Don't Until the Final Verb Wait: Reinforcement Learning for Simultaneous Machine Translation

Alvin Grissom II, He He, Jordan Boyd-Graber, John Morgan, and Hal Daumé III

http://www.umiacs.umd.edu/~alvin/simtrans

University of Colorado, Boulder
University of Maryland, College Park

Real-time Interpretation

Nuremberg Trials

Videoconference

Outline

- SOV-SVO Simultaneous MT
  - Why this is difficult
  - Verb Prediction
- Reinforcement Learning
- Translation System
- Experiments and Results
- Related Work
- Future Work

Simultaneous Translation

- Begin translating before a sentence ends.
- First introduced on a large scale with Nuremberg Trials.
- Requires judgments about when to translate fragments.
- Skill learned from experience.

Simultaneous Machine Translation

- Most prior approaches have been rule-based.
- We would like to use machine learning.
- Difficult for humans, who must learn from experience.

SOV-SVO Simultaneous Translation

- Many languages (e.g., German, Japanese) are verb-final (SOV); other (e.g. English) are SVO.
- Translator must wait for verb, or predict.

ich	bin	mit	dem	Zug	nach	Ulm	**gefahren**
I	am	with	the	train	to	Ulm	**traveled**
I	(. . . . . . *waiting*. . . . . .)						**traveled** by train to Ulm

Verb Prediction

- Predict the final verb to produce a grammatical sentence.
- We use language models to predict the main verb and next word in the sentence.

Verb Prediction

Apple ist zum wertvollsten Konzern aller Zeiten avanciert

Nein, mit dem Virus ist es noch lange nicht getan

Eine vielbefahrene Brücke in New Jersey wurde grundlos gesperrt

Mit Drohen und Interpretieren ist es nicht getan

Frankfurter Flughafen für Passagiere weitgehend gesperrt

Verb Prediction

Apple ist zum wertvollsten Konzern aller Zeiten <u>avanciert</u>

Nein, mit dem Virus ist es noch lange nicht <u>getan</u>

Eine vielbefahrene Brücke in New Jersey wurde grundlos <u>gesperrt</u>

Mit Drohen und Interpretieren ist es nicht <u>getan</u>

Frankfurter Flughafen für Passagiere weitgehend <u>gesperrt</u>

Verb Prediction

- Build a language model for each verb.
- For any input text, x, we make a verb prediction:

$$\arg\max_{v} p(v) \prod_{i=1}^{t} p(x_i \mid v, x_{i-n+1:i-1})$$

Verb Prediction

- Most predictions will be incorrect.
  - Leads to terrible translations.

$$\arg\max_{v} p(v) \prod_{i=1}^{t} p(x_i \mid v, x_{i-n+1:i-1})$$

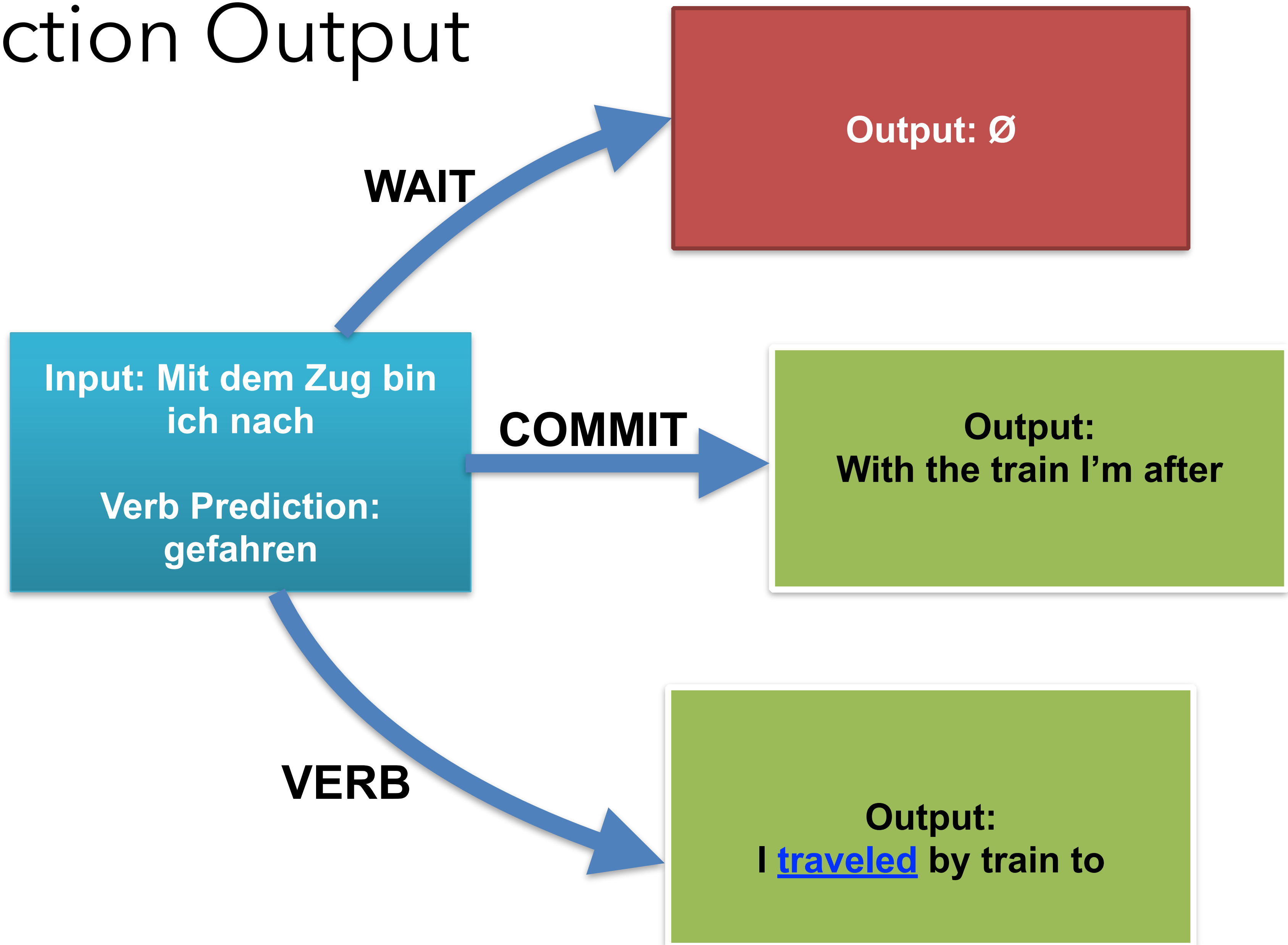
Learn When to Trust Predictions

- *Learn* under which circumstances to trust predictions.
- Translate when confident and wait for more information otherwise.
- Learn a **policy**, π, to do this.

Reinforcement Learning for Simultaneous Machine Translation

- State:
  - Observations (words), predictions (next word and verb) and prediction scores.
- Actions:
  - **WAIT** for more words.
    - Input: Word. Output: None.
  - Translate with predicted **VERB.**
    - Input: Word, verb prediction: Output: translated segment with verb.
    - We can also predict the **next word**.
  - **COMMIT** to partial translation.
    - Input: word. Output: translated segment.

Action Output



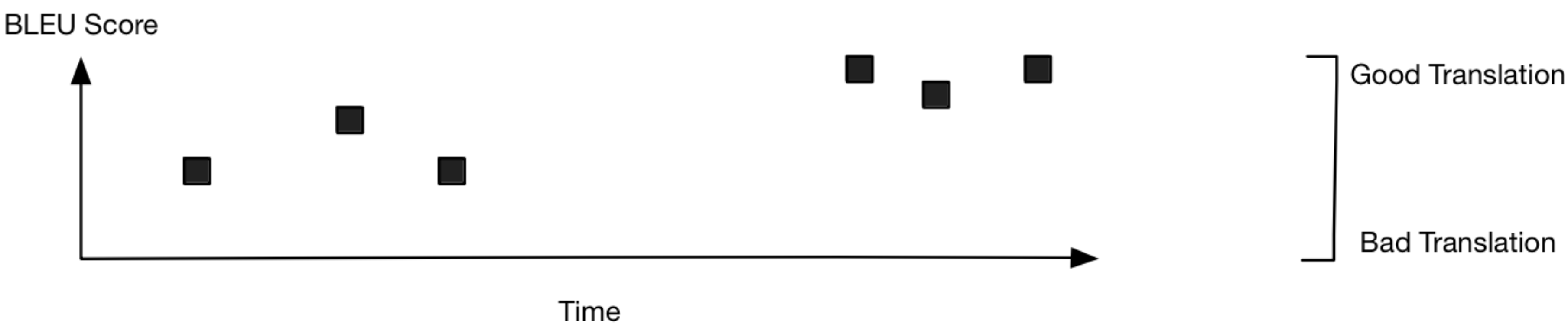
Reward

- We want to capture translation *quality* and translation *latency*.
  - Optimal translations are both *accurate* and *prompt*.
- Incrementalize (sentence-level) BLEU.
  - Score partial translations; sum their scores.

Reward from Many Translations

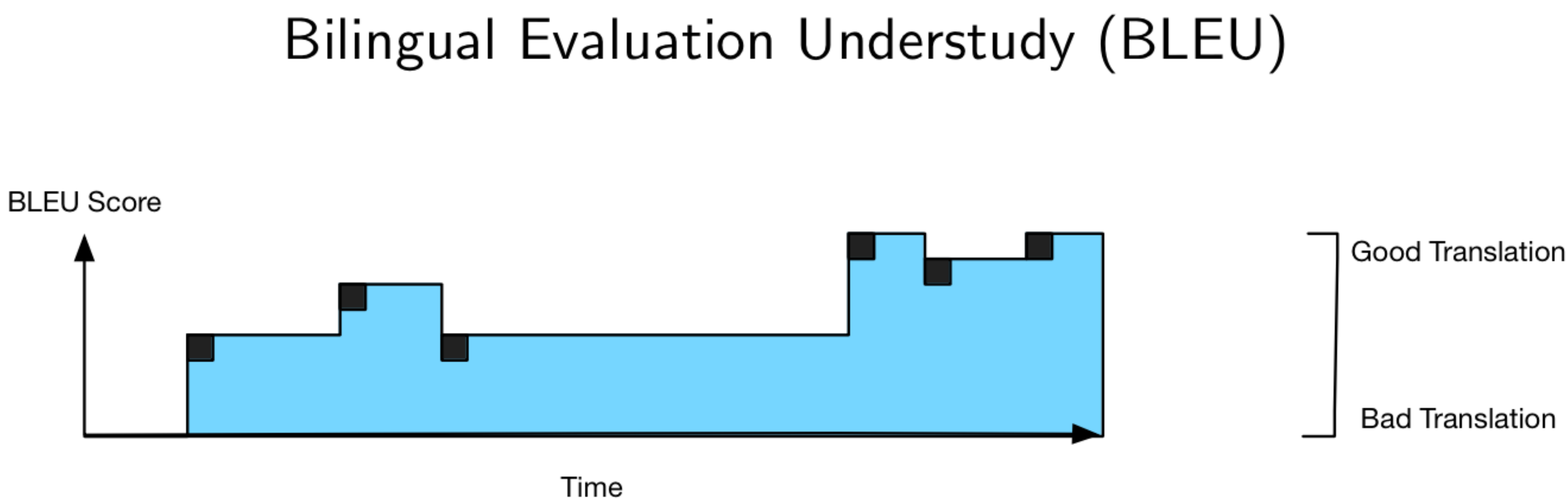
- Calculate BLEU score each time an action is executed.

Bilingual Evaluation Understudy (BLEU)



Reward from Many Translations

- Sum weighted translations over course of sentence.



Policy Comparison

Source Sentence

Er

Psychic

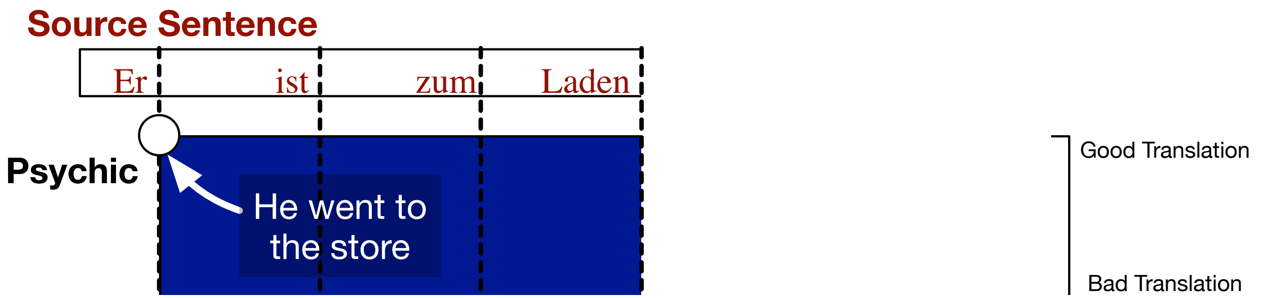
Good Translation

Bad Translation

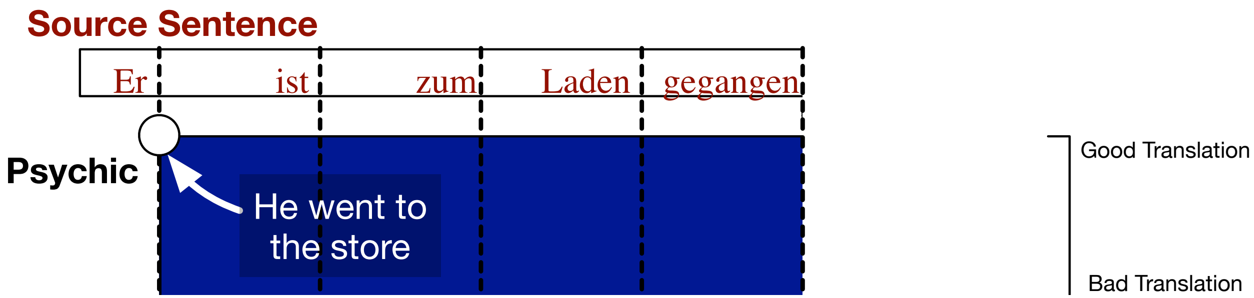
Policy Comparison

Policy Comparison

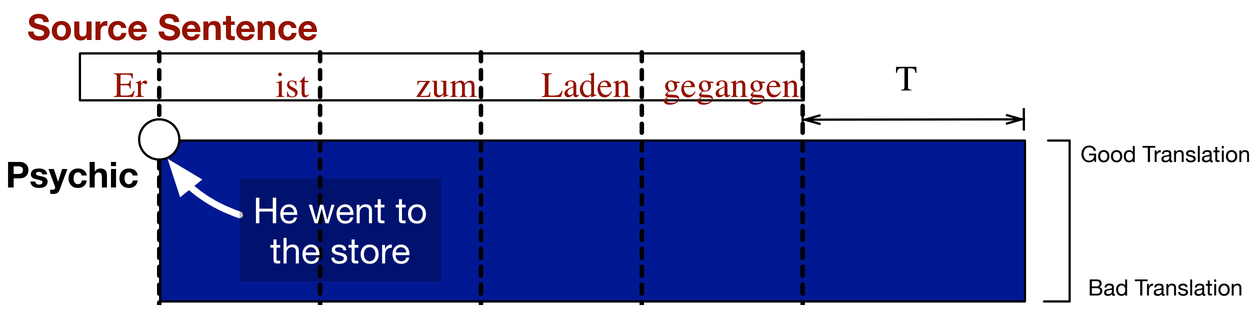
Policy Comparison



Policy Comparison

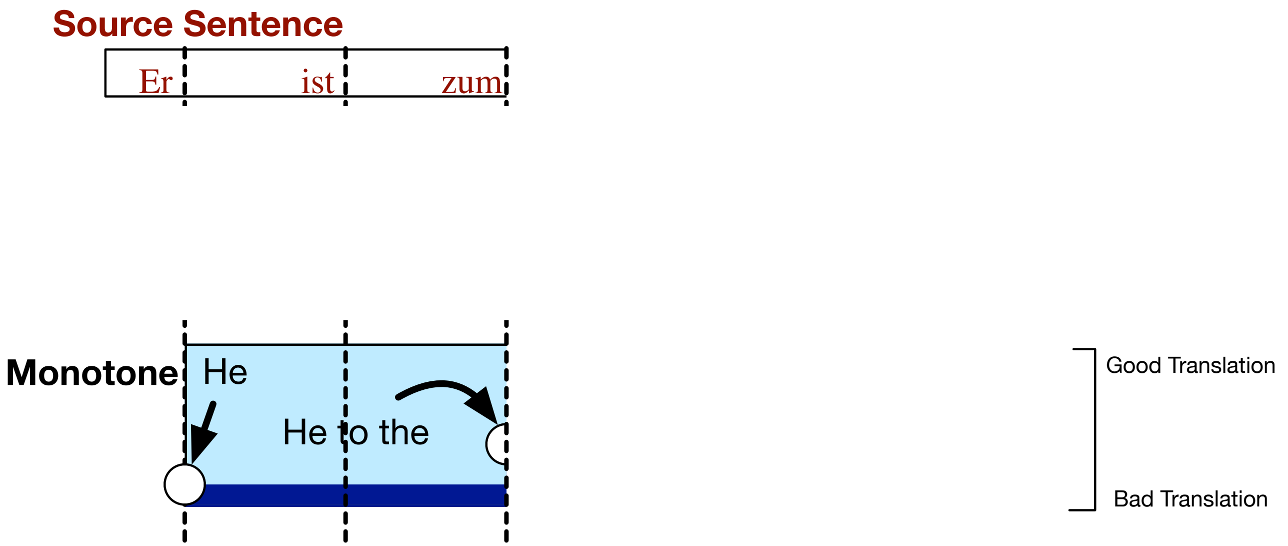


Policy Comparison

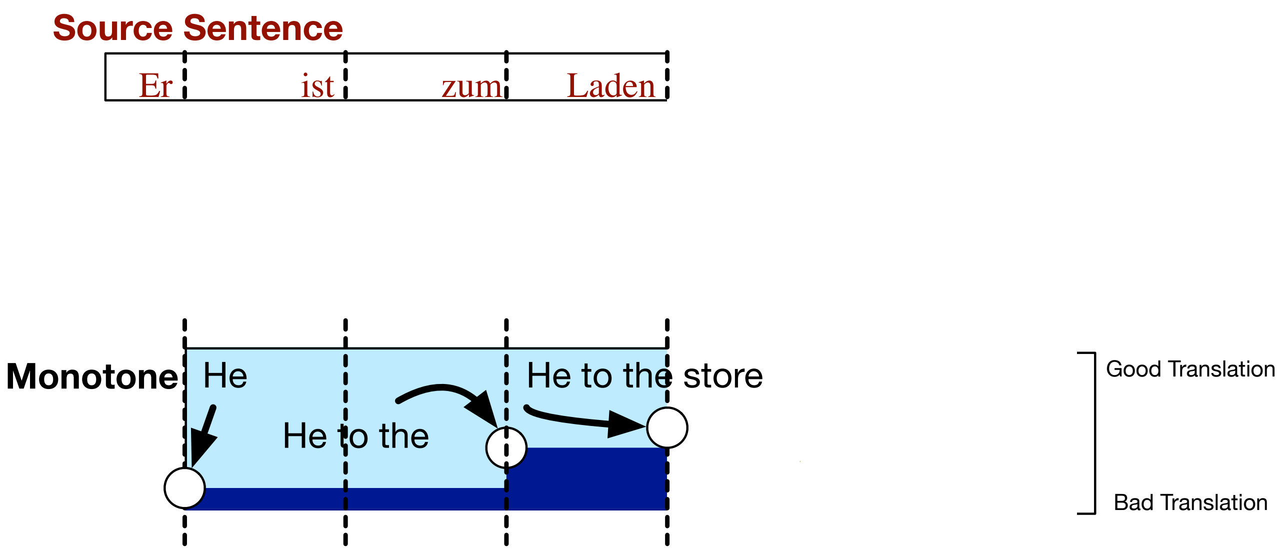


Policy Comparison

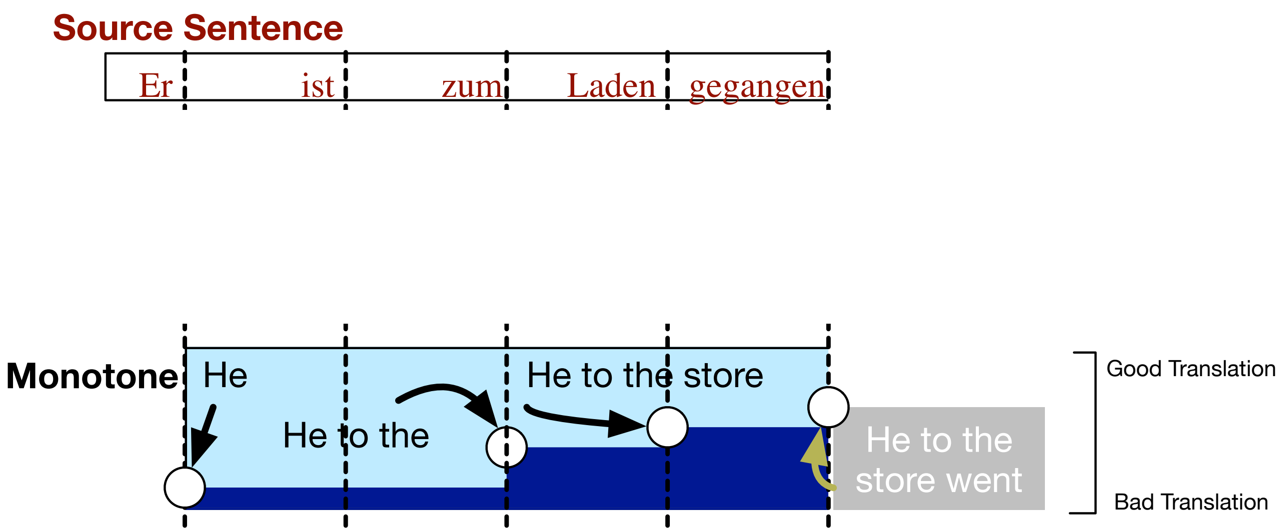
Policy Comparison



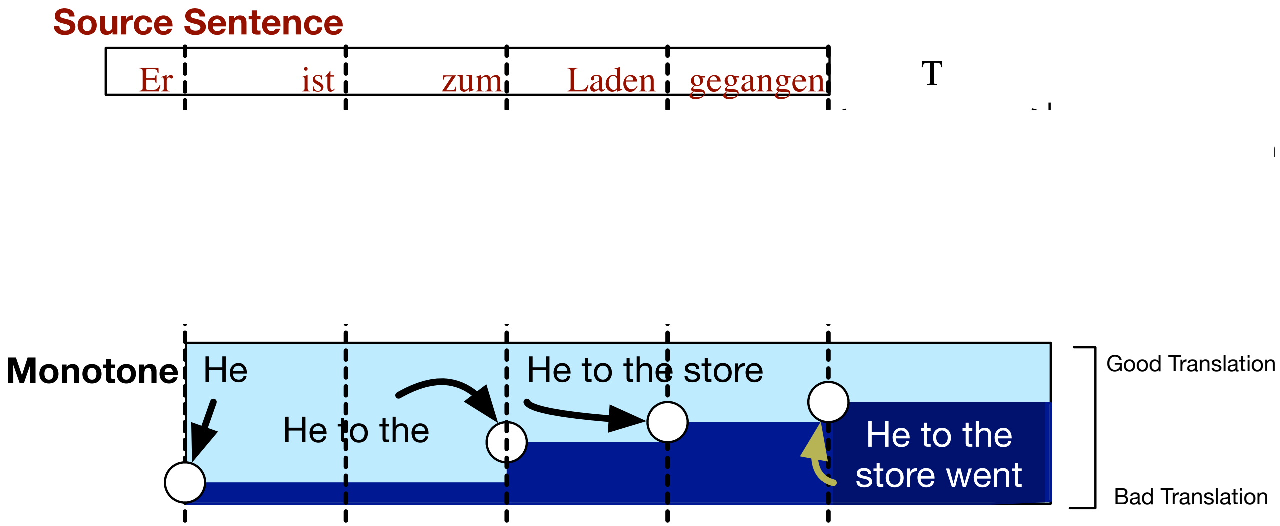
Policy Comparison



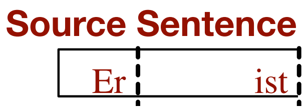
Policy Comparison

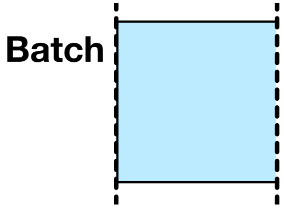


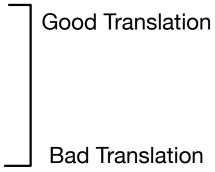
Policy Comparison



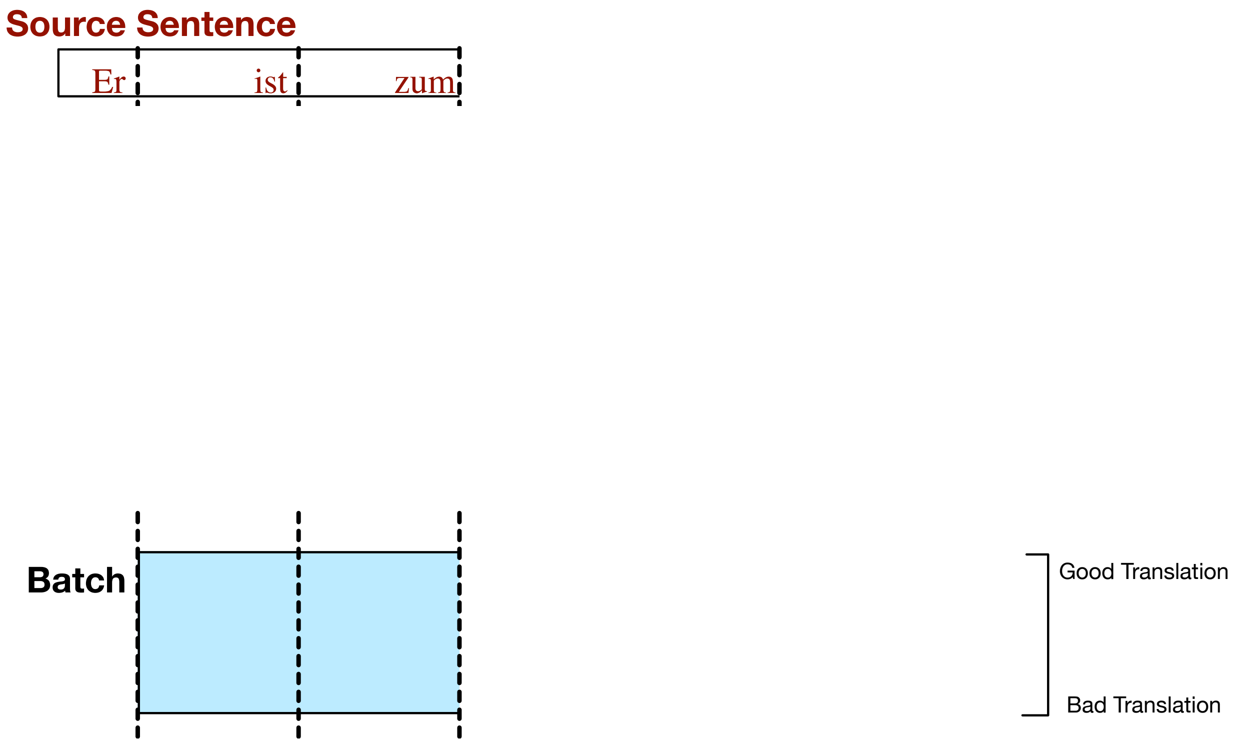
Policy Comparison



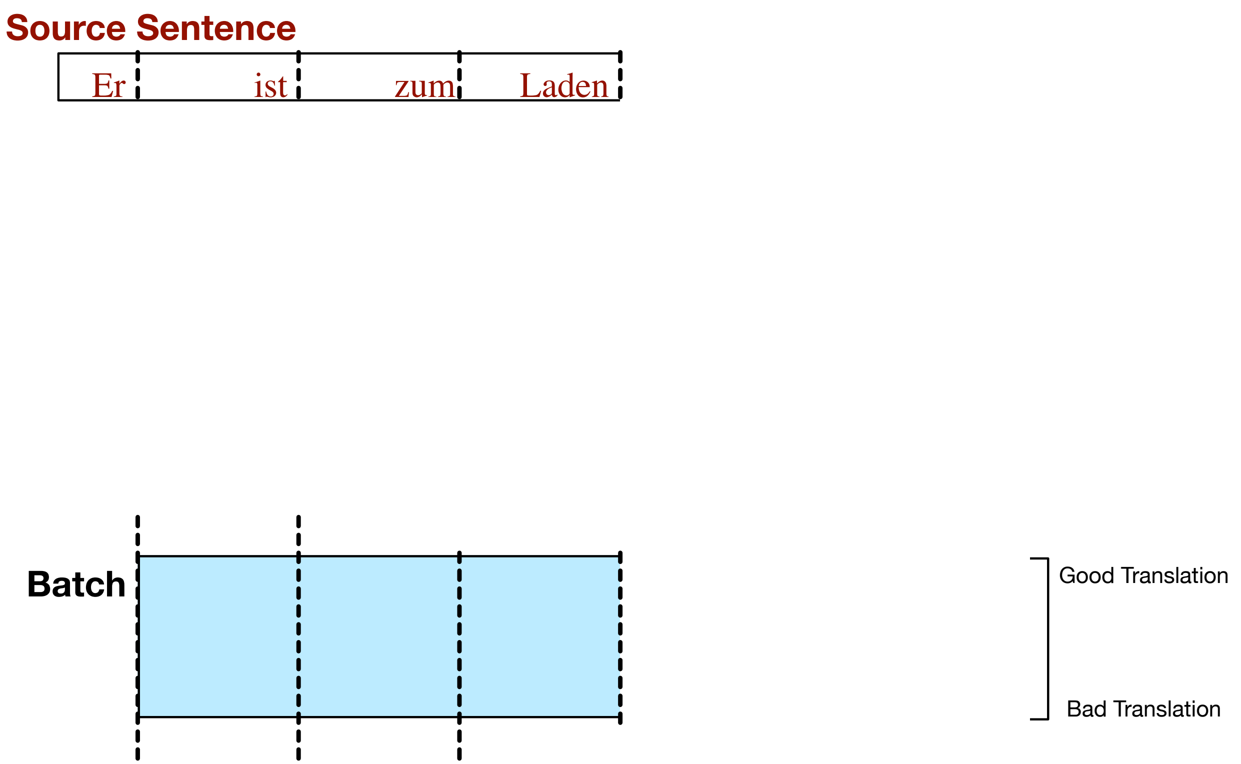




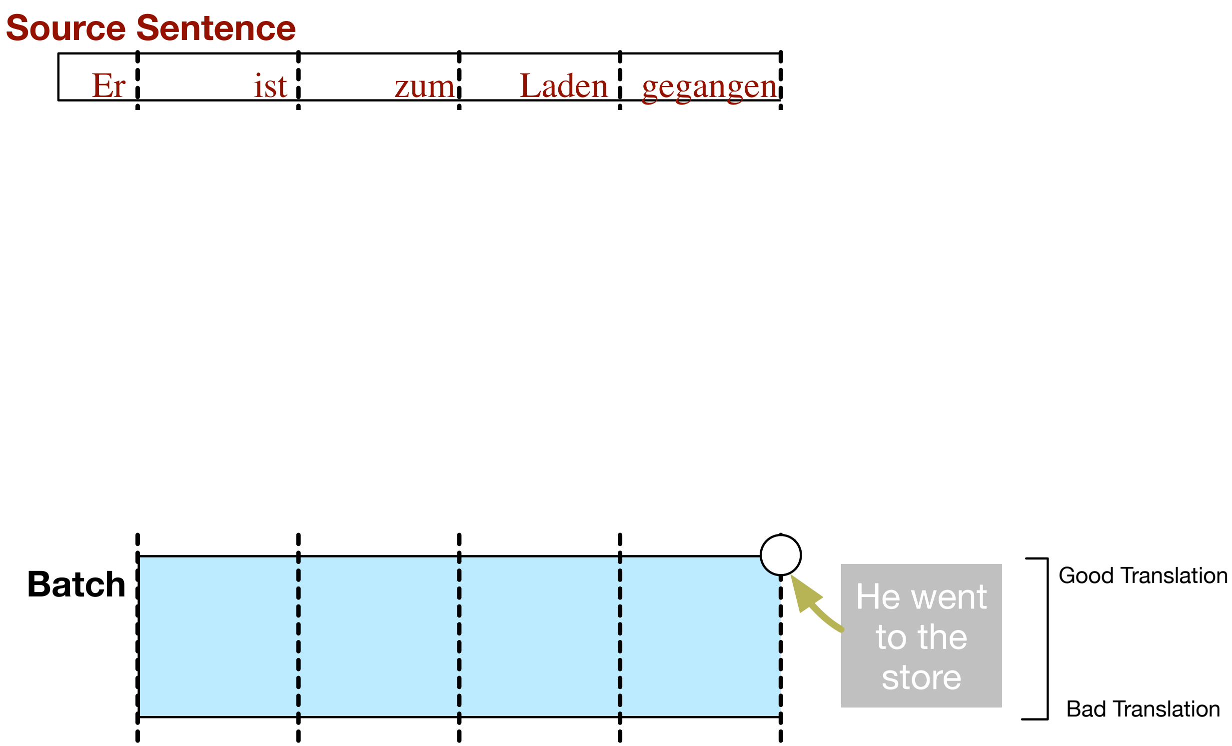
Policy Comparison



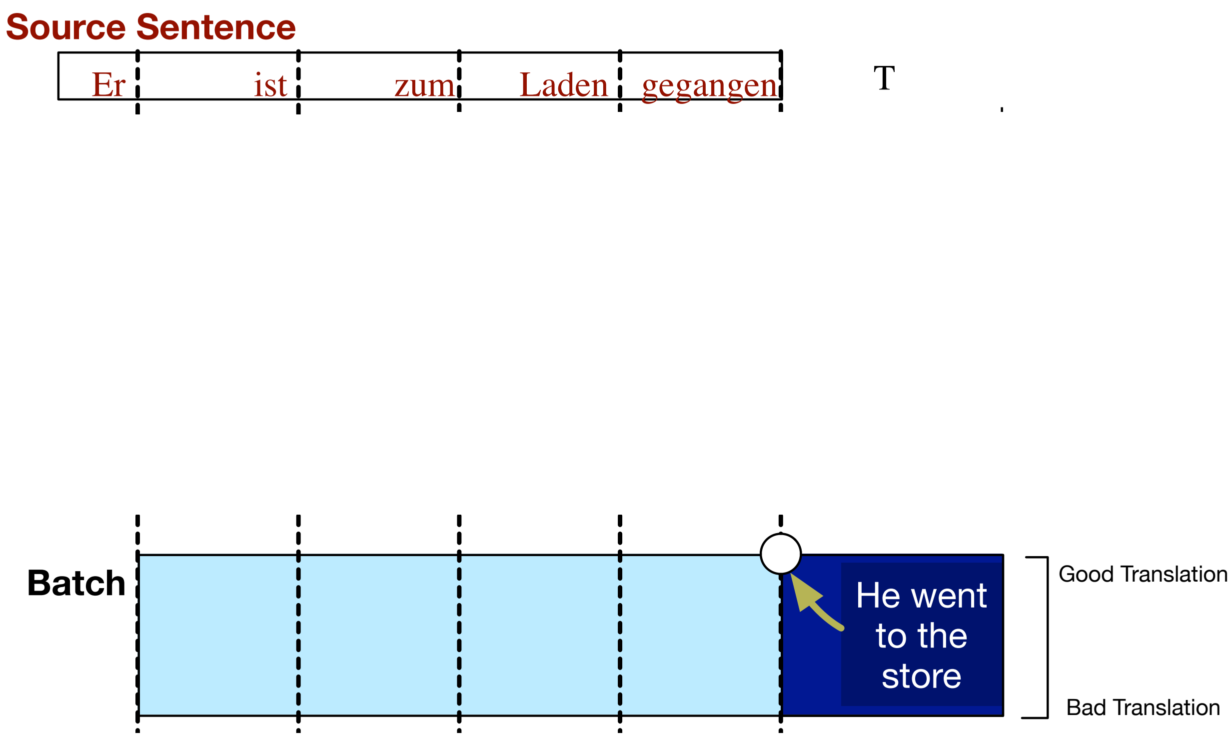
Policy Comparison



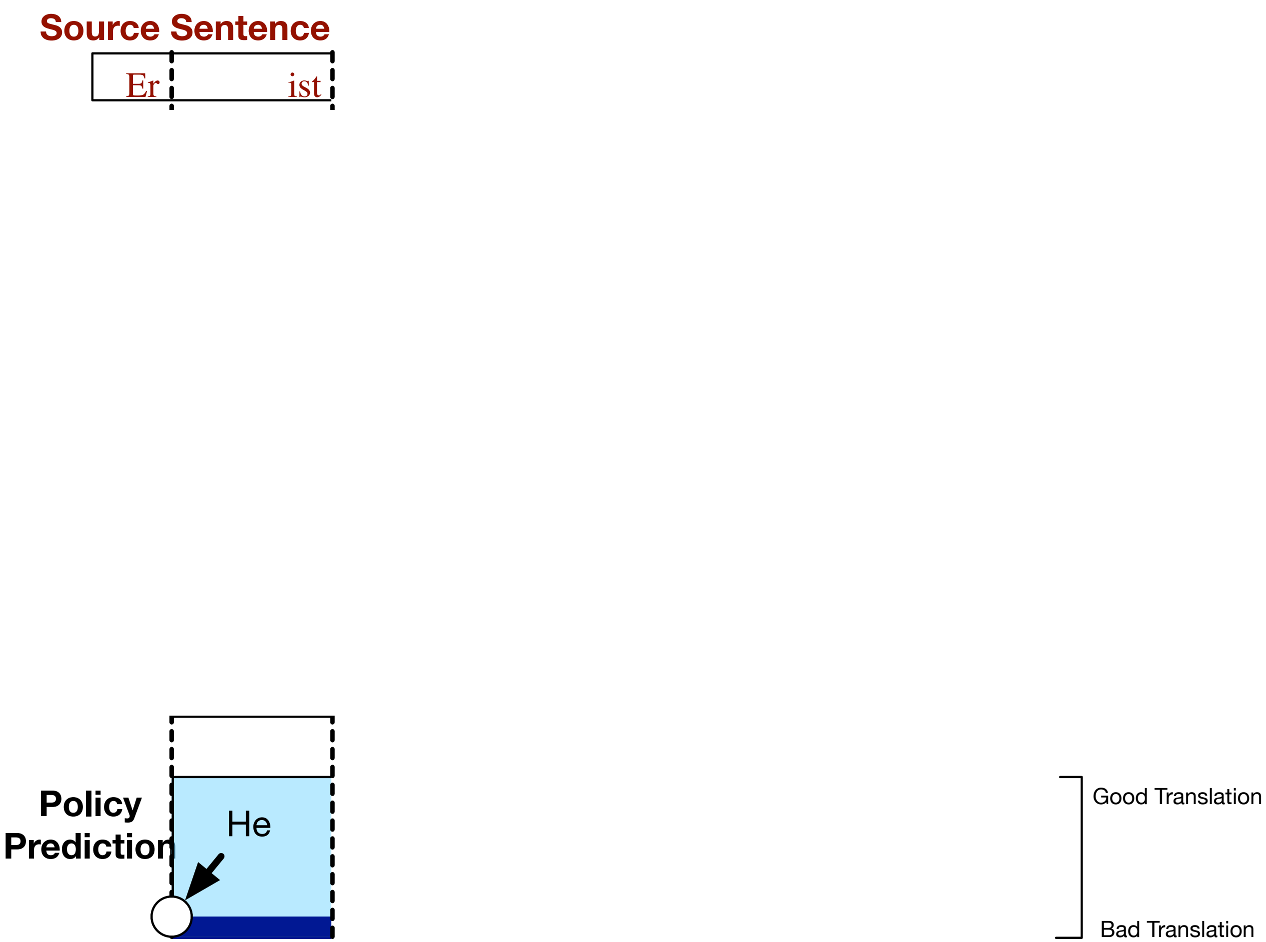
Policy Comparison



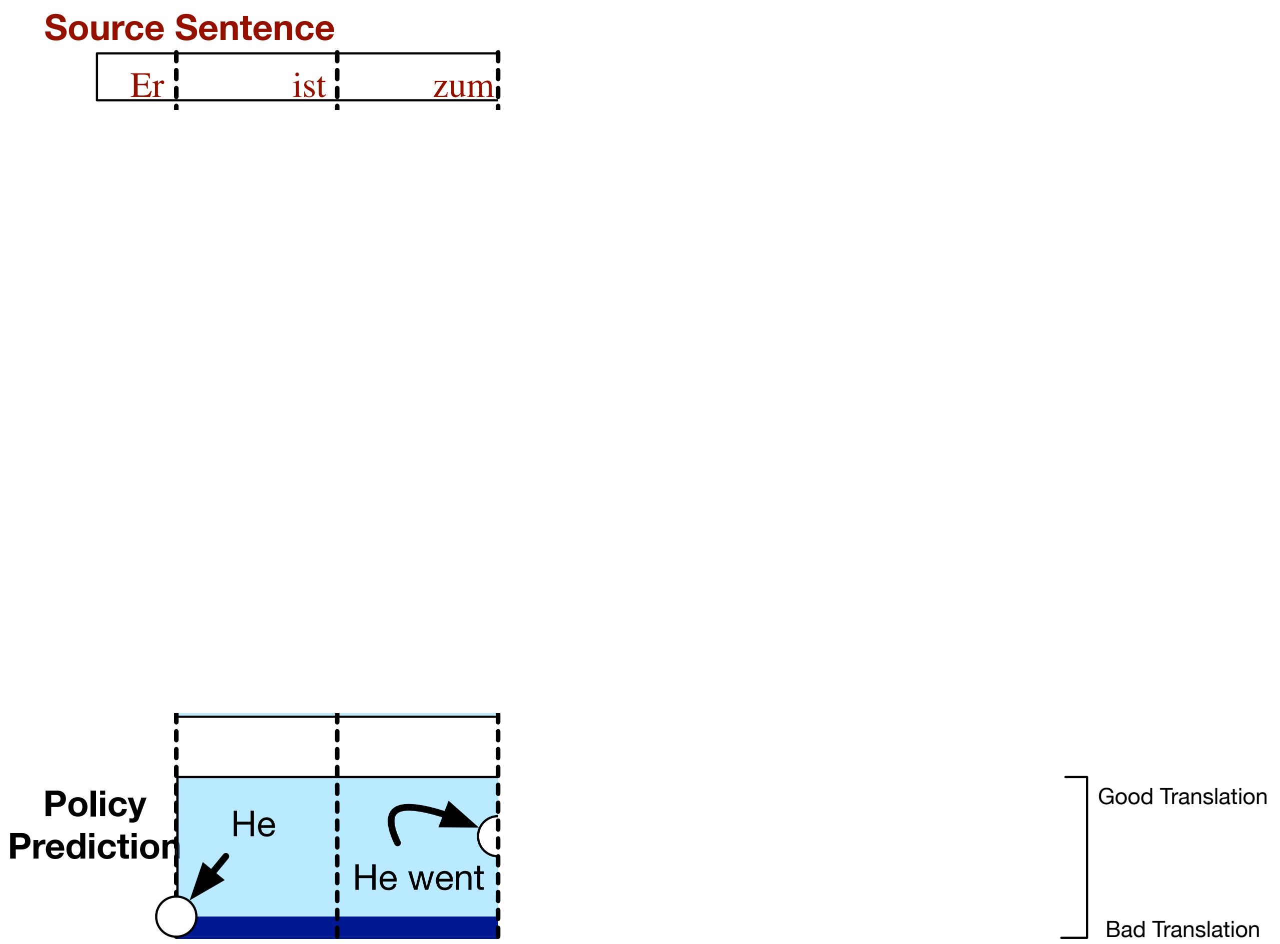
Policy Comparison



Policy Comparison

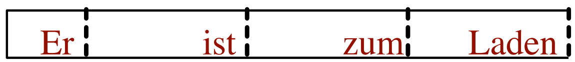


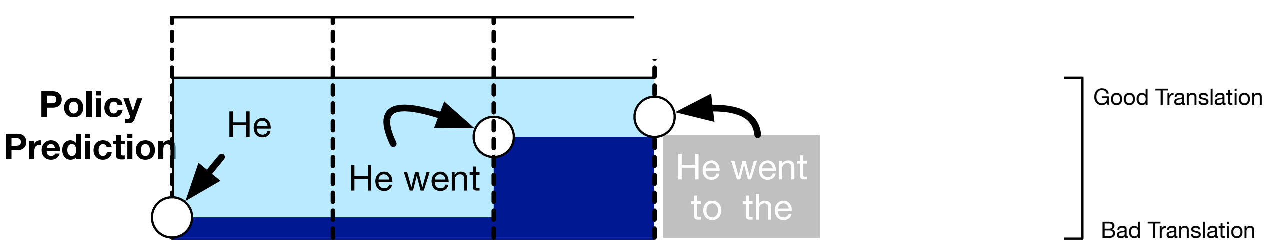
Policy Comparison



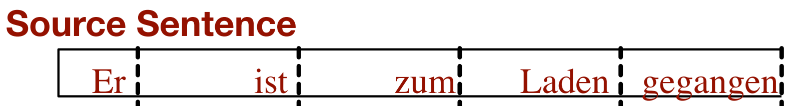
Policy Comparison

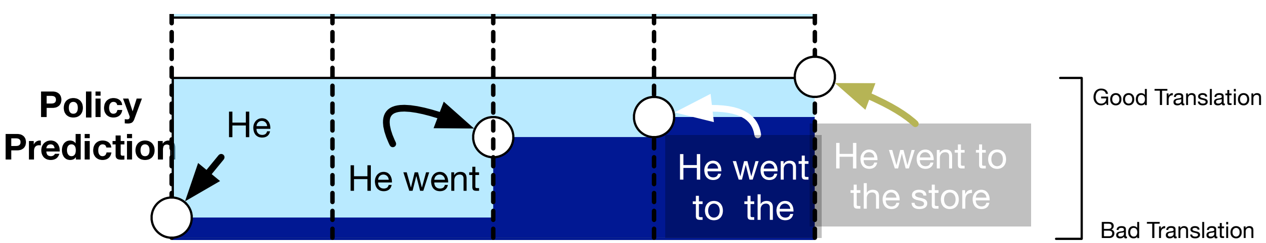
Source Sentence



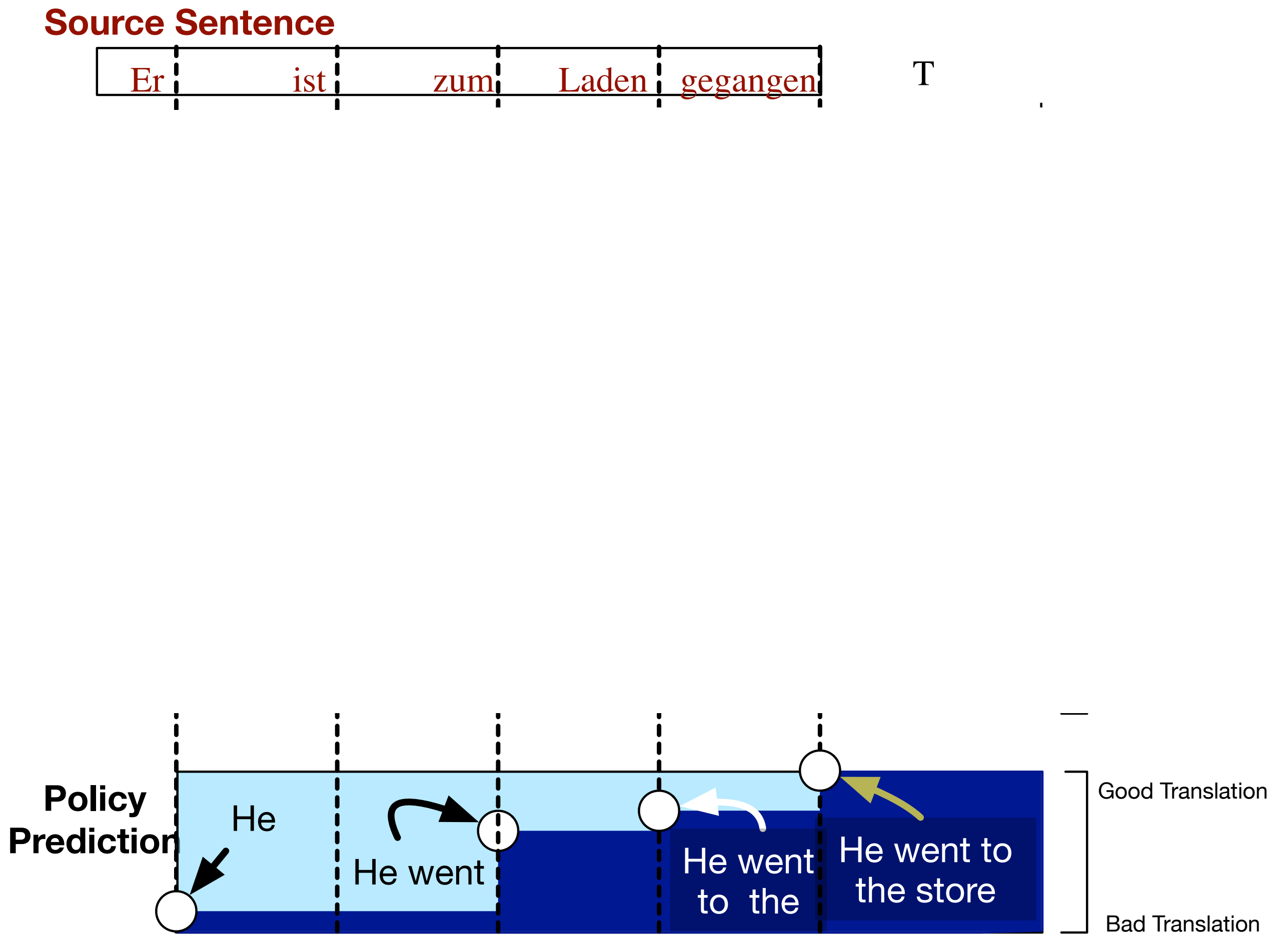


Policy Comparison

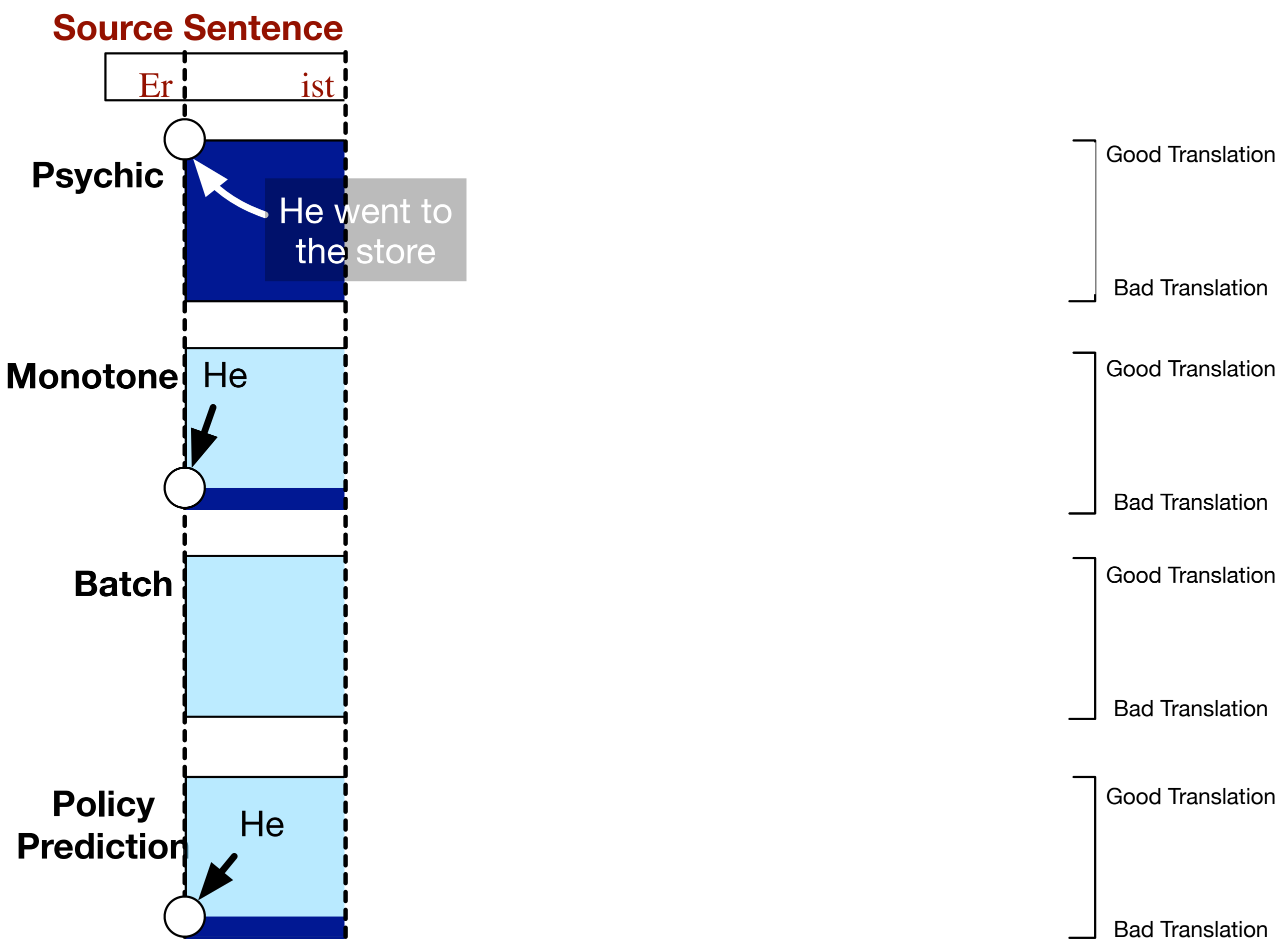




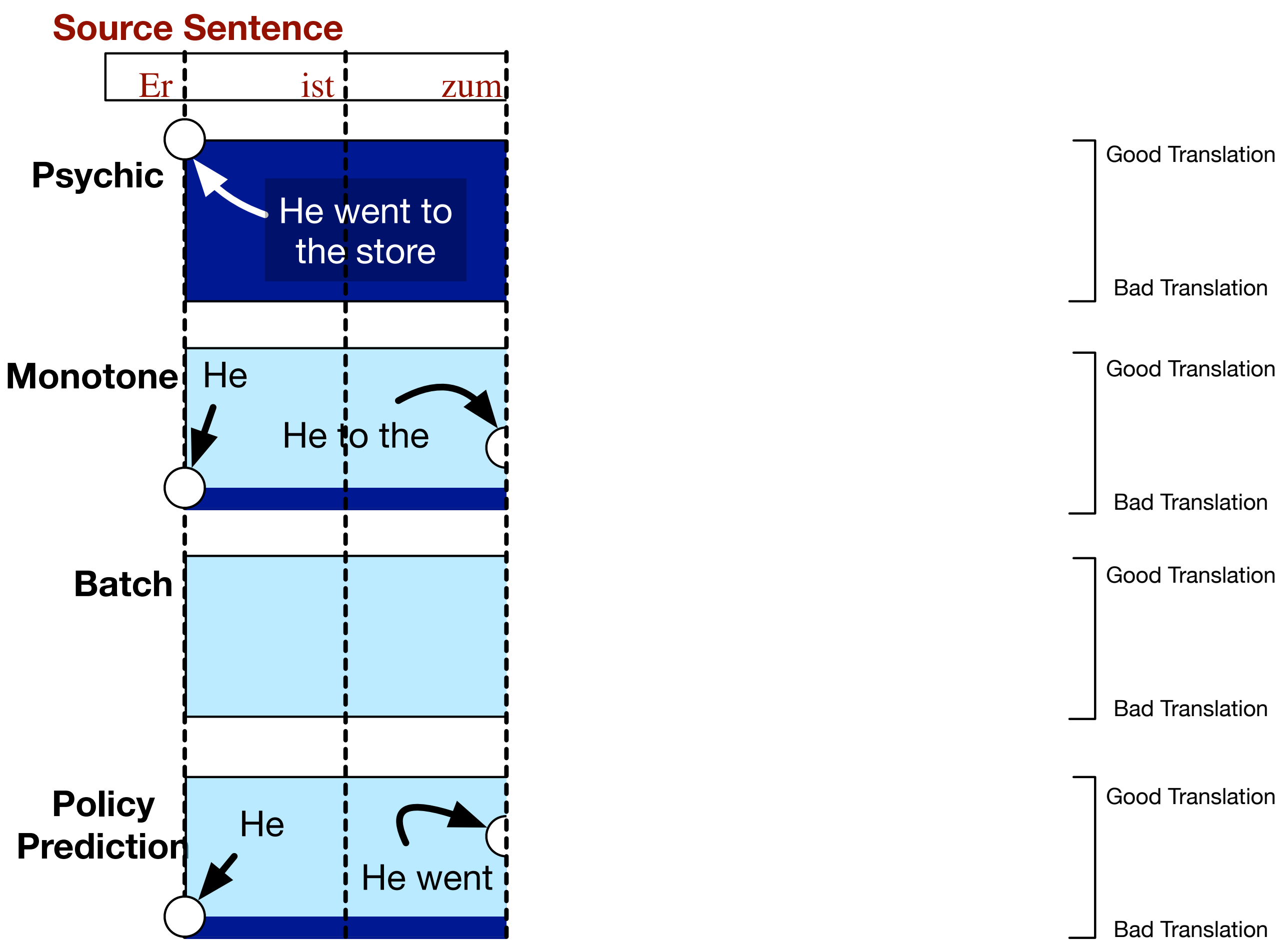
Policy Comparison



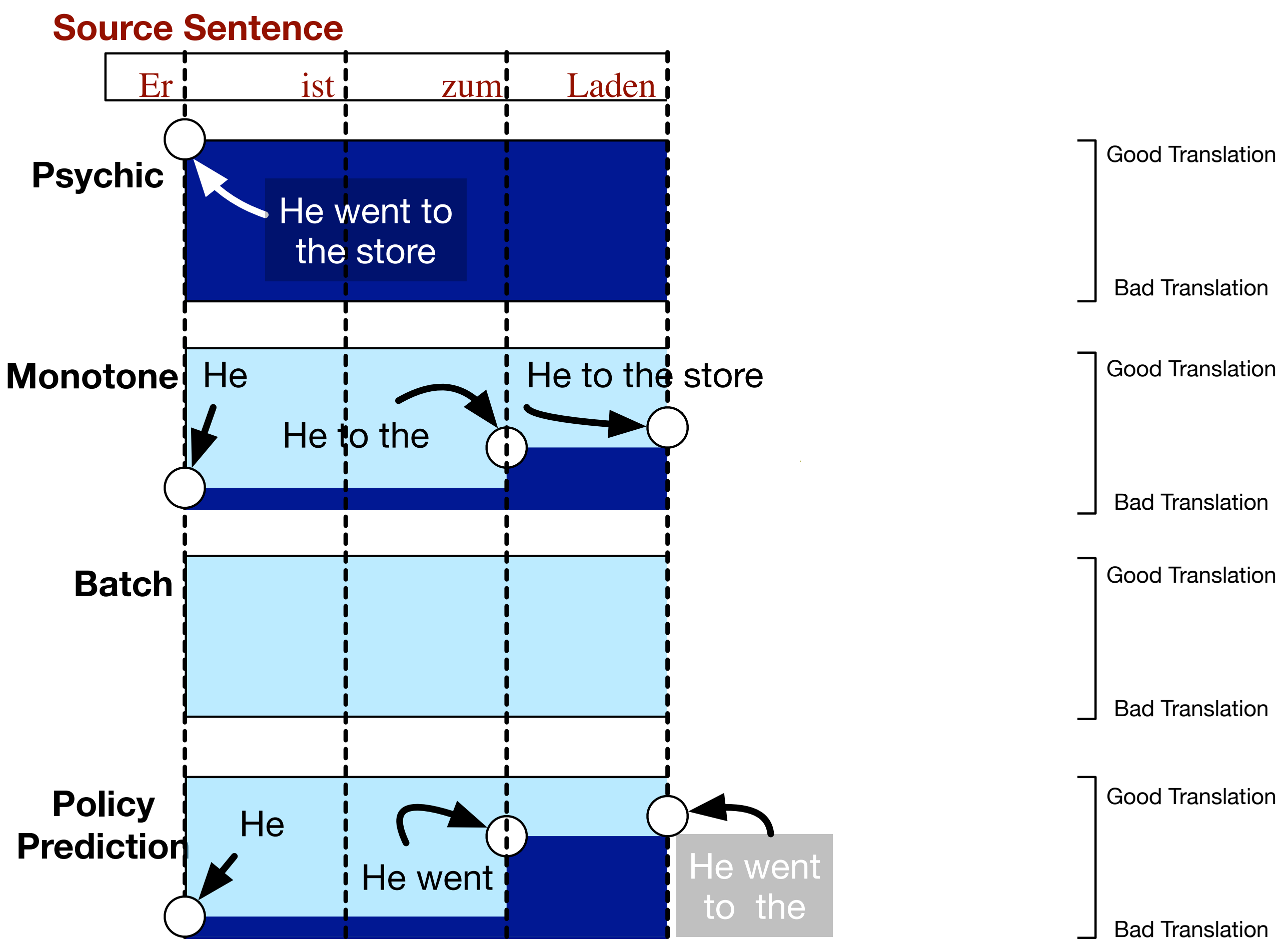
Policy Comparison



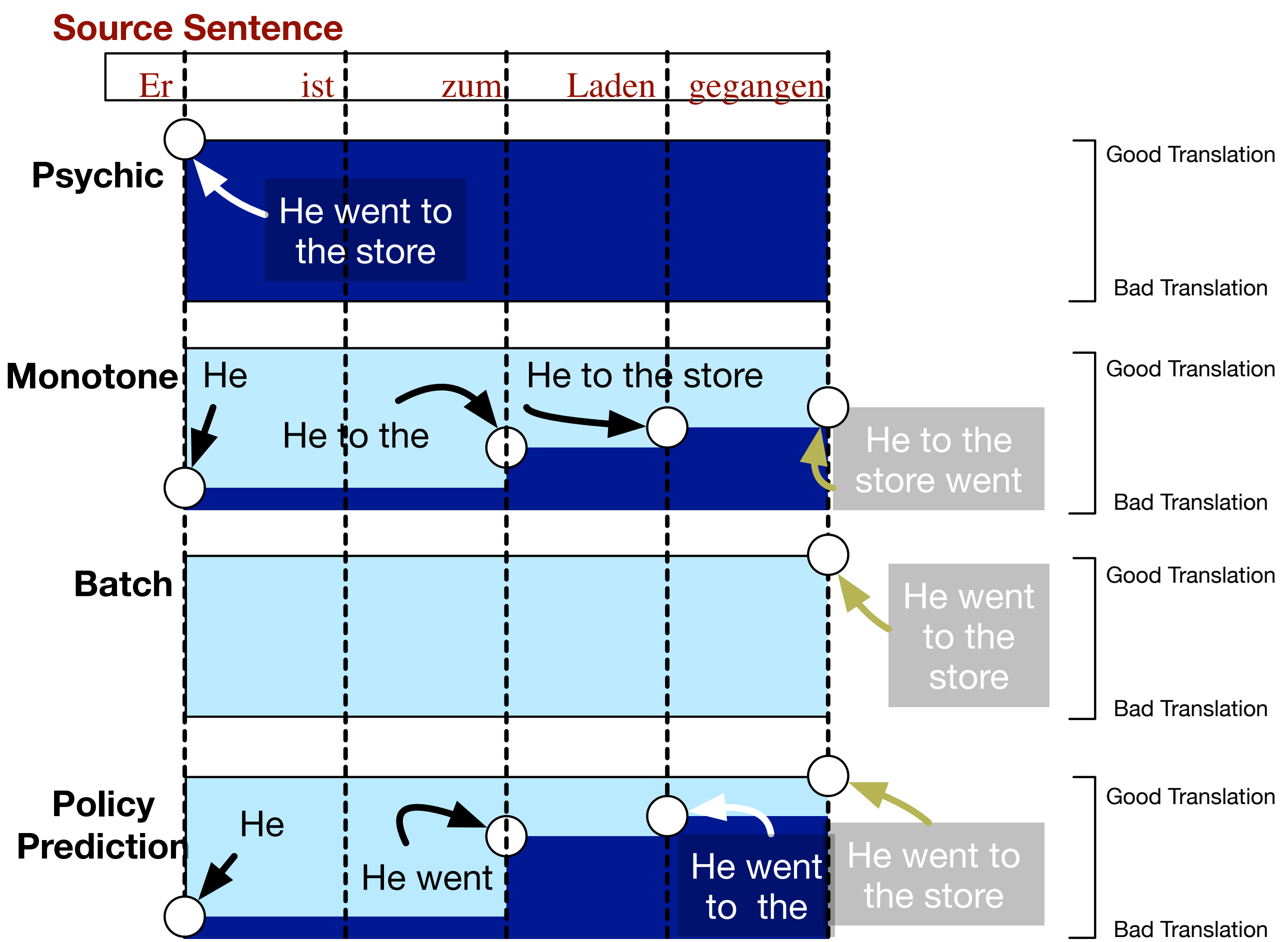
Policy Comparison



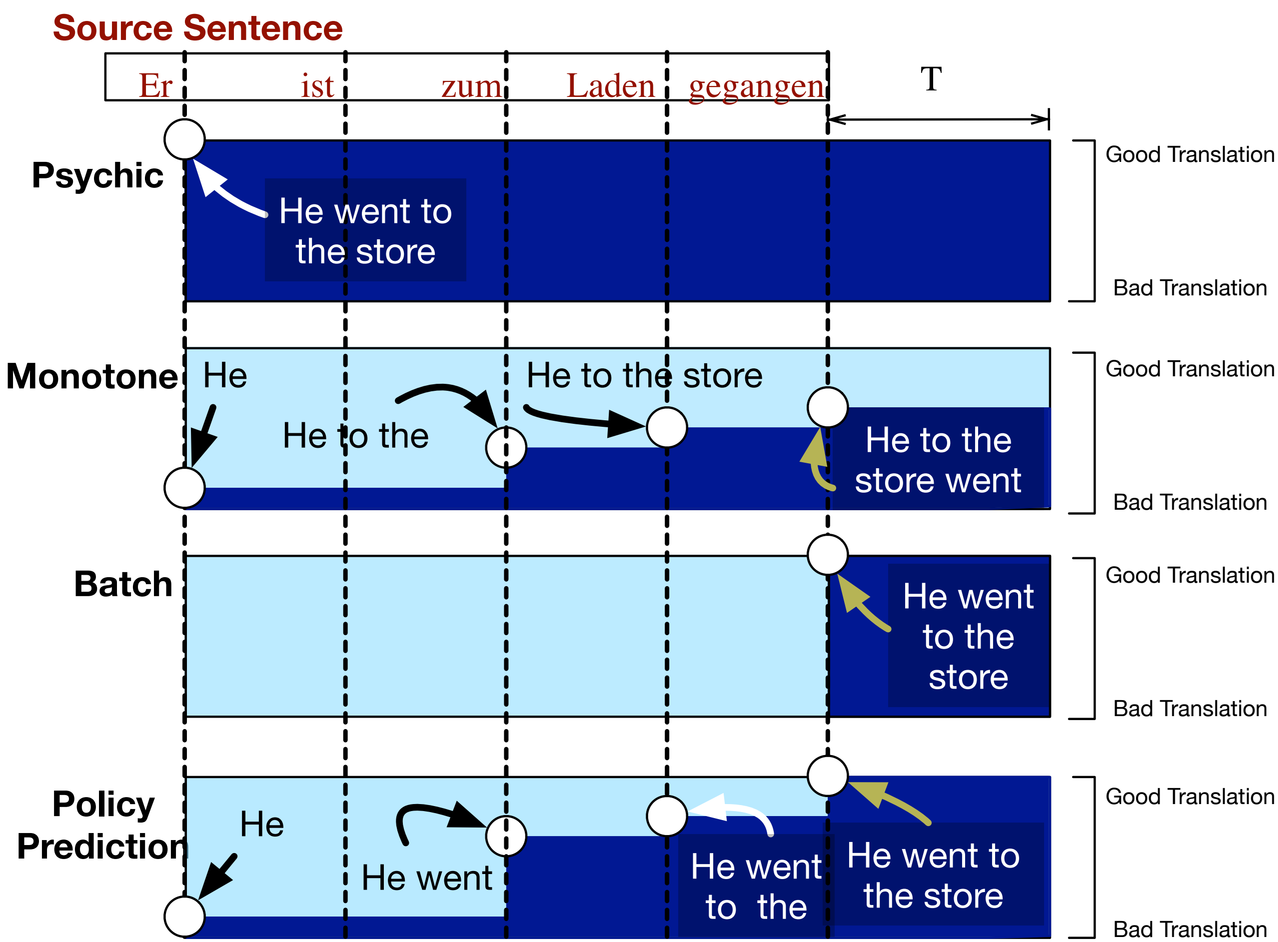
Policy Comparison



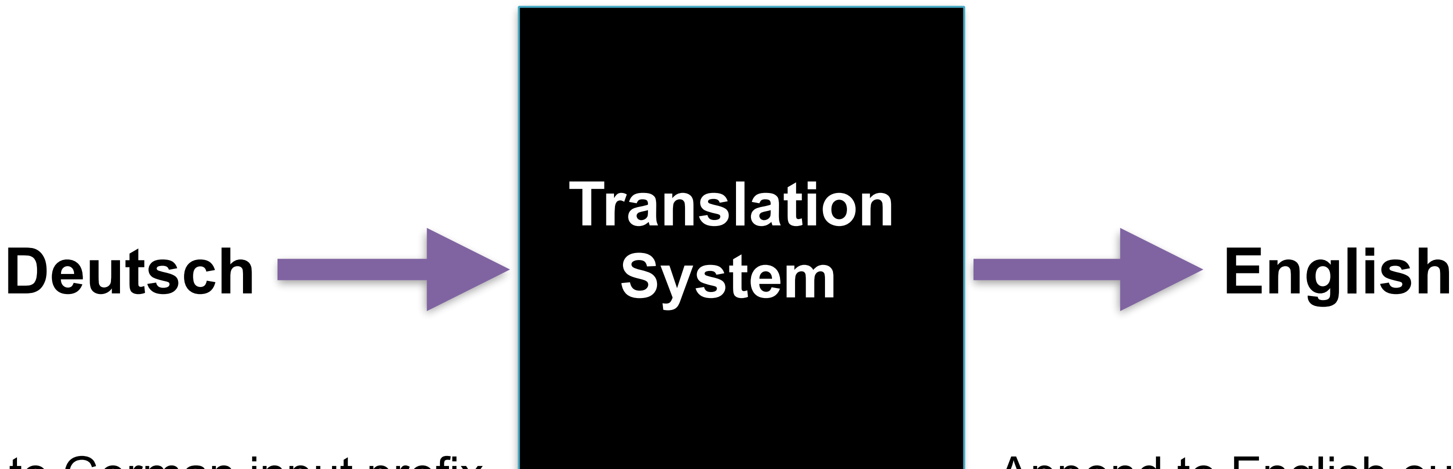
Policy Comparison



Policy Comparison



Translation System



Append to German input prefix.

Append to English output prefix.

Translation System

1 He

He

Translation System

2 It was designed

1 He

He

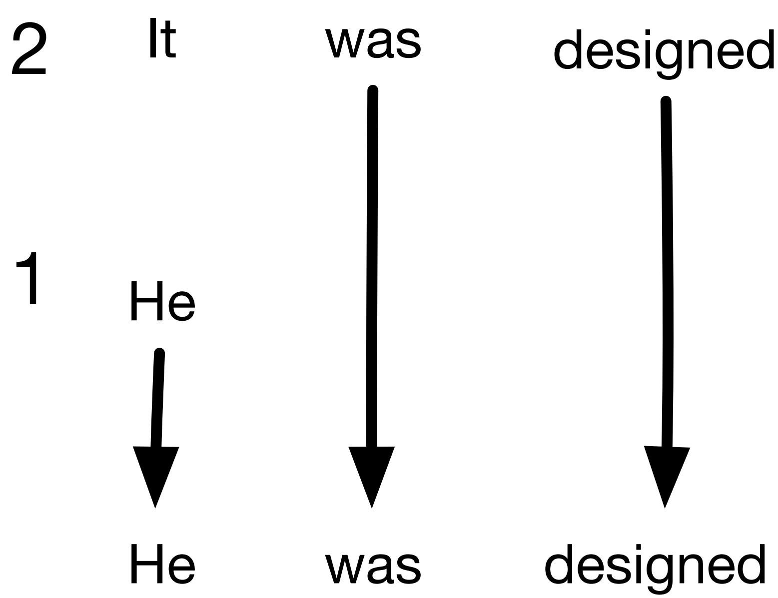
Translation System

2 It was designed

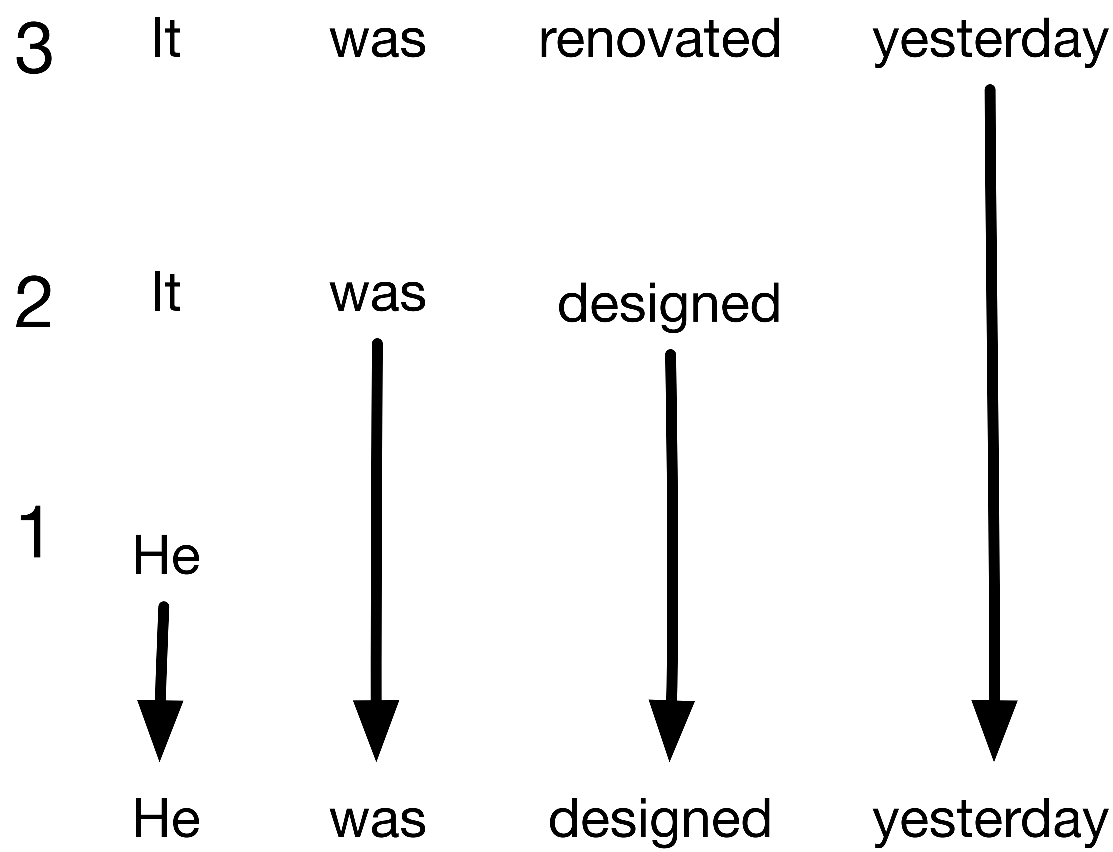
1 He

He

Translation System



Translation System



Action Sequence Learning

Observation

1. Mit dem Zug

state

Verb: **gewesen**
Next: **und**

Action Sequence Learning

Observation

1. Mit dem Zug

Wait

state

Verb: **gewesen**
Next: **und**

Action Sequence Learning

Observation		Observation
1. Mit dem Zug	Wait →	2. Mit dem Zug bin ich
state: Verb: **gewesen** Next: **und**		Verb: **geliefert** Next: **nur**

Action Sequence Learning

Observation

1. Mit dem Zug

Wait

Observation

2. Mit dem Zug bin ich

Wait

state

Verb: **gewesen**
Next: **und**

Verb: **geliefert**
Next: **nur**

Action Sequence Learning

Observation		Observation		Observation
1. Mit dem Zug	Wait →	2. Mit dem Zug bin ich	Wait →	3. Mit dem Zug bin ich nach
Verb: **gewesen** Next: **und**		Verb: **geliefert** Next: **nur**		Verb: **gefahren** Next: **Berlin**

state

Action Sequence Learning

Observation
1. Mit dem Zug

Wait

Observation
2. Mit dem Zug bin ich

Wait

Observation
3. Mit dem Zug bin ich nach

Predict

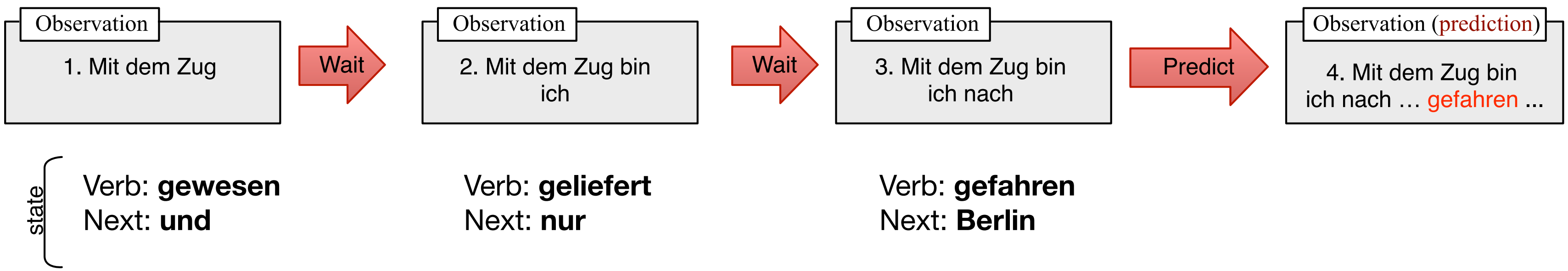
state

Verb: **gewesen**
Next: **und**

Verb: **geliefert**
Next: **nur**

Verb: **gefahren**
Next: **Berlin**

Action Sequence Learning



Action Sequence Learning

Observation
1. Mit dem Zug

Wait

Observation
2. Mit dem Zug bin ich

Wait

Observation
3. Mit dem Zug bin ich nach

Predict

Observation (prediction)
4. Mit dem Zug bin ich nach … gefahren ...

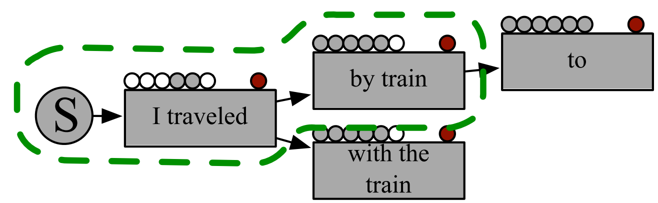
state
Verb: **gewesen**
Next: **und**

Verb: **geliefert**
Next: **nur**

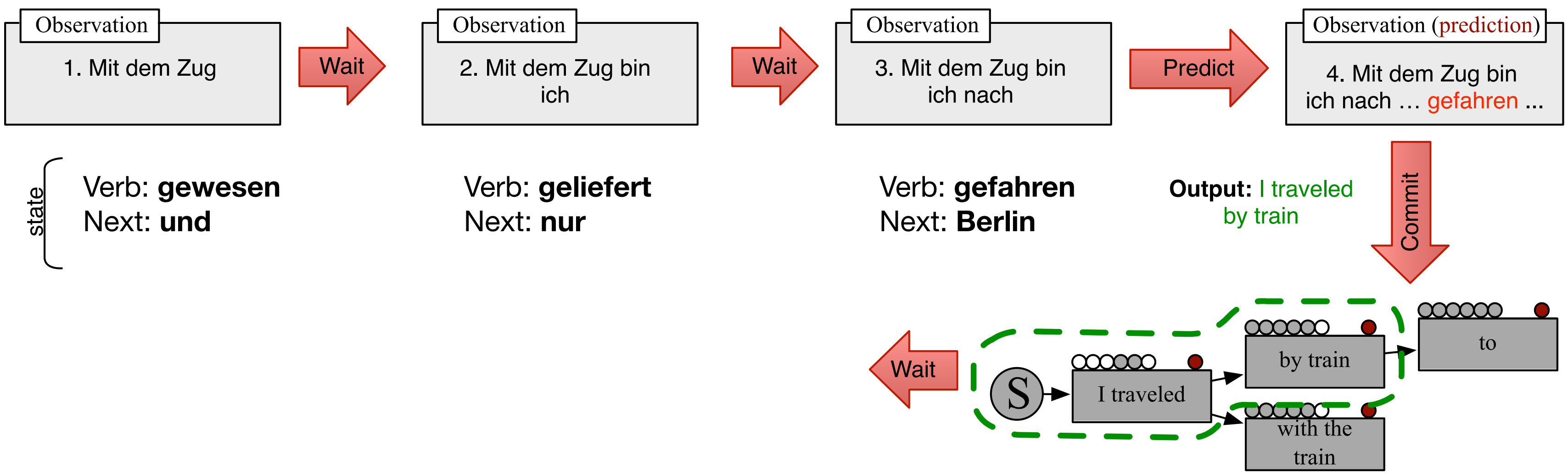
Verb: **gefahren**
Next: **Berlin**

Output: I traveled by train

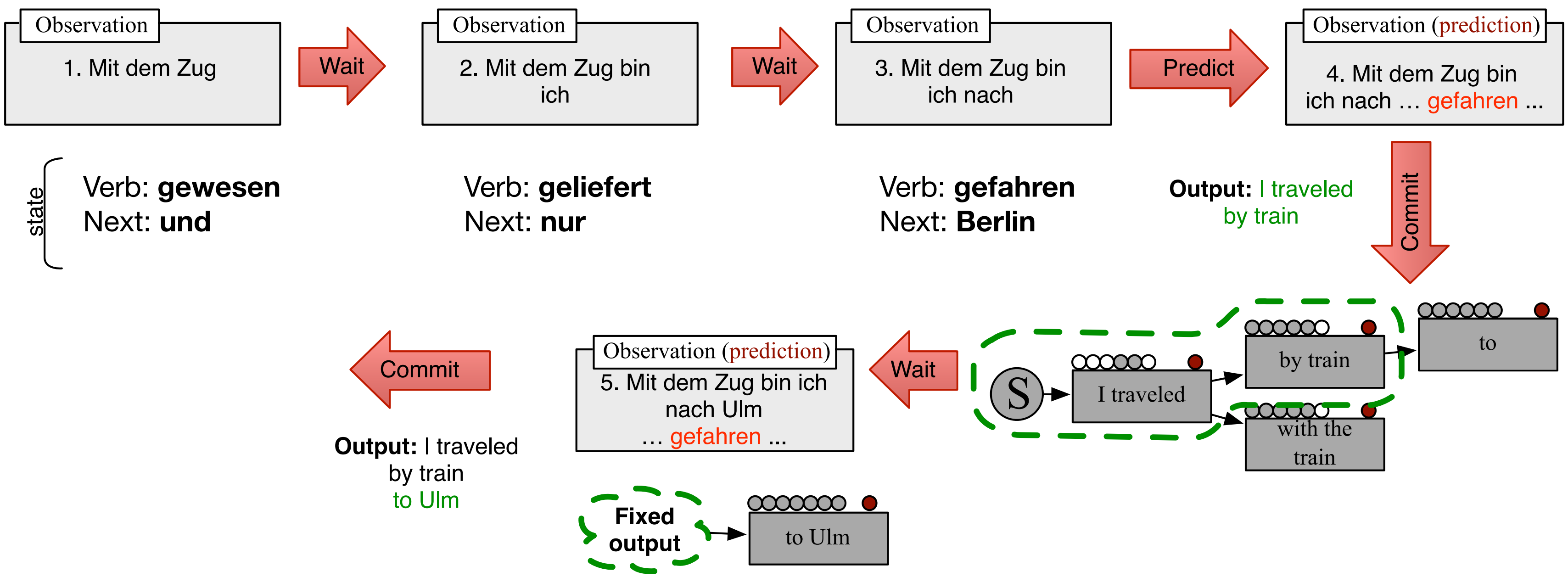
Commit



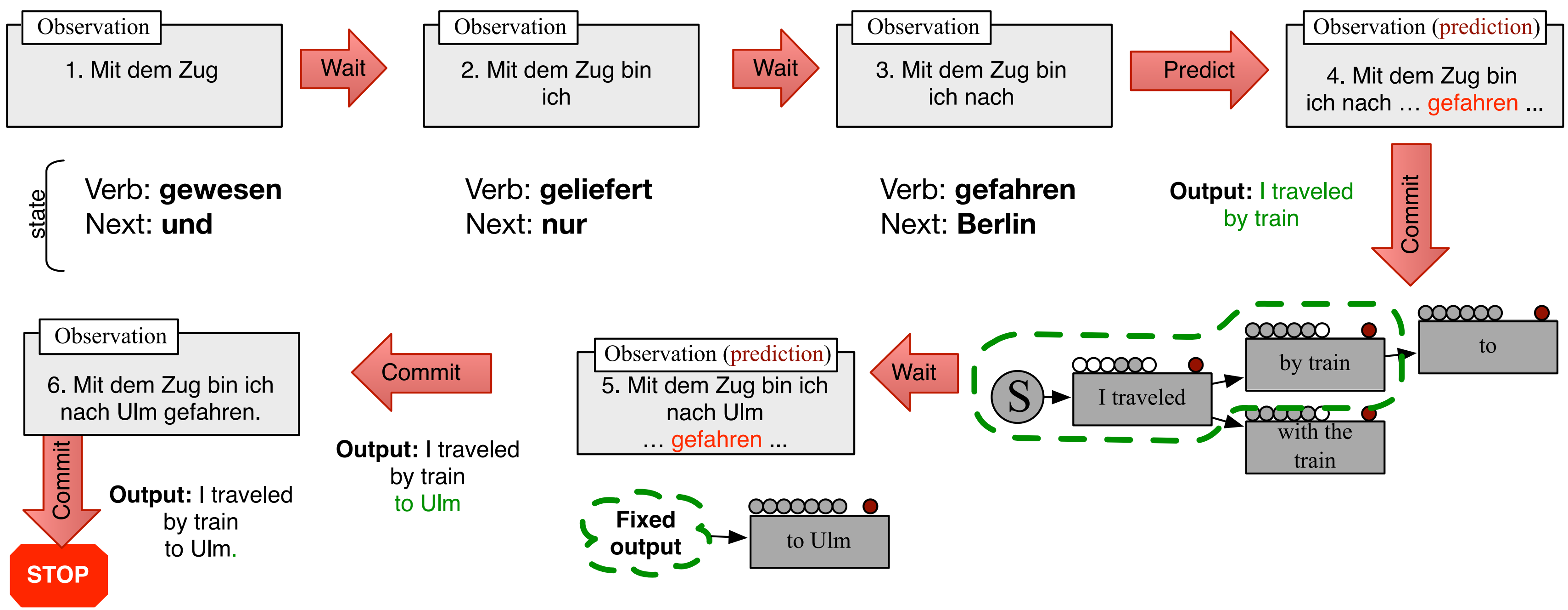
Action Sequence Learning



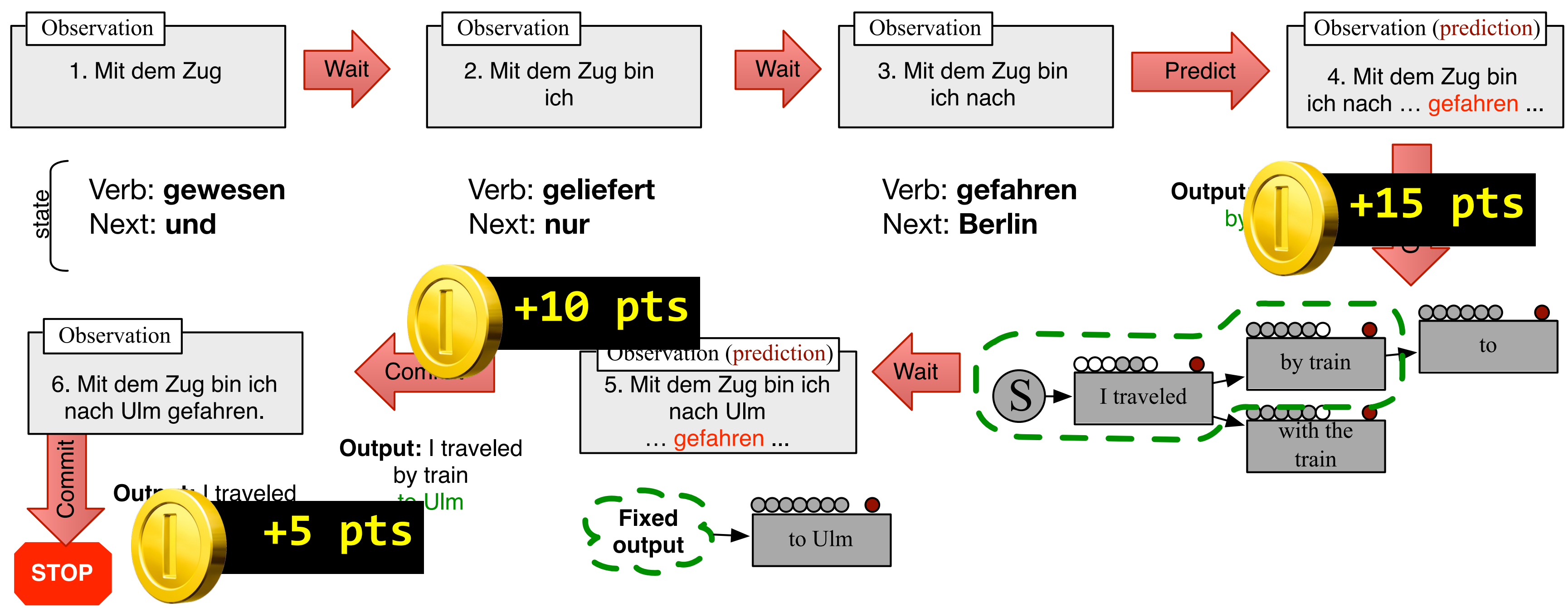
Action Sequence Learning



Action Sequence Learning

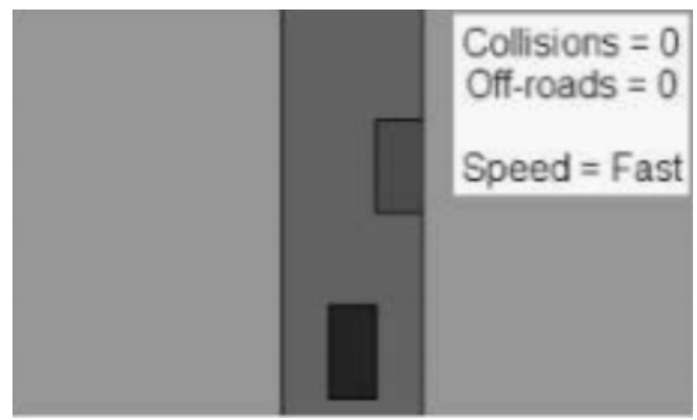


Action Sequence Learning

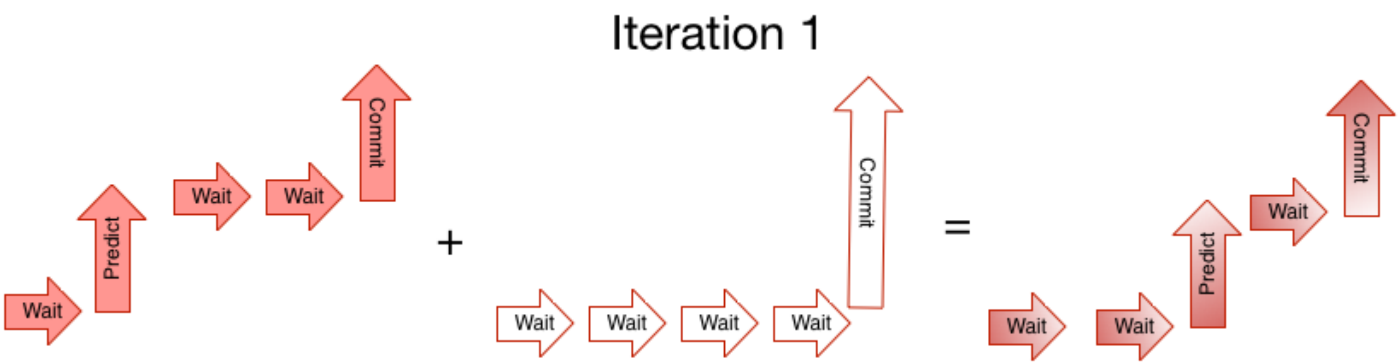


Learning When and How

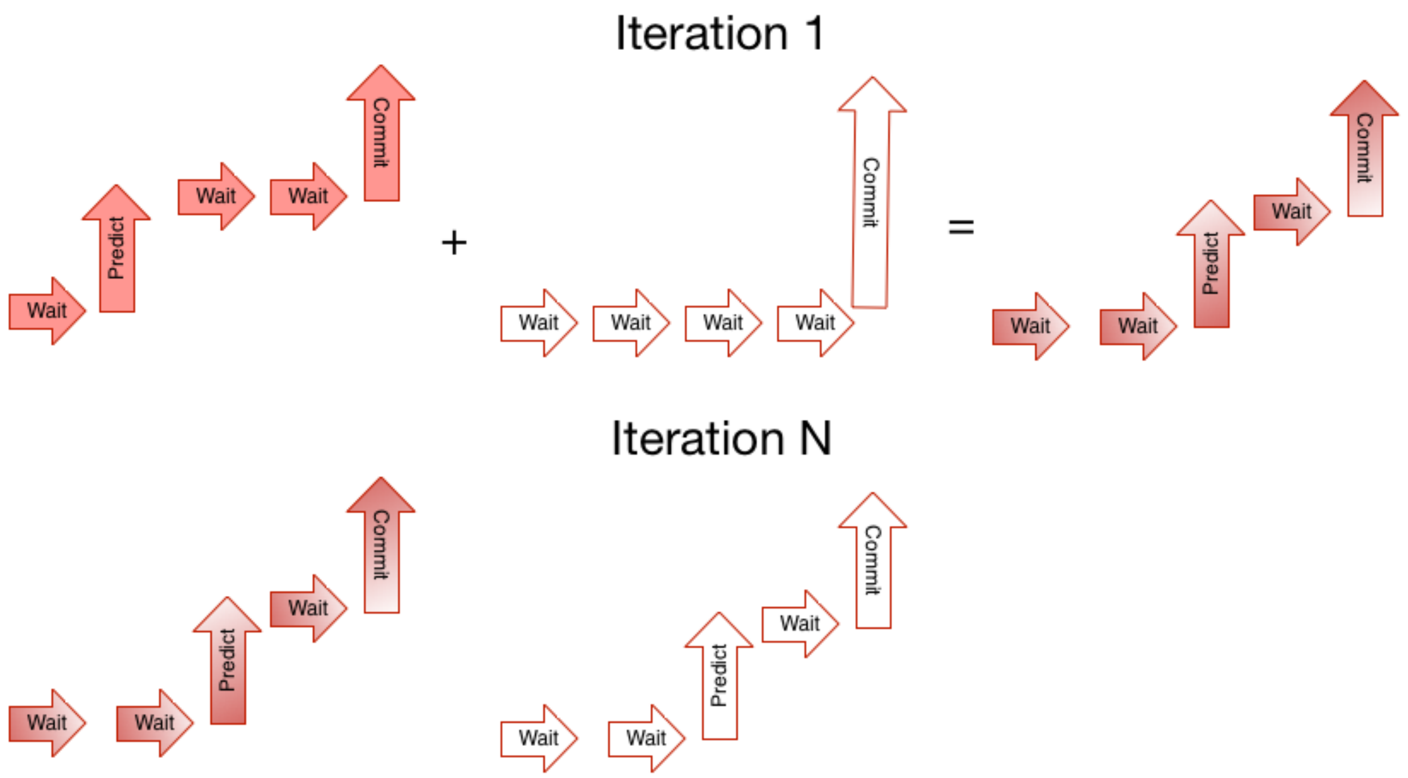
- Imitation learning
  - Special case of reinforcement learning.
  - Given predictions and prediction-informed translations, discover optimal policies in hindsight.



Learning When and How



Learning When and How



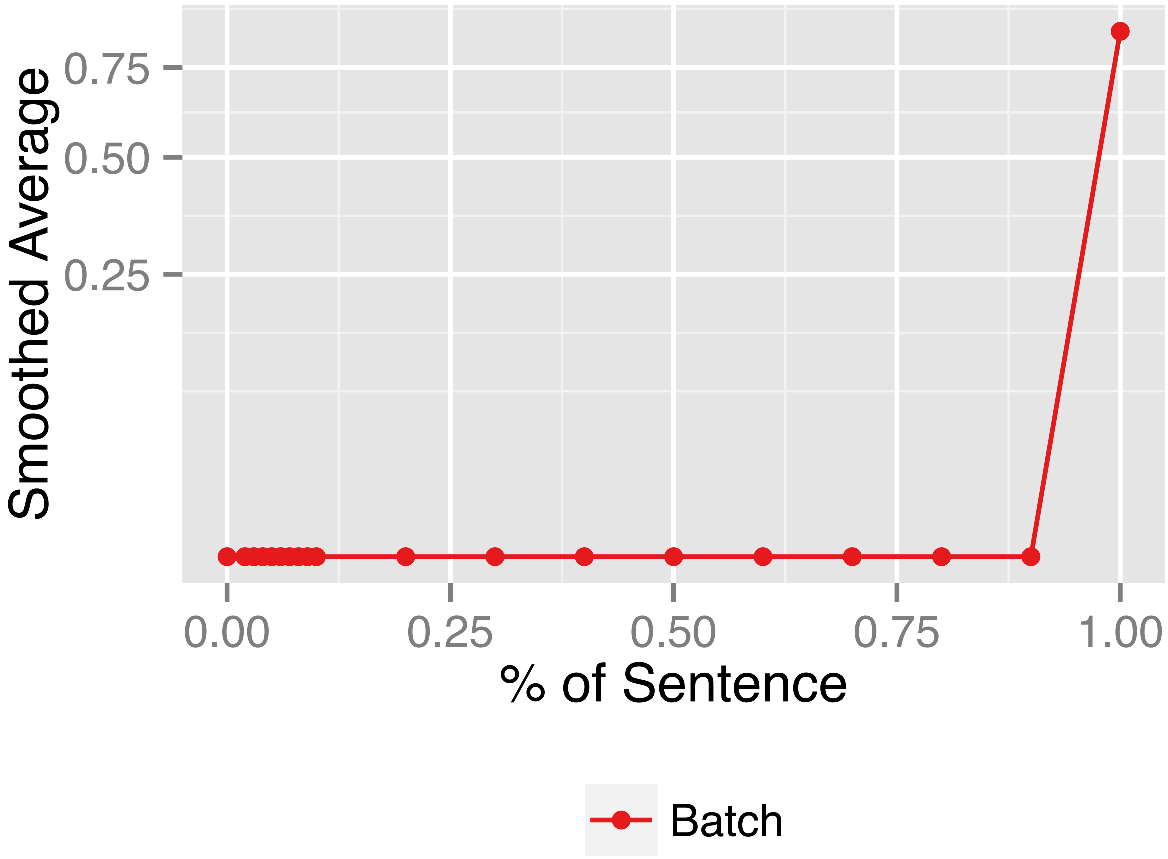
Learning When and How

- Find optimal policy, π^*.
- Set *initial* policy to optimal policy: $\pi_0 \equiv \pi^*$
- Until convergence:
  - Generate examples of state-action pairs: $(\pi_t(s), s)$
  - Generate a classifier (apprentice policy) mapping states to actions: $h_t : f(s) \mapsto A$
    - $\mathcal{C}(a_t, \boldsymbol{x}) \equiv Q(\boldsymbol{x}, \pi^*(x_t)) - Q(\boldsymbol{x}, a_t(x_t))$
- Interpolate learned classifier with previous iteration's policy: $\pi_{t+1} = \lambda \pi_t + (1 - \lambda) h_t$
- Searn (Daumé III et al., 2006)

Experiments

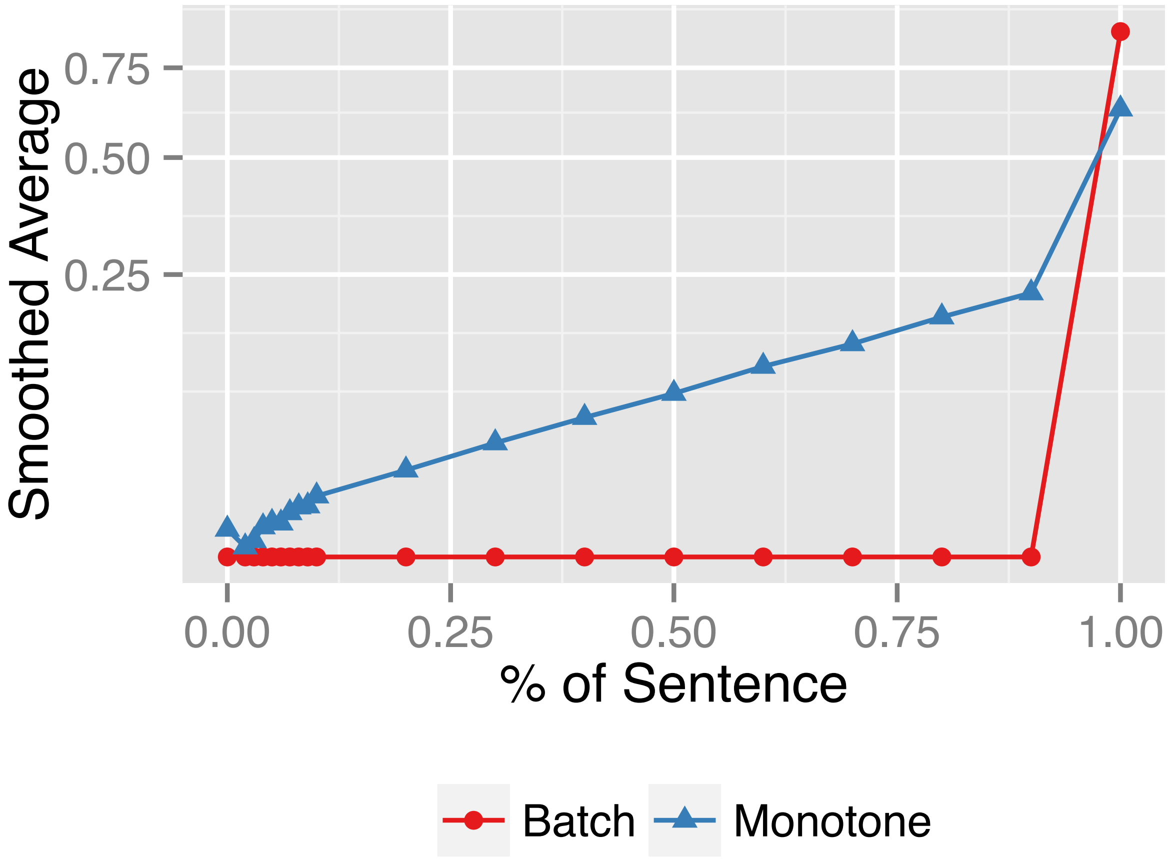
- "de-news" corpus
  - Daily News transcriptions (1996-2000).
  - German-English parallel data.
- Only verb-final sentences.
- Some compromises with translation system (to ensure verbs appeared).

Results



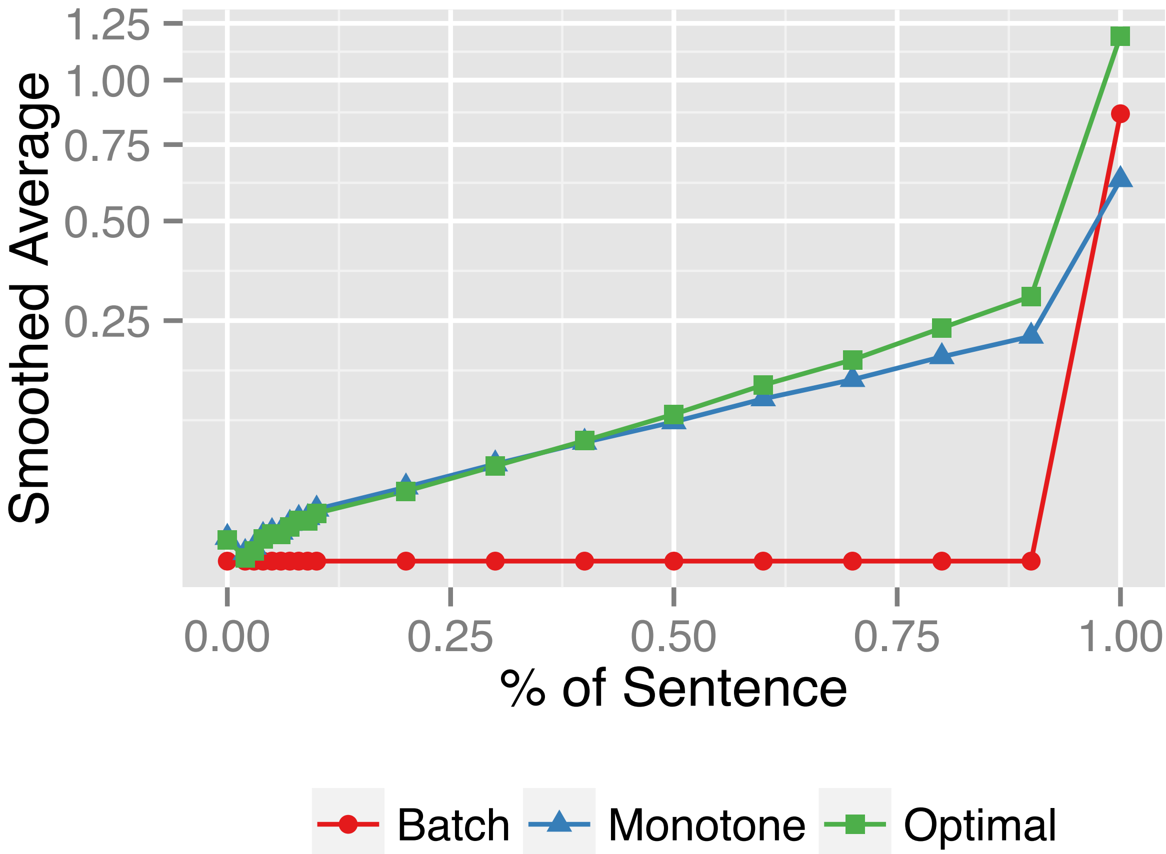
Smoothed Average
0.75
0.50
0.25
0.00
0.25
0.50
0.75
1.00
% of Sentence
Batch

Results



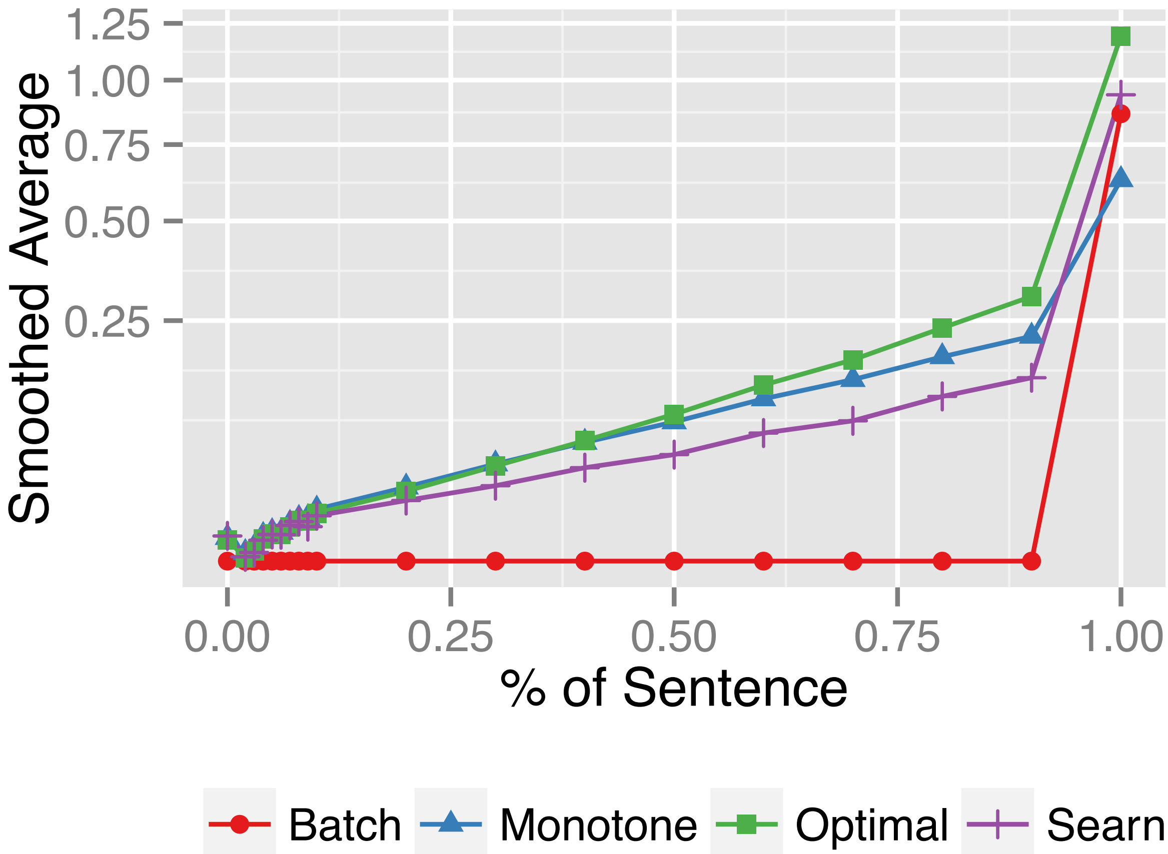
Smoothed Average
0.75
0.50
0.25
0.00
0.25
0.50
0.75
1.00
% of Sentence
Batch
Monotone

Results



Smoothed Average
1.25
1.00
0.75
0.50
0.25
0.00
0.25
0.50
0.75
1.00
% of Sentence
Batch
Monotone
Optimal

Results



Smoothed Average
1.25
1.00
0.75
0.50
0.25
0.00
0.25
0.50
0.75
1.00
% of Sentence
Batch
Monotone
Optimal
Searn

Results: Example

INPUT		OUTPUT
bundesumweltministerin **Merkel**	NEXT	federal minister of the environment angela merkel
bundesumweltministerin merkel hat den entwurf **gezeigt**	VERB	federal minister of the environment angela merkel shown the draft
bundesumweltministerin merkel hat den entwurf eines umweltpolitischen programms vorgestellt	COMMIT	federal minister of the environment angela merkel shown the draft of an ecopolitical program

“presented”

Related Work

- Oda et al. (2014) learned segmentations for simultaneous interpretation with greedy search and dynamic programming.
- Previous rule-based approaches using parsing (Mima et al., 1998; Ryu et al., 2006), rule-based decisions (Wahlster, 2000), phrase-table probabilities (Fujita et al., 2013), pauses in speech (Sakamoto et al., 2013), and word alignments (Ryu et al., 2012)
- Verb Prediction for simultaneous machine translation
  - Matsubara et al. (2000) used pattern matching to predict English verbs for Japanese-English simultaneous machine translation.

Future Work

- Improved verb predictions with more robust models.
- Improved translation system.
- Incorporate richer feature space.
- Predict other components aside from verbs.
- Use on other languages, e.g., Japanese.
    - Verb/semantic-centric scoring metrics. (e.g., MEANT).