Spotlight on Four-steps SRL decomposition

Alessandro Moschitti[1], Ana-Maria Giuglea[1],
Bonaventura Coppola[2], Roberto Basili[1]

[1]University of Rome "Tor Vergata"
[2]ITC-Irst and University of Trento

CoNLL-2005 Shared Task

Two-steps SRL decomposition

- Many systems exploit a traditional decomposition of the SRL task into Argument Identification and Argument Classification (i+c)

- In the Shared Task 2005, 11 systems out of 19 exploit an i+c strategy (Carreras and Màrquez, CoNLL 2005)

- Underlying assumption: verb arguments (including adjuncts) share common properties which are not dependent on their semantic role.

- Benefit: reduced computational effort in Argument Classification (one versus all)

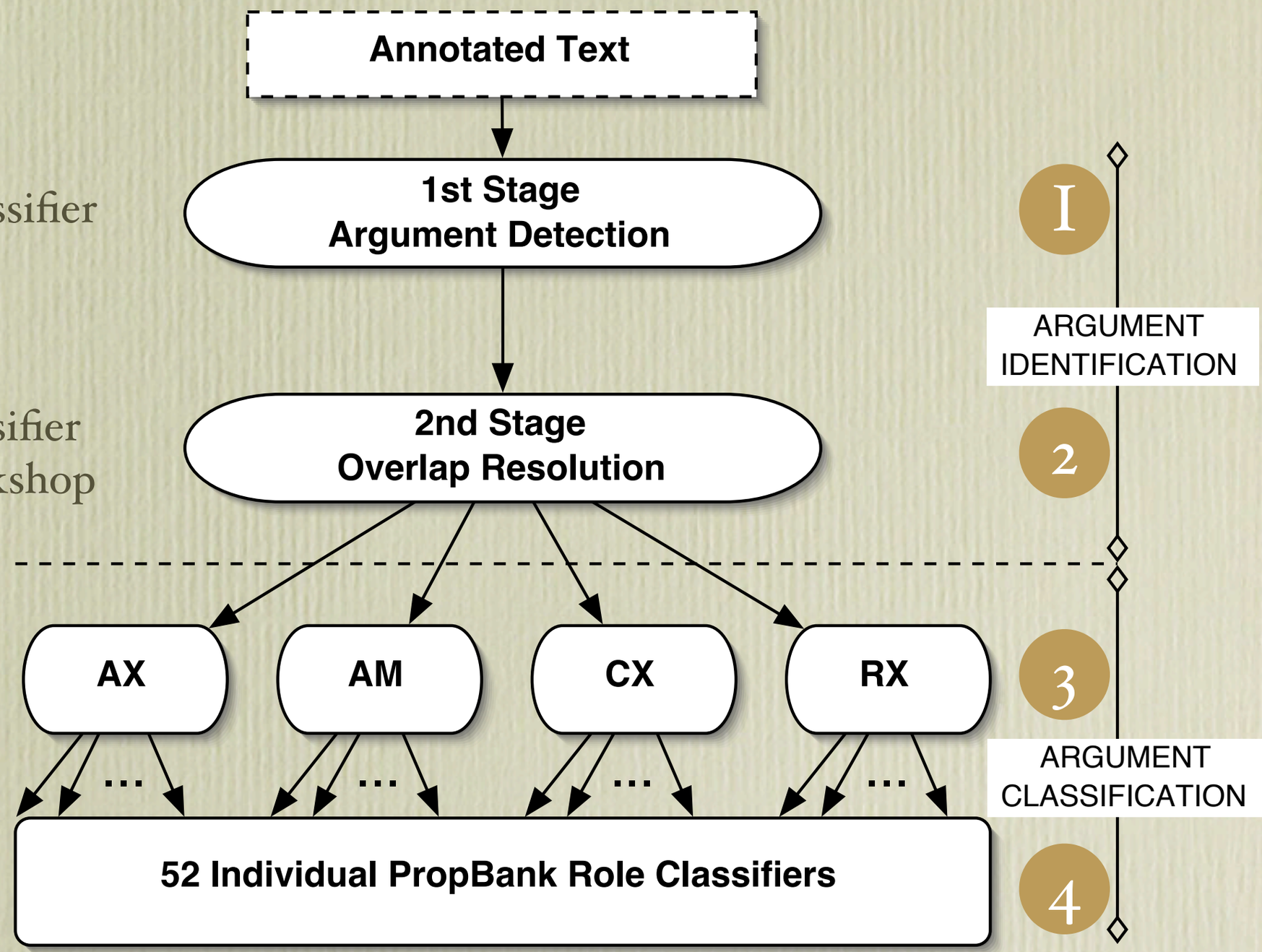
Four-steps SRL decomposition

- Both Identification and Classification are further split

High-recall binary SVM classifier

Tree-kernel-based PAST classifier (Moschitti et al., ACL05 Workshop on Feature Engineering)

Two-stages Hierarchical Argument Classifier



Impact of Hierarchical Classification

- Simple and hierarchical role classifiers compared over gold boundaries. Development set (sec. 24) used for testing. F1 reported.

	AX	AM	CX	RX
Train Instances	172457	59473	2954	7928
Test Instances	5930	2132	105	284
Step 3: binary class.	97.29	97.35	70.86	93.15
Step 3: combination	95.99			
Step 4 multi-c (over gold 3)	92.50	85.88	91.43	91.55
Steps 3+4 comb	88.15			
Basic 2-steps system	88.61			

Discussion of Results

- Introducing an intermediate classification layer (step 3) doesn't noticeably reduce performance
- Hierarchical Classification allows for independent tuning of AXs, AMs, RXs, and CXs classifiers with respect to:
  - Parameter optimization
  - Feature selection
- Benefit: Further reduction of processing time
- Future Work: Possible exploitation of Tree Kernels

Questions for Public Discussion

- Did you find/analyze different behaviors of argument classes with respect to individual features?
- Did you attempt any per-class feature selection?
- Can you report on previous experience in splitting multi-classification problems into several layers?