A Coactive Learning View of Online Structured Prediction in SMT

Artem Sokolov[*], **Stefan Riezler**[*], **Shay B. Cohen**[‡]
[*]Heidelberg University, [‡]University of Edinburgh

Online learning protocol

1. observe input structure x_t
2. predict output structure y_t
3. receive feedback (gold-standard or post-edit)
4. update parameters

A tool of choice in SMT

- memory & runtime efficiency
- interactive scenarios with user feedback

Usual assumptions

- convexity (for regret bounds)
- reachable feedbacks (for gradients)

Reality

- SMT has latent variables (non-convex)
- most references live outside the search space (nonreachable)
- references/full-edits are expensive (= professional translation)

Intuition

- light post-edits are cheaper
- have better chance to be reachable

Question

Should editors put much effort into correcting SMT outputs anyway?

Goals

- demonstrate feasibility of learning from weak feedback for SMT
- propose a new perspective on learning from surrogate translations
- **note: the goal is not to improve over any full-information model**

Contributions

- Theory
  - extension of the coactive learning model to latent structure
  - improvements by a derivation-dependent update scaling
  - straight-forward generalization bounds
- Practice
  - learning from weak post-edits does translate to improved MT quality
  - surrogate references work better if they admit an underlying linear model

[Shivaswami & Joachims, ICML'12]

- rational user: feedback $\bar{y}_t$ improves some utility over prediction y_t

$$U(x_t, \bar{y}_t) \geq U(x_t, y_t)$$

- regret: how much the learner is 'sorry' for not using optimal y_t^*

$$\mathrm{REG}_T = \frac{1}{T}\sum_{t=1}^{T} U(x_t, y_t^*) - U(x_t, y_t) \qquad \rightarrow \min$$

- feedback is α-informative if

$$U(x_t, \bar{y}_t) - U(x_t, y_t) \geq \alpha(U(x_t, y_t^*) - U(x_t, y_t))$$

- no latent variables

Algorithm

Feedback-based Structured Perceptron

Initialize $w \leftarrow 0$
for $t = 1, \ldots, T$ **do**
    Observe x_t
    $y_t \leftarrow \arg\max_y w_t^\top \phi(x_t, y)$
    Obtain weak feedback $\bar{y}_t$
    if $y_t \neq \bar{y}_t$ **then**
        $w_{t+1} \leftarrow w_t + \big(\phi(x_t, \bar{y}_t) - \phi(x_t, y_t)\big)$

Algorithm

Feedback-based Latent Structured Perceptron

```
Initialize w ← 0
for t = 1, ..., T do
    Observe x_t
    (y_t, h_t) ← argmax_(y,h) w_t^⊤ φ(x_t, y, h_t)
    Obtain weak feedback ȳ_t
    if y_t ≠ ȳ_t then
        h̄_t ← argmax_h w_t^⊤ φ(x_t, ȳ_t, h)
        w_{t+1} ← w_t + Δ_{h̄_t,h_t} (φ(x_t, ȳ_t, h̄_t) − φ(x_t, y_t, h_t))
```

Initialize $w \leftarrow 0$

for $t = 1, \ldots, T$ **do**

  Observe x_t

  $(y_t, h_t) \leftarrow \arg\max_{(y,h)} w_t^\top \phi(x_t, y, h_t)$

  Obtain weak feedback $\bar{y}_t$

  if $y_t \neq \bar{y}_t$ **then**

    $\bar{h}_t \leftarrow \arg\max_h w_t^\top \phi(x_t, \bar{y}_t, h)$

    $w_{t+1} \leftarrow w_t + \Delta_{\bar{h}_t, h_t} \big(\phi(x_t, \bar{y}_t, \bar{h}_t) - \phi(x_t, y_t, h_t)\big)$

Under the same assumptions as in [Shivaswami & Joachims'12]:

- linear utility: $U(x_t, y_t) = {w_*}^\top \phi(x_t, y_t)$
- w_* is the optimal parameter, known only to the user
- $||\phi(x_t, y_t, h_t)|| \leq R$
- some violations of α-informativeness are allowed

$$U(x_t, \bar{y}_t) - U(x_t, y_t) \geq \alpha(U(x_t, y_t^*) - U(x_t, y_t)) - \xi_t$$

Convergence

Let $D_T = \sum_t^T \Delta^2_{\bar{h}_t, h_t}$. Then

$$\mathrm{REG}_T \leq \frac{1}{\alpha T} \sum_{t=1}^{T} \xi_t + \frac{2R\,||w_*||}{\alpha} \frac{\sqrt{D_T}}{T}$$

- standard perceptron proof [Novikoff'62]
- better than $\mathcal{O}(1/\sqrt{T})$ if D_T doesn't grow too fast
- [Shivaswami & Joachims'12] is a special case of $\Delta_{\bar{h}_t, h_t} = 1$

Generalization

Let $0 < \delta < 1$, and let $x_1, \dots, x_T$ be a sequence of observed inputs. Then with probability at least $1 - \delta$,

$$\mathbb{E}_{x_1,..,x_T}[\mathrm{REG}_T] \leq \mathrm{REG}_T + 2||w_*||R\sqrt{\frac{2}{T}\ln\frac{1}{\delta}}.$$

- how far the expected regret is from the empirical regret we observe
- proof uses the results of [Cesa-Bianchi'04]
- see the paper for more

- LIG corpus [Potet et al.'10]
  - ➡ news domain, FR→EN
  - ➡ (**FR input**, MT output, **EN post-edit**, EN reference), 11k in total
  - ➡ split
    - train 7k — online input data
    - dev 2k — to get w_* for simulation/checking convergence
    - test 2k — testing
- Moses, 1000-best lists
- cyclic order

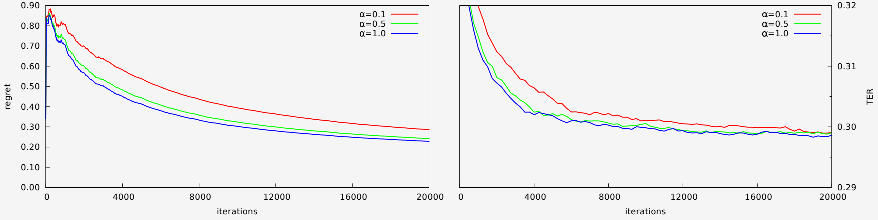
User simulation:

- scan the n-best list for derivations that are α-informative
- return the first $\bar{y}_t \neq y_t$ that satisfies

$$U(x_t, \bar{y}_t) - U(x_t, y_t) \geq \alpha(U(x_t, y_t^*) - U(x_t, y_t)) - \xi_t$$

(with minimal ξ_t, if no $\xi_t = 0$ found for a given α)

Regret and TER for α-informative feedback



- convergence in regret when learning from weak feedback of differing strength
- simultaneous improvement TER (on test)
- stronger feedback leads to faster improvements of regret/TER
- setting $\Delta_{\bar{h}_t,h_t}$ to Euclidean distance between feature vectors leads to even faster regret/TER improvements

Feedback from Surrogate Translations

- so far the feedback was simulated
- what about real post-edits?
- main question: how do the practices for extracting surrogates from user post-edits for discriminative SMT match with the coactive learning?

1. `oracle` – closest to the post-edit in the full search graph

$$\bar{y} = \operatorname*{arg\,min}_{y' \in \mathcal{Y}(x_t; w_t)} \mathsf{TER}(y', y)$$

2. `local` – closest to the post-edit from the n-best list [Liang et al.'06]

$$\bar{y} = \operatorname*{arg\,min}_{y' \in n\text{-best}(x_t; w_t)} \mathsf{TER}(y', y)$$

3. `filtered` – first hyp in the n-best list w/ better TER than the 1-best

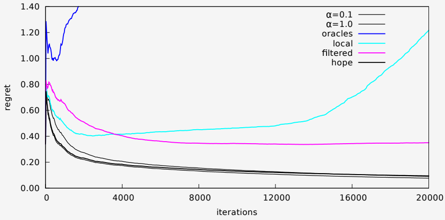
$$\mathsf{TER}(\bar{y}, y) < \mathsf{TER}(y_t, y)$$

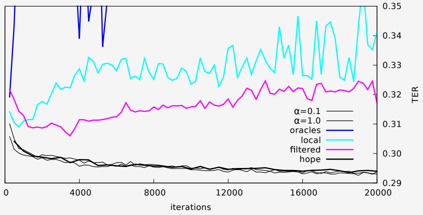
4. `hope` – hyp that maximizes model score and negative TER [Chiang'12]

$$\bar{y} = \operatorname*{arg\,max}_{y' \in n\text{-best}(x_t; w_t)} (-\mathsf{TER}(y', y) + w_t^\top \phi(x_t, y', h))$$

Degrees of model-awareness

- `oracle` – model-agnostic
- `local` – constrained to the n-best list, but ignores the ordering
- `filtered` & `hope` – letting the model score/ordering influence the surrogate





- regret diverges when learning with model-unaware surrogates
- convergence in regret when learning with model-aware surrogates

	% strictly α-informative
local	39.46%
filtered	47.73%
hope	**83.30%**

Conclusions

- regret & generalization bounds
  - ➡ latent variables
  - ➡ changing feedback
- concept of weak feedback in online learning in SMT
  - ➡ still can learn without observing references
  - ➡ surrogate references should admit an underlying linear model

- regret & generalization bounds
  - ➡ latent variables
  - ➡ changing feedback
- concept of weak feedback in online learning in SMT
  - ➡ still can learn without observing references
  - ➡ surrogate references should admit an underlying linear model

Thank you!