Segmentation Data Model

December 9, 2015

Let D be the match data. Let Ω be a set of discrete elements that we shall call labels. Later these labels will be used to label segments of the data set, for example a label might be 'hero 1' or 'laning'. Let $S \subset \mathbb{R}^3$ be the space composed of 2D space and 1D time relevant to our match data set, i.e., a cuboid in 3D bounded by the (x, y) coordinates of the map and the start and end times of the match.

We are interested in partitioning the space S into segments and assigning labels to each of these segments. We also wish to pass each segment the data necessary to perform finer grained analysis relevant to that segment. This motivates us to define a function f mapping S (the space in which the data exists in the game), Ω (the set of labels we have defined) and D (the data set) into n segments:

$$f : (S, \Omega, D) \to \{(S_1, \Omega_1, D_1), (S_2, \Omega_2, D_2), \ldots, (S_n, \Omega_n, D_n)\}$$

where $P = \{S_1, S_2, \ldots, S_n\}$ is a partition of S, Ω_i is a set of labels and $D_i \subseteq D$ for $i = 1, 2, \ldots, n$. Each segment (S_i, Ω_i, D_i) may then also be partitioned into segments to provide finer grained labelling of each segment, and so on, creating a tree structure which we call the segment hierarchy. For example, the top level might be segmented into n segments according to which team is winning and labeled with 'dire winning' or 'radiant winning', this segmentation is carried out according to some criteria that we define such as 'gold advantage', 'xp advantage', 'map control'. Since each of these states change depending on events that happen in the game i.e., a hero kills another hero and gets gold, this feeds up the tree so that we can make statements like 'the dire team take the lead because hero x killed hero y and earned z amount of gold'.

A hero narrative is the sequence of segments that a hero is part of. Each hero can only be part of one segment at each level of the segment hierarchy.

0.1 Definitions

A family of sets $P = \{S_1, S_2, \ldots, S_n\}$ is a partition of a set S if and only if

$$\emptyset \notin P, \quad \bigcup_{i=1}^{n} S_i = S, \quad S_i \cap S_j = \emptyset \text{ if } i \neq j.$$