Automated text recaps: A new layer of content delivery for esports broadcasters

Introduction

Esports matches are now watched by around half a billion people worldwide, with well over half of that number considered to be casual viewers rather than esports enthusiasts.

As broadcasters compete for their share of this fast-growing market, they face the challenge of providing coverage that makes sense of the complex gameplay for relative newcomers, but without compromising on the in-depth insights demanded by more seasoned viewers.

Much of the responsibility for this balancing act lies with the casters, the expert presenters and commentators whose job it is to convey in words the excitement and significance of the gameplay action as it unfolds. We believe that these words form a potentially valuable resource of esport-specific linguistic data.

In 2020 we conducted research to explore whether a Natural Language Generation (NLG) system could make use of this data to create an accurate and readable text description of an esports match and, if so, how such a text description could be used to improve a viewer's understanding of the action.

Our findings point the way to a new layer of content delivery for esports broadcasters that could enhance the overall viewing experience for fans of all types.

The here and now: Overcoming the language barrier

The power of automated storytelling has already been harnessed by media organisations outside the esports world. Companies such as Yahoo and Associated Press, for example, have used this technology to generate content such as weather forecasts and fantasy football updates. However, prior to our research, there have been no studies on the application of data-to-text tools to the world of esports.

One of the challenges of automated text generation is that vocabulary, grammar and meaning can differ greatly depending on the subject matter and intended audience, a factor which is especially true of the esports domain. For this reason, we focused our research on the use of an NLG system in which the rules governing word choice and grammar are derived from a specific corpus of text, in this case the oral commentaries and post-match analysis provided by professional esports casters.

For the purposes of our study, we selected key passages of play from 55 professional Dota 2 matches in 2019 and 2020, and created transcriptions of the relevant sections of audio commentary.

We then used a Natural Language Toolkit to parse each sentence of commentary and map it onto the relevant non-linguistic data for the same passage of play.

The idea was to preserve the eloquence, authority, tone and substance that suffuses the words of the professional casters, and so generate text descriptions for esports fans that come across almost as human-written: accurate, natural to read, and easy to understand.

The future: Demonstrating the benefits of automated text recaps for esports broadcasting

To test the quality of our automated text descriptions, we began by asking 127 active Dota 2 players to watch a two-minute video clip of a key piece of action and then read a text description of the same passage of play.

When asked to rate the text for accuracy, 86% of respondents judged it to be 'very' or 'extremely' accurate (64% and 22% respectively). The ratings for fluency were also positive: the results were more evenly split between 'moderately', 'very' and 'extremely', and the highest proportion of respondents (38%) opted for 'very fluent'. This marks the first demonstration that the challenge of domain-specific text generation can be met by applying NLG technology to a curated corpus of linguistic source material.

To evaluate the role played by text recaps in the understanding of a passage of play, we split our respondents into three different groups: one group watched a video of the action; a second was shown only a text description; and a third was shown both. All were then asked the same five questions to assess how well they could recall the chain of events.

We found that the group shown only the text recap registered a median score of three out of five, while the other two groups scored a median of four. On the one hand, this demonstrates that, when it comes to understanding a passage of play, reading a text description of it does not equate to watching and listening to it for yourself. On the other hand, it shows that a text description can convey a significant proportion of what is going on.

This suggests that text-only recaps could be a valuable way to reach fans who do not have the time or resource to view a particular match because they are at work, in transit, or trying to keep up with multiple matches during a tournament.

While no significant difference was found between the median score of those who watched the video only and those who saw both video and text, the scores of the latter group were more consistently at a high level. This indicates that text descriptions do not noticeably enhance a viewer's understanding of a passage of play, but that there may be scope for them to play a complementary and supportive role.

Examples could include adding an explanatory text-recap feature to existing graphics-led esports apps and production tools such as WEAVR and Echo. The technology could also prove useful to the hoped-for integration of esports coverage into mainstream television broadcasting. Perhaps, for example, by providing instant text content for a scrolling news-style ticker below the video feed.

While our study has focused on professional-level matches, we can also envisage using our NLG system to create automated professional-level text recaps of matches played in lower leagues, which could prove to be a valuable labour-saving device at a level where fewer resources are available.

Summary and next steps

Our research shows that esports commentaries are a viable linguistic resource for generating accurate, readable and readily understandable text recaps.

Further work is required to streamline the corpus transcription and curation process but we believe our work lays the foundations for the integration of state-of-the-art speech-to-text machine learning models, and the introduction of a new layer of content delivery within the esports industry.

We now plan to focus on the potential for the NLG system to produce personalised text recaps by making lexical choices according to the skill and preference of each individual viewer.

Read the academic paper:
https://pure.york.ac.uk/portal/en/publications/automatic-generation-of-text-for-match-recaps-using-esport-caster-commentaries(017163d4-a6bb-4acb-b496-02b018cb81fe).html