Active Imitation Learning with Noisy Guidance

Kianté Brantley[1], Amr Sharaf [1], Hal Daumé III[1,2]
[1]*University of Maryland*, [2]*Microsoft Research*

Main ideas

Using a *disagreement classifier* to reduce annotation cost in structured prediction problems

We seek an algorithmic scheme that:

- address the annotation cost problem
- minimizes Type II error (which leads to biased labels)

Our approach:

- only queries for labels when its uncertain
- uses expert and heuristic function annotations
- trains the difference classifier using Apple Tasting
- uses the difference classifier to decide who to query
- reduce expert annotations without affecting performance

Annotation cost problem

In *imitation learning* training proceeds by producing structured outputs one piece at a time and, *at every step*, asking the expert "what would you do here?" and learning to mimic that choice.

Do we have to ask for annotation *at every step*?

Apple Tasting for One-Sided Learning

The *difference classifier* is trained on one-sided feedback (it only observes errors when it predicts "disagree" between expert and heuristic function)

if it predicts "*disagree*" then it gets to see the truth, but if it predicts "*agree*" it never finds out if it was wrong

We use "*STAP*" an Apple Tasting[1] Algorithm that randomly samples from the difference classifier predictions that are predicted "disagree" and changing them to "agree".

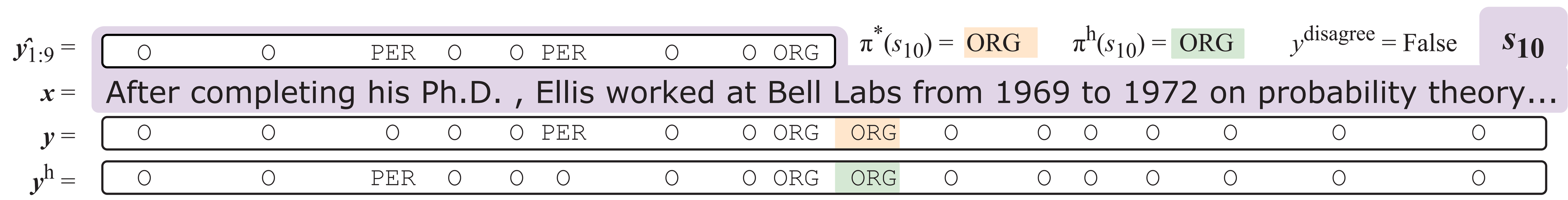
Formally, STAP switches the difference classifier "disagree" predictions with probability:

$$\sqrt{(m+1)/t}$$

(m is the number of mistakes)
(t is the number of "agree" predicted so far)

(Helmbold et al., *Information and Computation* 2000)[1]

Our approach: Named entity recognition example



At state s_{10}

- state representation is x_{10} combined with $\hat{y}_{1:9}$
- expert label at $t = 10$ is y_{10} =ORG
- heuristic label at $t = 10$ is y^h_{10} =ORG

if agent is certain at s_{10}

- we *do not* query for a label

if agent is uncertain at s_{10}

- query *difference classifier* $\hat{d}_i = h_i(S_{10})$
- query AppleTaste-STAP(s_{10}, $\pi^h(s_{10})$, $\hat{d}_i$)
- if *agree* we query the heuristic function
  - Aggregate data to train the agent
- if *disagree* we query the expert
  - Aggregate data to train the agent
  - Aggregate data to train the difference classifier

Experiments Details

Task	Named Entity Recognition	Keyphrase Extraction	Part of Speech Tagging
Dataset	CoNLL'03	SemEval 2017 Task 10	Universal Dependencies
State Features	English BERT	SciBERT	M-BERT
Heuristic	offline gazetteer	unsupervised keyphrase model	Dictionary from Wiktionary
Heur Quality	P 88%, R 27%, F 41%	P 20%, R 44%, F 27%	10% coverage, 67% acc

Experiments Results

The **top row** shows performance *(f-score or accuracy)* with respect to the number of queries to the expert.
The **bottom row** shows the *number of queries* as a function of the number of words seen.