Spoken Language Understanding in dialogue systems, using a 2-layer Markov Logic Network: improving semantic accuracy

Ivan V. Meza-Ruiz, Sebastian Riedel, and Oliver Lemon*
School of Informatics
University of Edinburgh
I.V.Meza-Ruiz,S.R.Riedel@sms.ed.ac.uk, olemon@inf.ed.ac.uk

Abstract

We describe a two layer Markov Logic Network (MLN) model for the Spoken Language Understanding (SLU) task in dialogue systems. We augment the set of features used in Meza-Ruiz et al. (2008) with the help of off-the-shelf resources. We show that this setup increases the performance of the previous MLN models, which also outperform the state-of-the art "Hidden Vector State" (HVS) model of He and Young 2006. In particular the 2 layer approach produces more accurate sets of slot-values for user utterances (9% improvement).

1 Introduction

The Spoken Language Understanding (SLU) task in dialogue systems consists in producing semantic representations for user utterances. In this work, our approach is trained on slot-value representations which are a common choice in the development of dialogue systems. Table 1 shows an example of slot-values as a semantic representation.

USER:*what flights are there arriving in Chicago on continental airlines after 11pm*
$GOAL$ =FLIGHT $TOLOC.CITY\_NAME$ =Chicago $AIRLINE\_NAME$ =continental_airlines $ARRIVE\_TIME.TIME\_RELATIVE$ =after $ARRIVE\_TIME.TIME$ =11pm

Table 1: Slot-values as a semantic representation.

*This work is partially funded by EPSRC grant number EP/E019501/1 and the EC FP7 project "CLASSiC" (ICT-216594). www.classic-project.org.

In particular, we are exploring robust statistical models of the SLU task using the Markov Logic Network (MLN) framework(Richardson and Domingos, 2007). An MLN is a collection of weighted First Order Logic (FOL) formulae that serves as a template to instantiate complex Markov Networks (MNs). MLNs are particularly interesting for language modelling because they are easy extensible with new features and allow the use of complex relations between nodes of the networks. Figure 1 shows a MN for a slot-value representation. In this case, the lighter nodes represent the hidden variables (i.e., the slots to produce) while the darker nodes represent the observable variables (i.e., the words of the utterance).

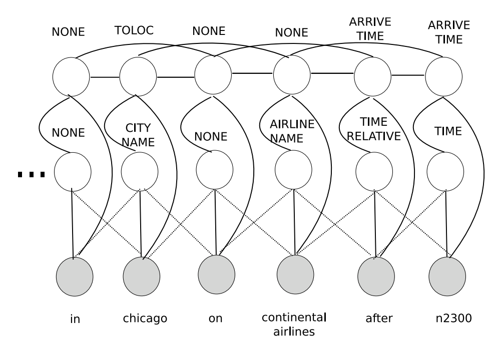

Figure 1: Markov Network as slot-values

In this paper, we focus on two aspects. First, the use of a two-layer MLN model to represent the slots. And second, the use off-the-shelf resources to extend the set of features available (e.g. POS taggers).

2 The MLN model

We split the SLU task into two. The first task consists into modelling the $GOAL$ element of the slot-values. Figure 2 shows a MN for the goal of our example. You can notice, that the $GOAL$ element, depends on the whole utterance. The second task consists of modelling the rest of the slots. Figure 1 shows a two-layer MN for the slots of our example. The first layer can be seen as a named entity, while the second layer represents a modifier/function for those named entities.

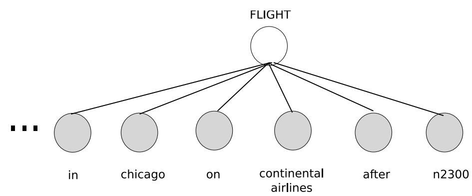

Figure 2: Markov Network for the goal slot

With the two layer model we aim to capture the relations between elements which constitute the slots. This is achieved by specifying the edges between the two layers using the FOL of the MLN. This model also includes the 1^{st} and 2^{nd} order Markov assumptions for the second layer. With these we aim to capture any dependency in the sequence of slots.

2.1 Feature extensions

In MLNs it is possible to add more observable variables which will be related to the input words. For this purpose, we use off-the-shelf resources to generate (i) POS tags for the words of the utterances using the TnT tagger (Brants, 2000), and (ii) syntactic chunks for the words of the utterances using the CASS chunker (Abney, 1996). With this information we define the following features for the slot model: Orthography and POS task of the word for a window of two previous and next words; a binary feature if the word is a number or the word is unknown ; and the head words of syntactic chunks.

3 Experiments and Results

For our experiments we use the Air Travel Information System corpus extended by (He and Young, 2006). This version is composed of slot-value labellings of the ATIS-2 and ATIS-3 training sets, and the ATIS-3 $NOV93$ testing set. We measured the *global score* and *exact match* metrics. This first metric measures the amount of slot-values recovered in the whole experiment, while the second one measures the exact set of slot-values recovered for each utterance.

The experiments tested the two layer models described in the previous sections. Table 2 presents the final results. Our baseline corresponds to the previous best result for the task, which outperforms the HVS model of (He and Young, 2006) on the labelling task (Meza and et al., 2008). In this case, the baseline is our starting point. The $MLN_{2-layer}$ model uses the features described in the previous section, plus a two layer model. The difference bewteen both models is statistically significant with $\rho < 0.05$.

$MLN_{baseline}$	Global	91.56%
	Exact	69.89%
$MLN_{2-layer}$	Global	92.99%
	Exact	78.97%

Table 2: f-scores for baseline and two layer model

4 Summary and discussion

We have shown an improvement on the performance of the SLU task by using a two layer model and by augmenting extra features in our previous model. In particular, the improvement of 9.08% in the *exact match* metric is interesting. This is because it shows that not only was the model able to identify more slot-values, but the set of slot-values for each utterance were more accurate.

References

Steven Abney. 1996. Partial Parsing via Finite-State Cascades. In *Natural LanguageEngineering*, volume 2(4), 337-344.

Thorsten Brants. 2000. TnT – a statistical part-of-speech tagger. In *ANLP*.

Yulan He and Steve Young. 2006. Spoken Language Understanding using the Hidden Vector State model, *Speech Communication*, volume 48(3-4), 262-275.

Ivan Meza, Sebastian Riedel and Oliver Lemon. 2008. Accurate Statistical Spoken Language Understanding from limited development resources, In *ICCASP08*.

Matt Richardson and Pedro Domingos. 2007. Markov Logic Networks. *Machine Learning*, volume 62, 107-136.