

# Research Questions

Deep syntax/semantics & Deep Learning

Jan Hajič

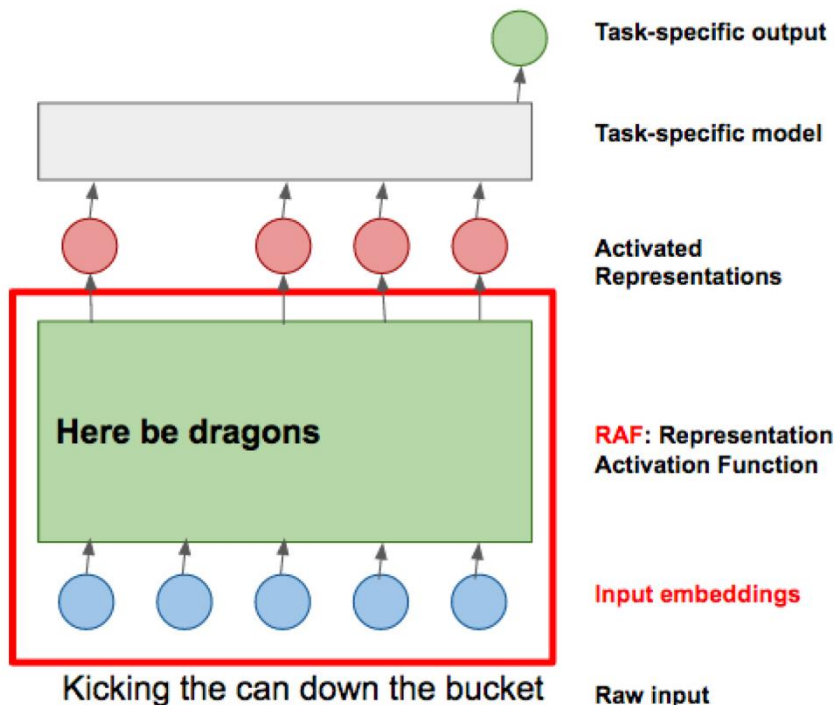
# Is there a place for linguistics in Deep Learning?

- Related questions
  - Where does “language system” ends and the “real world” starts?
  - What is the shared among languages?
    - Representation (> grammars / grammar formalisms)
- Reverse question:
  - Can Deep Learning help us to understand how language works?
    - Discovering features, categories, representations

# The “Red Box” (TM by H. Schütze et al.)

- Dagstuhl Jan. 2017 representation WG:

## Red Box: Key considerations

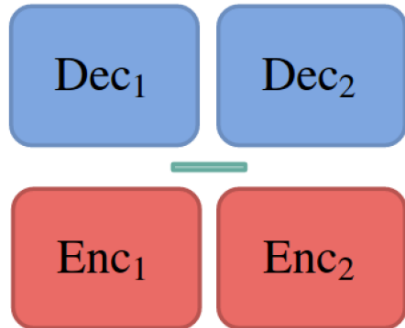


1. Activated representations = standard embeddings + some context + some syntax + some semantics
2. Aim: facilitate learning task-specific models
3. RA function equivalent to NLP pipeline?
4. **Training:** multi-task setup with joint core up to AR.
5. **RAF eval objective:** minimise sample complexity for downstream tasks

# The reverse question

- Shared representation as the “neural interlingua”

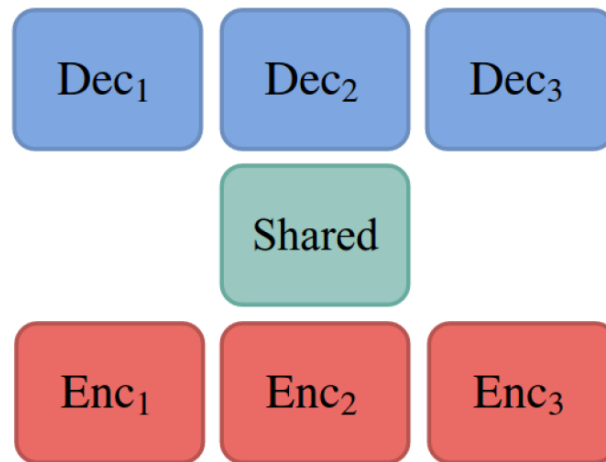
Fixed Length  
Representation



Luong et al. 2015

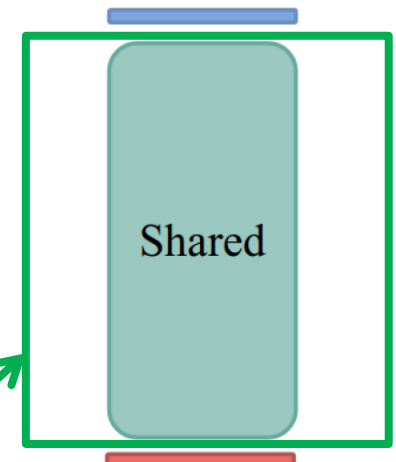
Drawing © by Orhan Firat

Shared Machinery:  
Attention Module



Firat et al. 2016a

Shared Machinery:  
Attention, Encoder, Decoder



How to  
analyze this?

Lee et al. 2016  
Johnson et al. 2016  
Ha et al. 2016