FrameNet-Based Modeling of the Domains of Tourism and Sports for the Development of a Personal Travel Assistant Application

Alexandre Diniz da Costa, Maucha Andrade Gamonal, Vanessa Maria Ramos Lopes Paiva, Natália Duarte Marção, Simone Rodrigues Peron-Corrêa, Vânia Gomes de Almeida, Ely Edison da Silva Matos, Tiago Timponi Torrent
Federal University of Juiz de Fora – FrameNet Brasil
Rua José Lourenço Kelmer, s/nº – Campus Universitário – 36036-900 – Juiz de Fora/MG – Brazil
{alexandre.costa, ely.matos, tiago.torrent}@ufjf.edu.br, {duarte.natalia, vania.almeida2017}@letras.ufjf.br, {mauchaandrade, speronjf, vanessaletrasufjf}@gmail.com

Abstract
This paper presents an enriched frame-based multilingual lexicon covering the domains of Tourism and Sports, which supports a personal travel assistant application – m.knob – developed to help tourists get recommendations of attractions and activities, as well as to communicate with other tourists and service providers, in the context of major international sports events, such as the Summer Olympics. Recommendations are provided through frame-based automated categorization of tourist attractions based on semantic information extracted from tourists' comments on online platforms, which are then matched to semantic information extracted from the input the user inserts in a conversational user interface.

Keywords: Tourism and Sports Modeling, Algorithmic Categorization, m.knob, FrameNet Brasil.

1. Introduction

Events such as the Summer Olympics provide the meeting of people from different parts of the world, who have different interests related to tourist attractions and sports, as well as speak different languages. Therefore, major international events like this one call for multilingual tools that can assist tourists in their choices related to places to eat or visit, sports events to attend, and so on.

Also, planning a trip or leisure activity requires different types of information about a tourist attraction or event. Many travel guides can assist in bringing information about places, how to get there, what to do, or even the temperature and weather conditions at any given time of the year. Likewise, these tools often focus on prominent attractions or more general information that aid in the basic planning for a trip. However, travel guides do not provide specific information that many tourists may need when planning a trip, such as which attraction is better for a rainy day or which museum is interesting for children. This information is either subjective and subject to change or is scattered around the text. While this kind of information may be available on online platforms in the form of comments and reviews posted by users, reading them all is a task incompatible with the dynamism of a trip.

Considering this context, an automatic analysis of these comments could generate more useful information to the tourist, especially if they are made available in an interactive and dynamic platform. It is not only a matter of extracting if the general impression about a certain attraction is positive or negative, an already classic task in Natural Language Processing (NLP), but also to go beyond such classification, bringing more specific information that helps the user make decisions. In addition, this specific information can also help the tourist to choose the sports disciplines, considering the context of the Olympic games, and to find the places where the competitions take place, since sports are also a type of leisure activity searched by tourists in this context.

This work is developed under the m.knob (Multilingual Knowledge Base) project of the FrameNet Brasil Computational Linguistics Laboratory at the Federal University of Juiz de Fora. Such a project is developing a personal travel assistant in the form of a chatbot with which tourists can interact using natural language to get recommendations for attractions, places to eat and leisure activities.

In this context, this paper aims (a) to show how the modeling was carried out, and (b) to present an automated categorization methodology for tourist attractions based on semantic information extracted from comments posted to online platforms. Such a methodology provides for the existence of an analyzer that extracts the semantic information from the comments and translates it into a cluster of frames. The system also generates clusters from the user's inputs and later maps the similarities between the clusters, suggesting attractions and tourist activities that can adhere to the user's interests.

2. Frame Semantics and FrameNets

Frame Semantics is an approach to lexical semantics whose main assumption is that meanings are relativized to scenes (Fillmore, 1977), that is, to frames. Fillmore (1985) proposes an approach to semantics based on language understanding, analyzing the linguistic choices made to produce utterances so that they convey beliefs about the world, experiences, and the way speakers see things. Frames are defined as a system of concepts related in such a way that "to understand one of them, it is necessary to understand the whole structure in which it fits" (Fillmore 1982, p. 111).

The main application of Frame Semantics is FrameNet, a project started in the International Computer Science

Institute (ICSI), by Charles Fillmore, with the purpose of providing, through the exposition of Lexical Units (LUs), the frames evoked by these LUs, identified by the Frame Elements (FEs) that constitute them. By FEs, we mean any semantic role specifically defined in the frame. FEs provide additional information to the semantic structure of the sentence. LUs, in turn, are pairings of lemmas and the frames they evoke (Fillmore, 1982). The analyzes performed on the LUs, therefore, provide us with a description of their syntactic valence properties (grammatical functions and syntagmatic types that co-occur in the syntactic locality of the lexical item) and semantics (frame elements instantiated by these valents).

Figure 1 shows the Attracting_tourists frame, its FEs and LUs. There's also a definition of the frame, as well as one for its core FEs – ATTRACTION, PLACE and TOURIST – and their definitions as well.

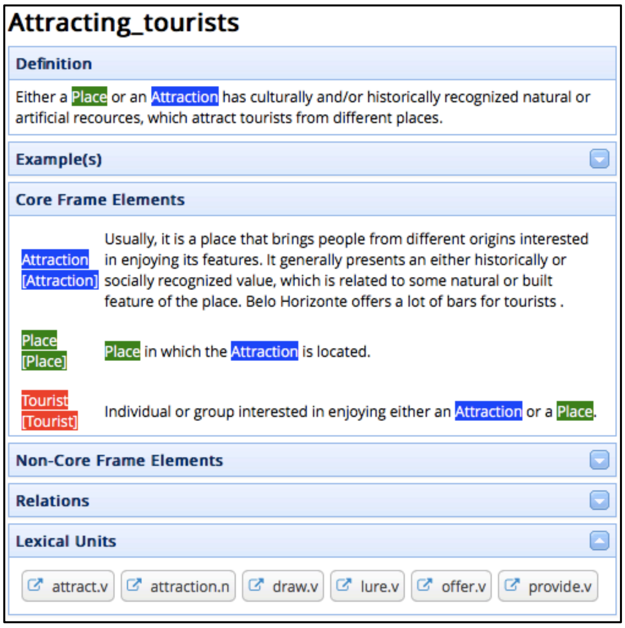

Figure 1: The Attracting_tourists frame.

LUs evoking this frame include *attract.v, draw.v, lure.v, offer.v and provide.v.* Sentences containing these LUs are annotated in a multiple layer fashion (Frame Element, Grammatical Function and Phrase Type), and show clear examples of basic combinatorial possibilities (valence patterns) for each target LU. Note that, although some of these lemmas may also take part in LUs evoking different frames – such as Cause_motion, Manipulate_into_doing, Offering and Supply, respectively – their sense in the context of sentences (1-5), extracted from travel guides in the FrameNet Brasil corpus, takes the Attracting_tourists frame as a background, not the other frames mentioned above, as indicated by the color code matching the linguistic material in each sentence to the FEs shown in Figure 1.

(1) The mighty Songhua River, running through Harbin from west to east, inevitably ATTRACTS tourists.

(2) Manzanillo initially DREW the interest of international visitors for its excellent fishing.

(3) Alicante Swaying palms and luminous skies, along with some of Spain's best restaurants and tapas bars, LURE visitors to the provincial capital of Alicante.

(4) Few countries OFFER so much to visitors as Brazil.

(5) Kuta, and its progressively upscale neighbors to the north ' Legian , Seminyak , and Kerobokan (as well as Tuban, to the south) PROVIDE an enormous selection of hotels, restaurants, pubs, and shopping choices . INI

Based on FrameNet, lexical resources are being developed for different languages such as German (Boas et al., 2006), Japanese (Ohara et al., 2004), Spanish (Subirats & Petruck, 2003), Chinese (You & Liu, 2005), Swedish (Borin et al., 2010) and Brazilian Portuguese (Salomão, 2009). Similarly to Berkeley FrameNet, FrameNet Brasil follows the same methodology with a team of linguists and computer scientists who are involved in various fields of research, from the construction of lexical resources to the development of applications for natural language understanding. We now turn to one of such applications developed by FrameNet Brasil: m.knob.

3. Multilingual Knowledge Base

Multilingual Knowledge Base (m.knob) is a travel assistant app that offers personalized information to tourists about the specific domains of Tourism and Sports. The alpha version of the app was released during the Rio 2016 Summer Olympics and has been redesigned to include other functions in its beta version.

The app covers three languages – Brazilian Portuguese, English and Spanish – and has two main functions, (i) a chatbot providing recommendations on tourist attractions and activities; and (ii) a semantically enhanced sentence translator algorithm based on frames and qualia relations (Pustejovsky, 1995).

The Tourism domain was modeled in a previous application: the 2014 World Cup Dictionary. Torrent et al. (2014) developed a frame-based trilingual electronic dictionary for the 2014 World Cup, covering the domains of Football, Tourism and the World Cup in the same three languages. The modeling carried out for the Tourism domain (Gamonal, 2013; Gomes, 2014; Souza, 2014) included, at first, 40 frames. For m.knob, it has been revised and improved to cover other aspects of the travel experience, and currently features 58 frames, 16 of which already existed in the Berkeley FrameNet Data Release 1.7. As for the Sports Domain, Costa & Torrent (2017) created 29 new frames and used 4 frames from Berkeley FrameNet 1.7. Currently, the m.knob lexicon comprises a total of 5,152 LUs: 1,671 for Brazilian Portuguese, 2,551 for English, 930 for Spanish.

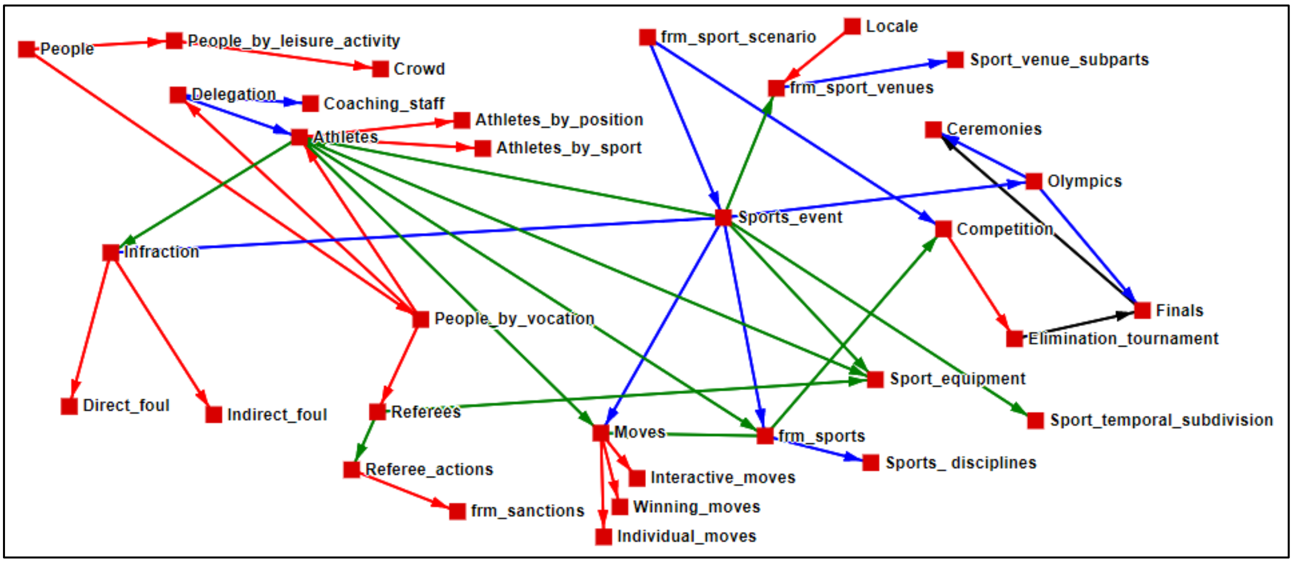

Figure 2: Frames and relations in the Sports domain. Arrow colors indicate the types of relations: Inheritance (red), Using (green), Subframe (blue) and Precedes (black).

The process of modeling the Tourism and Sports domains, besides creating new frames, also led to the enrichment of FrameNet Brasil to the extent that it incorporated new relations to the database. This process is discussed next.

3.1 Modeling the Tourism and Sports Domains

The process of creating and modeling the frames for Tourism and Sports adopted a bottom-up approach and started with the compilation of trilingual corpora related to the domains. Texts were extracted from travel guides and blogs, governmental portals on tourism and on the Olympics, as well as from sports manuals and websites of associations of each Olympic sport. The corpus compilation tool used was Sketch Engine (Kilgarriff et al., 2014).

Next, candidate terms in the corpus were extracted using TermoStat (Drouin, 2003) and the context in which they occur is analyzed to both (i) validate the term as evoking a frame related to the relevant domains, and (ii) expand the list of candidate terms. Example sentences were then analyzed to provide the basis for the proposition of the frames. Finally, the resulting proto-frames were then refined – based on the literature on tourism (Gamonal, 2013) and on the rules of the Olympic sports –, and related to one another in a network, using the frame-to-frame relations originally defined by Berkeley FrameNet (Ruppenhofer et al., 2016). The resulting model for the Sports domain is presented in Figure 2.

Besides the frames and LUs modeling the specific terminology of Sports and Tourism, the m.knob lexicon also contains domain general frames and LUs relevant to the description of tourist and sports attractions. Such frames and LUs were selected from the Berkeley FrameNet data release 1.7 and expanded into Brazilian Portuguese and Spanish. This selection was based on a pilot study in which a corpus of 3,495 comments written in English about 939 tourist locations in San Francisco was analyzed semi-automatically in a three-step procedure:

- first, candidate LUs were automatically extracted from the corpus, by comparing the word forms in the comments to those associated to LUs – and, therefore, frames – in Berkeley FrameNet;
- second, frames were ranked from the most to the least frequent, regardless of the LU evoking them;
- third, annotators in the FrameNet Brasil team manually checked which frames were actually relevant and which of them were irrelevant to the domains.

Among the examples of relevant frames are Stimulus_focus (evoked by LUs such as *great.a, beautiful.a, interesting.a*), Expensiveness (*expensive.a, cheap.a*), Kinship (*son.n, grandfather.n*), People_by_age (*child.n, senior.a*), Locales_by_use (*museum.n, church.n*), Natural_features (*LUs such as beach.n and valley.n*) and so on. Frames were judged as irrelevant mostly when the word forms triggering their recognition by the system should actually point to another frame, or to no frame at all in the context of the comments. The parade examples are the Performers_and_roles (evoked by *be.v*) and the Sex (evoked by have.v) frames. Both *be.v* and *have.v* are very frequent in the comments, but not in the senses of playing some character or having sex, respectively.

The pilot study resulted in the incorporation of 250 Berkeley FrameNet frames to the m.knob lexicon. English LUs evoking those frames were imported into the database from the data release 1.7. Brazilian Portuguese and Spanish LUs are being created in those frames through the regular expand process used in FrameNet Brasil (Torrent & Ellsworth, 2013).

The conceptual structure represented by the m.knob lexicon is a graph. Nodes in this graph include lemmas, LUs, frames and FEs. The arcs in this graph are the several relations between those nodes, such as the frame-to-frame relations currently used by most – if not all – framenets, but also new ones, which were created by FrameNet Brasil, such as FE-to-frame, qualia and metonymy relations.

Because the m.knob lexicon is meant to be used as the basis for a recommendation system and a sentence translator, new relations were added to the database apart from those originally created by Berkeley FrameNet – illustrated in Figure 2 – either to provide more specific links – connecting LUs instead of frames –, or to account for the definition of the entities participating in an event and for the possible metonymic relations between those entities.

The first set of new relations, those connecting LUs, was adapted from Pustejovsky's (1995) qualia (Costa & Torrent, 2017). So far, three different qualia were implemented in the m.knob database: formal, constitutive and telic. The formal quale is used to indicate that a given LU has the same ontological type of another, more generic LU. It is a *is-a* relation and is used to indicate, for example, that *taphouse.n*, *sports bar.n* and *pothouse.n* are a *bar.n*. The constitutive quale indicates that the referent of a given LU functions as a part or content of the referent of another LU. It indicates for example that *bleachers.n* and *field.n* are parts of a *stadium.n*. Finally, the telic quale is used, in m.knob, to indicate either the inherent purpose of an object or the actions prototypically performed by an agent. It is used to indicate, for example, that the *ace.n* in a soccer team usually scores a *goal.n*, but not an *ace.n*, which is prototypically performed by a *tennis player.n*.

The second set of new relations models the fact that participants in a frame can be defined in terms of other (entity) frames, and also that, in some cases, they can be represented metonymically. Using the Attracting_tourists frame (Figure 1) as an example, an FE-to-frame relation models that the PLACE FE may be defined in terms of the Locale frame, while the TOURIST FE may be defined in terms of the People frame (Figure 3). Additionally, inside the People frame, a FE-to-FE Metonymy relation indicates that the non-core FE ORIGIN, may stand for the core FE PEOPLE (Gamonal, 2017).

Changes as the one just described, allow m.knob to extract, from (6), that the Attracting_tourists frame was evoked in the sentence, because:

- first, an FE-to-frame relation links *city.n*, in the Political_locale frame to the FE PLACE, via the Locale frame;
- second, the Metonymy relation creates a link between *Brazilian.a* and *people.n* – or any other LU in the People frame;
- third, an FE-to-frame relation links *Brazilian.a* to the FE TOURIST, via the People frame.

(6) The city lures Brazilians with beautiful beaches and nice shops.

This kind of structure is then key to m.knob's recommendation system, which will be presented in section 3.2.

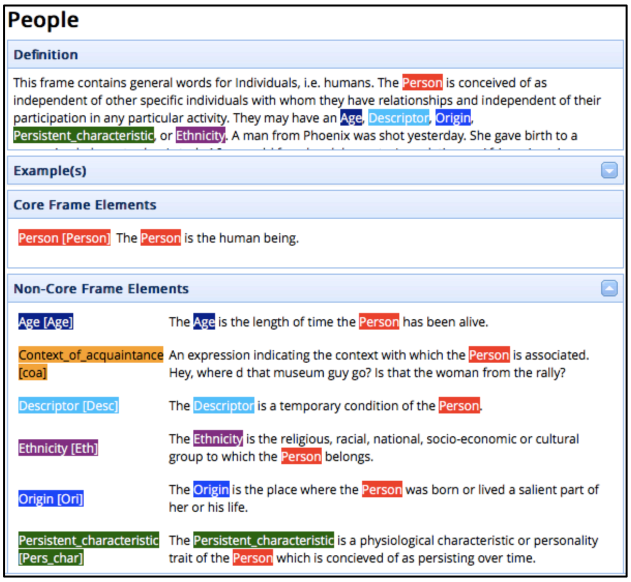

Figure 3: The People frame

3.2 Automated Categorization of Attractions

Although the collaborative culture of the internet has made subjective assessments of tourist attractions available through diverse tools, this is still not enough for the user to take advantage of this information, given the impossibility of reading all the comments when planning a trip. The application described in this work overcomes these limitations through a categorization algorithm that uses the m.knob lexicon to generate detailed semantic representations of attractions and events.

Based on the algorithmic categorizer, the system parses comments posted to online platforms and extracts the meaning of the candidate words. In a first stage, the set of frames evoked in the comments is gathered. Then, the evoked frames are weighed as to their frequency in the data. In a third step, the frame clusters representing each place are derived and stored in the m.knob database, as well as additional information about the place itself, such as its name, opening hours, location and, very important, its type in the online platform. Such types are stored also in the m.knob lexicon, in the form of LUs such as *bar.n, park.n, beach.n* and so on. Place types are usually the dominant node of formal quale relations, as the ones exemplified in section 3.1.

On the other end, a conversational user interface, namely a chatbot, provides the user with the possibility of entering, in one of the three languages covered by the resource, what she'd like to do. In the final stage, the system provides the tourist with recommendations resulting from a cluster-matching process between the semantic representation generated for the user's input and those generated for the attractions from the analysis of the comments.

As an example, consider that one user enters sentences (7), (8), (9) or (10) to the chatbot system.

(7) Quero passear com a minha família.
I want to go out with my family.
(8) Quero passear com a minha família à noite.
I want to go out with my family tonight.
(9) Quero ter contato com a natureza.
I want to be close to nature.
(10) Quero passear com a minha família em contato com a natureza.
I want to go out with my family to be close to nature.

First, the system extracts the LU candidates from the sentences and finds in the m.knob lexicon the correspondences shown in Table 1.

Br-Pt LU	En Gloss	Frame
passear.v	*go out*	Going_places
família.n	*family*	Kinship
contato.n_1	*be in contact with*	Contacting
contato.n_2	*be close to*	Spatial_contact
natureza.n	*nature*	Natural_features

Table 1: LUs found in sentences (7-10) and the frames they evoke in the m.knob lexicon

Second, using the relations between frames, FEs and LUs described in section 3.1, the system disambiguates the lemmas pointing to more than one LU. In this example, *contato.n 'contact'* is an ambiguous lemma, since it could refer to both an LU in the Contacting frame and one in the Spacial_contact frame. However, in the user input, it appears close to *natureza.n 'nature'*, which evokes the Natural_features frame. Based on that, the system infers that Spatial_contact is more likely, because the distance – in terms of the relations described in 3.1 and also those common to FrameNet, such as Inheritance, Perspective and so on – between this frame and Natural_features is shorter than that between Contacting and Natural_features (see Torrent et al., 2014 for a description of the frame disambiguation system).

Third, the system generates a semantic cluster to represent the user query. In this process, it takes two other kinds of linguistic information into account, besides the LUs found in the query: words that do not evoke frames, but appear both in the user input and in the comments – such as *noite.n 'night'*, for example –, and other LUs evoking the frames in the query – such as *filho.n 'son', pai.n 'father', mother.n 'mãe'*, in the Kinship frame, and *montanha.n 'mountain'* in the Natural_features frame. That way, the system, once again, makes use of the network-like infrastructure of FrameNet to broaden the linguistic bases used for recommendation.

Next, the cluster representing the query is to be matched to those representing places to be recommended. This is made possible by: first, turning the cluster into a graph in which LUs, frames, and other words are nodes and the relations connecting them in the m.knob lexicon are arcs, and, second, by applying spreading activation techniques to this graph to find which of the places in the database is the best fit for the user query (see Matos et al., 2017 for a description of the spreading activation process used in FrameNet Brasil).

For the sake of exemplification, let's assume that the m.knob database has six places which are potentially relevant to queries (7-10). By applying the first three steps described for the analysis of the user query to the comments written about those places – namely, LU candidate extraction, frame disambiguation and semantic cluster generation –, the system derives a semantic cluster representing each place, as shown in Table 2.

Such clusters are also represented as graphs, whose nodes will be activated in the cluster matching process. In the end, the places the system will recommend to the user are those with the highest activation levels achieved based on the user input and how it matches to the semantic representation of the place.

Place_#	LUs	Frames	Other
Place_1	*contato.n_2* *natureza.n*	Spatial_contact Natural_features	
Place_2	*contato.n_2* *natureza.n*	Spatial_contact Natural_features	
Place_3	*passear.v* *família.n*	Going_places Kinship	
Place_4	*passear.v* *família.n*	Going_places Kinship	
Place_5	*passear.v* *família.n*	Going_places Kinship	*noite.n*
Place_6	*contato.n_2* *natureza.n* *passear.v* *família.n*	Spatial_contact Natural_features Going_places Kinship	

Table 2: LUs, frames and other relevant words in the clusters describing Places 1-6 in the m.knob database

Hence, given, for example, the user input in (7), the system would recommend Places 3, 4, 5 and 6, all of them with an activation level of 1.9368, as shown in Figure 4.

Note that the activation process starts by setting the activation value of each LU in the query to 1.000. Then, every time the activation spreads to another node via an arc, this value is reduced. When a node is activated by more than one path, activation values are added up in the final node.

For the user input in sentence (8), once again Places 3, 4, 5 and 6 are activated. However, Place_5 has a higher activation value [1.9611], as shown in Figure 5, and would then be recommended as the best-fit option to the user query. This is so because both the query in (8) and Place_5 feature the word *noite.n*, demonstrating that additional information provided by the user may help the system provide better recommendations.

As for sentence (9), the activation process yields Places 1, 2 and 6 as equally good recommendations. However, if the user input is (10), then all places are activated, but Place_6 gets a higher activation score [3.8710], as shown in Figure 6.

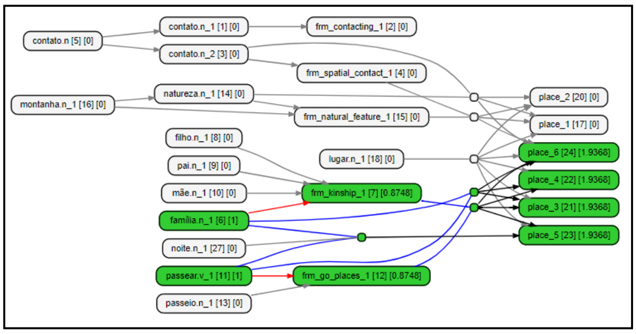

Figure 4: Graph representation of the cluster-matching process between sentence (7) and Places 1 to 6. Nodes in green are activated and numbers in the second pair of square brackets indicate their level of activation.

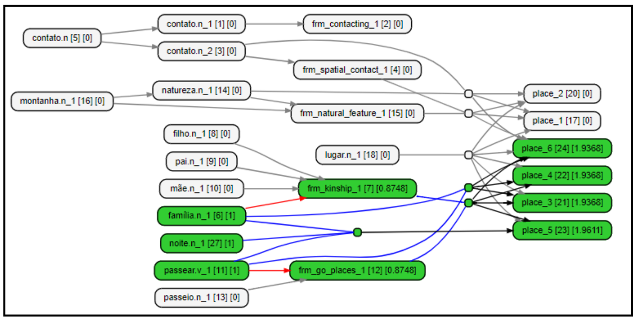

Figure 5: Graph representation of the cluster-matching process between sentence (8) and Places 1 to 6. Nodes in green are activated and numbers in the second pair of square brackets indicate their level of activation.

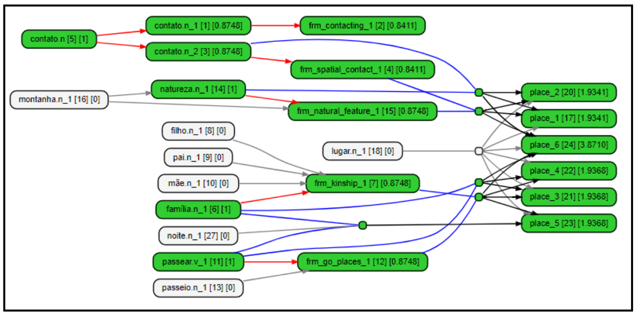

Figure 6: Graph representation of the cluster-matching process between sentence (10) and Places 1 to 6. Nodes in green are activated and numbers in the second pair of square brackets indicate their level of activation.

4. Currently Limitations and Outlook

In this paper, we demonstrated how a domain-specific framenet for Tourism and Sports can be used for providing recommendations for tourists by applying spreading activation techniques to graphs representing the semantics of user inputs to a conversational interface. Currently limitations of the system refer to both the lexicon and the algorithm.

On the side of the lexicon, there's, first, the need to balance the number of LUs for each language. Currently, the number of LUs in Spanish is half of that in Brazilian Portuguese, which, in turn, is 50% lower than that of English LUs. Second, the consistency of the newly created relations in the database must be checked.

On the side of the algorithm, the clusterization process operating on the comments uses n-grams to delimit the scope of the lemma disambiguation process. This is not ideal, since n-grams do not capture the structural relations between the lexical items, and, the m.knob lexicon, on the other hand, models plenty of those relations. In the future, we plan to substitute the use of n-grams by the constructional parser being developed by FrameNet Brasil (Matos et al., 2017).

5. Acknowledgements

The m.knob project is funded by CNPq Grant No. 448990-2014/8 and by FAPEMIG Grant No. CHE-APQ-00471/15.

6. Bibliographical References

Boas, H. C., Ponvert, E., Guajardo, M., and Rao, S. (2006). The current status of German FrameNet. In SALSA workshop at the University of the Saarland (Vol. 14), Saarbrucken, Germany, june.

Borin, L., Dannélls, D., Forsberg, M., Gronostaj, M. T., and Kokkinakis, D. (2010). The past meets the present in Swedish FrameNet++. In *Proceedings of the 14th EURALEX international congress*, pages 269–281, Reykjavik, Icenland, may. European Language Resource Association (ELRA).

Costa, A. D., and Torrent, T. T. (2017). A Modelagem Computacional do Domínio dos Esportes na FrameNet Brasil. In *Proceedings of the 11th Brazilian Symposium in Information and Human Language Technology*, pages 201–208, Uberlândia, Brazil, nov. Sociedade Brasileira de Computação (SBC).

Drouin, P. (2003). Term Extraction Using Non-technical Corpora as a Point of Leverage. *Terminology*, 9(1): 99–115.

Fillmore, C. J. (1977). The case for case reopened. In P. Cole & J. Saddock (eds). *Grammatical Relations*. New York: Academic Press, pp. 59--81.

Fillmore, C. J. (1982). Frame Semantics. In: The Linguistic Society of Korea (org.). *Linguistics in the morning calm*. Seoul: Hanshin, pp. 111--137.

Fillmore, C. J. (1985). Frames and the semantics of understanding. *Quaderni di Semantica*, 6(2):222–254.

Gamonal, M. A. (2013). *COPA 2014 FrameNet Brasil: Diretrizes para a Constituição de um Dicionário Eletrônico Trilíngue a partir da Análise de Frames da Experiência Turística*. M.A. Thesis in Linguistics. Universidade Federal de Juiz de Fora. Juiz de Fora.

Gomes, D. S. (2014). *Frames do Turismo Esportivo no Dicionário Copa 2014_FrameNet Brasil*. M.A. Thesis in Linguistics. Universidade Federal de Juiz de Fora. Juiz de Fora.

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovvář, V., Michelfeit, J., Rychlý and P., Suchomel, V. (2014). The Sketch Engine: ten years on. *Lexicography*, 1: 7–36.

Matos, E. E., Torrent, T. T., Almeida, V. G., Silva, A. B. L., Lage, L. M., Marção, N. D., and Tavares, T. S. (2017). Constructional Analysis Using Constrained Spreading Activation in a FrameNet-Based Structured Connectionist Model In: *The AAAI 2017 Spring Symposium on Computational Construction Grammar* *and Natural Language Understanding Technical Report SS-17-02*, pages 222–229, Palo Alto, CA, mar. Association for The Advancement of Artificial Intelligence (AAAI).

Ohara, K. H., Fujii, S., Ohori, T., Suzuki, R., Saito, H., and Ishizaki, S. (2004). The Japanese FrameNet Project: An Introduction. In *Proceedings of LREC-04 Satellite Workshop Building Lexical Resources from Semantically Annotated Corpora (LREC 2004)*, pages 9-11, Lisbon, Portugal, may. European Language Resource Association (ELRA).

Pustejovsky, J. (1995). *The Generative Lexicon*. Cambridge, USA, MIT Press.

Salomão, M. M. M. (2009). FrameNet Brasil: um trabalho em progresso. *Calidoscópio*, 7(3):171–182.

Souza, B. C. P. (2014). *Frames de turismo como negócio no Dicionário Copa 2014_FrameNet Brasil*. M.A. Thesis in Linguistics. Universidade Federal de Juiz de Fora. Juiz de Fora.

Subirats, C., and Petruck, M. (2003). Surprise: Spanish FrameNet. In *Proceedings of CIL*, Vol. 17, Prague, Czech Republic, jun.

Torrent, T. T., and Ellsworth, M. (2013). Behind the Labels: Criteria for Defining Analytical Categories in FrameNet Brasil. *Veredas*, 17(1): 44–65.

Torrent, T. T., Salomão, M. M., Campos, F. A., Braga, R. M., Matos, E. E., Gamonal, M. A., Gonçalves, J., Souza, B. C., Gomes, D. S., and Peron-Correa, S. R. (2014). Copa 2014 FrameNet Brasil. In *Proceedings of COLING 2014*, pages 10–14, Dublin, Ireland, aug. Association for Computational Linguistics (ACL).

You, L., and Liu, K. (2005). Building Chinese FrameNet database. In *Proceedings of 2005 IEEE International Conference on Natural Language Processing and Knowledge Engineering*, pages 301–306, Whuan, China, oct-nov. Institute of Electrical and Electronic Engineers (IEEE).