CAUTION: These lecture notes are under construction. You may find parts that are incomplete.

7 PLAYER EVALUATION AND IN-GAME STRATEGY

From the previous chapter, we have a way to calculate the expected value $v(s)$ of any game state $s \in \mathcal{S}$, using value iteration. In this chapter, we discuss two applications of this expected value. The first application is player evaluation, in which we assign credit to players for the change in expected value on the plays for which they are responsible. The second application is in-game strategy, in which we choose the strategy which yields the highest expected value.

7.1 PLAYER EVALUATION

Generally, measuring player contributions to team success in team sports is a challenging but important task for front offices that need to make decisions on player trades and free agent contracts. One commonly used strategy is to isolate individual actions taken by players and to measure the expected value of the game state before and after the action. This tends to work well in highly structured sports like baseball and football and less well in less structured sports like basketball and soccer.

Suppose that the game is in state s and that an action taken by a player causes the game to transition to state s'. On this transition, the expected value has updated from $v(s)$ to $v(s')$, with a reward of $r(s, s')$ on the play. If we believe that this state transition is the sole responsibility of the player who took the action, then we may credit this player with the difference $r(s, s') + v(s') - v(s)$ in expected value. Across a set of such plays, we generally calculate the average change in state value as the player's performance. If a player has n actions with $\{s_1, ..., s_n\}$ denoting the game state prior to each action and $\{s'_1, ..., s'_n\}$ denoting the game state after each action, then the player's contribution to the expected outcome is

$$\frac{1}{n} \sum_{i=1}^{n} (r(s_i, s'_i) + v(s'_i) - v(s_i)).$$

7.1.1 EXAMPLES

In baseball, we often model the game state of an inning as $s \in \{0, 1\} \times \{0, 1\} \times \{0, 1\} \times \{0, 1, 2\}$, denoting which of the three bases are occupied and how many outs there are. In addition to the 24 non-terminal base-out states, there are four terminal end-of-inning states $s = \in \{0, 1, 2, 3\}$, with r denoting the number of runs scored on the final play of the inning. We use the reward function $r(s, s')$ to denote the number of runs scored on the transition from state s to state s', and the value function $v(s)$ describes the expected number of runs scored starting from the state s until the end of the inning. Suppose that a batter steps up to the plate with the bases loaded and two outs, i.e. $s = (1, 1, 1, 2)$. If the batter strikes out and no runs score, then the state transitions to $s' = (0)$, and we credit the batter with the change in expected value:

$$r((1, 1, 1, 2), (0)) + v((0)) - v((1, 1, 1, 2)) = -v((1, 1, 1, 2)).$$

Measuring batter performance with this method is known as RE24 in public baseball analytics work.

In football, we often model the game state of a drive using the yard line, the down, and the distance to go for a first down, i.e. $s \in \{1, ..., 99\} \times \{1, 2, 3, 4\} \times \{1, ..., 99\}$. For example, a first-and-10 from a team's own 20 yard line would be denoted by $s = (20, 1, 10)$. In addition to the non-terminal states, there are six terminal end-of-drive states $s \in \{-2, 0, 3, 6, 7, 8\}$ representing the number of points scored on the drive (–2 denoting a safety and 2 points for the defensive team). In this case the reward function $r(s, s')$ is zero except for transitions from non-terminal states to terminal states, in which case the reward is equal to the value of the terminal state. If a first-and-10 play from a team's own 20 yard line results in a sack and a loss of 5 yards, then the change in expected value on this play is:

$$r((20, 1, 10), (15, 2, 15)) + v((15, 2, 15)) - v((20, 1, 10)) = v((15, 2, 15)) - v((20, 1, 10)).$$

This quantity $v(s') - v(s)$ is known as EPA (Expected Points Added) in public football analytics. Attribution to individual players is less clear than in baseball because multiple players on both sides of the ball are involved in every play.

7.2   IN-GAME STRATEGY

When a coach/manager faces a tactical decision, they do so starting from a game state s and corresponding value $v(s)$. Their decision might lead deterministically to a new state s' and corresponding value $v(s')$, or it might lead stochastically to a new state, meaning that there is a probability distribution over the new state. In the deterministic case, the analysis is very simple. If $v(s') \geq v(s)$, then the decision is a good one; otherwise it leads a state of less value. The stochastic case is only slightly more complex. Assume the decision leads to state s' with probability $p(s')$. Then we compare the expected value of the resulting state with the original state: If

$$\sum_{s' \in \mathcal{S}} p(s') v(s') \geq v(s),$$

then the decision is a good one; otherwise it leads (in expectation) to a state of less value. We can similarly compare multiple potential decisions, each leading to a stochastic outcome.

7.2.1   EXAMPLES

In baseball, we return to the Markov chain model for the progression of an inning. We can imagine a situation in which there is a runner on second base and no outs, i.e. $s = (0, 1, 0, 0)$. The manager might consider instructing their next batter to bunt in this situation. Although bunt attempts are not guaranteed outcomes, we can simplify the problem by thinking about the bunt attempt as leading deterministically to a runner on third base with one out, i.e. $s' = (0, 0, 1, 1)$. If $v((0, 0, 1, 1)) > v((0, 1, 0, 0))$, then the bunt attempt increases the expected number of runs scored in the inning. Otherwise, it decreases expected runs. This comes with two important caveats:

1. This analysis assumes that a league-average batter is at the plate. The worse the batter, the better the decision to bunt.

2. More important that maximizing run expectancy is maximizing win probability. We could estimate a different model with a reward function based on winning the game rather than scoring runs. In particular situations (depending on the inning and score differential), we will get a different result from the analysis based on run expectancy.

In football, we can imagine a situation in which a team has possession of the ball at the 50 yard line, facing a fourth-and-1, i.e. $s = (50, 4, 1)$. The coach needs to decide whether to punt or to go for a first down. In this case, the score expectation model is insufficient to support this decision because it does not factor in the downside risk of turning the ball over on downs (as opposed to punting). We would need to create a win probability model (with the reward based on reaching the terminal "win" state), factoring in the score differenc and the time remaining in the game in order to adequately support this decision. Even given a win probability model, the outcome of the "go for it" decision is a stochastic one, not a deterministic one, complicating the problem further.