

NLG for Fantasy Sports News

Vaibhav Singh and John Maxwell

In fantasy sports users take on the role of team manager to draft athletes from real life to form their fantasy team. Relevant statistics about a player's performance in real life games are translated into virtual points. Each week team managers match their players against an opponent's to decide who wins. Users make decisions about which players to start or trade based on their statistics, but a lot of users also enjoy reading weekly recaps or news blurbs about what players have done. There are a lot of players to write blurbs for and the blurbs are fairly well structured making it a good candidate for natural language generation.

The aim of this project is to use NLG and player statistics to automatically generate news blurbs for players. We intend to accomplish this by using news blurbs and player stats from rotoworld.com, a fantasy sports news site. We'll use the news blurbs from rotoworld.com to generate a set of templates for future reuse and combination. We'll also create rules for what types of values should be put into which slots in the templates.

The second phase of our project will generate new news blurbs by deciding which templates to use based on the significance of a player's statistics in relation to that player's past performance and the performance of other players of a similar type. This could mean we generate nothing for players which have done nothing noteworthy, a smaller blurb for players that have something noteworthy, or a larger blurb for players that have many noteworthy stats. The templates we generate and use for a particular blurb will depend on domain rules, syntax rules, and context.

We will test our approach by asking human judges to guess whether a particular news blurb was generated by us or from rotoworld.com. If human judges aren't able to consistently distinguish between the two we'll consider our approach a success. We aim for our generated text to be an accurate representation of players' stats, natural sounding text free of grammatical errors, and enjoyable to read.

John Maxwell - taking CS 583 (Data Mining and Text Mining)

Vaibhav Singh - taken CS 583 (Project: Aspect-based Sentiment Analysis)