EACL 2021

FrameForm: An Open-source Annotation Interface for FrameNet

Büşra Marşan, busra.marsan@boun.edu.tr
Olcay Taner Yıldız olcay.yildiz@ozyegin.edu.tr

Abstract

In this study, we introduce FrameForm, an open-source annotation tool designed to accommodate predicate annotations based on Frame Semantics (Fillmore et al., 1976). FrameForm is a user-friendly tool for creating, annotating and maintaining computational lexicography projects like FrameNet and has been used while building the Turkish FrameNet (Marşan et al., 2021). Responsive and open-source, FrameForm can be easily modified to answer the annotation needs of a wide range of different languages.

FrameNet in a Nutshell

FrameNet (Lowe, 1997; Baker et al., 1998; Fillmore and Atkins, 1998; Johnson et al., 2001) is a growing NLP resource developed by the International Computer Science Institute in Berkeley, California. Having its theoretical background in Fillmore's Frame Semantics notion (Fillmore et al., 1976), FrameNet is a coherent and exhaustive computational lexicography that provides in-depth semantic information regarding the argument structure and thematic relations of a predicate.

In FrameNet, predicates are annotated into their respective frames. A frame refers to a schematic representation that brings lemmas together based on their semantic properties and syntactic features (for a more detailed definition and discussion of the frame notion, see (Fillmore et al., 1976).

Our Motivation

Thus far, mostly Indo-European languages followed suit with the English FrameNet. These languages are relatively poorer in morphology compared to agglutinative languages like Turkish. That is why the annotation tools used in the creation processes of other FrameNet projects above do not offer a morphologic analyser component which is essential for morphologically richer languages like Turkish. Our annotation tool, FrameForm allows adding a morphological analyser and introducing a dictionary and/or WordNet. That is why different languages can utilise FrameForm for their annotation processes simply by making some minor adjustments in the back-end.

Features of the FrameForm

FrameForm saves every annotation pertaining to a Lexical Unit in a single file: Morphologic analysis, predicate analysis (shows which word or group of words is the predicate), semantic analysis (maps the related meaning to the word), frame information and frame elements. This way, one can find all the necessary information regarding a Lexical unit with one click.

What Can be Annotated Using FrameForm?

FrameForm is a very potent tool for annotation. It allows the user to:

- Create new frames,
- Transfer data between the frames,
- Manually edit or change sample sentences,
- Delete Lexical Units,
- Do morphologic analysis, semantic annotation, predicate marking and frame element annotation.

Different Interfaces of the FrameForm

FrameForm has 4 different screens for each step of the annotation process: Morphologic analysis screen (see Figure 1), semantic annotation screen (see Figure 2), predicate marking screen (see Figure 3) and frame element annotation screen (see Figure 4).

Morphological analysis interface

Using a morphological analysis tool or library, FrameForm offers an automatic morphologic analysis to speed up the process.

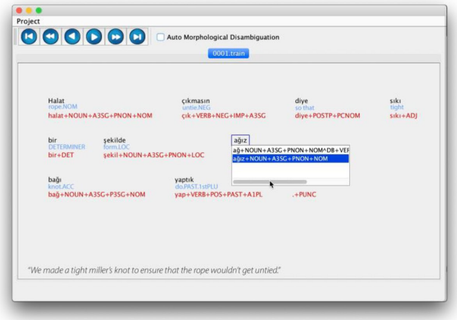

Figure 1 Morphologic analysis interface

Semantic annotation interface

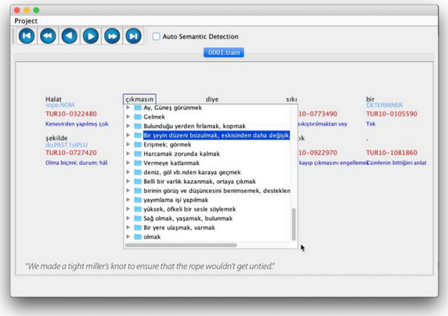

Figure 2 Semantic annotation interface

In this step, the annotator should select the correct meaning of the Lexical Unit in regard to the frame, and appropriate meanings of the other elements in the sample sentence as well. For the semantic annotation step, FrameForm refers to a dictionary or WordNet.

Predicate marking interface

Here the annotator can mark the predicate/Lexical Unit in the sample sentence.

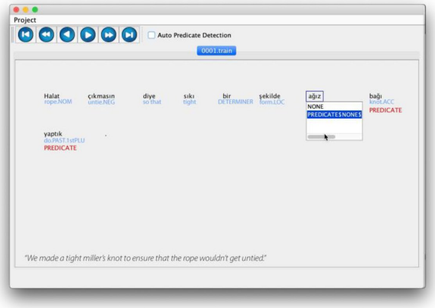

Figure 3 Predicate marking interface

Frame element annotation interface

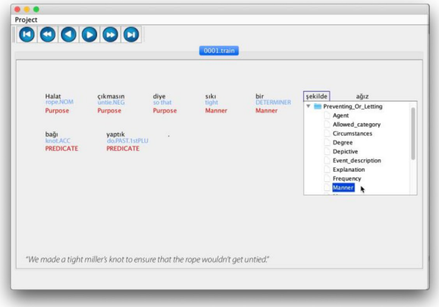

Figure 4 Frame element annotation interface

Final step is annotating the Frame Elements where the annotator can see all the FEs within that frame and match them with related sentence elements.

Conclusion

It is possible to create a FrameNet from scratch, grow it and maintain it using only FrameForm. FrameForm can be downloaded freely on GitHub. Being easy to access and distribute, a crowded team of annotators can use Frame-Form for their annotation needs. Since annotators can see progress made by other members of the team, FrameForm makes it easier to ensure inter-annotator agreement.

References

Collin F. Baker, Charles J. Fillmore, and John B. Lowe. 1998. The Berkeley FrameNet project. In 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics, Volume 1, pages 86–90, Montreal, Quebec, Canada. Association for Computational Linguistics.

Charles J Fillmore and Beryl TS Atkins. 1998. Framenet and lexicographic reference. In First International Conference on language resources & evaluation: Granada, Spain, 28-30 May 1998, pages 417– 426. European Language Resources Association.

Charles J Fillmore et al. 1976. Frame semantics and the nature of language. In Annals of the New York Academy of Sciences: Conference on the origin and development of language and speech, volume 280, pages 20–32. New York.

Christopher Johnson, Charles J Fillmore, E Wood, Josef Ruppenhofer, Margaret Urban, MIRIAM Petruck, and COLLIN Baker. 2001. The framenet project: Tools for lexicon building. Manuscript. Berkeley, CA, International Computer Science Institute.

Büşra Marşan, Neslihan Kara, Merve Özçelik, Bilge Nas Arıcan, Neslihan Cesur, Aslı Kuzgun, Ezgi Sanıyar, Oğuzhan Kuyrukçu, and Olcay Taner Yildiz. 2021. Building the Turkish FrameNet. In Proceedings of the 11th Global Wordnet Conference, pages 118–125, University of South Africa (UNISA). Global Wordnet Association.