Common Sense Inference Using Verb Valency Frames

Zuzana Nevěřilová and Marek Grác

NLP Centre, Faculty of Informatics, Masaryk University
Botanická 68a, 602 00 Brno, Czech Republic
{xpopelk,xgrac}@fi.muni.cz

Abstract. In this paper we discuss common-sense reasoning from verb valency frames. While seeing verbs as predicates is not a new approach, processing inference as a transformation of valency frames is a promising method we developed with the help of large verb valency lexicons. We went through the whole process and evaluated it on several levels: parsing, valency assignment, syntactic transformation, syntactic and semantic evaluation of the generated propositions.

We have chosen the domain of cooking recipes. We built a corpus with marked noun phrases, verb phrases and dependencies among them. We have manually created a basic set of inference rules and used it to infer new propositions from the corpus. Next, we extended this basic set and repeated the process. At first, we generated 1,738 sentences from 175 rules. 1,633 sentences were judged as (syntactically) correct and 1,533 were judged as (semantically) true. After extending the basic rule set we generated 2,826 propositions using 276 rules. 2,598 propositions were judged correct and 2,433 of the propositions were judged true.

1 Introduction

The cookbook "story" *fry the onion till it looks glassy* means peel a fresh, uncooked onion, chop it, put grease into a cooking pot and heat it, put the onion into the pot and wait until the onion looks glassy. In NLP systems we have to deal with implicit information to resolve "stories" such as: fry the onion till it looks glassy, reduce heat and cover. Where the heat comes from? What to cover?

Texts in natural languages usually contain "facts" (also known as common sense propositions or common sense facts) that are considered to be true in "normal" situations (also referred as stereotypical information [1]), e.g. fried onion looks glassy. This information is obvious for humans therefore rarely mentioned. The problem of the implicit information has been recognized since the beginning of the AI research. There are many approaches including frames [2] or scripts [3].

In this paper we concentrate on inferring new propositions from verb valency frames. The technique is based on transformations[1] on syntactic level and evaluation

[1] The word *transformation* is not linked to Chomsky's transformational grammar, but to Sowa's broader definition of logic as "any precise notation for expressing statements that can be judged true or false". In the same context an inference rule is defined as "a truth-preserving transformation: when applied to a true statement, the result is guaranteed to be true" [4].

P. Sojka et al. (Eds.): TSD 2012, LNCS 7499, pp. 328–335, 2012.

on semantic level. The system works with syntactic units – noun phrases (NP)[2] and verb phrases (VP), but during the evaluation the meaning of the proposition is examined. A corpus of cooking recipes was created and the work is thus related to the cooking domain. This domain is quite strictly delimited and moreover it contains verbs describing mostly actions (fry, pour, cut etc.).

The aim of this inferencing prototype is to answer questions such as "do this meal contain gluten?" or "do I need a blender to cook this meal?" While the answer to the former question is not yet reachable, the answer to the latter can be provided already.

The paper structure follows: Section 2 depicts using verbs as predicates w.r.t. Czech. Section 3 focuses on recognizing textual entailment. In Section 4 we describe the whole process of inferring new propositions in detail. We start with describing the nature of the cooking recipes language, the processes of annotating the corpus and building inference rules. Afterwards, the inference algorithm is provided with example outputs. The number of rules increased automatically using verb valency lexicon VerbaLex. Section 5 provides evaluation and discussion respectively. Section 6 proposes further development directions.

2 Verb Frames and Semantics

Verbs mostly describe an action or state. Since the verb "is the hook upon which the rest of a sentence hangs" [5], it is often seen as a predicate (for example $tastelike(x, y)$ means that x tastes like y). *Verb valency* then refers to the number of arguments of a verbal predicate. *Syntactic valencies* describe the syntactic properties (such as subject or object) of an argument. In Czech (as well as most other Slavic languages) syntactic properties are expressed by the case and possibly a preposition (e.g. syntactic subject is in nominative).

Semantic valencies assign semantic roles to arguments of a verbal predicate. "A semantic role is the underlying relationship that a participant has with the main verb in a clause" [6].

VerbaLex [7] is a large valency lexicon of Czech verbs and their arguments (in frame lexicons often called slots). It captures the syntactic information (prepositions and cases of the arguments in VerbaLex) as well as semantics (reference to semantic roles and Princeton WordNet [8] (PWN) hypernym.

3 Relationship to Recognizing Textual Entailment

Recognizing Textual Entailment (RTE) is a sub-problem of NLP. Its focus is in determining if a statement (called *hypothesis*) can be inferred from a given text. RTE systems consist of syntactic parsing, role labeling, named entities recognition, logical representation and other modules.

Apart from the ad-hoc and shallow approaches the sound approaches (e.g. [9]) use tree transformation operations that generate the hypothesis from the given text and knowledge based operations.

[2] For the purpose of this paper we take noun phrases and prepositional phrases together and later call it NPs.

Since our work is linguistically-motivated we use syntactic parsing but the result is not a complex dependency tree but rather one or more small trees or better a *bush*. This notion is discussed in Section 4.2.

At the first stage we did not use a *knowledge base* for recognizing the entities in the text. However the assumption that knowledge base will improve the results is expressed at the end of this paper. On the other hand the set of manually created inference rules itself is a knowledge base. It adds information about arguments of the verbs and relations between different verb frames. It also introduces new entities such as ingredients and cookware.

4 Common Sense Inference in Cooking Recipes

4.1 The Language of Cooking Recipes

The language of cooking recipes differs from the general language in the following attributes:

- use of imperative. In Czech cooking recipes most cooking recipes authors use first person plural (literally "we fry the onion...") instead of imperative. Sometimes, infinitive or imperative forms are used. In all verbs occurring in cooking recipes 6 % were imperatives, 51 % were indicatives in first person plural, 11 % were infinitives (some of these infinitives are bare, i.e. are together with another verb such as "let the onion fry"). The remaining verbs were 3rd person indicatives (such as "the onion looks glassy").
- frequent use of phrasal coordinations of NPs and of VPs: in cooking recipes corpus there are approx. three times more coordinations than in a corpus of blog texts.

4.2 Building Annotated Corpus

The annotation method was that of the BushBank project [10]. The corpus was annotated on several language levels: tokens (words and sentence boundary marks), morphology (lemma and morphologic tag for tokens), syntactic structures (NPs, VPs, coordinations and clauses), relations between syntactic structures (dependencies). The annotation of tokens was done purely by annotators' intuition since it is straightforward for humans to detect word and sentence boundaries. Detecting boundaries of NPs and VPs (or better detecting errors in NPs' and VPs' boundaries) was also quite an easy task. For searching dependencies verb valency lexicon VerbaLex was used. However, we did not find a way of using VerbaLex automatically because of high semantic ambiguity of verbs. Therefore annotators only consulted VerbaLex during their work.

Data for annotation were obtained purely by automatic tools (desamb [11], SET [12]) and validity of syntactic structures and their relations were confirmed during manual annotation. This means that structures that were not identified by automatic tools could not be added by annotators.

This was done contrary to traditional requirements in which we tried to obtain *completeness* of the annotation. BushBank ideas put greater emphasis on *simplicity* of the annotation (without definition of all border-line cases), *usability* (proved by the evaluation described in Section 5) and *rapid-development* (annotation itself was done in 40 (wo)man-hours). As we are working on a concept, data were manually checked by just one annotator. We plan using at least two annotators in the future with measuring their agreement.

4.3 Inference Rules

The inference rule for a particular sentence contains an input verb phrase *Vinput*, the output verb phrase *Voutput*, information about the grammatical polarity[3] preservation n, information on how the arguments participate in the inference process (syntactic rules $\mathcal{S}$) and inference type t (see below), in short the rule I is a tuple $I = (Vinput, Voutput, n, \mathcal{S}, t)$. The grammatical polarity preservation allows to formulate rules that result in sentences with opposite grammatical polarity (e.g. "x cooks y" effects in "y is *not* raw").

Each syntactic rule $S \in \mathcal{S}$ is a pair of syntactic properties of the *Vinput* dependent *SPinput* and syntactic properties of the *Voutput* dependent *SPoutput*, in short $S = (SPinput, SPoutput)$. Syntactic property is a pair of the appropriate case of the dependent and a preposition. Prepositions can be either none (direct case) or prepositions agreeing to a case. Case is marked by a number[4].

The inference itself is a process of filling up an output verb frame with definite arguments and creating another verb frame with some of the previous arguments. The algorithm is described in Section 4.4.

```
<inference type="effect" verb="dochutit" mean="to_flavour">
  <ruleset id="taste_like" inf_verb="chutnat" negation="False">
    <rule case="c4" prep="" inf_case="c1" inf_prep=""/>
    <rule case="c7" prep="" inf_case="c6" inf_prep="po"/>
  </ruleset>
</inference>
```

Fig. 1. Example of the inference rule notation: *toflavour*(x, y, z) (has effect) *totaste*(y, z) means that "x flavours y with z" has effect "y tastes like z" (x is not part of the inference and therefore is not mentioned). Translation: *dochutit*=to flavour, *chutnat*=to taste, *po*=preposition (here meaning "like").

The system covers the following inference types t: effect (66 rules), precondition (47 rules), near synonymy (75 rules), conversion between active and passive verb forms (24 rules).

The inference rules were created by expert linguists according to their introspection and experience. Figure 1 shows an example of the inference rule description.

After creation the inference algorithm (described below) was applied and the resulting propositions were evaluated syntactically. This evaluation lead back to *a*) inference rules correction *b*) improvements of syntactic rules for sentence generation. Table 1 shows progressive improvements after each such cycle.

[3] The distinction of affirmative and negative.

[4] 1 – nominative, 2 – genitive, 3 – dative, 4 – accusative, 6 – locative, 7 – instrumental.

Table 1. Manual creation and evaluation of inference rules. After 3 cycles the system did significantly improve its outputs. At the same time the number of applied rules increased.

cycle	# of generated propositions	# of correct propositions	# of rules used
1	1,792	1,016	168
2	1,734	1,415	174
3	1,783	1,633	175

4.4 The Inference Algorithm and Example Outputs

For the verb phrase present in the sentence $Vinput \in Sinput$ find all inference rules that contain $Vinput$ as a input verb phrase.

For each $I = (Vinput, Voutput, n, \mathcal{S}, t)$:

1. find all dependents $D_1, \ldots, D_n$ of $Vinput$ in *Sinput*.
2. transform $Vinput$ to $Voutput$. Since the Czech language uses declination we have to find the right form of the output verb. This is done by *a*) determining the morphological tag of $Vinput$ (person, number, tense, polarity), *b*) generating an appropriate verb form for $Voutput$. For this purpose we have used the morphological analyzer/generator majka [13]. In case of passive verbs we have to transform the grammatical categories of the passive participle (number and gender) as well. This transformation depends on grammatical categories of the inferred subject ($SPoutput_i$ where case is nominative).
3. determine all dependents $\{D_1, \ldots, D_k\}$ that have syntactic properties of *SPinput*
4. transform all the dependents selected in the previous step according to their corresponding rule $S = (SPinput_i, SPoutput_i)$. The NP has to change its form according to the new case. The case was determined during the annotation, the preposition is changed (to that of $SPoutput_i$) and a new form of the NP is generated using the morphological analyzer/generator majka [13].
5. generate a sentence *Soutput* from $Voutput$ and transformed dependents. The new proposition is generated as a sequence of $SPoutput_1, \ldots, SPoutput_l$ and $Voutput$. Since Czech is a nearly free word order language, at least the NPs' order can be interchanged without worries.

The ideal case happens when all dependents in *Sinput* are transformed to dependents of *output*. Some inferences were incomplete. Examples of both types are shown in Table 2.

4.5 Adding New Rules Using VerbaLex

It is obvious that manual creation of inference rules cannot lead to significant results in short-term. We extended the existing rule set using VerbaLex. We extracted all verb frames that contain verbs $Vinput$ and one of their functors is `SUBS`(tance). Afterwards, we have created rules by replacing $Vinput$ by its synonyms in the particular verb frame. Using this procedure we added aspectual pairs, writing variants and synonyms with the same syntactic valencies.

The number of inference rules increased from 212 to 599. Since in VerbaLex, only 3% of verbs are biaspectual, about half of the new rules employ different aspect of the

Table 2. Example (syntactically and semantically correct) outputs

Sinput	*t*	*Soutput*
Rozpustíme máslo (Melt the butter)	effect	máslo může být tekuté (butter can be liquid)
Kuřecí prsa naklepeme natenko (Tenderize chicken breast thinly)	precondition	kuřecí prsa jsou druh masa (Chicken breast is a meat)
Nahrubo nastrouháme všechny sýry (Grate coarsely all sort of cheese)	precondition	na všechny sýry vezmeme struhadlo (We take a grater suitable for all sort of cheese)
Občas trochu podlít vodou (Occasionally baste with water)	equals	vodu nalíti (Pour water)

same verb (e.g. vysušit/perfective – vysoušet/imperfective – both meaning to dry up). The rest of the rules was generated from synonyms.

Afterwards, we did the same evaluation as described in Table 1. N.B. that the number of rules used in inference did not increase accordingly to the increase of the rules. It is caused by the fact that some of the late rules lead to sentences that are correct but unnatural, e.g. *tvarovat těsto (to shape a pastry)* is usual but *modelovat těsto (to form a pastry)* is not.

5 Evaluation

We have picked up 164 verbs occurring in Czech cooking recipes. For these verbs 212 inference rules were manually created. The inference process was tested on a corpus of 37 thousands tokens (2,400 sentences). As the result 1,783 new sentences were generated and evaluated.

Evaluation proceeded on two levels: syntactic and semantic. At each level annotators had to decide whether or not the new sentence is correct. We then observed and classified the types of errors.

From 1,783 sentences 1,633 were evaluated syntactically correct (using 175 rules). Afterwards, within 1,633 sentences 1,533 was evaluated semantically correct (the new proposition was judged true given the original proposition).

Next step went with extending the rule set automatically. By adding synonyms from VerbaLex, we increased the number of inference rules nearly three times – up to 599 rules. From these 599 rules, 276 were used, 2,826 propositions were generated and 2,598 of them were evaluated syntactically correct. From these 2,598 propositions 2,443 were evaluated as semantically correct.

Inference most often results in incomplete propositions. This is caused by frequent occurrence of verb coordinations and other ellipses. A working solution would be to generate sentences for the verb coordinations, e.g. from “chop onions, stir and cover” generate “chop onions”, “stir onions” and “cover (the pot containing) onions”. Ellipses could also be solved by anaphora resolution, e.g. “chop onions and put it into the pot” will lead to “chop onions” and “put onions into the pot”. These features should be implemented as a preprocessing module to the parser.

Second, parsing was not successful on NPs or VPs containing unknown words. For this reason we plan to use named entity taxonomies (e.g. from Wikipedia pages) as a preprocessing prior to parsing. Unknown words were exclusively from the food domain, e.g. amasaké (amazake drink), feta (feta cheese), mascarpone, žervé (fresh cheese). These words very often are not inflected.

6 Conclusion and Future Work

We made a prototype application for automatic inference. We went the whole way from syntactic parsing, NPs and VP detection, verb valency assignment and generating new sentences in Czech.

The results can be improved by a good preprocessing tools such as clause generator from verb coordinations and (domain specific) named entities recognizer. These tools should improve the syntactic parsing significantly but currently are not ready to use with Czech language. Involving taxonomies (such as Czech WordNet [14]) in the process seems to be a good choice, however currently we do not have one that is rich enough and contains Czech literals at the same time.

Second, we plan to create even more rules since we know that the coverage even on the restricted domain is still low. To obtain more inference rules automatically we plan to use verb grouping according to semantic role patterns (e.g. verbs such as pour, spill, sprinkle have the same semantic roles in their slots) [15].

The results have shown that we have to concentrate on automatic valency assignment instead of manual annotation. This next step can lead to the last one on a way from a prototype to a working application – an user interface that will allow users to ask questions on cooking recipes.

Acknowledgments. This work has been partly supported by the Ministry of Education of CR within the LINDAT-Clarin project LM2010013 and by the Czech Science Foundation under the project P401/10/0792.

References

1. Lehmann, D.J.: Stereotypical reasoning: Logical properties. CoRR cs.AI/0203004 (2002)
2. Minsky, M.: A framework for representing knowledge. Technical report, Massachusetts Institute of Technology, Cambridge, MA, USA (1974)
3. Schank, R.C., Abelson, R.P.: Scripts, Plans, Goals, and Understanding: An Inquiry Into Human Knowledge Structures (Artificial Intelligence), 1st edn. Lawrence Erlbaum Associates, Hardcover (July 1977)
4. Sowa, J.: Fads and fallacies about logic. IEEE Intelligent Systems, 84–87 (2007)
5. Kingsbury, P., Kipper, K.: Deriving verb-meaning clusters from syntactic structure. In: Proceedings of the HLT-NAACL 2003 workshop on Text meaning, vol. 9. HLT-NAACL-TEXTMEANING 2003, pp. 70–77. Association for Computational Linguistics, Stroudsburg (2003)
6. Loos, E.E., Anderson, S., Dwight, H., Day, J., Jordan, P.C., Wingate, J.D.: Glossary of linguistic terms (January 2004), http://www.sil.org/linguistics/GlossaryOfLinguisticTerms/

7. Hlaváčková, D., Horák, A.: Verbalex – new comprehensive lexicon of verb valencies for Czech. In: Proceedings of the Slovko Conference (2005)
8. Fellbaum, C.: WordNet: An Electronic Lexical Database (Language, Speech, and Communication). MIT Press, Hardcover (May 1998)
9. Stern, A., Dagan, I.: A confidence model for syntactically-motivated entailment proofs. In: Proceedings of the International Conference Recent Advances in Natural Language Processing 2011, RANLP 2011 Organizing Committee, Hissar, Bulgaria, pp. 455–462 (September 2011)
10. Grác, M.: Case study of BushBank concept. In: PACLIC 25th Pacific Asia Conference on Language, Information and Computation, pp. 353–361 (2011)
11. Šmerk, P.: Towards Morphological Disambiguation of Czech. Ph.D. thesis proposal, Faculty of Informatics, Masaryk University (2007)
12. Kovář, V., Horák, A., Jakubíček, M.: Syntactic Analysis Using Finite Patterns: A New Parsing System for Czech. In: Vetulani, Z. (ed.) LTC 2009. LNCS, vol. 6562, pp. 161–171. Springer, Heidelberg (2011)
13. Šmerk, P.: Fast morphological analysis of Czech. In: Proceedings of the RASLAN Workshop 2009, Masarykova univerzita (2009)
14. Pala, K., Smrž, P.: Building Czech Wordnet (7), 79–88 (2004)
15. Nevěřilová, Z.: Semantic Role Patterns and Verb Classes in Verb Valency Lexicon. In: Sojka, P., Horák, A., Kopeček, I., Pala, K. (eds.) TSD 2010. LNCS, vol. 6231, pp. 150–156. Springer, Heidelberg (2010)