

Goal analysis

These are the main goals I set for this project and their state:

1. Programming an automatic obtention of a box score (the main statistics source) of a match from play-by-play data

Done. The box score can be between two match timestamps, which gives the project more interest, as this feature cannot be found anywhere. I have additionally developed filtering methods for the box scores, so the user can filter by values, categories or players.

2. Programming of automatic match features extraction. They can be more direct, such as, the longest time without scoring by a team, or more abstract, such as trying to determine whether a timeout made a change in a team performance

Done. I have provided about five match characteristics extractions (greatest streak, longest time without scoring, etc). I have also programmed the extraction of information about the playing time of the players. I will not provide interpretations of the information, as I later on considered it is better to leave them for the user.

3. Design and implementation of visualisations describing the performance of a team, player or top X players in a match.

Work in progress. At this very moment, I am figuring out how to deal with the visual part.

4. Generation of a dynamic visualisation of the evolution of a match (a visual play-by-play)

Work in progress. I have programmed the backend of the functionality and, for the moment, I have offered a simple front end (based on time-delayed printing).

5. Making the tool able to read play-by-play from other/any sources (specially to be able to deal with alternative languages to English)

Cancelled. Once getting involved with the information sources, I realised that the previous information about the input format makes the information treatment much easier. So, if I want to benefit from how the data is formatted, I cannot do a generalised tool, as it would imply setting too many conditions. Moreover, the main benefit of making it general is to be able to treat other leagues. But I consider that staying in the NBA environment is perfectly valid. Moreover, I found a play-by-play source for the ACB (the Spanish league), but it is in Spanish. It is the only one I found. That means that if I wanted to treat the Spanish league, I would have to deal with another language, which either means more conditions or the development of an NLP method (which might be too complex without meaning good results)

Tasks analysis

These are the project tasks, including former and new tasks:

1. Creation of an intermediate language ("standard play-by-play") to use as an input for the feature extraction

Done. I initially did not include this as a task, as I considered it was part of the box score obtention. However, the standard play-by-play file is used in every feature computation, so its design should be mentioned independently.

2. Simple box score obtention

Done

3. Advanced box score obtention

Done. I included several efficiency standard parameters not usually included in box scores and added variables designed on my own

4. Design of visualisations and tables showing statistics, either of the match or of the evolution during a season:

Work in progress: At this moment, there are several statistics given. There are a few more match statistics to be computed. There is also a shooting efficacy visualisation planned to be designed if possible.

5. Design of a visual play-by-play

Work in progress: I have programmed its backend and decided to delay the visual part. This is because of two things:

- The consideration that there are features which are more important to develop than the visual play-by-play
- The fact that I still have to wonder what my possibilities are due to the application I will use

So, for the moment, I have programmed a very simple visual play-by-play based on time-spaced printing.

6. Interface design

Work in progress: I have started to develop an application in GUI tool and I am studying its suitability for my objective.

7. Exploration of a method (such as transformer or clustering) that adapts to other play-by-play syntactic formats and to other languages

Cancelled: Already explained in the goals section. If I have the time, I might manually design a translation from a Spanish league source in Spain to my standard language.

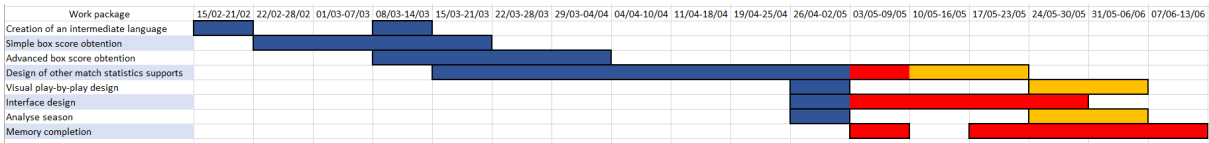
8. Programming of season analysis (method to obtain the evolution of some statistics from a desired team and a desired season)

New. Work in progress: I have designed a basic program that creates a chart describing the evolution of a desired statistic. I still have to adapt it to describe the evolution of a box score category. This section will be secondary, as I consider that finishing the other features is preferable, in order to try to make the match analysis almost complete.

9. Memory completion

Work in progress: I am starting to write the initial sections of the memory.

Gantt



The blue sections express sections in the past. The red colour expresses future sections that I have to do in order to complete the project and yellow sections are the optional sections I described in the tasks section.