), missed field-goal (from kick) or punt

Fixed probabilities are assigned to the outcome of each action.

3 Dynamic Programming Formulation

3.1 Optimal Solution

3.2 Approximation with Neuro-Dynamic Programming

4 Simulation

The approximations of the reward-to-go function $J_{u_k}(i^*)$ are obtained from simulated sample trajectories. The simulation routine takes as inputs a policy μ and a starting state i^* and then generates N sample drives. Each generated sample drive represents a realization of the following probabilistic model (probabilities in brackets):

- (R)un attempts result in either

    - (0.05) fumble
    - (0.95) run with movement $D_r - 2$ with $D_r \sim \text{Poisson}(6)$

- (P)ass attempts result in either

    - (0.05) pass interception with movement $D_p - 2$ with $D_p \sim \text{Poisson}(12)$
    - (0.05) sack with movement D_s with $D_s \sim \text{Poisson}(6)$
    - (0.45) pass incomplete
    - (0.45) pass complete with movement $D_p - 2$ with $D_p \sim \text{Poisson}(12)$

- P(U)nt attempts always result in turn-over with movement $6D_p + 6$ with $D_p \sim \text{Poisson}(10)$

- (K)ick attempts result in either

    - (max(0,(0.95-0.95x/60))) successful field goal
    - (otherwise) missed field goal

The drive terminates in any of the following events:

The simulation returns N sample drives. Each drive l is described by its status sequence (i_t^l) for $t = 1, ..., T^l$ and reward $g_{T^i}^l$.

4.1 Expected Reward

Given a set of N simulated sample trajectories for a specific policy μ, we can estimate the expected reward of this policy from a starting state i^*:

$$\widetilde{J}_{u_k}(i^*) = \frac{1}{N} \sum_{l=1}^{N} g_{T^i}^l$$

where $g_{T^i}^i$ denotes the reward of the ith sample trajectory with drive length T^i.

4.2 Heuristic Benchmark

Rhe simulations and the expected reward function can be used to yield a heuristic bench-mark for the optimal play-selection policy. We use the suggested heuristic policies from the paper in order to establish the correctness of the simulation algorithm. In particular, we consider the following heuristic policies:

1. If $d = 1$ (first down): (**P**)ass

2. If $d = 2$ (second down):

    (a) If $y < 3$ (less than 3 yards to next first down): (**R**)un
    (b) If $y \geq 3$ (3 or more yards to next first down): (**P**)ass

3. If $d = 3$ (third down):

    (a) If $x < 41$ (less than 41 yards to goal):
        i. If $y < 3$ (less than 3 yards to next first down): (**P**)ass or (R)un
        ii. If $y \geq 3$ (3 or more yards to next first down): (**P**)ass or (R)un
    (b) If $x \geq 41$ (less than 41 yards to goal):
        i. If $y < 3$ (less than 3 yards to next first down): (P)ass or (**R**)un
        ii. If $y \geq 3$ (3 or more yards to next first down): (**P**)ass or (R)un

4. If $d = 4$ (forth down):

    (a) If $x < 41$ (less than 41 yards to goal):
        i. If $y < 3$ (less than 3 yards to next first down): (P)ass or (**R**)un or (K)ick
        ii. If $y \geq 3$ (3 or more yards to next first down): (P)ass or (R)un or (**K**)ick
    (b) If $x \geq 41$ (less than 41 yards to goal):
        i. If $y < 3$ (less than 3 yards to next first down): (P)ass or (**R**)un or P(U)nt
        ii. If $y \geq 3$ (3 or more yards to next first down): (P)ass or (R)un or P(**U**)nt

We estimated the expected reward for each of the $2^4 * 3^4 = 1296$ heuristic policy combinations from starting position $i^* = (x = 80, y = 10, d = 1)$ (one of the most likely starting positions in football) and found the best heuristic expected reward-to-go $J_\mu(i^*) = -1.26$. The associated policy to this reward is highlighted in bold. Note that the reward-to-go matches the heuristic result of the underlying paper, thus establishing the correctness of the simulation algorithm.
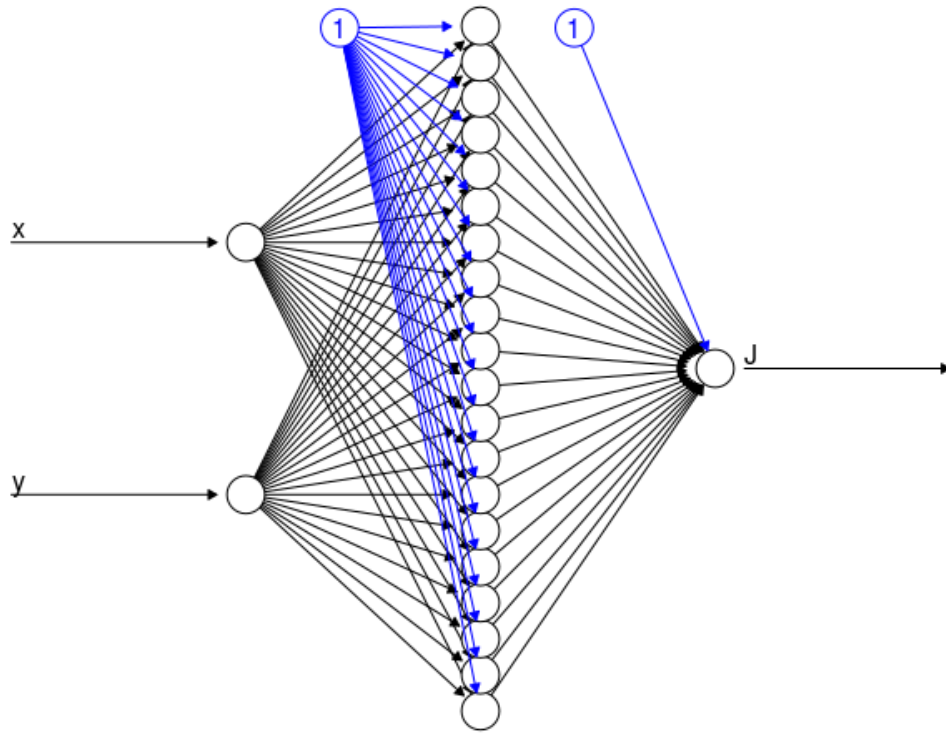
Figure 1: Architecture of neural network with one hidden layer ($R = 20$ nodes)

5 Approximate and Optimistic Policy Iteration

5.1 Multilayer Perceptron (MLP)

5.2 Policy Update

6 Results