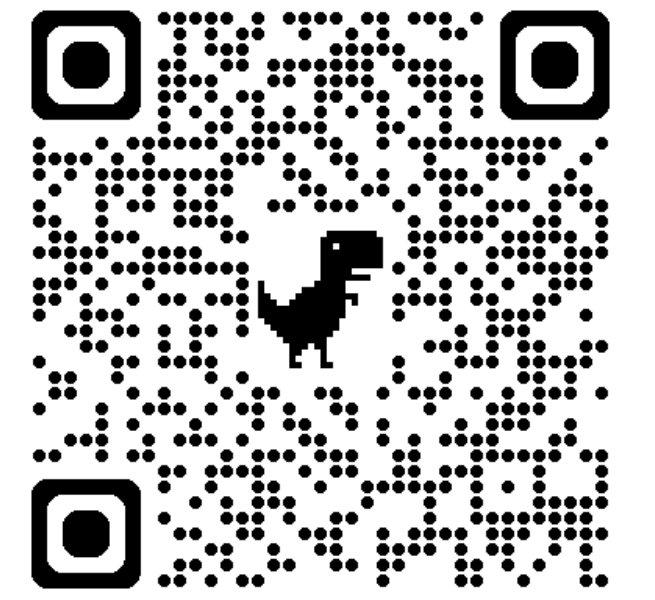


Structured Document Translation via Format Reinforcement Learning

Haiyue Song, Johannes Eschbach-Dymanus, Hour Kaing,
Sumire Honda, Hideki Tanaka, Bianka Buschbeck, Masao Utiyama
haiyue.song@nict.go.jp

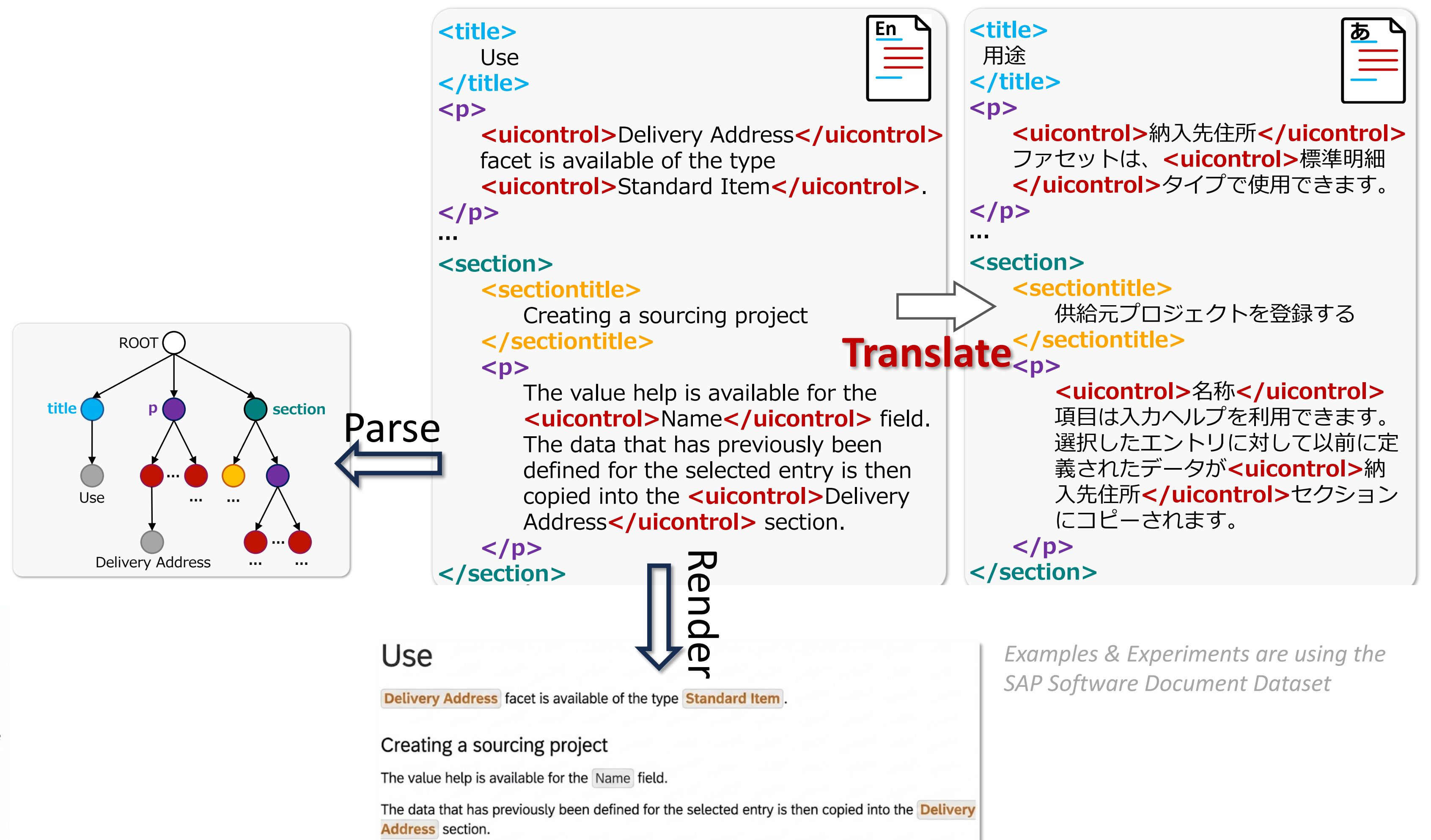
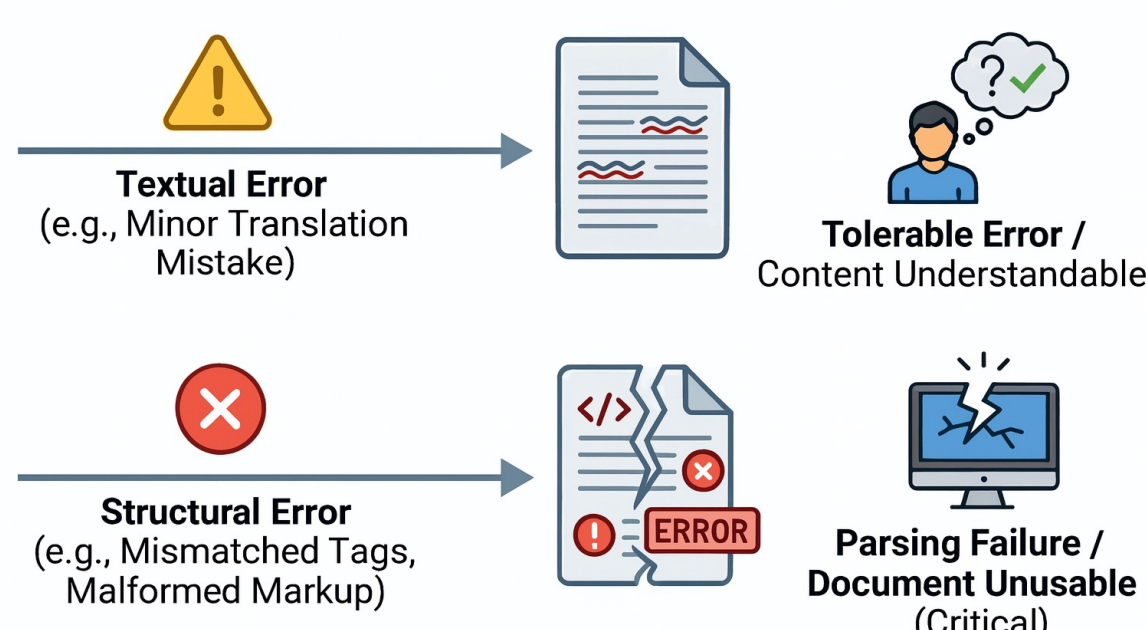


Structured Document Translation

- e.g. XML, HTML pages →
- It is more than just putting sentence-level translations in the correct position
 - Languages have different word order
 - Document Translation relies on context
- LLM can handle this task in an end-to-end way!

Challenge

- Minor translation errors are tolerable
- Structural errors cause parsing failure

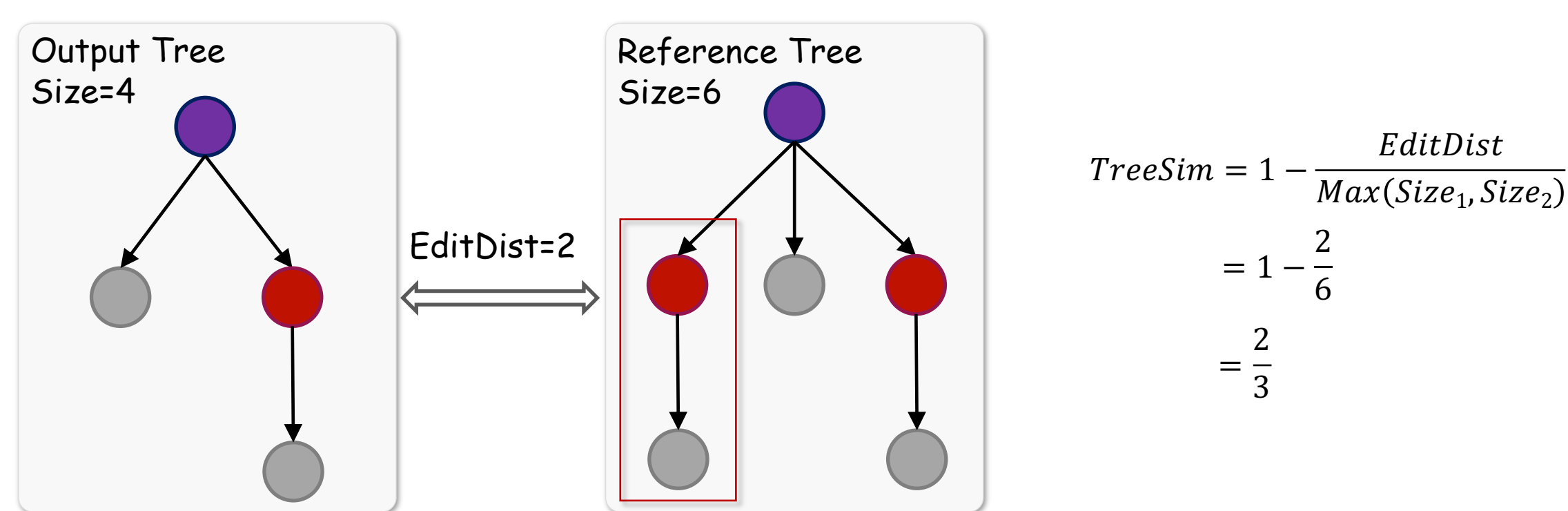


Examples & Experiments are using the SAP Software Document Dataset

Format Reward Design

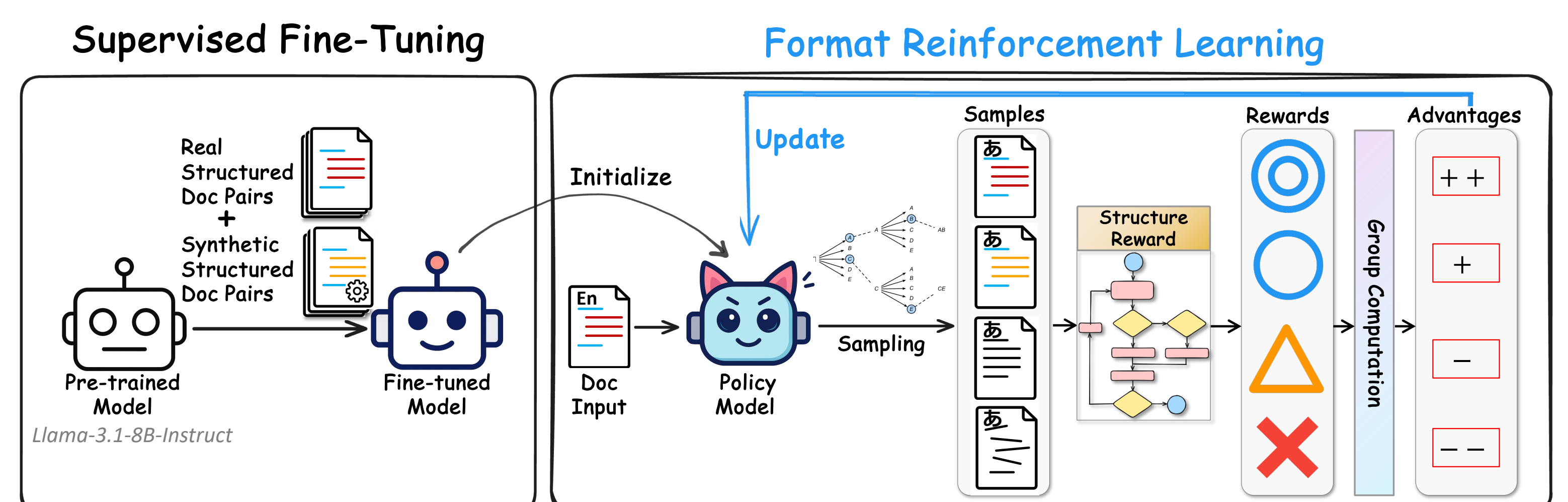
Reinforce structure preservation ability through rewards

- TreeSim: Tree similarity between output tree and reference tree (More rewards in Analysis)



FormatRL Two-stage Training

- Stage 1: SFT Learn basic document translation ability
- Stage 2: RL Reinforce structure preservation ability using GRPO



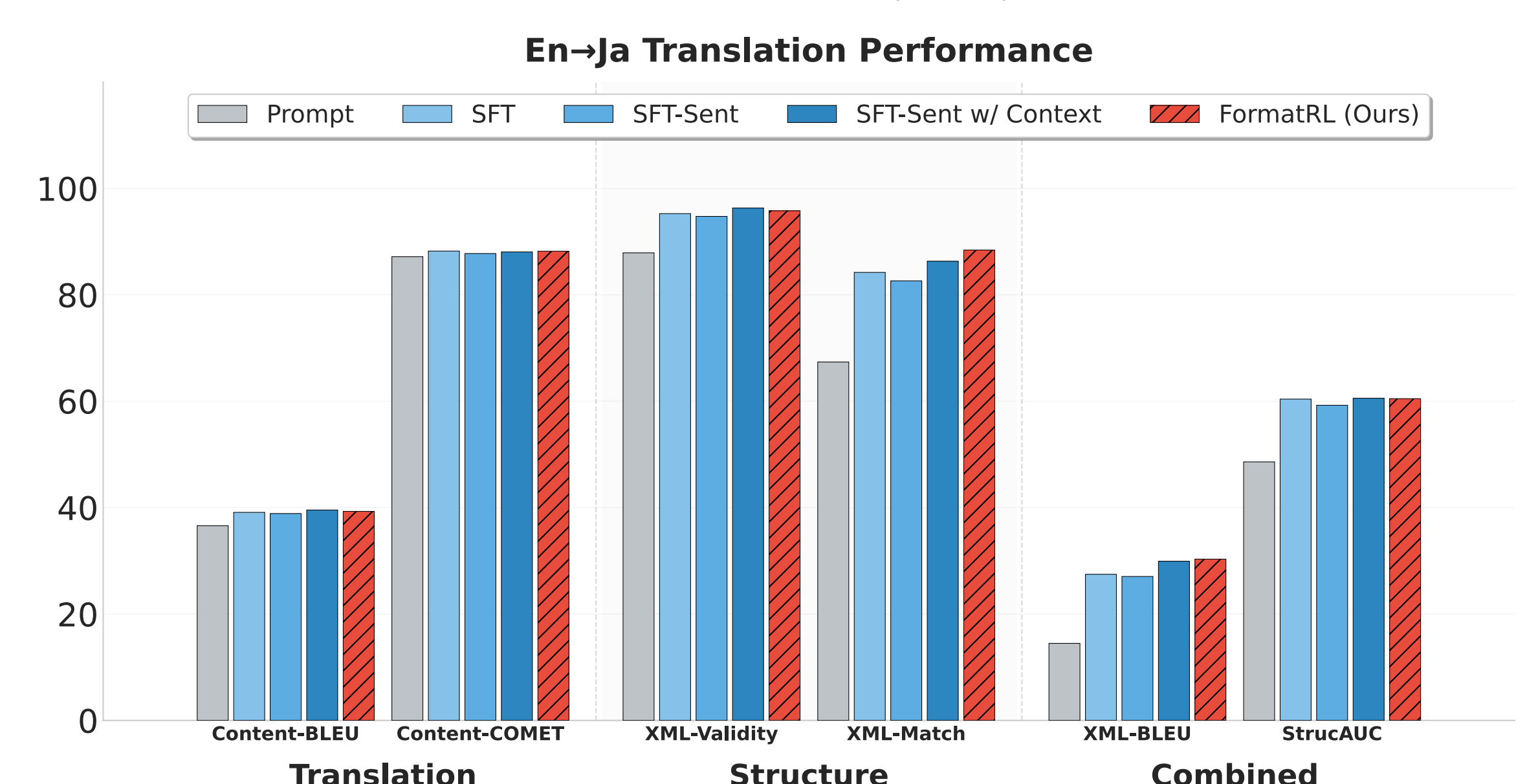
Results

- Structural accuracy improved.
- Combined metrics improved. They take account both translation and structure accuracy

Src→Tgt	Method	Translation		Structure		Combined		Avg.
		Content-BLEU	Content-COMET	XML-Validity	XML-Match	XML-BLEU	StrucAUC	
En→Zh	Prompt	49.88	86.16	91.05	76.84	27.50	57.75	64.86
	SFT	49.66	86.47	94.21	85.26	36.38	63.57	69.26
	FORMATRL	49.88	86.48	95.26	87.37	38.07	64.12	70.20
Zh→En	Prompt	48.82	85.25	93.16	82.11	26.34	71.39	67.84
	SFT	56.41	85.34	94.74	83.68	27.58	71.66	69.90
	FORMATRL	56.28	85.25	95.26	86.84	29.14	72.84	70.94
En→Ja	Prompt	36.60	87.17	87.89	67.37	14.49	48.60	57.02
	SFT	39.11	88.22	95.26	84.21	27.47	60.40	65.78
	FORMATRL	39.30	88.20	95.79	88.42	30.32	60.48	67.09
Ja→En	Prompt	44.14	85.93	90.53	80.00	22.38	65.25	64.70
	SFT	52.19	85.96	95.26	82.11	24.15	67.92	67.93
	FORMATRL	52.79	86.01	94.74	87.37	26.67	69.82	69.57

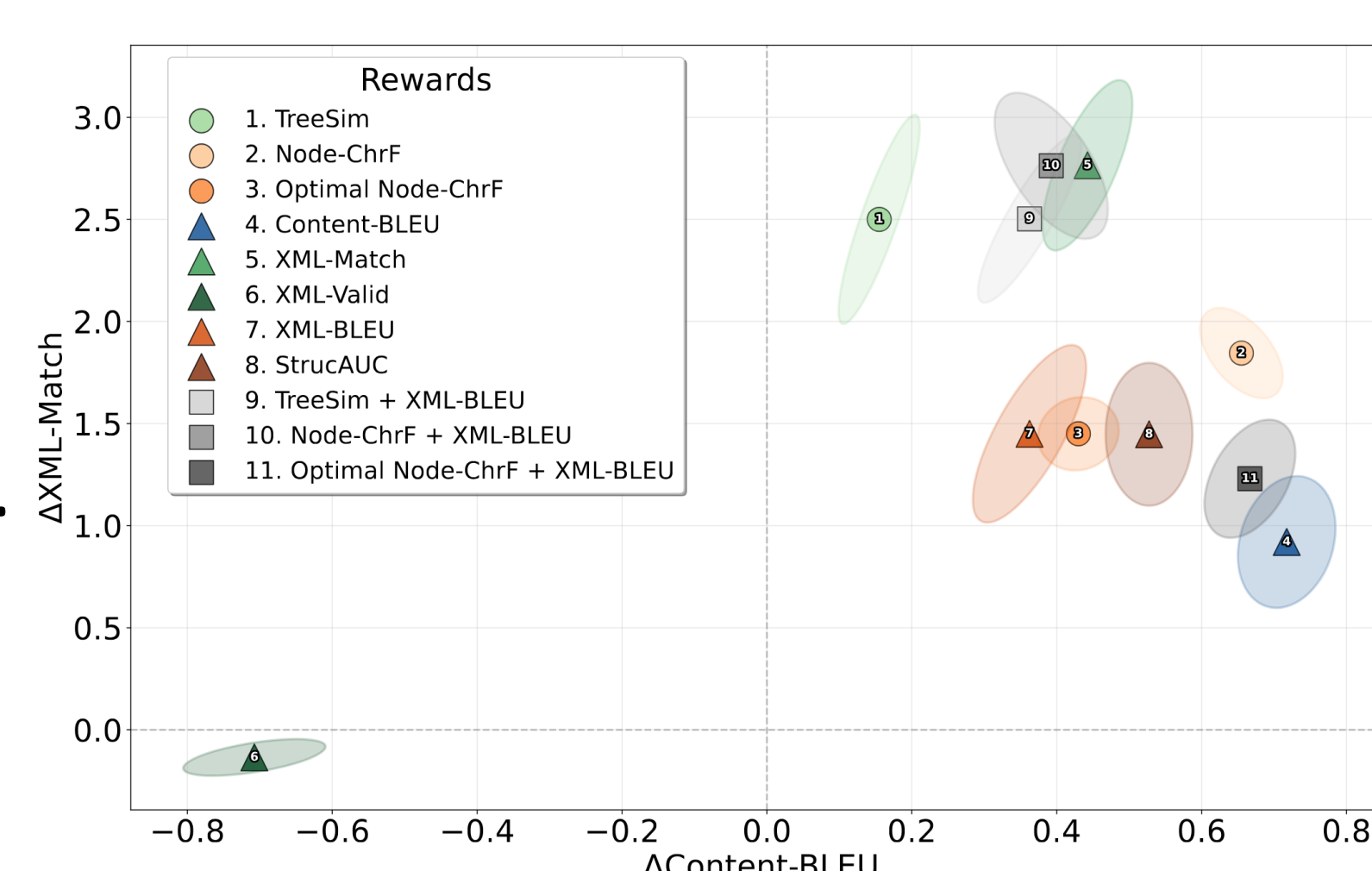
Better than Non-end-to-end Methods

- SFT-Sent: parse-and-assemble, sentence-level translation w/o context
- SFT-Sent w/ Context: parse-and-assemble, sentence-level translation w/ Doc as context (slow)



Analysis

- Structural rewards (1. TreeSim) improve structural acc.
- Content translation rewards (4. BLEU) improve translation acc.
- Achieve different goals by simply changing rewards!



Takeaways

- Structural document translation is a joint text and format task
- Format reinforcement learning with TreeSim reward can boost structural accuracy
- Reward design is simple & efficient for fine-grained needs