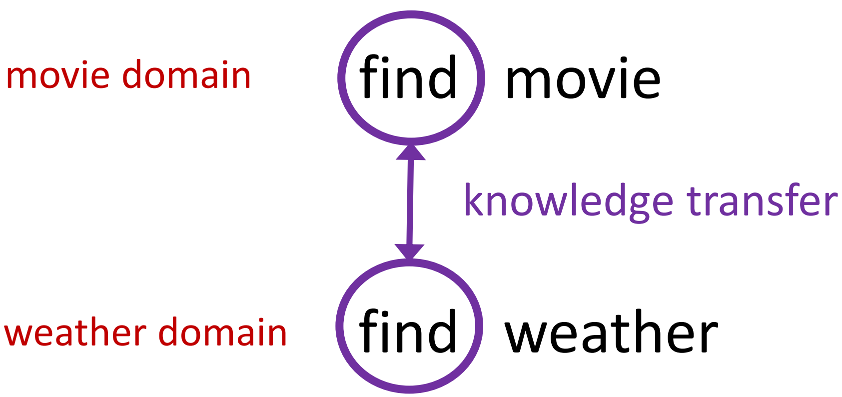
Learning Bidirectional Intent Embeddings by Convolutional Deep Structured Semantic Models for Spoken Language Understanding

Yun-Nung (Vivian) Chen, Dilek Hakkani-Tür, and Xiaodong He

- Motivation: ***Inflexible Intent Schema***
  - Intents are usually *predefined* and *inflexible* to expand and transfer across domains and genres
  - Re-designing a semantic schema requires manual annotation and model re-training.
- Approach: ***Learning Intent Representation***
  - Learn *high-level semantic representations* to bridge the semantic relation across domains and across genres

✓ Cross-domain: **intent expansion**

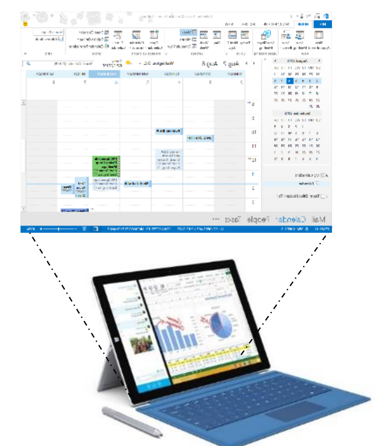


✓ Cross-genre: **actionable item detection**

Will Vivian come here for the meeting?

human-human

Find Calendar Entry



human-machine

CarnegieMellon

Microsoft Research

Learning Bidirectional Intent Embeddings by Convolutional Deep Structured Semantic Models for Spoken Language Understanding

Yun-Nung (Vivian) Chen, Dilek Hakkani-Tür, and Xiaodong He

➢ Model Architecture

Semantic Layer: y

Semantic Projection Matrix: W_s

Max Pooling Layer: l_m

Max Pooling Operation

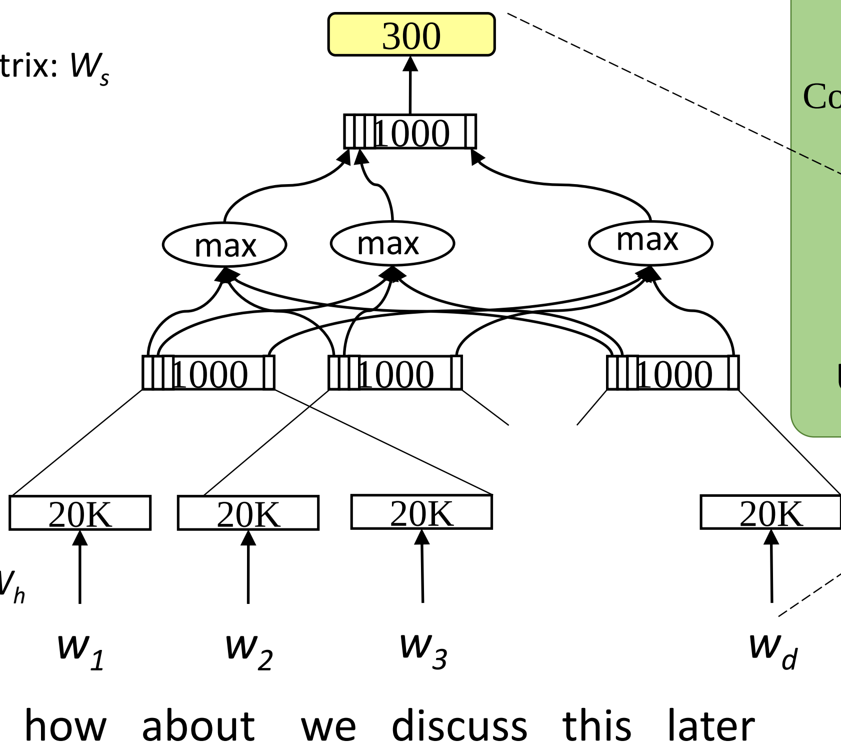
Convolutional Layer: l_c

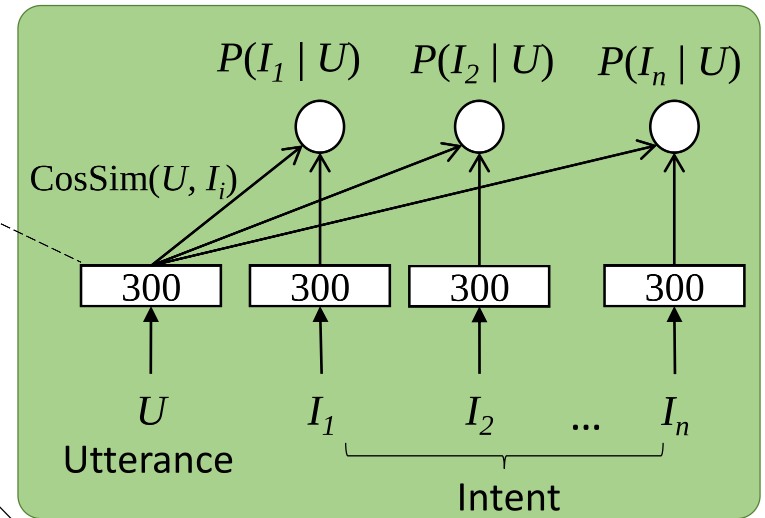
Convolution Matrix: W_c

Word Hashing Layer: l_h

Word Hashing Matrix: W_h

Word Sequence: x





$$P(I \mid U) = \frac{\exp(\text{CosSim}(U, I))}{\sum_{I'} \exp(\text{CosSim}(U, I'))}$$

➢ Training Procedure

$$\Lambda(\theta_1) = \log \prod_{(U, I^+)} P(I^+ \mid U)$$

→ **Predictive Model:** maximizes the likelihood of associated intents given utterances

$$\Lambda(\theta_2) = \log \prod_{(I, U^+)} P(U^+ \mid I)$$

→ **Generative Model:** maximizes the likelihood of generated utterances given user intents

Microsoft Research