

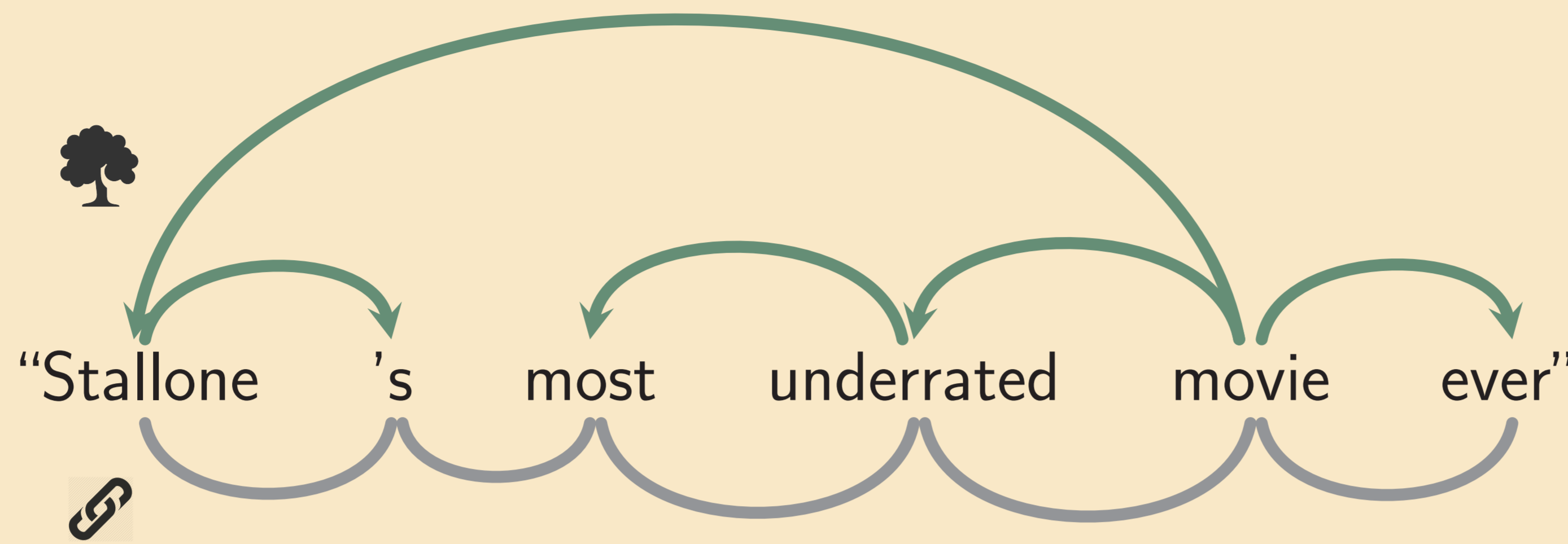
Extending Hidden Markov (tree) models for word representations



university of
 groningen

Simon Šuster and Gertjan van Noord
 Language Technology, University of Groningen, The Netherlands

PROBLEM



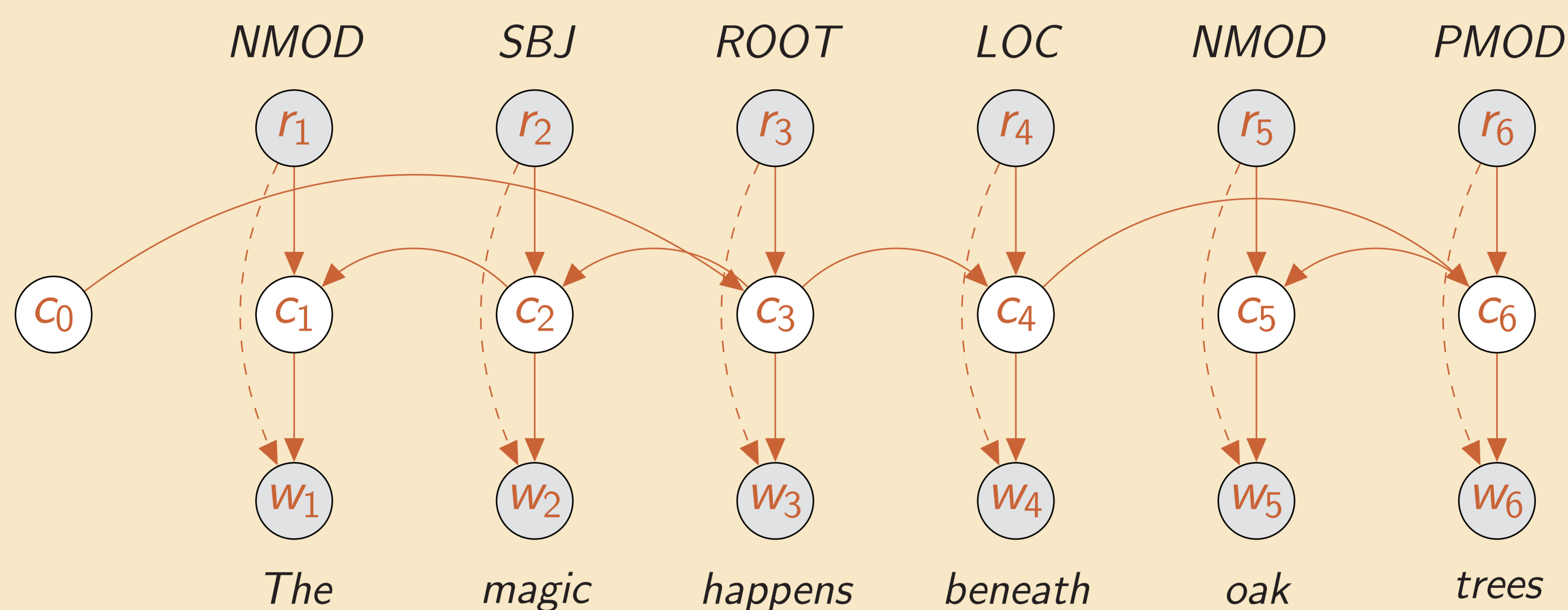
🌳 more precise than 🔗, but:
 🌳 models typically insensitive to underlying syntactic functions (SUBJ, OBJ etc.)

WHY

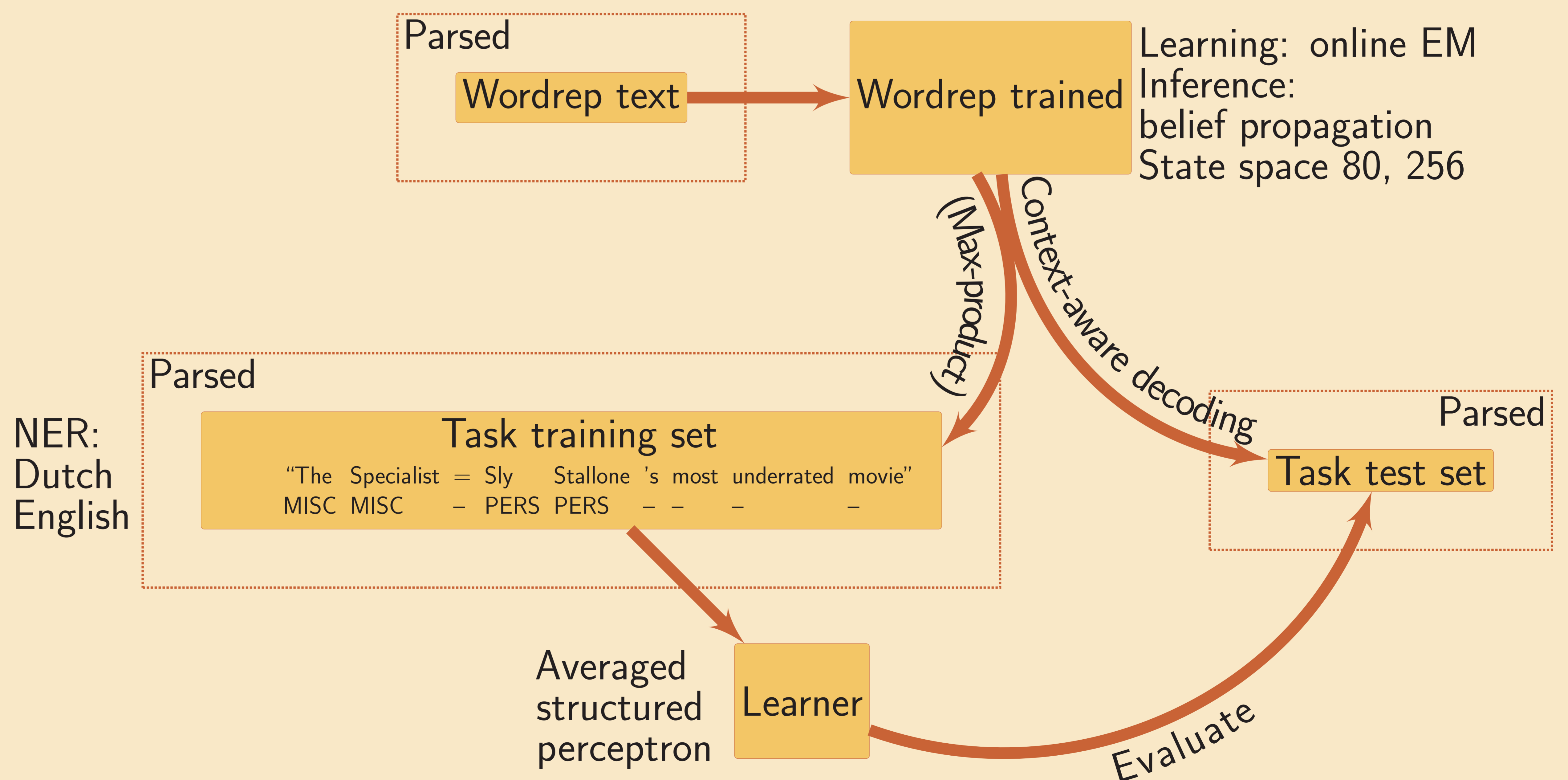
Word representations crucial for generalization in NLP
 Capture similarity in structure and meaning

IDEA

Latent variable tree model (\approx Input-Output HMM)



APPROACH



PROGRESS

Advantage of simple 🌳 over 🔗 HMMs not yet confirmed
 Performance (no syntactic functions) \approx Brown clusters

REFINE

state-splitting | pre-initialization | prefer sparse solutions | continuous instead of one-hot