

Ontologia "Futebol Português"

Sérgio Manuel Rodrigues Caldas - A57779
Processamento e Representação de Conhecimento
Departamento de Informática
Universidade do Minho

May 2016

1 Introdução

No âmbito da disciplina de Processamento e Representação de conhecimento (PRC), inserida no perfil de Processamento de Linguagens e Conhecimento (PLC) do curso Mestrado Integrado em Engenharia Informática (MIEI), foi-nos proposto, como projeto final da disciplina, desenvolver uma ontologia para um tema à nossa escolha, bem como povoá-la e por fim mostrar a sua utilidade através de queries exemplo.

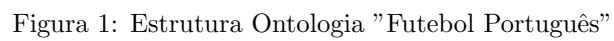
Uma ontologia, é um modelo de dados, que relaciona um conjunto de conceitos dentro de um domínio (no meu caso o futebol português). Com uma ontologia é possível inferir conhecimento relativo aos objetos do domínio. As ontologias são usadas em diferentes áreas das ciências da computação, como forma de representação da informação.

Para o desenvolvimento da minha ontologia, utilizei a aplicação protégè, versão 5.0.0. Nesta aplicação criei todas as classes, relacionamentos e propriedades necessárias, bem como o povoamento da mesma. Depois de completa a ontologia, ativei o reasoner, para inferir o conhecimento em falta, procedendo depois a exportação desta ontologia para uma outra já completa, com todos os dados (inferidos e não inferidos). Por fim utilizei a mesma aplicação para o desenvolvimento das queries que provam a utilidade da mesma.

2 Estrutura da Ontologia

Antes de passar ao desenvolvimento da ontologia propriamente dita, tive de estudar bem o domínio da mesma, bem como a estrutura das classes e suas respetivas relações.

Depois de uma análise detalhada de como o "Futebol Português" está organizado cheguei ao esquema que pode ser consultado na figura 1. Este esquema mostra a estrutura da ontologia por mim desenvolvida, aqui podemos ver todas as classes e respetivas relações que a constituem.



Como podemos analisar, ao observarmos a figura 1, o "Futebol Português" está organizado por *Associações*, as associações tem *Divisões*, as divisões tem *Campeonatos*, os campeonatos ou têm *Equipas* ou têm *Series*, se estes forem constituídos por series, então cada série tem *Equipas*, cada equipa tem um *Presidente*, um *Plantel*, e uma *Equipa Técnica*. O plantel tem *Jogadores* e a *Equipa Técnica* tem um ou mais *Treinadores*.

Todas as relações representadas na figura, possui uma relação inversa, assim sendo por exemplo ao dizer que uma *Equipa* tem um *Plantel*, inversamente, o *Plantel* pertence a uma *Equipa*. As restantes relações inversas podem ser consultadas facilmente na figura 1.

3 Povoação da Ontologia

Para povoar a ontologia, procedi à consulta e respetiva extração dos mesmos no website de futebol *zerozero*[1].

Devido à grande quantidade de dados, e ao facto de nem todos os dados estarem organizados da mesma maneira, apenas utilizei como exemplo completo o campeonato português, isto é a "Liga NOS", para este exemplo encontram-se disponiveis todos os dados referentes a esta competição desde as equipas, aos jogadores, presidentes e treinadores. Contudo toda a ontologia está preparada para receber toda a informação relativa ao futebol sénior existente em Portugal, desde o primeiro escalão até às competições distritais.

3.1 Dados Estatísticos da Ontologia

Em baixo podemos consultar alguns dados estatísticos da ontologia por mim desenvolvida:

- 1 Associação
- 3 Divisões
 - Primeira Divisão
 - Segunda Divisão
 - Terceira Divisão
- 3 Campeonatos
 - Liga NOS
 - Ledman LigaPro
 - Campeonato Nacional de Seniores

Referente à Liga NOS (a única que tem a informação completa), temos os seguintes números de dados:

- 18 Equipas

- 18 Planteis
 - * 513 Jogadores
- 18 Equipas Técnicas
 - * 18 Treinadores
- 18 Presidentes

Todos estes dados e respetivas relações constituem um total de:

- 35482 Triplos

4 Conclusão

Depois de concluída a ontologia, considero que o resultado final é bastante positivo. Como já foi visto anteriormente esta ontologia é constituída por 35482 triplos, apenas com a primeira liga portuguesa. Como se pode ver pelos dados em cima apresentados, esta ontologia já é uma ontologia de uma dimensão bastante aceitável.

Quanto à prova da utilidade desta ontologia, desenvolvi um conjunto de queries (10), que provam a utilidade da mesma, estas queries podem ser consultadas no ficheiros *queries.txt* (ficheiro com as queries para a ontologia sem o conhecimento inferido) e no ficheiro *queries_inf.txt* (ficheiro com as queries para a ontologia com o conhecimento inferido). Ambos estes ficheiros fazem parte do ficheiro *zip* enviado ao professor.

Ao longo do desenvolvimento desta ontologia deparei-me com algumas dificuldades, nomeadamente no que toca à extração e organização da informação. Iniciei este processo com o desenvolvimento de *scripts* para a extração e organização dos dados, contudo os dados no website não estavam organizados da maneira que eu esperava, e como tal tive imensas dificuldades em acertar essas *scripts* de forma a organizar os dados como eu desejava. Isto também explica o porquê de eu não ter conseguido povoar a ontologia com todos os dados do "Futebol Português", desde a primeira liga até às competições distritais, mas como já referi anteriormente considero bastante positivo o resultado final mesmo tendo em conta estas adversidades.

Como trabalho futuro, pretendia completar na totalidade esta ontologia. Isto é, povoa-lá com toda a informação, relativa ao escalão sénior do "Futebol Português". Depois de povoada toda a ontologia considero que esta iria ficar completa, robusta e com maior utilidade.

Referências

- [1] Zerozero. <http://www.zerozero.pt/compet.php?idpais=1continente=1esc=0gen=0mod=0stats=0>.