Made for Each Other: Broad-coverage Semantic Structures Meet Preposition Supersenses

Jakob Prange[1,*] Nathan Schneider[1] Omri Abend[2]

*jp1724@georgetown.edu [1]Georgetown University [2]The Hebrew University of Jerusalem

Motivation & Background

In the wild, we find a great variety of **preposition and case** usages. We can use meaning representations (MRs) to categorize these

- by similar and diverging **senses** and **structures**,
- comprehensively,
- and with a level of abstraction that can be applied to many genres and languages.

This can benefit **linguistic analysis**, as well as downstream **NLP tasks** like paraphrasing or MT.

UCCA (Abend and Rappoport, 2013)

- **predicate-argument structure**
- typologically-motivated
- 15 coarse categories (no semantic roles)

SNACS (Schneider et al., 2018)

- token-based **sense disambiguation** of a wide range of adpositional expressions
- incl. certain infinitivals and conjunctions
- 50 hierarchical categories, incl. semantic roles

Both meaning representations are **comprehensive**, **abstract**, and **language-agnostic**

Data

English Web Treebank
Genre: online reviews ★★★★★

Previously annotated with SNACS and UCCA, independently (Hershcovich et al., 2019)

4k sentences, 56k tokens, 8:1:1 train:dev:test

We **release** a new version of this corpus, annotated with our **integrated** representation.

Learning Architectures

Baseline: Transition-based UCCA parser with BiLSTM + MLP action classifier (Hershcovich et al., 2018);
SVM-based SNACS classifier (Schneider et al., 2018)

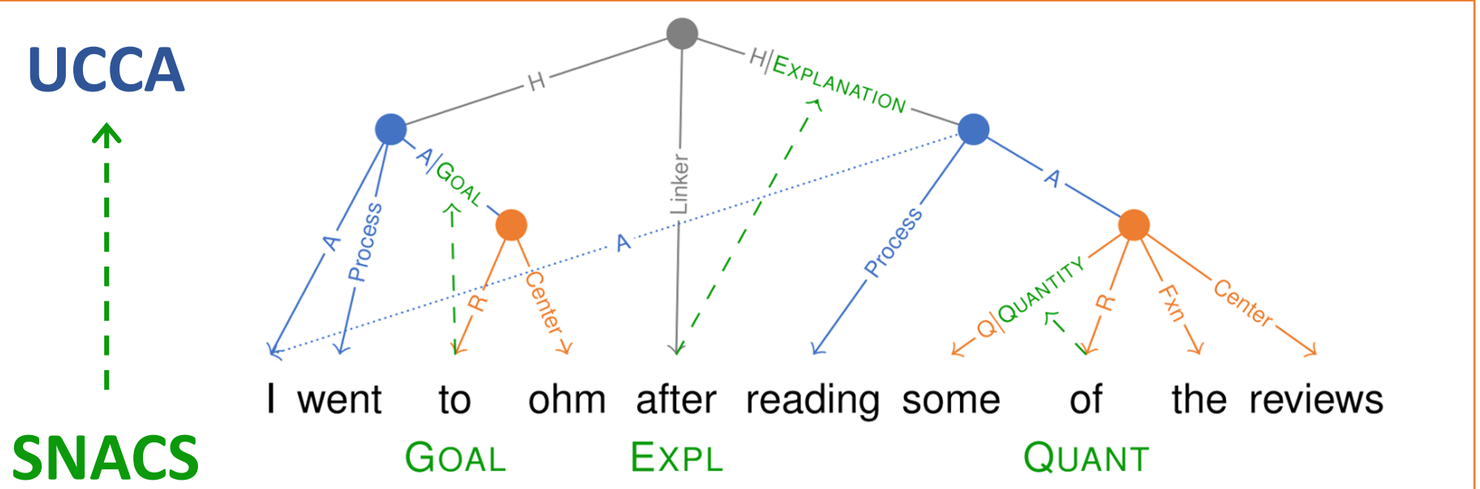
Pipeline: SNACS labels obtained from separate classifier and used as features in the UCCA parser

Independent MTL: A SNACS tagger and an UCCA parser are optimized using Multitask Learning (sharing part of their hidden layers)

Dependent MTL: The SNACS classifier is integrated into the UCCA parser, and the two are trained together, again under the MTL paradigm

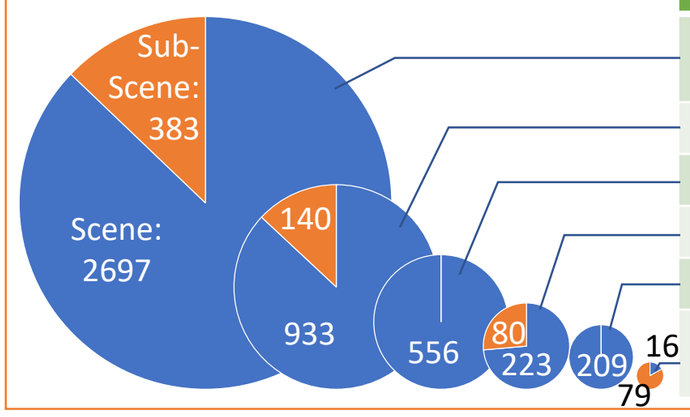
Joint: A single parser with a tag set that consists of concatenated UCCA+SNACS categories

We brought together two MRs…



…using 6 classes of **heuristic rules** based on SNACS, UCCA and UD

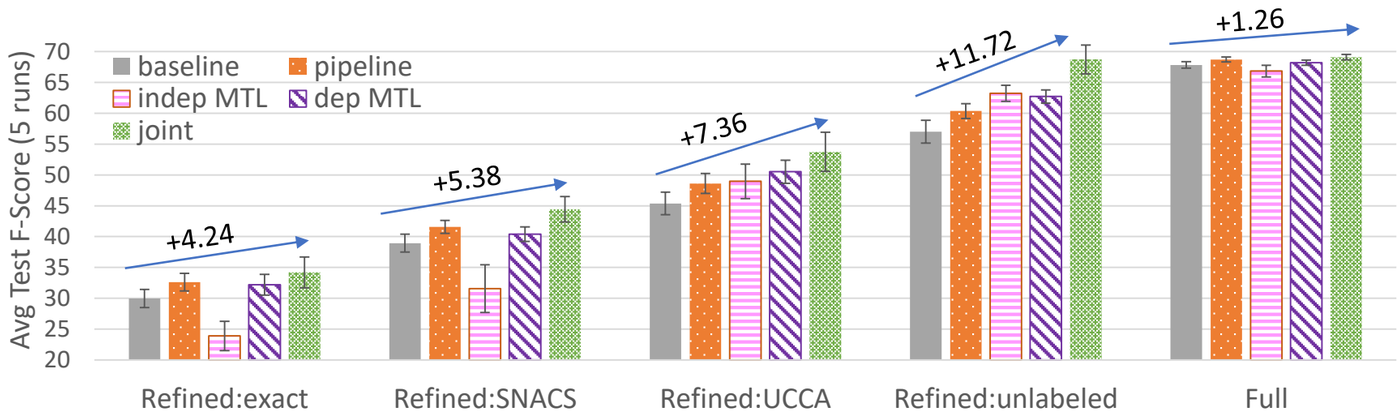
…with **high coverage**: 98% (train) – 99% (dev, test)



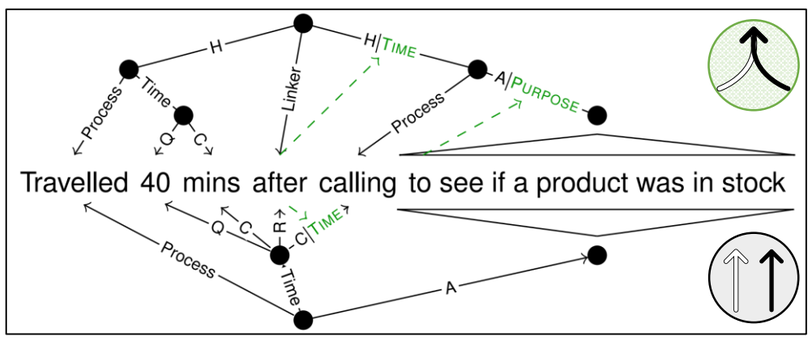
Class	Examples
Relator	I went **to** *ohm* (scene); Cheapest drinks **in** *Keene* (sub-scene)
Possessive	***Her*** speech was excellent; We got ***our*** food
Linkage	Travelled … **after** *calling*
Intransitive	Drive … **down** to Stevens Creek
Predication	They were **back** quickly
Infinitival	Travelled … **to** *see* if a product was in stock; A good place **to** *eat*

…and found that they are made for each other

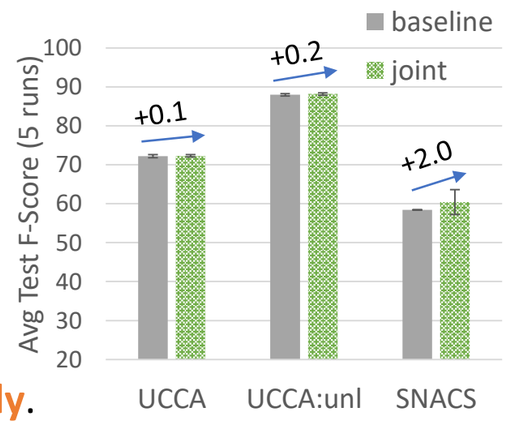
We test this by running combined UCCA and SNACS prediction experiments, varying the **degree and form of interaction** between the two schemes.



Our methods **outperform the baseline** in terms of UCCA F1 on units that are **refined** with SNACS labels.



In particular, the additional lexical information helps recover the (unlabeled) **structure**.



Jointly optimizing for both representations yields the biggest **performance boost**.

We observe slight increases in overall F1-score not only for the **combined** parsing task, but even for the UCCA and SNACS prediction tasks **individually**.

nert

GEORGETOWN UNIVERSITY

THE HEBREW UNIVERSITY OF JERUSALEM

https://arxiv.org/abs/1909.08796