Towards Universal Semantic Tagging

Abzianidze, Lasha; Bos, Johan

Towards Universal Semantic Tagging

Lasha Abzianidze
CLCG, University of Groningen
l.abzianidze@rug.nl

Johan Bos
CLCG, University of Groningen
johan.bos@rug.nl

Abstract

The paper proposes the task of universal semantic tagging—tagging word tokens with language-neutral, semantically informative tags. We argue that the task, with its independent nature, contributes to better semantic analysis for wide-coverage multilingual text. We present the initial version of the semantic tagset and show that (a) the tags provide semantically fine-grained information, and (b) they are suitable for cross-lingual semantic parsing. An application of the semantic tagging in the Parallel Meaning Bank supports both of these points as the tags contribute to formal lexical semantics and their cross-lingual projection. As a part of the application, we annotate a small corpus with the semantic tags and present new baseline result for universal semantic tagging.

1 Introduction

Part-of-speech (POS) tagging represents one of the most popular Natural Language Processing (NLP) tasks, especially when it comes to syntactic parsing. It is proven by practice that the information about POS-tags makes syntactic parsing easier. An independent nature of the task and its lower complexity (compared to syntactic parsing) make POS tagging a perfect preprocessor for syntactic parsing.

But to what extent is POS-tag information useful for semantic parsing—obtaining semantic representations of natural language texts? Trying to answer this question in favor of POS-tags, we take a stand of a semantic parsing approach that heavily relies on them. One of such approaches is the formal compositional semantics driven by syntactic derivations of Combinatory Categorial Grammar (CCG, Steedman 2001), where a meaning representation is derived by composing formal meaning representations of lexical items (Bos et al., 2004; Lewis and Steedman, 2013; Mineshima et al., 2015).[1] Since lexical items come with fully fledged semantics, obviously, assigning correct lexical semantics is crucial for this approach. This is the place where POS-tags come into play by providing lexical information helping to determine lexical semantics. For example, given a POS-tag `NN` (singular or mass noun) or `JJ` (adjective)[2], it is possible to assign the desired lexical semantics to the modifiers in (1). But there are cases where POS-tags fall short of providing sufficient information for lexical semantics. For instance, regardless of their semantics, quantifiers get the same tag `DT` (determiner). Hence one needs to check the lemma of a determiner in order to define its semantics in (2).

$$\frac{\text{beer}^{\texttt{NN}}\ /\ \text{transparent}^{\texttt{JJ}}}{\begin{array}{c} N/N \\ \lambda px.\, beer(x) \wedge p(y) \wedge for(y,x) \text{ if pos=NN} \\ \lambda px.\, transparent(x) \wedge p(x) \text{ if pos=JJ} \end{array}} \quad \frac{\text{bottle}^{\texttt{NN}}}{\begin{array}{c} N \\ \lambda x.\, bottle(x) \end{array}} \quad (1)$$

$$\frac{\text{no}^{\texttt{DT}}\ /\ \text{every}^{\texttt{DT}}}{\begin{array}{c} NP/N \\ \text{if lemma=`no'}\ \lambda pq.\, \neg\exists x\big(p(x) \wedge q(x)\big) \\ \text{if lemma=`every'}\ \lambda pq.\, \forall x\big(p(x) \rightarrow q(x)\big) \end{array}} \quad \frac{\text{man}^{\texttt{NN}}}{\begin{array}{c} N \\ \lambda x.\, man(x) \end{array}} \quad (2)$$

Formal semantics of a content word usually involves a symbol corresponding the lemma. This is the case for each lexical item in (1) and (2) except for the quantifiers. But when dependence of lexical semantics on a lemma is beyond a simple substitution, i.e., one needs to verify a lemma to define lexical semantics, then this case fails to generalize across different languages. For example, assigning lexical

[1] Similarly, the semantic parsing based on dependency structures (Reddy et al., 2016, 2017) also rely heavily on POS-tags.

[2] Throughout the paper the Penn Treebank POS-tags (Marcus et al., 1993), widely accepted in the NLP community, will be assumed unless otherwise stated.

semantics to quantifiers based on their lemma does not scale up for multilingual semantics. On the other hand, the treatment of common nouns in (1) and (2) generalizes for a multilingual case by using a simple assignment:

$$[\![\langle w,\ \text{pos} = \texttt{NN},\ \text{category} = N\rangle]\!] = \lambda x.\, \texttt{SYM}(x) \tag{3}$$

where SYM is a lexical predicate, usually a lemma, corresponding to the word w.

In order to compensate the shortcomings of POS tagging for semantic parsing, we propose a new NLP task, called *Universal Semantic Tagging* or *Semantic Tagging* in short.[3] The task represents a standard sequence tagging problem where each word token gets a language-neutral semantic tag, in short *sem-tag*. Sem-tags carry information that better characterizes lexical semantics than POS-tags do. We will show that the semantic tagging not only improves over POS tagging but also subsumes the task of Named Entity (NE) classification. We argue that importance of the task for (cross-lingual) semantic parsing is comparable to the one POS tagging has for syntactic parsing.

The rest of the paper is organized as follows. First, we further motivate the idea behind semantic tagging—how it includes semantic virtues of POS-tags and Named Entity (NE) classes; Then we present the current version of the semantic tagset. To show application of semantic tagging in semantic parsing, we describe its use in the Parallel Meaning Bank (PMB) project[4] (Abzianidze et al., 2017), where the sem-tags help to determine formal lexical semantics. We also present a baseline result for semantic tagging on a small annotated corpus. In the end, the paper discusses possible directions of future research on semantic tagging.

2 Motivation for Semantic Tagging

The information about POS-tags and NE classes do contribute to determine lexical semantics to some extent, but they are not sufficiently informative. One of the goals of the semantic tagging is to incorporate semantic virtues of these two tasks and fill gaps in semantic modeling by adding new categories.

In a tagging task, a sequence of characters that takes a tag is called a *word token*, or simply a *token*. Definition of a token may vary depending on a tagging task and its application. We find the concepts of token for POS tagging and semantic tagging somewhat different. For example, *"20-year-old man from New Zealand"* represents five tokens for POS tagging while we consider six token version *"20 year old man from New␣Zealand"* more suitable for semantic analysis.[5] Hereafter, when talking about semantic tagging, tokens should be understood as meaningful atoms.

In addition to the examples from the previous section, POS-tags fail to disambiguate lexical semantics of series of word tokens. For example, reflexive and emphasizing pronouns get the same POS-tag PRP. The conjunctions *and*, *or* and *but* are all POS tagged as coordinating conjunctions (CC).[6] A comma can have several semantic functions, e.g., Arivazhagan et al. (2016) distinguishes nine semantic roles including apposition, location or listing. Both infinitival and prepositional uses of *to* are POS tagged as TO (Santorini, 1990, p. 5). Semantics of the determiner *any* needs to be disambiguated in context. The auxiliary verbs (e.g., *do* and *have*) and content verbs obtain similar POS-tags based on their syntactic features. This complicates to determine whether a verb introduce an event entity or not. The relative pronouns *which* and *that* both get the WDT POS-tag regardless of their restrictive or non-restrictive behaviour. It is natural to distinguish semantics of intersective adjectives (e.g., *ill* and *dead*) from subsective ones (e.g., *skillful* and *professional*), but this is impossible to do with the single POS-tag JJ. The above-mentioned partial list clearly shows that POS-tags are not sufficient for fine-grained (formal) lexical semantics.

[3] Since semantics of linguistic expressions is language independent to a large extent, we find *universal* redundant from a semantic perspective. On the other hand, from an NLP perspective, we would like to emphasize the universal (i.e., cross-lingual) nature of the task.

[4] http://pmb.let.rug.nl

[5] In general, we assume a fixed multiword expression as one token if it is semantically non-compositional and has an obscure syntactic structure. Such multiword tokens include proper names (e.g., *Alfred␣Nobel* and *European␣Union*), numerical expressions (e.g., *ten␣thousand* and *10␣000*), and function phrases like *as␣well␣as*, *each␣other*, and *so␣that*.

[6] Moreover, there are at least two semantic usages of *and* one might want to distinguish: distributive and collective readings.

For wide-coverage semantic analysis one needs to identify NEs, detect their type, and model their semantics appropriately. The information extraction community has been actively working on the problem of NE classification and designed annotation schemas. For example, the named entity task at MUC-7 (Chinchor and Robinson, 1998) distinguished three general classes of NEs, where each of them contain several types: entity names (person, organization, location), temporal expressions (date and time) and number expressions (money and percentage). These types of NEs are motivated by downstream applications of information extraction. For a fine-grained semantic analyses, one might go beyond this standards. For example, one of such moves, following to Doddington et al. (2004), is to distinguish the locations without political or social groups (e.g., seas, parks and mountains) from those with them, i.e. geo-political entities such as villages, cities, countries, etc. Also one can introduce new NE classes, for instance, the classes for events (e.g., *9/11* and *World War II*) and artifacts (e.g., *Ubuntu 12.04 LTS*) or generalize existing ones, for example, go beyond monetary currency and percentage and cover the measure words like *meter* and *kilogram*.

In the next section we present an inventory of the universal semantic tags which incorporates semantic merits of POS-tags and named entity classes, fill the gaps in semantic annotation, and represents one unified tagset aiming to facilitate cross-lingual semantic parsing.

3 The Universal Semantic Tagset

The universal semantic tagset aims to provide general cross-lingual description for lexical semantics of all sorts of word tokens. It significantly differs from POS tagset, which is not semantically motivated, and generalizes over NE classes as the latter only covers the words of a particular type. The current version of the semantic tagset (v0.7) is given in Table 1, a revised version of the tagset (v0.6) presented in Bjerva et al. (2016).[7] The sem-tags are organized into 13 coarse-grained semantic classes each having its own meta-tag. This division is informal as many sem-tags easily qualify for several classes. We designed the tagset in a data-driven fashion while bearing in mind formal semantic properties of tokens. The employed corpus consists of several parallel corpora of various genres spanning over four languages (see Sec. 4).

Before we characterize the sem-tags, let us explain how the tags can or cannot be interpreted. A sem-tag of a token describes a semantic contribution of the token with respect to the meaning of the source expression. In this way, the principle of semantic compositionality underlies the semantic tagging. Later, in Sec. 4, an application shows how to interpret a sem-tag as an unspecified semantic schema/recipe. In general, sem-tags are not responsible for encoding a syntactic function of a token; For example, concrete quantities and colors get `QUC` and `COL` regardless of being a nominal modifier or a head of a noun phrase.[8] Moreover, currently sem-tags do not separate adjectives and adverbs but treat them as properties. The information about thematic roles are not also provided by the sem-tags. In principle, sem-tags provide the semantic information that complements thematic roles, syntax and lemma. Due to the abstraction from syntactic and lemma-related information, sem-tags are suitable for cross-lingual application.

The semantic classes ATT, COM, NAM, EVE and UNE cover both open and closed class words while the rest of the classes focus on the closed class words. The sem-tags that model closed class words make two major contributions: (i) semantically disambiguate highly ambiguous words that usually belong to closed class words, and (ii) act as an umbrella term for cross-lingual variants of a word and opens the door to multilingual semantic tools.

Let us first describe the groups of sem-tags covering closed class words. The anaphoric tags encompass definite articles and types of pronouns. They distinguish emphasizing pronouns (`EMP`) from reflexive

[7]Major revisions concern the classes of named entity (NAM), attributes (ATT), events (EVE), deixis (DXS) and tense (TNS). In contrast to the tagset v0.6, the current tagset excludes 15 tags and includes 13 new ones. More details about the changes are explained below.

[8]In contrast to this, depending on a syntactic context a color can get the `NN` or `JJ` POS-tag in the Penn Treebank (Santorini, 1990, p. 12): *The plants are dark green/*`JJ` vs *The plants are a dark green/*`NN`. It is also needless to say that sem-tags do not distinguish singular or non-3rd person verb forms, unlike the POS-tags.

Table 1: The Universal Semantic Tagset v0.7: 73 sem-tags grouped into 13 meta-tags. The sem-tags are accompanied with the examples where several highly ambiguous tokens come with a context. The new sem-tags of v0.7 wrt v0.6 are marked with an asterisk.

Meta-tag	Sem-tag	Description and examples
ANA anaphoric	PRO	anaphoric & deictic pronouns: *he, she, I, him*
	DEF	definite: *the, lo*^{IT}*, der*^{DE}
	HAS	possessive pronoun: *my, her*
	REF	reflexive & reciprocal pron.: *herself, each_other*
	EMP	emphasizing pronouns: *himself*
ACT speech act	GRE	greeting & parting: *hi, bye*
	ITJ	interjections, exclamations: *alas, ah*
	HES	hesitation: *err*
	QUE	interrogative: *who, which, ?*
ATT attribute	QUC*	concrete quantity: *two, six_million, twice*
	QUV*	vague quantity: *millions, many, enough*
	COL*	colour: *red, crimson, light_blue, chestnut_brown*
	IST	intersective: *open, vegetarian, quickly*
	SST	subsective: *skillful* surgeon, *tall* kid
	PRI	privative: *former, fake*
	DEG*	degree: 2 meters *tall*, 20 years *old*
	INT	intensifier: *very, much, too, rather*
	REL	relation: *in, on, 's, of, after*
	SCO	score: *3-0*, grade *A*
COM comparative	EQU	equative: *as* tall as John, whales *are* mammals
	MOR	comparative positive: *better, more*
	LES	comparative negative: *less, worse*
	TOP	superlative positive: *most, mostly*
	BOT	superlative negative: *worst, least*
	ORD	ordinal: *1st, 3rd, third*
UNE unnamed entity	CON	concept: *dog, person*
	ROL	role: *student, brother, prof., victim*
	GRP*	group: John *{,}* Mary *and* Sam gathered, a *group* of people
DXS deixis	DXP*	place deixis: *here, this, above*
	DXT*	temporal deixis: *just, later, tomorrow*
	DXD*	discourse deixis: *latter, former, above*
LOG logical	ALT	alternative & repetitions: *another, different, again*
	XCL	exclusive: *only, just*
	NIL	empty semantics: *{.}, to, of*
	DIS	disjunction & exist. quantif.: *a, some, any, or*
	IMP	implication: *if, when, unless*
	AND	conjunction & univ. quantif.: *every, and, who, any*
MOD modality	NOT	negation: *not, no, neither, without*
	NEC	necessity: *must, should, have* to
	POS	possibility: *might, could, perhaps, alleged, can*
DSC discourse	SUB	subordinate relations: *that, while, because*
	COO	coordinate relations: *so, {,}, {;}, and*
	APP	appositional relations: *{,}, which, {(}, {—}*
	BUT	contrast: *but, yet*
NAM named entity	PER	person: *Axl_Rose, Sherlock_Holmes*
	GPE	geo-political entity: *Paris, Japan*
	GPO*	geo-political origin: *Parisian, French*
	GEO	geographical location: *Alps, Nile*
	ORG	organization: *IKEA, EU*
	ART	artifact: *iOS_7*
	HAP	happening: *Eurovision_2017*
	UOM	unit of measurement: *meter, $, %, degree_Celsius*
	CTC*	contact information: *112, info@mail.com*
	URL	URL: http://pmb.let.rug.nl
	LIT*	literal use of names: his name is *John*
	NTH*	other names: table *1a*, equation *(1)*
EVE events	EXS	untensed simple: to *walk*, is *eaten, destruction*
	ENS	present simple: we *walk*, he *walks*
	EPS	past simple: *ate, went*
	EXG	untensed progressive: is *running*
	EXT	untensed perfect: has *eaten*
TNS tense & aspect	NOW	present tense: *is* skiing, *do* ski, *has* skied, *now*
	PST	past tense: *was* baked, *had* gone, *did* go
	FUT	future tense: *will, shall*
	PRG*	progressive: has been *being* treated, *aan_het*^{NL}
	PFT*	perfect: has *been* going/done
TIM temporal entity	DAT*	full date: *27.04.2017, 27/04/17*
	DOM	day of month: *27th* December
	YOC	year of century: *2017*
	DOW	day of week: *Thursday*
	MOY	month of year: *April*
	DEC	decade: *80s, 1990s*
	CLO	clocktime: *8:45_pm, 10_o'clock, noon*

ones (REF). Other types of determiners like indefinite articles, demonstratives, and quantifiers (most of which get the DT pos-tag) are covered by existential (DIS), universal (AND), place deixis (DXP) and vague quantity (QUV) sem-tags. Besides place deixis, there are sem-tags for temporal and discourse deixis (Löbner, 2013, Ch. 4). In addition to the sem-tags for subordinated (SUB) and coordinated (COO) discourse relations, there are separate tags APP and BUT for appositional and contrasting relations. Phrasal conjunctions and other discourse relations that have relatively transparent formal logical semantics are singled out by the logical sem-tags DIS, IMP, and AND. Tokens with vacuous semantics are tagged with NIL. Such tokens might include punctuations, infinitival *to*, and *of* from pseudo-partitives, e.g., *two liters of water*. The LOG class also includes the tags ALT and XCL covering words with semantics involving inequality. Given these sem-tags, a comma might be tagged with NIL, APP, AND, or DIS depending on its semantic contribution. Relative pronouns of restrictive and non-restrictive relative clauses get AND and APP respectively.

Accounting for modal words in semantics is crucial as they often block certain entailments. For this reason, the tagset has dedicated tags for tokens with modal functions, including a tag NOT for negative lexical items. In contrast, the Penn Treebank POS-tagset distributes most of negative items among adverbs, prepositions and determiners. Unlike the POS-tags, the sem-tags distinguish tense and aspect marking auxiliary verbs (TNS) from content (EVE) and modal (MOD) ones.

Since date and time expressions play an important role is downstream applications and have been a target of several shared tasks, the tagset has fine-grained sem-tags for them: DAT and CLO are designed for the full date and time formats while the rest marks (unspecified) components of the date format. We also design special sem-tags for speech acts.

The attributive and comparative classes mostly cover words like adjectives, adverbs, quantities and words derived from them. Since both adjectives and adverbs can be seen as modifiers (of entities and events) from a semantic perspective, sem-tags do not differentiate them.[9] The sem-tags distinguish concrete QUC and vague QUV quantities (which were previously merged in a single quantity sem-tag in v0.6). There are separate tags for intersective, subsective and privative adjectives. The adjective like *alleged* that are neither subsective nor privative are tagged with the modal sem-tag POS. Adverbs are usually tagged as intersective. DEG marks adjectives that subcategorize for degrees , e.g., *10cm long* or *2 years old*, as they are not subsective. From comparative sem-tags, we would like to mention EQU which covers words with interpretation of (tense-free) equality.

The tagset makes fine-grained distinction of proper names. In addition to the standard NE tags PER and ORG, following LDC (2008), geographical locations are divided into geo-political entities (geographical regions defined by political and/or social groups, GPE) and the rest of geographical entities (GEO). To link individuals to the NEs they originate from, we use GPO. Units in measure phrases are tagged with UOM as they act like NEs.[10] UOM generalizes over the standard NE class for currency and percentage.

The event sem-tags account for semantics of content verbs that introduce Davidsonian event entities. EXS marks a content verb without tense or aspect (including nominalizations and gerunds) while the other sem-tags in EVE additionally encode tense or aspect. Currently, the sem-tags in EVE are motivated by English, German, Dutch and Italian. The sem-tags for unnamed entities divide nouns into concepts (CON), roles (ROL), and collective/group nouns (GRP). Moreover, ROL also covers relational nouns while GRP marks collective operators too.

The examples of semantically tagged tokenized texts are given below. In (4), *tall* is marked with DEG as it is not affirmative—the question is not asking weather the green monster is tall. The sem-tags in (6) and (7) disambiguate existential and universal semantics of *any* and *a*. Notice that the latter is tagged with AND as *$ 100 a day* is semantically equivalent to *$ 100 each day*. More examples of semantically tagged text can be accessed online via the PMB Explorer.[11]

How[QUE] tall[DEG] is[NOW] the[DEF] green_monster[ART] at[REL] Fenway[GEO] ?[QUE] (4)

My[HAS] sister[ROL] went[EPS] to[REL] the[DEF] United_States[GPE] to[SUB] study[EXS] English[CON] .[NIL] (5)

Any[AND] contribution[CON] was[PST] appreciated[EXS] but[BUT] we[PRO] have[NOW] n't[NOT] got[EXT] any[DIS] .[NIL] (6)

He[PRO] himself[EMP] can[POS] earn[EXS] $[UOM] 100[QUC] a[AND] day[UOM] .[NIL] (7)

4 Applications and Results

The idea of the universal semantic tagging was originally motivated by the goals of the PMB project (Bos, 2014; Abzianidze et al., 2017): (i) compositionally derive formal meaning representations for wide-coverage English text (Bos, 2009), and (ii) project the meaning representations to Dutch, German and Italian translations via word alignments (Evang and Bos, 2016). These requirements challenge semantic competence and cross-lingual scalability of the universal semantic tagging.

A high quality large-scale semantic lexicon is crucial for the PMB as both projection and derivation of

[9]Moreover, some languages like Dutch and German make little grammatical distinction between adverbs and adjectives.

[10]It seems unnatural to treat them as predicates and therefore license entailments using the WordNet hypernymy relations: *he ran five kilometers* ⇒ *he ran five metric linear units*.

[11]http://pmb.let.rug.nl/explorer/

meaning representations starts from lexical items. The semantic tagset plays a crucial role in development and organization of the lexicon. In particular, in the PMB, Boxer (Bos, 2008, 2015) interprets a sem-tag as a mapping from CCG categories (augmented with thematic roles) to a formal semantic schema which is further specified by a token-related predicate/constant symbol and thematic roles (if any). The function behind the `EXS` tag is partially depicted in (8):

$$\texttt{EXS} = \left\{ \begin{array}{rcl} S\backslash_{R_1} NP & \mapsto & \lambda P\, r.\, P\big(\lambda x.\, \boxed{\begin{array}{l} e \\ \hline \mathrm{SYM}(e)\quad R_1(e,x) \end{array}}\, ; r(e)\big) \\ (S\backslash_{R_1} NP)/_{R_2} NP & \mapsto & \lambda Q\, P\, r.\, P\Big(\lambda x.\, Q\big(\lambda y.\, \boxed{\begin{array}{l} e \\ \hline \mathrm{SYM}(e)\quad R_1(e,x)\quad R_2(e,y) \end{array}}\, ; r(e)\big)\Big) \end{array} \right\} \tag{8}$$

In order to collect large semantically annotated data via bootstrapping, we prepared initial silver and gold datasets for system training and testing respectively. The datasets are part of the PMB, where the gold data is manually checked and consists of 2.4K English sentences (14.6K tokens) while the silver data (457K tokens) consists of the PMB documents that are tagged by the neural semantic tagger of Bjerva et al. (2016) and have some manual corrections. For the data collection via bootstrapping, we initially employ the tri-gram based TnT tagger (Brants, 2000) rather than data-hungry neural models. After training TnT on the silver data, it correctly tagged 86.89% of tokens in the gold data: almost 5% improvement over the most frequent tag per-word baseline (82.18%). This accuracy seems promising for bootstrapping application.[12]

Besides the application in the PMB, Bjerva et al. (2016) showed that using sem-tags as auxiliary information significantly improves English Universal Dependencies POS tagging. Given that Boxer and similar semantic parsing scenarios are commonly used (Mineshima et al., 2015; Beltagy et al., 2016; Lewis and Steedman, 2013), semantic tagging will help those researches to shift to a cross-lingual level. Additionally, multilingual semantic parsing approaches might also benefit from semantic tagging. For example, sematic tags can help UDEPLAMBDAReddy et al. (2017) to decrease efforts of looking up lexical information for several words, e.g., quantifiers and negation markers.

5 Conclusion

We have proposed a novel NLP task that contributes to wide-coverage cross-lingual semantic parsing. Tagging tokens with universal semantic tags represents an independent task that unifies and generalizes over semantic virtues of POS-tagging and NE recognition. The expressive semantic tagset allows disambiguation of various semantic phenomena. Besides their application in semantic parsing, already demonstrated in the PMB project, sem-tags can contribute to other NLP tasks, e.g. POS tagging, or research lines rooted in compositional semantics.

In contrast to POS tagsets (Marcus et al., 1993; Petrov et al., 2012) augmented with morphological/universal features (Sylak-Glassman, 2016; Nivre et al., 2016), the semantic tagset is less expressive from a morphological perspective. On the other hand, the tagset is leaner and models several semantic phenomena, e.g., roles (`ROL`), subsectives (`SST`), privatives (`PRI`), and degrees (`DEG`), that are beyond morphology. Compared to the standard NE classes (Sang and Meulder, 2003), the named entity class (NAM) of the tagset is broader. The annotations of temporal expressions at TempEval (UzZaman et al., 2013) and MUC-7 (Chinchor and Robinson, 1998) differ from semantic tagging in terms of granularity: they annotate entire time expressions (e.g., *August of 2014*) while the semantic tagset opts for a more compositional analysis.

In future research, we plan to annotate more data with the help of human annotators, automatically tag large monolingual data via bootstrapping, further improve cross-lingual projection of sem-tags, and prepare an annotation guideline for semantic tagging. Elaboration of compositional semantics in the PMB might lead to an additional refinement of the semantic tagset. For example, one can distinguish

[12]This result of the TnT tagger is not directly comparable to the result (83.6%) of the neural semantic tagger reported by Bjerva et al. (2016) since the experiments differ in terms of training/test data and the semantic tagset.

genders or animacy for better pronoun resolution or mark plurality information for better semantic analysis.

Acknowledgements

This work has been supported by the NWO-VICI grant "Lost in Translation Found in Meaning" (288-89-003). We also wish to thank the three anonymous reviewers for their helpful comments.

References

Abzianidze, L., J. Bjerva, K. Evang, H. Haagsma, R. van Noord, P. Ludmann, D.-D. Nguyen, and J. Bos (2017, April). The parallel meaning bank: Towards a multilingual corpus of translations annotated with compositional meaning representations. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, Valencia, Spain, pp. 242–247. Association for Computational Linguistics.

Arivazhagan, N., C. Christodoulopoulos, and D. Roth (2016). Labeling the semantic roles of commas. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February 12-17, 2016, Phoenix, Arizona, USA.*, pp. 2885–2891.

Beltagy, I., S. Roller, P. Cheng, K. Erk, and R. J. Mooney (2016). Representing meaning with a combination of logical and distributional models. *Computational Linguistics 42*(4), 763–808.

Bjerva, J., B. Plank, and J. Bos (2016). Semantic tagging with deep residual networks. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, Osaka, Japan, pp. 3531–3541.

Bos, J. (2008). Wide-coverage semantic analysis with boxer. In J. Bos and R. Delmonte (Eds.), *Semantics in Text Processing. STEP 2008 Conference Proceedings*, Research in Computational Semantics, pp. 277–286. College Publications.

Bos, J. (2009). Towards a large-scale formal semantic lexicon for text processing. In C. Chiarcos, R. Eckart de Castilho, and M. Stede (Eds.), *From Form to Meaning: Processing Texts Automatically. Proceedings of the Biennal GSCL Conference 2009*, pp. 3–14.

Bos, J. (2014). Semantic annotation issues in parallel meaning banking. In *Proceedings of the Tenth Joint ACL-ISO Workshop on Interoperable Semantic Annotation (ISA-10)*, Reykjavik, Iceland, pp. 17–20.

Bos, J. (2015). Open-domain semantic parsing with Boxer. In B. Megyesi (Ed.), *Proceedings of the 20th Nordic Conference of Computational Linguistics (NODALIDA 2015)*, pp. 301–304.

Bos, J., S. Clark, M. Steedman, J. R. Curran, and J. Hockenmaier (2004). Wide-coverage semantic representations from a CCG parser. In *Proceedings of the 20th International Conference on Computational Linguistics (COLING 2004)*, Geneva, Switzerland, pp. 1240–1246.

Brants, T. (2000). Tnt: A statistical part-of-speech tagger. In *Proceedings of the Sixth Conference on Applied Natural Language Processing*, ANLC '00, Stroudsburg, PA, USA, pp. 224–231. Association for Computational Linguistics.

Chinchor, N. and P. Robinson (1998). Appendix e: Muc-7 named entity task definition (version 3.5). In *Seventh Message Understanding Conference (MUC-7): Proceedings of a Conference Held in Fairfax, Virginia, April 29 - May 1, 1998.*

Doddington, G., A. Mitchell, M. Przybocki, L. Ramshaw, S. Strassel, and R. Weischedel (2004, May). The automatic content extraction (ace) program tasks, data, and evaluation. In *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC-2004)*, Lisbon, Portugal. European Language Resources Association (ELRA).

Evang, K. and J. Bos (2016). Cross-lingual learning of an open-domain semantic parser. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, Osaka, Japan, pp. 579–588.

LDC (2008). *ACE (Automatic Content Extraction) English Annotation Guidelines for Events* (Version 6.6 2008.06.13 ed.).

Lewis, M. and M. Steedman (2013). Combined distributional and logical semantics. *Transactions of the Association of Computational Linguistics 1*, 179–192.

Löbner, S. (2013). *Understanding Semantics, Second Edition*. Understanding Language. Taylor & Francis.

Marcus, M. P., M. A. Marcinkiewicz, and B. Santorini (1993). Building a large annotated corpus of english: The penn treebank. *Computational Linguistics 19*(2), 313–330.

Mineshima, K., P. Martínez-Gómez, Y. Miyao, and D. Bekki (2015, September). Higher-order logical inference with compositional semantics. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, Lisbon, Portugal, pp. 2055–2061. Association for Computational Linguistics.

Nivre, J., M.-C. de Marneffe, F. Ginter, Y. Goldberg, J. Hajic, C. D. Manning, R. McDonald, S. Petrov, S. Pyysalo, N. Silveira, R. Tsarfaty, and D. Zeman (2016, may). Universal dependencies v1: A multilingual treebank collection. In N. C. C. Chair), K. Choukri, T. Declerck, S. Goggi, M. Grobelnik, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, and S. Piperidis (Eds.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, Paris, France. European Language Resources Association (ELRA).

Petrov, S., D. Das, and R. McDonald (2012, may). A universal part-of-speech tagset. In N. C. C. Chair), K. Choukri, T. Declerck, M. U. Doan, B. Maegaard, J. Mariani, A. Moreno, J. Odijk, and S. Piperidis (Eds.), *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, Istanbul, Turkey. European Language Resources Association (ELRA).

Reddy, S., O. Täckström, M. Collins, T. Kwiatkowski, D. Das, M. Steedman, and M. Lapata (2016). Transforming Dependency Structures to Logical Forms for Semantic Parsing. *Transactions of the Association for Computational Linguistics 4*, 127–140.

Reddy, S., O. Täckström, S. Petrov, M. Steedman, and M. Lapata (2017). Universal semantic parsing. *CoRR abs/1702.03196*.

Sang, E. F. T. K. and F. D. Meulder (2003). Introduction to the conll-2003 shared task: Language-independent named entity recognition.

Santorini, B. (1990). Part-Of-Speech tagging guidelines for the Penn Treebank project (3rd revision, 2nd printing). Technical report, Department of Linguistics, University of Pennsylvania, Philadelphia, PA, USA.

Steedman, M. (2001). *The Syntactic Process*. Cambridge, Ma., USA: The MIT Press.

Sylak-Glassman (2016). The composition and use of the universal morphological feature schema (Unimorph schema). Technical report, Johns Hopkins University.

UzZaman, N., H. Llorens, L. Derczynski, J. Allen, M. Verhagen, and J. Pustejovsky (2013, June). Semeval-2013 task 1: Tempeval-3: Evaluating time expressions, events, and temporal relations. In *Second Joint Conference on Lexical and Computational Semantics (*SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013)*, Atlanta, Georgia, USA, pp. 1–9. Association for Computational Linguistics.