Using Le

School of C

Abstract. In natural languages the default specification of arguments of verbs is often omitted in the surface form. We address the need for commonsense knowledge in computational lexicons, and discuss the requisite lexical knowledge of computational lexicons in language-to-vision applications. We compare existing computational lexicons such as WordNet, FrameNet, LCS database, and VerbNet, and show how lexical knowledge in a generative lexicon can be used for disambiguation and commonsense inferencing to fill unspecified argument structures for the task of language visualisation. The possibility of lexical inference with WordNet is explored in order to extract default arguments of verbs.

Background: CONFUCIUS

Semantic analysis within CONFUCIUS' NLP module uses WordNet and the LCS database to fill underspecified arguments of verbs, and to

Lexical Knowledge of Verbs i

Minhua Ma and
f Computing & Intelligent Systems, Faculty of Enginee
{m.ma, p.mckevitt
http://www.infm.uls

Comparison of Computational Lexicons

WordNet does not have enough knowledge for compositional information of verbs, default instrument and functional information, which could be complemented by LCS database and VerbNet. However, the selection restrictions of the instrument argument in both lexicons are insufficient for language-to-vision applications.

Lexicons	*WordNet*	*FrameNet*	*LCS DB*	*VerbNet*
Semantic domains	all	limited	all	all
POS	all	all	verb	verb
Hypernymy (is_a)	+	+	+	+
Hyponymy (n.) troponymy (v.)	+	+	-	-
Metronymy constructive (n.) compositional (v.)	+ (n.) - cause (v.)	-	+ conceptual structure	+ decompose with time func
Instrument	-	-	? selection restrictions	? selection restrictions
Functional information (telic role)	-	+ used_by	n/a	n/a

in Language-to-Vision Appli

d Paul Mc Kevitt
neering, University of Ulster, Derry, BT48 7JL, Norther
tt}@ulster.ac.uk
lst.ac.uk/~eunice

"knife" is a default assignment of instrument for "cut". We propagate through its hypernym tree (Fig. 2) and find that all hyponyms of "edge tool" are acceptable instruments of cutting, the same for "cutter, cutlery, cutting tool" and "cutting implement". When propagating through one more level, not all children of "tool" are appropriate to serve as instruments of cutting. Some hypernyms of "tool" (e.g. "drill", "comb") cannot be used for cutting. Therefore "cutting implement" is the highest node of possible instruments for "cut" and should be stored as a default argument in its lexical entry.

```
knife
 => edge tool
  => cutter,cutlery,cutting tool
   => cutting implement
```

lications

nern Ireland

- Qualia structure: *formal*, *constitutive*, *telic* and *agentive* roles
- Lexical inheritance

The default/shadow arguments and telic roles in a generative lexicon can complement WordNet with regard to instrument and functional information. Previous research in language-to-vision also proves the necessity of such information in the lexicon. Hence, using a generative lexicon to make inferences on given sentences is potentially useful for language-to-vision applications.

Conclusion

- Language-to-vision conversion relies on lexical knowledge, e.g. default arguments of verbs, which may not be included in existing computational lexicons

perform semantic inference, disambiguation and coreference resolution. The current prototype of CONFUCIUS visualises single sentences which contain action verbs with *visual valency* of up to three (Fig. 1).

a. Key frame of "John left the gym"

b. Key frame of "Nancy gave John a loaf of bread"

Fig. 1. CONFUCIUS' 3D animation output

Previous Language-to-Vision Applications

Their approaches to lexical/commonsense knowledge storage and retrieval vary according to different application domains.

Virtual human Jack (Parameterized Action Representation)

Wordseye (WordNet and its own transduction rules and object vocabulary)

SONAS

Narayanan's iconic language visualisation (Schank's scripts and CD theory)

Extracting Default Arguments of Verbs From WordNet

Language visualisation requires lexical/ commonsense knowledge such as default instruments of action verbs, functional information and usage of nouns. In Table 2, the default instruments (or themes) are the highest nodes of the hypernymy (is_a) tree in WordNet [2], all of whose children are possible instruments. We start from a *default assignment* (of instrument), then propagate upward, and check if all the hyponyms of this lexical item are acceptable. Continue the propagation until we reach a level at which at least one of whose hyponyms is not acceptable.

Table 2. Default instruments of verbs

Verb	*Default instrument/theme (highest node of possible candidates in WordNet)*	*Example instrument/ theme*
cut	cutting implement	knife, scissors, lancet
bake	oven	oven
fry	pan	frying pan
boil	pot	kettle, caldron
drive	self-propelled vehicle	car, tractor, van
write	writing implement	pen, chalk
adorn	decoration, ornament, ornamentation	flower, jewel
kill	weapon	gun, bomb, bow

```
=> tool
 => implement
  => instrumentation
   => artifact, artefact
    => ...
```

Fig. 2. The hypernym tree of "knife" in WordNet

This approach provides a flexible specification of lexical knowledge with a proper grain size, avoiding too-particular specifications. This selection algorithm could be automated based on corpus data and linguistic ontologies. A generative lexicon with this knowledge provides the capability of visualising various activities without hardcoding them as part of the animation library.

Generative Lexicon

Pustejovsky's [5] generative lexicon contains information that is sometimes regarded as commonsense knowledge. A generative lexicon has:

- Argument structure: *true arguments*, *default arguments*, *shadow arguments*, *adjuncts*
- Event structure

- Computational lexicons, WordNet, FrameNet, LCS database and VerbNet are analysed and compared
- A selection algorithm based on WordNet is proposed for finding the highest hypernym of default instruments/themes
- The theory of generative lexicon shows adequacy to fill underspecified roles
- Future work will address the issue of finding appropriate hyponyms from lexical knowledge and context

References

1. Dorr, B.J., Jones, D.: Acquisition of Semantic Lexicons: using word sense disambiguation to improve precision. In Evelyne Viegas (Ed.), *Breadth and Depth of Semantic Lexicons*, 79-98, Norwell, MA: Kluwer Academic Publishers (1999)
2. Fellbaum, C. (Ed.): *WordNet: An Electronic Lexical Database*, Cambridge, MA: MIT Press (1998)
3. Jackendoff, R.: *Semantic Structures*. Cambridge, MA: MIT Press (1990)
4. Levin, B.: *English verb classes and alternations: a preliminary investigation*. Chicago, USA: The University of Chicago Press (1993)
5. Pustejovsky, J.: *The Generative Lexicon*. MIT Press (1995)