Exploring Crosslinguisitic FrameNet Alignment

Collin F. Baker, Arthur Lorenzi
International Computer Science Institute
1947 Center St. Suite 600
Berkeley, CA, 94704, U.S.A.
`{collinb,lorenzi}@icsi.berkeley.edu`

Abstract

The FrameNet (FN) project at the International Computer Science Institute in Berkeley (ICSI), which documents the core vocabulary of contemporary English, was the first lexical resource based on Fillmore's theory of Frame Semantics. Berkeley FrameNet has inspired related projects in roughly a dozen other languages, which have evolved somewhat independently; the current Multilingual FrameNet project (MLFN) is an attempt to find alignments between all of them. The alignment problem is complicated by the fact that these projects have adhered to the Berkeley FrameNet model to varying degrees, and they were also founded at different times, when different versions of the Berkeley FrameNet data were available. We describe several new methods for finding relations of similarity between semantic frames across languages. We will demonstrate ViToXF, a new tool which provides interactive visualizations of these cross-lingual relations, between frames, lexical units, and frame elements, based on resources such as multilingual dictionaries and on shared distributional vector spaces, making clear the strengths and weaknesses of different alignment methods .

Keywords: frame semantics, lexical semantics, semantic frames, multilingual lexical resources

1. Introduction

1.1. Frame Semantics and FrameNet

NLP researchers have long sought to develop tools and resources to build meaning representations beyond the word or syntax level, and many have looked to Charles J. Fillmore's theory of Frame Semantics (Fillmore, 1977b; Fillmore, 1977a) as part of the solution. Fillmore and his colleagues founded the FrameNet (FN) Project (Fillmore and Baker, 2010; Fontenelle, 2003) at the International Computer Science Institute (ICSI) in 1997 with the goal of establishing a general-purpose resource for frame semantic descriptions of English language text. FrameNet's lexicon is organized not around words, but **semantic frames** (Fillmore, 1976), which are characterizations of events, static relations, states, and entities. Each frame provides the conceptual basis for understanding a set of word senses, called **lexical units (LUs)**, that **evoke** the frame in the mind of the hearer; LUs can be any part of speech, although most are nouns, verbs, or adjectives. FrameNet now contains roughly 1,200 frames and 13,600 LUs.

FrameNet provides very detailed information about the syntactic-semantic patterns that are possible for each LU, derived from expert annotations on naturally occurring sentences. These annotations label the phrases that instantiate the set of roles involved in the frame, known as **frame elements (FEs)**. An example of a simple frame is **Placing**, which represents the notion of someone (or something) placing something in a location; this frame is **evoked** by LUs like *place.v, put.v, lay.v, implant.v*, and *billet.v* and also *bag.v, bottle.v*, and *box.v,* The core frame elements of **Placing** are the AGENT who does the placing (or the CAUSE of the placing), the THEME that is placed, and the GOAL. An example of a more complex frame is **Revenge**, which has FEs AVENGER, INJURED PARTY, INJURY, OFFENDER, and PUNISHMENT.

The FrameNet lexical database, in XML format, has been downloaded more than 3,000 times by researchers and developers around the world; the well-known NLP library NLTK (Loper and Bird, 2002) also provides API access to FrameNet (Schneider and Wooters, 2017). FrameNet's main publications have been cited over 2,500 times according to Google Scholar, and it has been an important basis for at least 14 PhD dissertations.

The wide use of FrameNet in NLP depends on the success of systems for automatic semantic role labeling (ASRL) of unseen text, trained on the FrameNet annotation data. ASRL then enables (or improves) downstream NLP applications, such as

- Question Answering (Shen and Lapata, 2007; Sinha, 2008)
- Information Extraction (Surdeanu et al., 2003)
- Text-to-scene generation (Coyne et al., 2012)
- Dialog systems (Chen et al., 2013)
- Social network extraction (Agarwal et al., 2014)
- Knowledge Extraction from Twitter (Søgaard et al., 2015)

In fact, automatic semantic role labeling has become one of the standard tasks in NLP, and several freely available FrameNet-based ASRL systems have been developed, including SEMAFOR (Das et al., 2010; Das et al., 2014) and open-sesame (Swayamdipta et al., 2018). The latter jointly exploits PropBank-based (Palmer et al., 2005) semantic role labeling and FrameNet to train a neural net (NN) to do frame and FE discrimination without run-time parsing. Other recent FrameNet-based ASRL systems have tried a variety of new approaches:

- FitzGerald et al. (2015) train a NN representing joint embeddings of PropBank and FrameNet roles,
- Kshirsagar et al. (2015) use structured features and the frame hierarchy,
- Roth and Lapata (2015) predict roles based both on the

entire document and on recent role assignments, and

- Peng et al. (2018) jointly model FrameNet roles and dependency parses.

1.2. Multilingual FrameNet

Since the beginning of Frame Semantics, the question has arisen as to whether semantic frames represent "universals" of human language or are language specific. Despite many language-specific patterns of expression, the conclusion from the FrameNet experience has been that many frames are applicable across different languages, especially those for basic human experiences, like eating, drinking, and sleeping. Even some cultural practices are similar across languages: e.g. in every culture, commercial transactions involve the roles BUYER, SELLER, MONEY, and GOODS (or services).

Since the Berkeley FrameNet (hereafter BFN) project began releasing its data, researchers in many countries have expressed interest in creating comparable resources for other languages; in fact, the BFN team is in contact with about a dozen FrameNets in languages other than English [1]. The methods used in building these FrameNets have differed, and each has created frames based on their own linguistic data, but all at least have an eye to how their frames compare with those created for English at ICSI (Boas, 2009).

Given that so much research has been conducted in building separate lexical databases for many languages, it is natural to ask whether these lexical databases can be aligned to form a multilingual FrameNet lexical database connecting all of the languages (as well as new FrameNets that may arise in the future), while also accounting for language-specific differences and domain-specific extensions to FrameNet. The results produced so far suggest that this is possible. It is also urgent to carry out this harmonization process as soon as possible, to take better advantage of the experience of each language project, to avoid duplication of effort, and to unify the representational format as much as possible. A number of FrameNet groups, led by FrameNet Brasil have also established the Global FrameNet[2] project to improve communication between FrameNets.

Despite differences among the FrameNet projects, all agree on the concept of semantic frames as their organizing principle and all have found the set of frames defined in BFN to be generally applicable to their language. For example, all languages have ways to express directed self-motion, which involves the frame elements MOVER, SOURCE, PATH and GOAL (although it is rare for all of these to be expressed in the same clause). Likewise, whenever a communicative act occurs, we can identify the FEs SPEAKER, ADDRESSEE, and TOPIC or MESSAGE, which are common to all the communication frames. Semantic frames thus should provide useful generalizations both over lexical units within a language and across languages.

However, the projects have adhered to the Berkeley FrameNet (BFN) model to different degrees: The Spanish, Japanese, and Brazilian Portuguese FNs have followed BFN rather closely, using BFN frames as templates, whereas the SALSA Project (for German), French FN, Swedish FrameNet++ and Chinese FN have diverged further from BFN, either adding many new frames and/or modifying the BFN-derived ones.[3]

More fundamentally, there is no reason to assume that cross-linguistic frame relations will be limited to equivalence. Frames in other languages can be broader or narrower than the nearest English frame, or similar situations may require a different point of view in different languages. For example, English *I like X*, where *like*.v is in the **Experiencer_focused_emotion** frame (along with *adore, dread* and *regret*), is regularly translated as Spanish *Me gusta X*, 'X pleases me' where *gustar*.v is in the **Experiencer_object** frame (along with *asombrar* 'astound', *chocar* 'shock', and *molestar* 'bother', cf. Subirats-Rüggeberg and Petruck (2003)). Other well-attested differences in information structure between languages are similarly reflected in differences in choice of frames, such as that between satellite-framed languages like English and German and verb-framed languages like Spanish and Japanese (Slobin, 1996). There will also be cultural differences, which may mean that equivalent frames do not exist, such as frames for religions or legal processes, which differ widely from country to country.[4] The Multilingual FrameNet project (Gilardi and Baker, 2018) is studying the relations between frames in different languages and will distribute a database of alignments between FrameNets. We have developed several approaches to calculating frame similarity to produce the cross-lingual alignments; these are described in Sections 2.1. through 2.4.. In order to compare these approaches and to evaluate their strengths and weaknesses under various settings of parameters, we have also built an interactive tool for visualization of frame alignments, called ViToXF (for "visualization tool across FrameNet"). We describe this tool in Sec. 3., and will demonstrate it at the workshop. Finally, Sec. 4. offers some qualitative evaluations of the alignment methods and discusses directions for future research.

2. Cross-lingual alignment and Visualization Techniques

Table 1 gives counts for frames and LUs for the six languages included in the preliminary version of the visualization tool; in some cases, these numbers may understate the total in each project, due to certain difficulties in importing the data.

2.1. Alignment by frame name/ID

At first glance, the alignment problem seems trivial: if the other FrameNets have largely used BFN frames, one might

---

[3] At this time, the MLFN effort is not trying to align the Italian, Arabic, or Hebrew data, for various reasons, including availability and coverage, among others.

[4] Of course, languages frequently develop, borrow, or calc terms and frames for concepts that are not "native" to the language, in order to discuss other cultures.
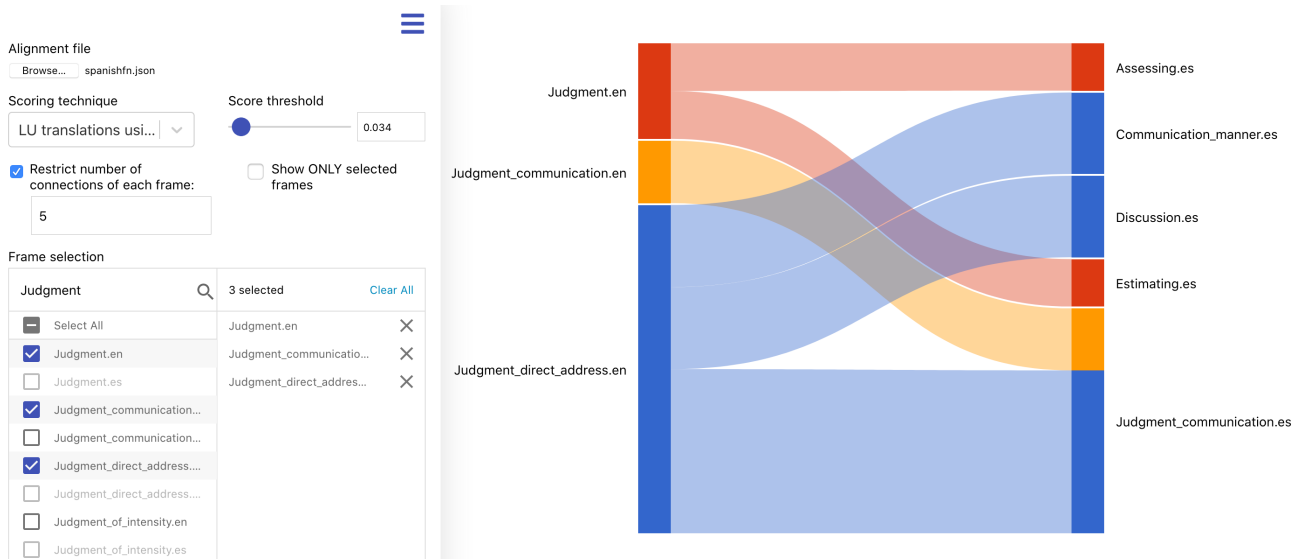
Figure 1: Visualization tool, showing user controls and resulting Sankey diagram of English Judgment frames aligned with Spanish Frames using LU translation by linking synsets at a high score threshold (see Sec. 2.2. for details)

.

Project	Frames	Lexical Units
FrameNet (ICSI) = BFN	1,224	13,675
Chinese FN	1,259	20,551
FN Brasil (PT)	1,092	2,896
French FN (Asfalda)	148	2,590
German FN	1,023	1,826
Japanese FN	984	3392
Spanish FN	1,196	11,352
Swedish FN	1,186	38,749

Table 1: Frame and LU counts of FrameNets now in ViToXF

just assume that a frame in another language with the same name as a BFN frame represents the same concept, and ignore any that don't have matching names. However, as might be expected, some of the other languages have used frame names in the target language, rather than English; this would mean aligning the frame names themselves across languages. In some cases, their frame data also includes a field for the BFN name or BFN ID, which can be used for alignment, even when the frame names are not in English. Furthermore, even when the names (or IDs) match, the non-English frame may be defined differently or have more or fewer core frame elements than the BFN frame, which, strictly speaking, makes it a different frame.

2.2. Alignment by LU translation

A second way of approaching alignment is to take all the lexical units from a source language frame and find translation equivalents in the target language. To the extent that frames are equivalent across languages, we would expect all the translations of LUs in one source language frame to fall into one target language frame. Of course, this depends on the accuracy of the translations. By definition, a lexical unit in a frame represents one sense of a lemma, so in theory that should greatly narrow the range of possible translations; however, exactly how to use information from frames and frame relations in the translation process is still to be determined.

The Open Multilingual WordNet (OMWN) (Bond and Foster, 2013) contains multilingual synsets, combining lemmas from WordNets for dozens of languages, data from Wiktionary, and the Common Locale Data Repository.[5] We are currently using it to find a set of translation equivalents between languages. The first step is to create a mapping $S(\ell)$ from each LU in each language to a set of OMWN synsets that represent its senses. That mapping is created by searching OMWN for synsets that contain the lemma (with the correct part of speech) of the LU. More formally, let e be a frame in the source language and f a frame in the target language; let L_e and L_f be the lists of the LUs in frames e and f respectively. Then any two LUs a and b in L_e and L_f (respectively) match if they occur together in at least one synset; this matching function can be expressed by Equation 1.

$$\mathrm{m}(L_e, L_f) = \{a \in L_e \mid b \in L_f : S(a) \cap S(b) \neq \emptyset\} \quad (1)$$

When evaluating the alignment between two frames, this function was used to calculate three different scores. The first is a metric that takes into consideration LUs from both frames (Equation 2); however, this gives frames containing more LUs more influence over the result. To avoid this problem, we decided to break the alignment into two other scores taking into account the direction of alignment, i.e., the score of the alignment from English to the target language can be different from the reverse. The basic formula for those scores is presented in Equation 3. (Note that the two scores can be obtained by simply swapping the arguments

---

[5] http://cldr.unicode.org/

in the equation.)[6]

$$s_1(L_e, L_f) = \frac{|m(L_e, L_f)| + |m(L_f, L_e)|}{|L_e| + |L_f|} \quad (2)$$

$$s_2(L_e, L_f) = \frac{|m(L_e, L_f)|}{|L_e|} \quad (3)$$

We also explored another alternative scoring method that is available in the visualization tool by selecting the "Synset count" scoring method. This score is calculated using Equation 3.

$$s_3(L_e, L_f) = \frac{\left|\bigcup_{a \in L_e} S(a) \cap \bigcup_{b \in L_f} S(b)\right|}{\left|\bigcup_{a \in L_e} S(a)\right|} \quad (4)$$

2.3. Alignment by frame element similarity

By definition, for two frames to be the same across languages, they must have the same number and type of frame elements (FEs). Some FrameNets (such as Spanish FN and Japanese FN) have simply copied the FEs from Berkeley FrameNet, so that their names and definitions are still identical to BFN. Others, such as Chinese FN, have translated or created both the names and the definitions in the target language; in those cases, we need to align the FEs by using the proximity of the names and definitions from the two languages in a shared vector space. French FN created FE names and definitions in English, even though many of their frames do not correspond to BFN. Swedish has FE names in English, but no definitions; since they state that the frames and FEs with English names are intended to be identical to the BFN frames and FEs of the same name, the English definitions should also apply to them. Finally, both Brazilian Portuguese and German (SALSA) have FEs in a mixture of English and the target language. In those two cases, we group the FEs according to whether they are in English or the target language,[7] calculate the similarity separately for the two groups, and then combine the scores.

2.4. Alignment by distributional similarity of LUs

Another approach to alignment is to use cross-lingual word embeddings to obtain translations equivalents. The current iteration of the visualization tool uses the FastText word embeddings from FaceBook Research, which were trained on Wikipedia data from various languages and aligned to a single embedding space (Bojanowski et al., 2017). The spaces were aligned by an unsupervised method that uses an adversarial approach, where the discriminator tries to predict the embedding origin and the generator aims to create transformations that the former is not able to accurately classify (Conneau et al., 2017). The transformed FastText vectors of many languages mapped to English space were made publicly available in the MUSE library.[8] We are currently using these pre-trained cross-lingual word embeddings for

two different scoring techniques. The first, "LU translations using MUSE", like those discussed above based on OMWN, uses the word embeddings as a way to obtain translation equivalents: we define the neighborhood around the vector embedding of a target language word as $n(\vec{v}, k, t)$, that is, the k-neighborhood of $\vec{v}$ in the target language with a cosine similarity greater than t. Then we define the alignment score between a pair of frames given their LU lists L_e and L_f by Equation 5.

$$\begin{aligned} s_4(L_e, L_f) = \\ \frac{|\{a \in L_e \mid b \in L_f : \vec{v}(b) \in n(\vec{v}(a), k, t)\}|}{|L_e|} \end{aligned} \quad (5)$$

The second scoring technique, "LU centroid similarity using MUSE", calculates the alignment between two frames by finding the average of the vectors of their LUs (i.e. the centroid vector of each frame) and computing the cosine similarity of those two centroids, similar to the approach of Sikos and Padó (2018).

3. Alignment Visualization Tool

3.1. Frame Alignment example

We will demonstrate the alignment of three related English frames with Spanish, **Judgment, Judgment_communication**, and **Judgment_direct_address**. The **Judgment** frame applies whenever a person (the COGNIZER) forms an opinion (good or bad) about someone or something (the EVALUEE). In the **Judgment_communication** frame, the COGNIZER, now called the COMMUNICATOR expresses that opinion, possibly to an ADDRESSEE. In the frame **Judgment_direct_address**, the ADDRESSEE is also the one being evaluated, so this frame contains LUs like *congratulate,harangue, scold, take to task* and *tell off.* The relations between these frames and their frame elements are spelled out in detail in FrameNet; the **Judgment_communication** frame **uses** two frames, **Judgment** frame and **Statement**, and **Judgment_direct_address** **inherits** from **Judgment_communication**.

3.2. Visualization modes

In its current iteration, the system has two visualization modes, one that uses a Sankey diagram to show alignments between frames and another that displays the translations between the LUs of a frame pair in a different type of graph.

Frame Alignment Visualization: Fig. 1 shows the main visualization mode of the tool. It is an interactive bipartite Sankey diagram where English frames are displayed on the left side and target language frames on the right. The width of each band in the diagram is proportional to the alignment score between the frame pair.

Due to the number of lexical units in the FrameNet projects, the resulting diagram can be very dense, making analysis difficult. To alleviate this problem, ViToXF allows both frame selection and band filtering. Frames to be shown can be selected from a list of all the English and target language frames, (marked with a two-letter suffix, *en* for English, *es* for Spanish and so forth). By default, the system will display any match that includes one of the selected frames.

---

[6]Both scores can also be visualized with the "LU translations using WordNet" options, cf.Sec. 2.2.

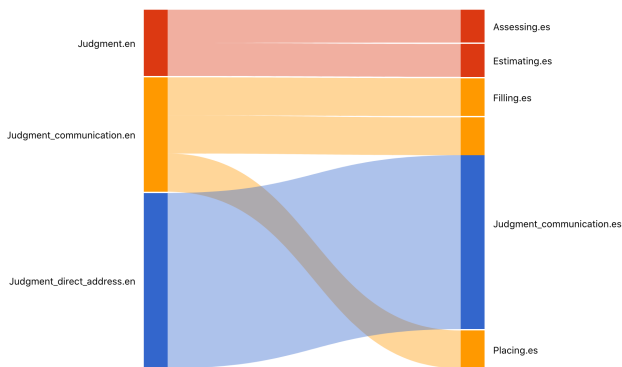[7]Using Michal Danilak's python library for language recognition tion https://pypi.org/project/langdetect/

[8]https://github.com/facebookresearch/MUSE

Figure 2: Frame Alignment: English ⇒ Spanish for Judgment-related Frames, matched using synsets at a lower score threshold

This can be further restricted by checking "Show ONLY selected frames"; in that case, it displays only connections where **both** frames are selected.

To filter which bands are displayed in the diagram, the user can also set an alignment score threshold (so that weaker alignments will not be shown) and/or set a limit on the number of connections from each frame. When the number of connections is restricted, those with the highest scores will always be displayed first.

When the "LU-based using MUSE vectors" is selected as the scoring technique for LU matching between frames, the parameters k (neighborhood size) and t (distance threshold) of the function $n(\vec{v}, k, t)$ described in Subsection 2.4., Eq. 5 can be modified, potentially changing alignment scores and hence, the graph displayed. Fig. 1 shows the sidebar, where all of these parameters can be controlled by the user.

Fig. 2 shows the same English-Spanish alignment of frames related to judgment, with a slightly lower similarity threshold than in Fig. 1. Note that alignments to two additional frames **Placing** and **Filling** have now appeared; this will be explained in the following section.

Lexical Unit Translation Visualization: This visualization mode is intended to demonstrate exactly how translations were found for the LUs of a frame pair, and can be accessed by clicking on any band in the Sankey diagram. ViToXF provides two methods for aligning frames and LUs across languages, one based on synsets, the other on vector embeddings; depending on which method is used for the Sankey diagram, the LU translation visualization will be somewhat different.

In both cases, the translation visualization is a tripartite graph with vertices organized in three columns: the left column is composed of the LUs of the BFN frame, the right column of the LUs of the target language frames. In the case of the synset-based LU translation method, the middle column lists the names of synsets, and edges are drawn between the synsets and the LUs in each language whose lemma+POS occurs in that synset. If an English LU and a

target language LU both match a lemma+POS in a synset, the name of that synset (or LU depending on the scoring method) is shown in green, and the overall matching score is raised. If an LU from one language matches the synset but not from the other language, the synset name or LU is yellow; this adds to the denominator of Equations 2, 3, 4 and 5 reducing the overall matching score. Synsets which match no LU in the source language are colored black; they do not influence the score.

When the vector embedding method is being used, the left and right columns are as described above, but the center column now represents wordforms; edges are connected to LUs in either language whose lemma lies within the neighborhood of the wordform in the embedding space. If the FastText vectors are used, this means that subparts of words play a role, and that may help connect the various wordforms of a lemma, but may also lead to false positives. Part of speech is not used. The meanings of the colors in the central column are as described above.

Continuing with our example of aligning from English to Spanish in the Judgment-related frames using LU translation via synsets, Fig. 3 shows how the LUs in each language link to the lemmas in the synsets of OMWN. Fig. 3 also shows what can go wrong: the lemma *charge*.v appears in the BFN **Judgment_communication** frame, but in OMWN it also appears in a synset with English *load*.v and Spanish *cargar*.v, defined as 'provide (a device) with something necessary'. Thus it links erroneously to the **Filling** frame in both languages; this problem is discussed further in the next section.

4. Discussion

Each new FrameNet constitutes an experiment in cross-linguistic Frame Semantics. Motivated by the fundamental research question "To what extent are semantic frames similar across languages?", ViToXF provides an intuitive, graphical, interactive tool to study a variety of methods for finding the relations between frames and lexical units across different languages. It also highlights some of the problems that need to be solved to create meaningful alignments that are useful for a wide range of NLP tasks. We do not yet have not results from testing these alignments against standard NLP tasks, but this section offers some qualitative evaluations of the methods and results so far.

4.1. Evaluating Synset-based methods

The alignment methods that depend on WordNet synsets have the merit that they take advantage of large-scale curated groupings of lemmas (by part of speech). However, they also make clear one problem with WordNets: for many common words, the number of senses given is simply too high. We noted in the preceding section some problems with the polysemy of *charge*.v; in fact, Princeton WordNet lists 31 senses for the verb alone! Some of the major divisions are clear:

- charge#1, bear down#3 (to make a rush at or sudden attack upon, as in battle) *He saw Jess charging at him with a pitchfork*
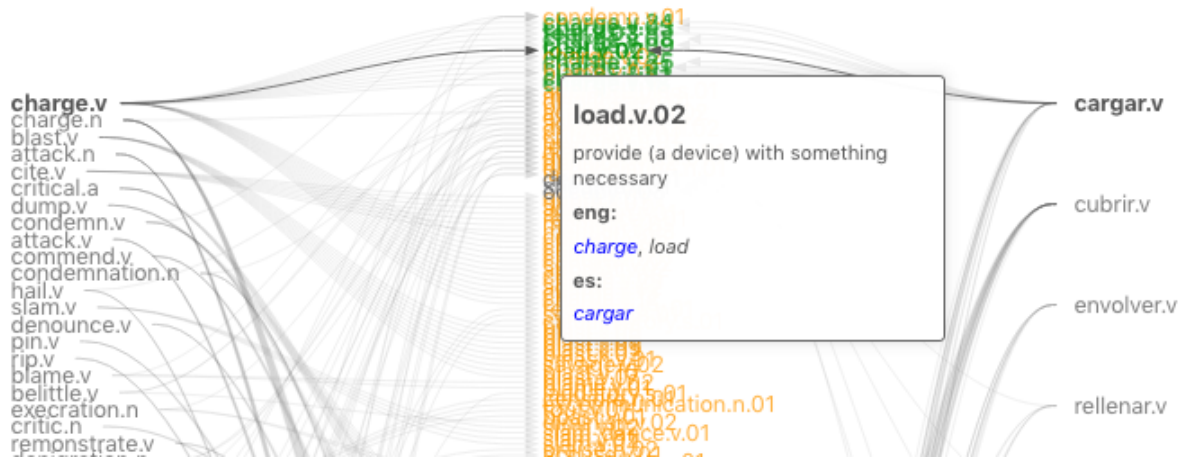
Figure 3: Lexical unit translation: English ⇒ Spanish for Judgment-related Frames, matched using synsets at a lower score threshold. In addition to the expected alignment to **Judgment_communication**, *charge*.v is also mapped to a synset in the **Filling** frame.

- charge#3, bill#1 (demand payment)*Will I get charged for this service?*
- appoint#2, charge#5 (assign a duty, responsibility or obligation to) *She was charged with supervising the creation of a concordance*
- charge#6, lodge#3, file#4 (file a formal charge against) *The suspect was charged with murdering his wife*
- charge#24 (energize a battery by passing a current through it in the direction opposite to discharge) *I need to charge my car battery*

But some senses are hard for humans to distinguish, let alone algorithms; compare for example:

- charge#2, accuse#2 (blame for, make a claim of wrongdoing or misbehavior against) *He charged the director with indifference* and
- charge#7 (make an accusatory claim) *The defense attorney charged that the jurors were biased.*

Are these separate from each other? How are they related to #6?

- charge#8 (fill or load to capacity) *charge the wagon with hay* and
- load#2, charge#16 (provide (a device) with something necessary) *He loaded his gun carefully.*

Are these the same as #24? Is #24 just a special case of charge#16?

The Ontonotes lexical resource (Pradhan et al., 2013), which is based on combining WordNet senses so that annotators can reliably distinguish the classes, may provide a coarser but more reliable list of senses for English and Chinese, but it does not include the other FrameNet language pairs.

4.2. Evaluating Vector-based methods

The alignment methods based on vector embeddings have the advantage of making it possible to measure distances between uses, distances which are arguably semantic; however these distances are not easily converted to "senses" that humans can understand. Also, the MUSE embeddings, like most distributional embeddings, are based on word forms, and do not generalize to the level of lexemes (e.g. most lexicographers would expect the verb *go* to be represented by a single vector that covers *go, went, gone, goes*, and *going*, rather than separate vectors for the five word forms). These embeddings also do not include Chinese and Japanese.

The major shortcoming of the current distributional embeddings, however, is that they provide only one vector per word form, with no distinction of senses. However, there have been encouraging results on finding embeddings for word **senses**, such as Upadhyay et al. (2016) who use multilingual corpora to learn sense-specific embeddings. They point out that often patterns of polysemy are similar across languages; continuing with the preceding example, English *charge* and Spanish *cargar* can both mean either 'file charges in court' or 'fill a battery with electricity'. However, adding an unrelated language such as Chinese often gives completely different translations:

4.3. Applications and Future Work

A major limitation of ViTOXF is simply that most FrameNets are rather small in comparison with other lexical resources, primarily because of the amount of human curation needed to produce them. However, interest in Frame Semantics continues to grow and new FrameNet projects are appearing frequently, so there may be a continuing interest in finding alignments for them. There are also numerous approaches to automatically or semi-automatically adding lexical units to FrameNets (e.g. Pavlick et al. (2015), Fossati et al. (2013),Hartmann and Gurevych (2013),Green (2004)), offering the prospect of much larger, if less precise, lexical inventories.

We expect that alignments produced by the methods outlined here and refined by the use of ViTOXF will prove useful to

- translators and second language learners seeking to understand cross-linguistic differences in framing;
- developers of MT systems, parsers, and grammars (especially for languages for which FrameNets already

exist) (e.g. Czulo et al. (2019)); and, of course,

- cognitive linguists and researchers creating new FrameNets.

Since FastText does not provide cross-linguistic embeddings for English-Japanese and English-Chinese, we will attempt to train some ourselves, to make that type of alignment available for them. We may be able to find ways to use the annotated sentences themselves to align frames, possibly using methods related to BERT vector embeddings, such as those of Zhang et al. (2020).

As just mentioned, instances of similar polysemy can usually be split apart by looking simultaneously at more languages, especially if the languages are unrelated. We therefore plan to look for frames which align well across three or more languages, making for highly robust alignments. Our immediate goal is to incorporate as many of the current FrameNet projects as possible.

We also plan to explore methods for creating sense-specific vectors in all the languages, and better techniques for finding translation equivalents; for example, a smaller number of translations from an MT system may prove more accurate than those from OMWN synsets. Finally, it should be clear that there are many ways to combine the similarity scores from the different methods to get an overall score between two frames. We plan to test the advantages and disadvantages of various weighted linear combinations of scores for different applications. The current code for the visualizer is essentially an alpha version; we welcome suggestions for improving the user interface. We will make the code for ViToXF available on Github; a demo version of the visualizer is available now at `https://icsi-berkeley.github.io/framenet-multilingual-alignment/`.

Acknowledgments

Bibliographical References

Agarwal, A., Balasubramanian, S., Kotalwar, A., Zheng, J., and Rambow, O. (2014). Frame semantic tree kernels for social network extraction from text. In *EACL*, pages 211–219.

Hans C. Boas, editor. (2009). *Multilingual FrameNets in Computational Lexicography: Methods and Applications*. Mouton de Gruyter.

Bojanowski, P., Grave, E., Joulin, A., and Mikolov, T. (2017). Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 5:135–146.

Bond, F. and Foster, R. (2013). Linking and extending an Open Multilingual Wordnet. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1352–1362, Sofia, Bulgaria, August. Association for Computational Linguistics.

Chen, Y.-N., Wang, W. Y., and Rudnicky, A. I. (2013). Unsupervised induction and filling of semantic slots for spoken dialogue systems using frame-semantic parsing. In *Automatic Speech Recognition and Understanding Workshop*, pages 120–125. IEEE.

Conneau, A., Lample, G., Ranzato, M., Denoyer, L., and Jégou, H. (2017). Word translation without parallel data. *arXiv preprint arXiv:1710.04087*.

Coyne, B., Klapheke, A., Rouhizadeh, M., Sproat, R., and Bauer, D. (2012). Annotation Tools and Knowledge Representation for a Text-To-Scene System. In *COLING 2012, 24th International Conference on Computational Linguistics, Proceedings of the Conference: Technical Papers*, pages 679–694.

Czulo, O., Torrent, T., Matos, E., Costa, A. D. d., and Kar, D. (2019). Designing a frame-semantic machine translation evaluation metric. In *Proceedings of the Human-Informed Translation and Interpreting Technology Workshop (HiT-IT 2019)*, pages 28–35, Varna, Bulgaria, September. Incoma Ltd., Shoumen, Bulgaria.

Das, D., Schneider, N., Chen, D., and Smith, N. A. (2010). Probabilistic Frame-Semantic Parsing. In *Proceedings of the North American Chapter of the Association for Computational Linguistics Human Language Technologies Conference*, Los Angeles, June.

Das, D., Chen, D., Martins, A. F. T., Schneider, N., and Smith, N. A. (2014). Frame-Semantic Parsing. *Computational Linguistics*, 40(1).

Fillmore, C. J. and Baker, C. F. (2010). A Frames Approach to Semantic Analysis. In Bernd Heine et al., editors, *Oxford Handbook of Linguistic Analysis*, pages 313–341. OUP. This is Chapter 13 in 1st edition, 33 in 2nd edition.

Fillmore, C. J. (1976). Frame semantics and the nature of language. *Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech*, 280(1):20–32.

Fillmore, C. J. (1977a). Scenes-and-frames semantics. In Antonio Zampolli, editor, *Linguistic Structures Processing*, number 59 in Fundamental Studies in Computer Science. North Holland Publishing.

Fillmore, C. J. (1977b). The need for a frame semantics in linguistics. In Hans Karlgren, editor, *Statistical Methods in Linguistics*. Scriptor.

FitzGerald, N., Täckström, O., Ganchev, K., and Das, D. (2015). Semantic role labeling with neural network factors. In *EMNLP*, pages 960–970.

Thierry Fontenelle, editor. (2003). *International Journal of Lexicography–Special Issue on FrameNet*, volume 16. Oxford University Press.

Fossati, M., Giuliano, C., and Tonelli, S. (2013). Outsourcing FrameNet to the Crowd. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 742–747, Sofia, Bulgaria, August. Association for Computational Linguistics.

Gilardi, L. and Baker, C. F. (2018). Learning to align across languages: Toward Multilingual FrameNet. In Tiago Timponi Torrent, et al., editors, *Proceedings of the International FrameNet Workshop 2018: Multilingual FrameNets and Constructicons*, pages 13–22. LREC.

Green, R. (2004). *Inducing Semantic Frames from Lexical Resources*. Ph.D. thesis, University of Maryland, College Park.

Hartmann, S. and Gurevych, I. (2013). FrameNet on the Way to Babel: Creating a Bilingual FrameNet Using Wiktionary as Interlingual Connection. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1363–1373, Sofia, Bulgaria, August. Association for Computational Linguistics.

Kshirsagar, M., Thomson, S., Schneider, N., Carbonell, J., Smith, N. A., and Dyer, C. (2015). Frame-Semantic Role Labeling with Heterogeneous Annotations. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 218–224, Beijing, China, July. Association for Computational Linguistics.

Loper, E. and Bird, S. (2002). NLTK: The natural language toolkit. In *In Proceedings of the ACL Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics. Philadelphia: Association for Computational Linguistics*.

Palmer, M., Gildea, D., and Kingsbury, P. (2005). The proposition bank: An annotated corpus of semantic roles. *Computational Linguistics*, 31(1):71–106.

Pavlick, E., Wolfe, T., Rastogi, P., Callison-Burch, C., Dredze, M., and Van Durme, B. (2015). FrameNet+: Fast Paraphrastic Tripling of FrameNet. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 408–413, Beijing, China, July. Association for Computational Linguistics.

Peng, H., Thomson, S., Swayamdipta, S., and Smith, N. A. (2018). Learning joint semantic parsers from disjoint data. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 1492–1502. Association for Computational Linguistics.

Pradhan, S., Moschitti, A., Xue, N., Ng, H. T., Björkelund, A., Uryupina, O., Zhang, Y., and Zhong, Z. (2013). Towards robust linguistic analysis using OntoNotes. In *Proceedings of the Seventeenth Conference on Computational Natural Language Learning*, pages 143–152, Sofia, Bulgaria, August. Association for Computational Linguistics.

Roth, M. and Lapata, M. (2015). Context-aware Frame-Semantic Role Labeling. *Transactions of the Association for Computational Linguistics*, 3:449–460.

Schneider, N. and Wooters, C. (2017). The NLTK FrameNet API: Designing for discoverability with a rich linguistic resource. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 1–6, Copenhagen, Denmark, September. Association for Computational Linguistics.

Shen, D. and Lapata, M. (2007). Using Semantic Roles to Improve Question Answering. In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*, pages 12–21.

Sikos, J. and Padó, S. (2018). Using embeddings to compare framenet frames across languages. In *Proceedings of the First Workshop on Linguistic Resources for Natural Language Processing*, pages 91–101. Association for Computational Linguistics. Collocated with COLING 2018.

Sinha, S. (2008). *Answering Questions about Complex Events*. Ph.D. thesis, EECS Department, University of California, Berkeley, Dec.

Slobin, D. I. (1996). Two ways to travel: Verbs of motion in English and Spanish. In M. S. Shibatani et al., editors, *Grammatical constructions: Their form and meaning*, pages 195–220. Clarendon Press, Oxford.

Subirats-Rüggeberg, C. and Petruck, M. R. L. (2003). Surprise: Spanish FrameNet! In Eva Hajičová, et al., editors, *Proceedings of the Workshop on Frame Semantics, XVII International Congress of Linguists (CIL)*, Prague. Matfyzpress.

Surdeanu, M., Harabagiu, S., Williams, J., and Aarseth, P. (2003). Using Predicate-Argument Structures for Information Extraction. In Erhard Hinrichs et al., editors, *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics*, pages 8–15.

Swayamdipta, S., Thomson, S., Lee, K., Zettlemoyer, L., Dyer, C., and Smith, N. A. (2018). Syntactic scaffolds for semantic structures. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3772–3782, Brussels, Belgium, October-November. Association for Computational Linguistics.

Søgaard, A., Plank, B., and Alonso, H. M. (2015). Using Frame Semantics for knowledge extraction from Twitter. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*.

Upadhyay, S., Faruqui, M., Dyer, C., and Roth, D. (2016). Cross-lingual models of word embeddings: An empirical comparison. *CoRR*, abs/1604.00425.

Zhang, Z., Wu, Y., Zhao, H., Li, Z., Zhang, S., Zhou, X., and Zhou, X. (2020). Semantics-aware BERT for language understanding. In *Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-2020)*.