Automatic prediction of aspectual class of verbs in context

Annemarie Friedrich and Alexis Palmer
Department of Computational Linguistics, Saarland University, Germany
ACL 2014, Baltimore, MD, USA

CLUSTER OF EXCELLENCE

Aspectual class of verbs in context

She **filled** the glass with juice. *(dynamic)*
'something happens'

The juice **fills** the glass. *(stative)*
'something is the case'

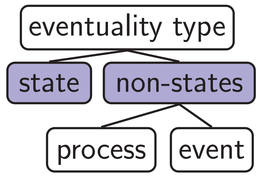
The glass was **filled** with juice.
(both readings possible)

Linguistic background & task

Vendler (1957): time schemata of **verbs** (*lexical aspect/aktionsart*)

states	*love, own*	**stative**
activities	*run*	**dynamic**
accomplishments	*write a letter*	
achievements	*realize*	

Bach (1986): **sentences**



Task: predict **fundamental aspectual class**

John will **love** this cake!	(*stative*)
John has **kissed** Mary.	(*dynamic*)
John **drives** to work.	(*dynamic*)

Siegel&McKeown (2000)

Method

LCS verb type seed sets
- *stative* (188): belong, cost,...
- *dynamic* (3760): alter, knock,...
- *both* (215): fill, stand, take,...

syntax-based distributional model (Thater et al. 2011)

The thrift holding company **said**	*dynamic*
it **expects** to obtain approval	*stative*
and **recognizes** the danger.	*both*

parsed, unlabeled background corpus (GigaWord)

distributional features (**Dist**)
avg. similarities with verbs in seed sets

linguistic indicators (**LingInd**)

past/present tense	frequency	perfect
progressive	negation	particles
in-PP	for-PP	no subject
temp./manner/evaluation/contin. adverbs		

verb type based

labeled clauses

instance-based features (**Inst**)

clause

Random Forest classifier

dynamic
stative
both

Type-based features

Linguistic indicators (LingInd)

verb type: `fill`
feature: `temporal-adverb`
value: 0.0085

⇒ 0.85% of the occurrences of *fill* are modified by one of the temporal adverbs.

Distributional features (Dist)

verb type: `fill`
sim. with *stative*: 0.31
sim. with *dynamic*: 0.27
sim. with *both*: 0.16

Instance-based features (Inst)

A little girl had just **finished** *her first week of school.*

`tense:`*past*	`progressive:`*false*
`pos:`*VBD*	`dobj:`*noun.time*
`perfect:`*true*	`particle:`*none*
`voice:`*active*	`subj:`*noun.person*

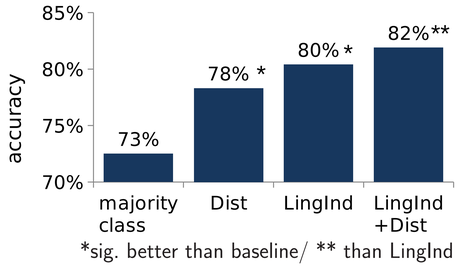
Experiments, results & conclusions

Data → freely available
- Asp-MASC: 6161 clauses (complete texts) excluding *be/have*, 2 annotators, $\kappa = 0.7$
- Asp-Ambig: 2667 sentences for 20 frequent ambiguous verbs, 2 annotators, $\kappa = 0.6$

Experiment 1: Seen verbs (labeled training data available)

Type-based features → same accuracy (84%) as only using **Lemma** of main verb as feature (= memorizing most frequent class per verb)

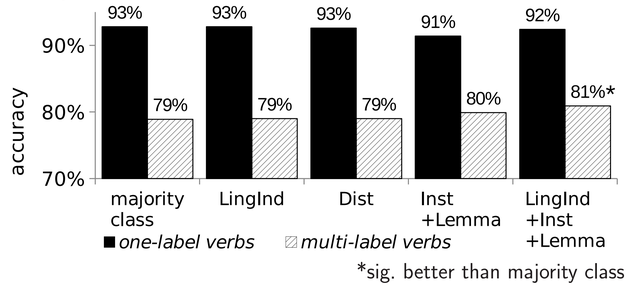
Experiment 2: Unseen verbs



*sig. better than baseline/ ** than LingInd

Type-based features generalize across verb types.

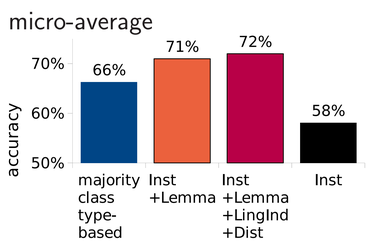
Experiment 3: One-label vs. multi-label

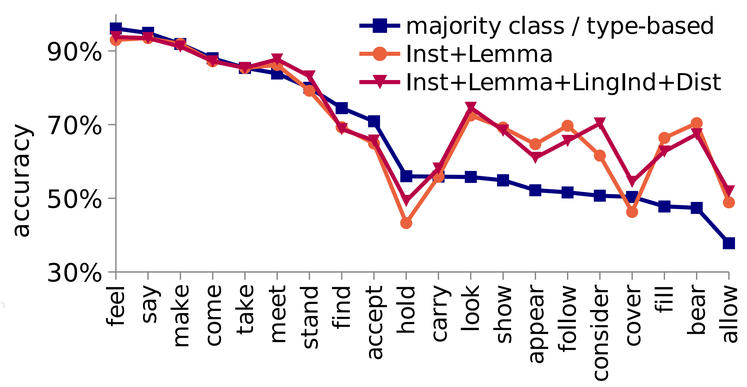


*sig. better than majority class

Instance-based features are essential for classifying ambiguous verbs.

Experiment 4: Instance-based
- Asp-Ambig





- Instance-based features do not generalize across verb types.
- The more ambiguous the verb, the more essential are instance-based features.
- Type-based features (bias) helpful? → depends on verb type

Future work

multi-stage approach: treat verbs according to their aspectual class distribution
→ bigger picture: improve (temporal) discourse processing

www.coli.uni-saarland.de {afried,apalmer}@coli.uni-saarland.de