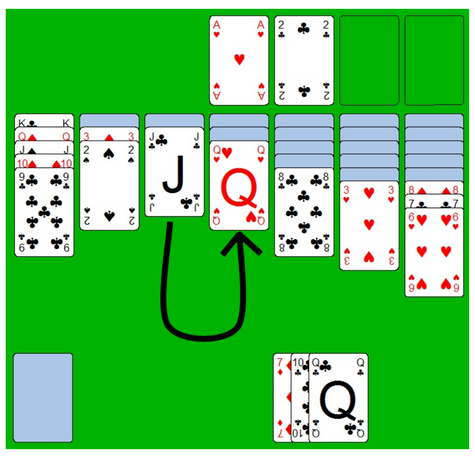
Weakly supervised concept tagging: Combining a generative and a discriminative approach

Janneke van de Loo[1], Guy De Pauw[1], Jort F. Gemmeke[2], Walter Daelemans[1]
[1]CLiPS, University of Antwerp [2]ESAT, KU Leuven

Application: a self-learning, adaptive vocal interface for physically impaired users, which learns the vocabulary and command structure of each individual user based on a set of example commands and associated actions / semantic frames. (Project ALADIN: http://www.aladinspeech.be/)

Data: transcribed Patience commands + associated semantic frames
1142 instances from 1 single speaker (from Patience corpus PATCOR)

Utterance: *Leg de klaveren boer op de harten dame*



Frame: MoveCard		
Slot		**Value**
<FromSuit>	FS	c
<FromValue>	FV	11
<FromColumn>	FC	3
<FromHand>	FH	-
<TargetSuit>	TS	h
<TargetValue>	TV	12
<TargetColumn>	TC	4
<TargetFoundation>	TF	-

Task: weakly supervised concept tagging

Concept tags: tags that refer to the slot values in the semantic frames.

- **Training data: transcribed utterances + associated semantic frames** (the semantic frames are unordered, redundant sets of concept tags)
- Task: given an unseen utterance, tag the utterance with concept tags
- Based on the concept tags, a semantic frame can be constructed

Leg	*de*	*harten*	*zes*	*op*	*de*	*schoppen*	*zeven*
O	O	I_FS=h	I_FV=6	O	O	I_TS=s	I_TV=7

Weak supervision: training material does **not** specify any alignments between words in the utterances and slots in the semantic frames.

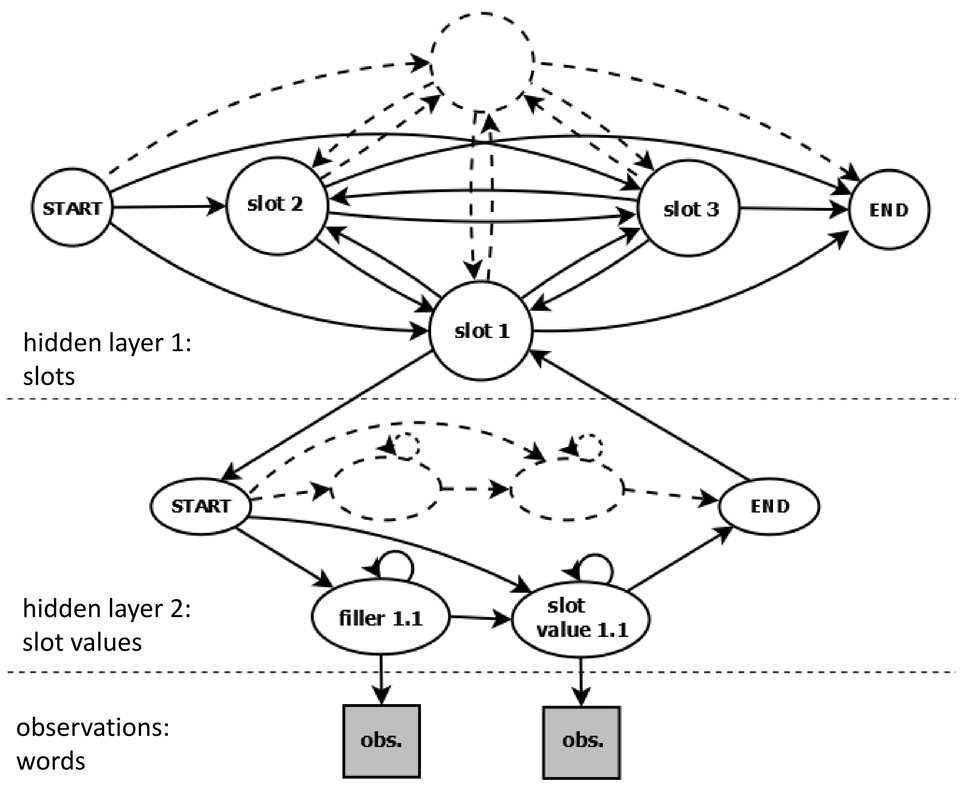
Combining a generative and a discriminative tagging approach

A. Train a generative tagger with weak supervision → tag training set **B.** Train a discriminative tagger with the tagged training set → tag test set

A. Generative, weakly supervised concept tagging: FramEngine
Based on hierarchical hidden Markov models (HHMMs)

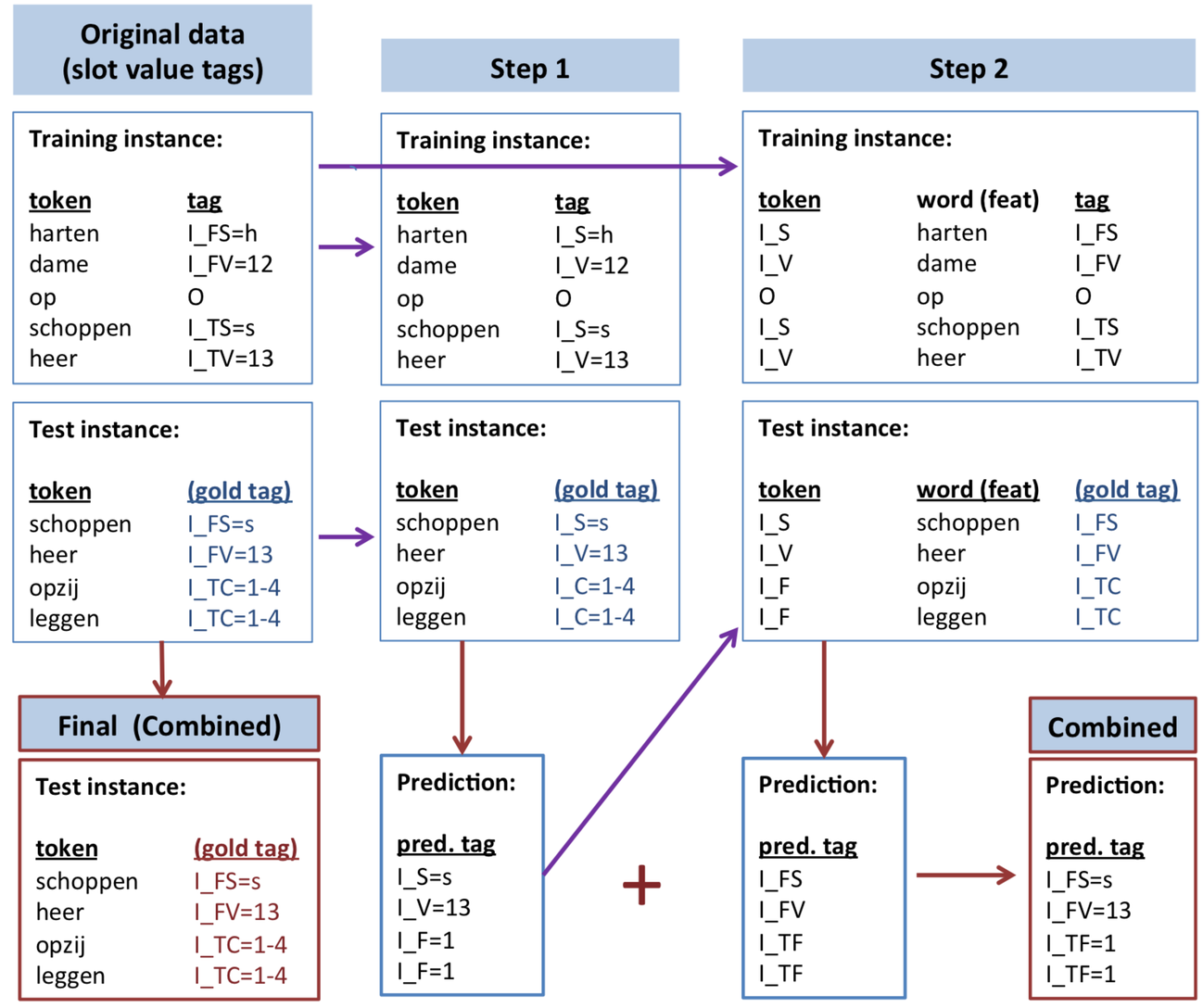
Apply generalisations by using parameter sharing techniques:

- Transition sharing: transition probs hold between slots rather than individual slot values: P(FS=h → FV=4) = P(FS=d → FV=5)
- Expression sharing: share the emission probability distributions of slot values that are likely to be expressed by the same words: P(FS=h → *harten* = P(TS=h → *harten*)



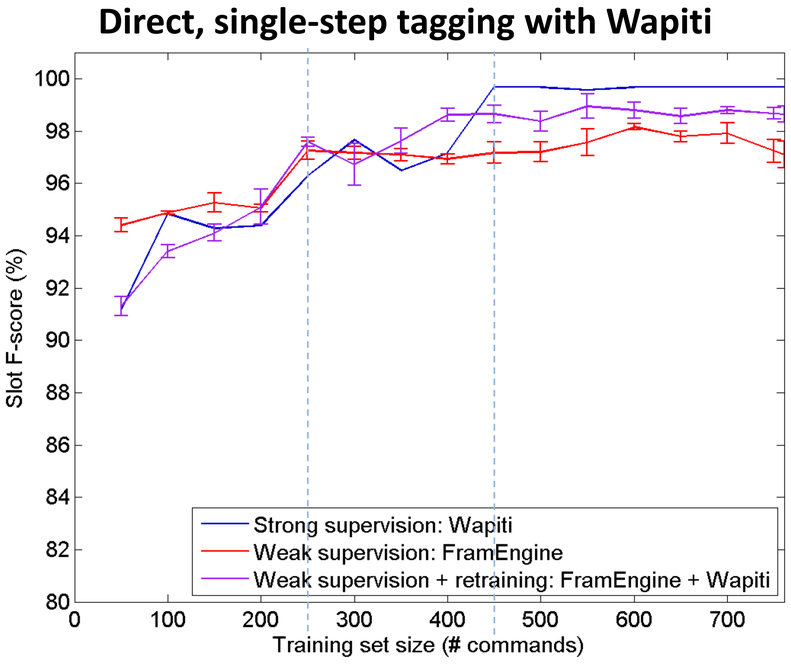
B. Discriminative, supervised concept tagging: Wapiti (Lavergne et al., 2010)
Based on Conditional Random Fields (CRFs)

Apply generalisations by using a two-step tagging approach:

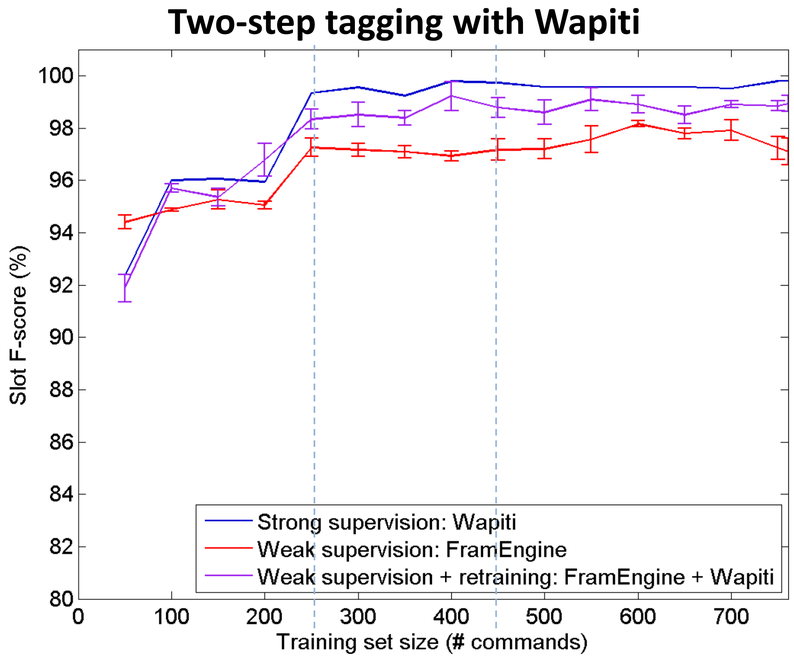


Experimental results (Test set: last 381 utterances = 1/3 of total)

Direct, single-step tagging with Wapiti



Two-step tagging with Wapiti



Main error cause at start of learning curves: until training utterance #200, only the word *koning* is used to refer to FV/TV=13, while in the test set, the synonym *heer* is used.
The word *heer* starts to appear as TV=13 from training size 250 and as FV=13 from training size 450.

Conclusions

For weakly supervised concept tagging:

- We can improve on FramEngine's tagging performance by combining FramEngine with a discriminative supervised concept tagger
- The main advantage is that the discriminative tagger can use right context as well as left context. This is especially beneficial for disfluent, noisy utterances → suitable for tagging speech-based data
- For small training set sizes, performance is improved by applying generalisation mechanisms in the supervised tagging step, i.e. by using a two-step tagging approach

E-mail: Janneke.vandeLoo@uantwerpen.be

Wapiti: T. Lavergne, O. Cappé, F. Yvon (2010). "Practical very large scale CRFs". In proceedings of ACL '10 , pp. 504-513.