Learning Bilingual Semantic Frames: Shallow Semantic Parsing vs. Semantic Role Projection

Pascale Fung, Zhaojun Wu, Yongsheng Yang, and Dekai Wu
Human Language Technology Center
Hong Kong University of Science and Technology
pascale@ee.ust.hk, {josephwu, ysyang, dekai}@cs.ust.hk

Abstract

To explore the potential application of semantic roles in structural machine translation, we propose to study the automatic learning of English-Chinese bilingual predicate argument structure mapping. We describe ARG_ALIGN, a new model for learning bilingual semantic frames that employs monolingual Chinese and English semantic parsers to learn bilingual semantic role mappings with 72.45% F-score, given an unannotated parallel corpus. We show that, contrary to a common preconception, our ARG_ALIGN model is superior to a semantic role projection model, SYN_ALIGN, which reaches only a 46.63% F-score by assuming semantic parallelism in bilingual sentences. We present experimental data explaining that this is due to cross-lingual mismatches between argument structures in English and Chinese at 17.24% of the time. This suggests that, in any potential application to enhance machine translation with semantic structural mapping, it may be preferable to employ independent automatic semantic parsers on source and target languages, rather than assuming semantic role parallelism.

1 Introduction

As statistical language learning technologies strain the limits of the relatively flat, simplistic structures of first-generation models, the need to incorporate representations that capture meaningful semantic patterns has become increasingly evident. Particularly for cross-lingual applications, techniques for multilingual semantic parsing and the acquisition of cross-lingual semantic frames have numerous potential applications. Error analysis suggests that a structured bilexicon containing a large inventory of cross-lingual semantic frame argument mappings—rather than merely word or phrase translations—would be invaluable toward attacking common types of errors in statistical machine translation, machine-aided translation, or cross-lingual information extraction or summarization models.

For example, inspection of recent contrastive error analysis data from a typical phrase-based SMT system shows that around 20% of the incorrect translations produced could have been avoided if the correct predicate argument information had been used (Och et al., 2003). Consider the following example from the error analysis data:

input 美国政府今天表示，有关美国要求澄清报导以色列意图在所占领的戈兰高地扩大犹太人的屯垦计划，以色列尚未给予满意的回答。

system The United States Government requested clarification of Israel's intention in the occupied Golan today, on the planned expansion of Jewish settlement, Israel has not yet given a satisfactory response.

reference The United States government said today that Israel had not provided a satisfactory answer to U.S. request for clarification about the reported plans to expand Jewish settlement in the occupied Golan Heights.

This example exhibits a typical mistake arising from the system's lack of awareness of the correct argument structure for the nominalized "*intention*" verb frame (as well as numerous other complements). Such errors of semantic role confusion are one of the most common sources of errors in current statistical systems that rely only on relatively flat representational structures and n-gram language models. Different languages realize semantic roles using different surface forms, and the language models and word reordering models in SMT are not always sufficient to discriminate between alternative hypotheses that may score equally well in fluency despite high variance in translation adequacy.

Bilingual frame semantics, if available, would provide an additional source of translation disambiguation leverage required to attack such problems. This necessitates the cross-lingual acquisition of a large inventory of *bilingual semantic frames*, which capture the needed role correspondence information in a manner independently of word reordering. Bilingual semantic verb frames specify the conventional patterns of alignment of semantic argument structures between a pair of semantic frames (or valency frames, qualia structures, etc.) for verbs in translation.

A challenge we faced is that (contrary to what one might first assume) even with semantic rather than syntactic arguments, the acquisition model still needs to be capable of dealing with the fact that predicate verb translations in English and Chinese often do *not* have the same semantic argument structure, due to cross-linguistic lexical and conceptual differences and translation idiosyncrasies. That is, the *ARG1* (say) in the Chinese semantic verb frame may not align to the *ARG1* in the frame for the corresponding English verb. This might seem surprising since, in principle, it would seem that semantic role labels for translatable verbs ought to be preserved more closely than syntactic roles across languages, since the agents, patients, and so forth seem more likely to remain constant in translation independent of verb alternations—whereas in contrast, surface syntactic labels (subject, object, etc.) often do not survive translation, due to language-specific verb alternations. However, we will describe experimental results indicating that even semantic roles are not preserved across Chinese and English 17.24% of the time.

Thus, our acquisition model cannot assume that the argument labels (*ARG0*, *ARG1*, ...) learned by our separately trained Chinese and English semantic parsers will necessarily correspond to each other cross-linguistically. To address this we introduce a cosine similarity model enabling our acquisition model to build and extract the bilingual semantic verb predicate-argument structure. We then compare this model to a semantic role projection model that uses syntactic constituent alignment, and which preserves semantic roles cross-lingually.

This paper is organized as follows. We begin by defining the bilingual semantic frame mapping problem. In section 3, we describe our findings from a manually aligned reference set of semantic structure mappings. Section 4 presents our new approach to semantic frame mapping, ARG_ALIGN, followed by the experimental results in section 5. In section 6, we then demonstrate experimentally how ARG_ALIGN outperforms a more conventional method based on semantic role projection, SYN_ALIGN.

2 Problem Definition

In recent years, researchers have shown that statistical machine translation models can be enhanced by incorporating structural information (Wu and Chiang, 2007). The attention, though, has thus far been largely focused on chunk or syntactic structures. Researchers only recently began seriously investigating whether incorporating *semantic* models can enhance statistical machine translation performance (Carpuat and Wu, 2005a; Carpuat and Wu, 2005b), and are only just beginning to show that semantic word sense disambiguation techniques can indeed improve accuracy (Carpuat et al., 2006; Carpuat and Wu, 2007). However, it remains an intriguing open question as to how semantic *structures*—semantic role mappings in bilingual semantic frames—can also be potentially leveraged to improve machine translation.

Thus, in order to overcome the immediate obstacle to exploring this potential, we are interested in learning the bilingual semantic structure given a predicate verb pair in English and Chinese, as in Figure 1. The predicate verb pair "*organized*/举办" have the operators *ARG0* "*African Environmental Centre*/非洲环境中心", and the operands *ARG1* "*Seminar on desertification*/沙漠化问题研讨会".

In the above example, the subject of the English sentence is *ARG1*, the operand, whereas the object is *ARG0*, the operator. On the other hand, the subject-object order is reversed in the Chinese sentence. The location "*Ivory Coast*" after the predicate verb and *ARG1*, at the end of the English sentence, whereas the Chinese translation is before the predicate verb, after *ARG0*, in the Chinese sentence. We are interested in learning and acquiring bilingual semantic frame mapping as illustrated in the above example, as an additional knowledge source for structural machine translation.

3 Findings in the Oracle Semantic Frame Mapping

To facilitate the development and evaluation of bilingual semantic frame acquisition methods, it was necessary for us to create an annotated gold standard reference corpus, containing parallel sentences whose semantic predicates and arguments are not only labeled but also mapped between Chinese and English.

Table 1: Reference Semantic Role Mappings

EN\CN	ARG0	ARG1	ARG2	ARG3
ARG0	326	77	7	1
ARG1	21	540	48	0
ARG2	3	28	39	2
ARG3	0	1	1	1

We aligned the semantic verb frames cross-lingually from a subset of the pre-release version of the Parallel Proposition Bank II for Chinese and English (Palmer et al., 2005). The Parallel Proposition Bank II for Chinese and English is derived from the Chinese Treebank English Parallel Corpus. Both the Chinese sentences and their English translations have been annotated syntactically in the Treebank format and semantically in the PropBank format.

We construct an oracle semantic role mapping based on manual semantic role alignment. The mapping matrix is shown in Table 1. Only the mapping between major core arguments (from $ARG0$ to $ARG3$ in the Proposition bank) are of interest at this stage. This is owing to the fact that, although the Chinese Propbank contains over 40 argument types and the English Propbank over 200, only core arguments $ARG0$ to $ARG5$ are responsible for representing the main semantic concepts, other argument types are served as adjunctive components (referred to as ARGM) that are used to provide additional information, for instance, ARGM-TMP for temporals. According to our observation, the occurrences of these core arguments diminish drastically after number 3.

As we can see from Table 1, around 82.74% of the mappings are direct mapping from ARG_i in English to ARG_i in Chinese. However, there remain a significant proportion of mappings that do not agree with direct mapping. Specifically, around 8.95% of the role mappings are from ARG_0 to ARG_1, 6.94% are from ARG_1 to ARG_2, and 0.27% are from ARG_2 to ARG_3. This type of cross-lingual role mismatch, also known as cross mapping, is also of particular interests since, if available, this knowledge source could be helpful

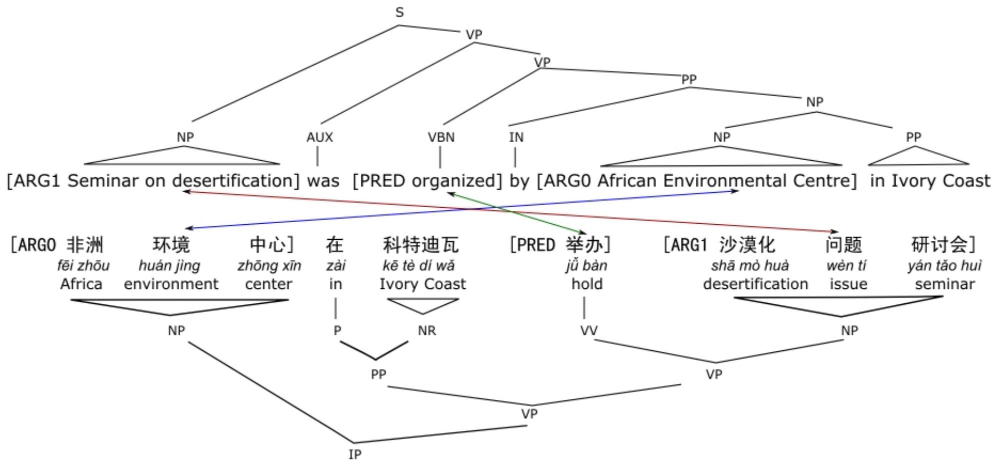

Figure 1: An example of bilingual semantic predicate argument mapping.

to MT systems.

One such cross-mapping example is shown below, where the "[$_{ARG1}$ world trade]" in English is mapped to "[$_{ARG0}$ 世界/world 贸易/trade]" in Chinese.

English Moreover , the report estimated that [$_{ARG1}$ world trade] [$_{ARGM-MOD}$ would] [$_{TARGET}$ grow] [$_{ARG2-EXT}$ by 9.4 %] [$_{ARGM-TMP}$ for 1997]

Chinese 此外， 报告 还 估计 [$_{ARGM-TMP}$ １９９７年] [$_{ARG0}$ 世界 贸易] [$_{TARGET}$ 增长] [$_{ARG1}$ 百分之九点四]

Gloss Moreover, report also estimate 1997 year world trade grow 9.4%

4 ARG_ALIGN: Learning Bilingual Semantic Frames via Chinese/English Shallow Semantic Parsing

We propose to first use shallow semantic parsers to annotate Chinese and English bilingual sentences with their semantic role boundaries and labels. Next, we propose to align these predicate-argument structures in the bilingual sentences by an automatic mapping approach.

Given all the candidate semantic roles parsed from the automatic semantic parsers, the automatic role mapping problem is cast as follows:

$$Z^* := \sum_{i=1}^{n} \min_{x} \sum_{j=1}^{m} x_{ij} c_{ij} \qquad (1)$$

$s.t.$

$$\sum_{i=1}^{n} x_{ij} = 1, j = 1, \cdots, m$$
$$x \geq 0$$

Z^* is the final role mappings we learned. x_{ij} is one element of the mapping matrix where argument i in Chinese is mapped to argument j in English, c_{ij} is one element of the cost matrix for aligning argument i in Chinese to j in English, n is the total number of arguments in a given source sentence and m is the total number of arguments in the target sentence.

To solve this bilingual predicate-argument role mapping problem, we propose an algorithm, ARG_ALIGN, as shown in Algorithm 1. In this algorithm, given S (source) and T (target) bi-sentence with semantic role annotation, we first match their predicate verbs based on a bilingual lexicon. Then, for each matched predicate verb pair *S-PRED* (source predicate) and *T-PRED* (target predicate), we extract their semantic arguments *S-ARG*s (source arguments) and *T-ARG*s (target arguments) and compute the cosine similarity score between all source and target arguments. We then extract the highest ranking matching pair of source and target constituents.

Algorithm 1 ARG_ALIGN

for each bilingual sentence pair **do**
  for each source predicate verb S-PRED **do**
    for each target predicate verb T-PRED **do**
      if S-PRED and T-PRED are translatable to each other, based on *bilingual lexicon* **then**
        S-ARGs $\Leftarrow ARG_0, \ldots, ARG_n$, given S-PRED
        T-ARGs $\Leftarrow ARG_0, \ldots, ARG_n$, given T-PRED
        for each ARG_i in S-ARGs **do**
          $\max(ARG_i) := 0$
          for each ARG_j in T-ARGs **do**
            align(ARG_i, $\hat{ARG_j}$)
            if $sim(ARG_i, ARG_j) \geq \max(ARG_i) \& sim(ARG_i, ARG_j) \geq threshold$ **then**
              $\max(ARG_i) := sim(ARG_i, ARG_j)$
              $\hat{ARG_j} := \text{argmax}\, ARG_i$
              where
              $sim(ARG_i, ARG_j) = \frac{ARG_i \cdot ARG_j}{|ARG_i||ARG_j|}$

4.1 Experimental Setup

Different sections of the Parallel Propbank corpus are used for algorithm development and evaluation. In order to determine the similarity threshold by which we can decide whether a pair of annotated bi-arguments match to each other, we randomly selected 497 sentence pairs as the test set and another set of 80 sentence pairs as the development data set.

Owing to the unavoidable errors through POS tagging, chunking or syntactic parsing, among the bilingual sentences, some Chinese and English sentences have no identifiable predicate verb, and are eliminated from further processing. Finally, 397 sentence pairs with automatic semantic parsing results are used in our predicate-argument mapping experiment.

In our proposed method, Chinese/English shallow semantic parsing is a prerequisite to achieving the task of bilingual semantic frame mapping. In recent years, there has been a lot of research on shallow semantic labeling or parsing both in English (Pradhan et al., 2004; Pradhan et al., 2005) and Chinese (Sun and Jurafsky, 2004; Xue and Palmer, 2005). In our experiments, we use the ASSERT semantic parser (Pradhan, 2005) to carry out the automatic semantic parsing on the English side and a similar SVM-based Chinese semantic parsing system (Wu et al., 2006) on the Chinese side. According to (Pradhan et al., 2005), their English semantic parser achieved 89.40 F-score with gold syntactic parse input, and 79.40 F-score with automatic syntactic parse input. Meanwhile, our SVM-based Chinese semantic parser yielded 89.89 F-score with gold syntactic parse input and 69.12 F-score with automatic syntactic parse input. Both of these parsers are among the-state-of-the-art shallow semantic systems in English and Chinese.

4.2 Experimental Results

Semantic role mapping output of our system is evaluated against the reference mappings described in the previous section, and measured with *Precision, Recall and F-score*[1] . In our evaluation strategy, a pair of arguments are considered correctly aligned to each other if the arguments are judged to be correct, and the mapping is judged to be correct.

The semantic role mapping result from our ARG_ALIGN algorithm is listed in Table 2 and the performance evaluation is listed in Table 3. 594 predicate-argument structure mappings are learned, with 219 unique Chinese verbs and 192 unique English verbs.

[1]F-score= $\frac{2 \times Precision \times Recall}{Precision + Recall}$

Table 2: Semantic Role Mappings from ARG_ALIGN

EN\CN	ARG0	ARG1	ARG2	ARG3
ARG0	259	8	7	0
ARG1	40	486	25	2
ARG2	3	26	15	0
ARG3	0	0	1	1

Table 3: Performance of Proposed Predicate-Argument Mapping

# words	[1,20]	<20,40>	[40,∞]	All
Precision	76.54	77.26	70.34	74.87
Recall	74.25	72.00	65.70	70.19
F-score	75.38	74.54	67.94	72.45

Many of these verbs are part of multiple context-dependent semantic structures. Human translation errors in the bilingual corpus, syntactic parsing and tagging errors account for some of the unmatched predicate-argument structures. Despite this, we obtained a fairly high F-score of 72.45% in bilingual semantic structure mapping, as evaluated against the mapping obtained from the oracle reference set.

5 Discussion of Results

Some of the mapping errors are due to errors in automatic syntactic and shallow semantic parsing. As a reference, we also evaluated the ARG_ALIGN algorithm directly on the Parallel Propbank data, by using the predicate-argument labels from manual annotation. The mapping accuracy in this case, free from parsing errors, is 98.9%.

Meanwhile, we observe that due to language differences and translation idiosyncrasies, predicate verb pairs in English versus Chinese do not always have the same argument structure. In this section, we present some interesting findings with examples in several categories.

5.1 Ellipsis

The ellipsis of some syntactic elements, such as the subject, occurred in either English or Chinese in the parallel sentences and might lead to some *NULL* argument mapping in the other language. As shown in the following example, [$_{ARG0}$ *PRO*] in Chinese is a filler constituent manually inserted in Chinese PropBank. However, the semantic role parser is not capable of generating this filler constituent automatically during the parsing. Thus, no *ARG*0 is labeled out in the automatic semantic parse result.

English Insiders feel that it would provide an excellent opportunity for [$_{ARG0}$ the economy and trade circles of China and South Korea] to [$_{TARGET}$ extend] [$_{ARG1}$ exchange and co-operation] .

Chinese 业内 人士 认为 ， 它 将 为 中 韩 两 国 经贸界 提供 一 次 [$_{ARG0}$ *PRO*] [$_{TARGET}$ 扩大] [$_{ARG1}$ 交流 与 合作] 的 良 机 。

Gloss Inside people believe , it will be China Korea two country economy and trade circles provide a extend communication and co-operation excellent opportunity

5.2 Parallel Structures in Chinese

When a Chinese sentence consisting of a parallel structure is translated into English, the parallel structure is consistently translated to clauses in English since these syntactic alternations are an effective translation technique to represent the same meaning of Chinese in one English sentence. Argument mapping is nevertheless correct despite this type of syntactic mismatching, as shown in the following example.

English [$_{ARG1}$ An office of Shanghai Customs posted at Chongming] , that was [$_{TARGET}$ approved] [$_{ARG0}$ by the China Customs Head Office] [$_{ARG2}$ to be set up] , was established a few days ago , and has already officially conducted business.

Chinese 经 [$_{ARG0}$ 中国 海关 总署] [$_{TARGET}$ 批准] 设立 的 上海 海关 驻 崇明 办事处 于 日前 成立 ， 并 正式 对 外 办理 业务。

Gloss Via China Customs Headquarters approval establish Shanghai Customs station Chongming office in current set up, and officially conduct business .

5.3 One-to-many Role Mapping

In our proposed algorithm, role mapping is based on individual ARG, not the ARG combination. However, in reality, it is possible for there to be one-to-many mappings. Thus, when this occurs, the one-to-many mapping is not possible to be identified. For example, in the following bi-sentence, *ARG1* and *ARG2* in English are mapped to *ARG1* in Chinese together.

English At present , about 150 thousand foreign-invested enterprises have opened accounts in the Bank of China , of which , [$_{ARG0}$ more than 20 thousand enterprises] have [$_{TARGET}$ received] [$_{ARG1}$ loan support] [$_{ARG2}$ from the Bank of China] .

Chinese [$_{ARGM-TMP}$ 目前] ， 约 有 十五万 家 外商 投资 企业 在 中国 银行 开立 帐户 ， 其中 [$_{ARG0}$ 二万多 家] [$_{TARGET}$ 获得] [$_{ARG1}$ 中国 银行 的 贷款 支持] 。

Gloss currently, about 150 thousand foreign merchant investment enterprise in China Bank open account, of which, 20 thousand more enterprise receive China Bank's loan support .

6 Role Mapping from Syntactic Constituent Alignment

To date, it is often casually assumed that semantic roles can be simply projected across language pairs by constituent alignment (Pado and Lapata, 2006). In such an approach, it is assumed that an English constituent is lexically translated into the Chinese constituent, in which case they must share the same role label. This sort of view is typically inspired by the many structurally-based statistical machine translation models that make use of some kind of syntactic constituent projection (Hwa et al., 2005).

Therefore it is worth investigating the possibility of projecting semantic role labels across matching syntactic constituents. To accomplish this, we implement a contrastive SYN_ALIGN algorithm that obtains semantic structure mapping based on Treebank syntactic parse projection. This model is similar in spirit to that of (Pado and Lapata, 2006), in which the authors proposed a semantic role projection model based on FrameNet rather than PropBank verb frames. While our semantic role projection model is inspired by (Pado and Lapata, 2006), we propose a novel solution to the Linear Assignment Problem in order to align syntactic constituents from both the English and Chinese sentences, and then project the semantic role labels from English across to Chinese. The reason why we project the semantic role from English to Chinese is because according to (Pradhan et al., 2005), their English semantic parser outperforms our Chinese one due to the larger training data available in English TreeBank and PropBank.

In this approach, we make a strong assumption that the English semantic roles can be projected directly to their corresponding entities in Chinese (although, obviously, this assumption does not always hold in reality), and then utilize the lexical and syntactic information from the syntactic parses to project the semantic roles from English to Chinese.

To decouple the effect of semantic parsing from syntactic parsing, we save the syntactic annotations on the bilingual sentences, but remove the semantic annotations from the Chinese sentences. Based on the "perfect constituent alignment" proposed in (Pado and Lapata, 2006), we then project English semantic role labels to their corresponding Chinese entities. Finally, an evaluation of the mapping results are carried out in reference to the gold standard mapping set.

6.1 Alignment Selection

Since most structural machine translation systems are based on tree alignments, we are interested in investigating semantic role mapping on top of such syntax tree alignments. In other words, we select syntactic constituent (i.e. chunk) as the alignment unit. Moreover, (Pado and Lapata, 2006) has also shown that the best semantic role projection is achieved with constituent based alignment.

6.2 Assignment Cost

Similar to (Pado and Lapata, 2006), we define the alignment cost between any pair of English and Chinese constituents as follows:

$$cost(e_c, c_c) = \frac{1}{sim(e_c(_{w1,w2\ldots}), c_c(_{w1,w2,\ldots}))} \tag{2}$$

where, e_c is an English constituent, c_c is a Chinese constituent, w_i belongs to the set of ($NP, PP, pronoun, numeral, quantifier$) and w_i is a content word. The purpose of this is to disregard any lexical items that would not be of interest to us in the ultimate task of argument mapping.

6.3 Constituent Alignment

(Pado and Lapata, 2006) proposed three alignment models for the constituent alignment: *total alignments*, *edge covers* and *perfect matchings*. We chose perfect matching for our experiment since (Pado and Lapata, 2006) reported superior performance using this model. "Perfect matching" is defined as follows: given all the constituents extracted from the Chinese and English parallel data, each constituent in Chinese must align to one and only one constituent in English, and vice versa. We observe that this problem can be cast as a Linear Assignment Problem, which of course is a fundamental combinatorial optimization problem. The Linear Assignment Problem can be described as follows:

$$Z^* := \min_x \sum_{i=1}^{n} \sum_{j=1}^{n} x_{ij} c_{ij} \tag{3}$$

$$s.t.$$

$$\sum_{j=1}^{n} x_{ij} = 1, i = 1, \cdots, n$$
$$\sum_{i=1}^{n} x_{ij} = 1, j = 1, \cdots, n$$
$$x \geq 0$$

Z^* is the solution of the linear assignment problem. x_{ij} is the assignment matrix where constituent i was assigned to constituent j, c_{ij} is the cost matrix for aligning constituent i to j.

Table 4: Role Mapping from Syntactic Projection

EN\CN	ARG0	ARG1	ARG2	ARG3
ARG0	248	0	0	0
ARG1	0	381	0	0
ARG2	0	0	22	0
ARG3	0	0	0	0

Table 5: Performance of Semantic Role Projection

# words	[1,20]	<20,40>	[40,∞]	All
Precision	54.45	45.10	39.35	44.57
Recall	59.78	50.14	41.98	48.90
F-score	56.99	47.49	40.62	46.63

Our semantic role projection algorithm, SYN_ALIGN, is described in Algorithm 2. Given the English and Chinese bi-parse, we first extract their constituents (chunks). These constituents are stored in two arrays. Then, for these two constituent arrays, we apply the classic Hungarian method (Kuhn, 1955) to solve the Linear Assignment optimization problem by using the cosine similarity score between two constituents as the assignment cost. Finally, we project the English semantic roles to the Chinese side based on the constituent alignment result.

The predicate-argument mapping learned from the constituent based semantic role projection is shown in Table 4 and the performance evaluation against the mapping learned from the gold standard is shown in Table 5.

6.4 Experimental Results

Again evaluating with respect to the gold standard reference mappings, the mapping F-score of SYN_ALIGN is only 46.63%. This mapping performance is significantly lower than achieved by our proposed ARG_ALIGN model, owing to the assumption that argument structures can be projected across syntactic constituents, which has hereby been shown to be brittle.

Algorithm 2 SYN_ALIGN

```
INPUT: Chinese and English parallel syntactic parse trees
let EN_Cons[] = source English constituents
let CN_Cons[] = target Chinese constituents
en_no = number of English constituents
cn_no = number of Chinese constituents
max_no = maximum(cn_no, en_no)
if cn_no < max_no then
   append max_no − cn_no with "dummy" constituents to CN_Cons[]
else if en_no < max_no then
   append max_no − en_no with "dummy" constituents to EN_Cons[]
for i = 1 to max_no do
   for j = 1 to max_no do
      similarity_score = cosine(CN_Cons[i], EN_Cons[j])
      if similarity_score == 0 then
         cost_matrix[i][j] = 1000.00
      else
         cost_matrix[i][j] = 1/similarity_score
alignment = hungarian_method(cost_matrix)
for all semantic roles in English semantic parsing result do
   project the semantic roles to Chinese side based on alignment solution
```

7 Conclusion

For machine translation purposes, it is meaningful to study the semantic structural mapping between the source and target language. We propose a new automatic algorithm, ARG_ALIGN, to extract the predicate-argument mappings from unannotated bilingual sentence pairs with 72.45% F-score, given an unannotated parallel corpus. We first identify and label the semantic structures using the Chinese and English shallow semantic parsers and then use ARG_ALIGN to find the mapping pairs.

Given bilingual sentence pairs with manually annotated semantic role labels, we record the semantic role mapping between bilingual argument structures if they are *lexically* aligned to each other. We observe that there are 17.24% of cross mapping between argument structures in English and Chinese. Among these, 8.95% are argument 0-1 mappings, 6.94% are 1-2 mappings, and 0.27% are argument 2-3 mappings. Referring to the manual gold standard mapping, the F-score of our proposed mapping between automatically annotated argument structures is 72.45%, showing promise for automatic semantic structure mapping in bilingual sentence pairs, applicable to machine translation and other multilingual and cross-lingual applications.

Contrary to a preconception that one sometimes hears, we show empirically that our model is superior to a semantic role projection model which assumes semantic parallelism in bilingual sentences. In the latter model, we propose using the Hungarian method in a syntax alignment algorithm we name SYN_ALIGN, to align syntactic constituents from both the English and Chinese sentences, and project the semantic role labels across. Compared to the gold standard mapping, the mapping F-score in this case is 46.63%.

Our results led us to believe that, since there is a non-negligible amount of cross argument mapping between English and Chinese translations, it maybe preferable to use automatic semantic role labeling in both the source and target languages, than to use direct projection of semantic role labels from one language to the other.

One obvious next step is to embed the shallow semantic parsers and the cross-lingual verb frame acquisition model in end-to-end machine translation systems or MT applications. We would also like to acquire cross-lingual semantic frames for other categories besides verbs.

Acknowledgments

This material is based upon work supported in part by the Defense Advanced Research Projects Agency (DARPA) under GALE Contract No. HR0011-06-C-0023, and by the Hong Kong Research Grants Council (RGC) research grant HKUST612806. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Defense Advanced Research Projects Agency.

We are grateful to Sebastian Pado, Martha Palmer, Sameer Pradhan, and Nianwen Xue for helpful discussions.

References

Marine Carpuat and Dekai Wu. 2005a. Evaluating the word sense disambiguation performance of statistical machine translation. In *Second International Joint Conference on Natural Language Processing (IJCNLP)*, pages 122–127, Jeju Island, Korea, Oct.

Marine Carpuat and Dekai Wu. 2005b. Word sense disambiguation vs. statistical machine translation. In *43rd Annual Meeting of the Association for Computational Linguistics (ACL-2005)*, pages 387–394, Ann Arbor, Jun.

Marine Carpuat and Dekai Wu. 2007. Improving statistical machine translation using word sense disambiguation. In *Conference on Empirical Methods in Natural Language Processing and Conference on Computational Natural Language Learning (EMNLP-CoNLL 2007)*, Prague, Jun.

Marine Carpuat, Yihai Shen, Xiaofeng Yu, and Dekai Wu. 2006. Toward integrating word sense and entity disambiguation into statistical machine translation. In *Third International Workshop on Spoken Language Translation (IWSLT 2006)*, pages 37–44, Kyoto, Nov.

Rebecca Hwa, Philip Resnik, Amy Weinberg, Clara Cabezas, and Okan Kolak. 2005. Bootstrapping parsers via syntactic projection across parallel texts. *Natural Language Engineering*, 11(03):311–325.

Harold W. Kuhn. 1955. The hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2(1):83–97.

Franz Josef Och, Daniel Gildea, Anoop Sarkar, Sanjeev Khudanpur, Kenji Yamada, Alexander Fraser, Libin Shen, Shankar Kumar, David Smith, Viren Jain, Katherine Eng, Zhen Jin, and Dragomir Radev. 2003. Website. `http://www.clsp.jhu.edu/ws2003/groups/translate/`.

Sebastian Pado and Mirella Lapata. 2006. Optimal constituent alignment with edge covers for semantic projection. In *Proceedings of ACL-COLING 2006*, pages 1161–1168, Sydney, Australia.

Martha Palmer, Nianwen Xue, Olga Babko-Malaya, Jinying Chen, and Benjamin Snyder. 2005. A parallel Proposition Bank II for Chinese and English. In *Proceedings of the Workshop on Frontiers in Corpus Annotations II: Pie in the Sky*, pages 61–67, Ann Arbor, Michigan, June. Association for Computational Linguistics.

Sameer Pradhan, Wayne Ward, Kadri Hacioglu, James H. Martin, and Dan Jurafsky. 2004. Shallow semantic parsing using support vector machines. In *Proceedings of NAACL-HLT 2004*.

Sameer Pradhan, Kadri Hacioglu, Valerie Krugler, Wayne Ward, James H. Martin, and Daniel Jurafsky. 2005. Support vector learning for semantic argument classification. *Machine Learning*, 60(1):11–39.

Sameer Pradhan. 2005. ASSERT: Automatic Statistical SEmantic Role Tagger. Website. `http://oak.colorado.edu/assert/`.

Honglin Sun and Daniel Jurafsky. 2004. Shallow semantic parsing of Chinese. In *Proceedings of Human Language Technology Conference/North American Chapter of the Association for Computational Linguistics (HLT/NAACL)*.

Dekai Wu and David Chiang, editors. 2007. *NAACL-HLT 2007 / AMTA Workshop on Syntax and Structure in Statistical Translation (SSST)*. Association for Computational Linguistics, Rochester, NY, USA, April.

Zhaojun Wu, Yongsheng Yang, and Pascale Fung. 2006. C-ASSERT: Chinese shallow semantic parser. Website. `http://hlt030.cse.ust.hk/research/c-assert/`.

Nianwen Xue and Martha Palmer. 2005. Automatic semantic role labeling for Chinese verbs. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence, Edinburgh, Scotland*.