

Structured Document Translation via Format Reinforcement Learning

Haiyue Song, Johannes Eschbach-Dymanus, Hour Kaing, Sumire
Honda, Hideki Tanaka, Bianka Buschbeck, Masao Utiyama



IJCNLP-AAACL
2025



Structured Document

- XML, HTML pages
 - Not only **textual content**
 - But also **<markup>** that defines the structure

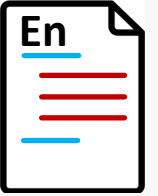
Use

Delivery Address facet is available of the type **Standard Item**.

Creating a sourcing project

The value help is available for the **Name** field.

The data that has previously been defined for the selected entry is then copied into the **Delivery Address** section.



<title>

Use

</title>

<p>

<uicontrol>Delivery Address**</uicontrol>**
facet is available of the type

<uicontrol>Standard Item**</uicontrol>**.

</p>

...

<section>

<sectiontitle>

Creating a sourcing project

</sectiontitle>

<p>

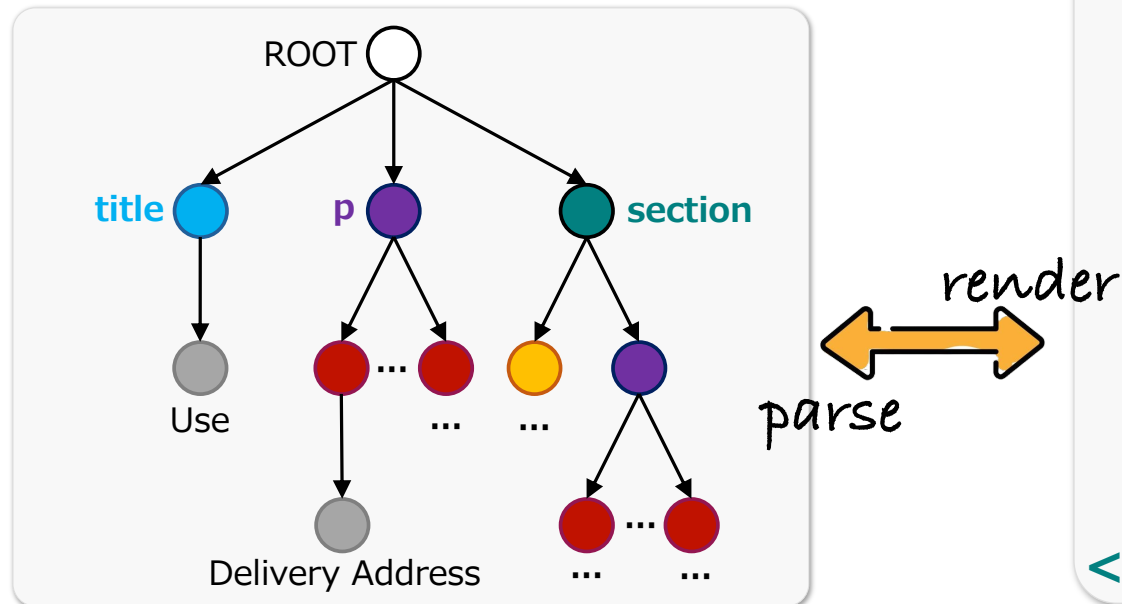
The value help is available for the
<uicontrol>Name**</uicontrol>** field.
The data that has previously been
defined for the selected entry is then
copied into the **<uicontrol>**Delivery
Address**</uicontrol>** section.

</p>

</section>

Structured Document

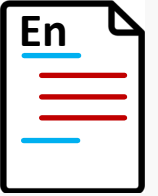
- XML, HTML pages
 - Not only **textual content**
 - But also **<markup>** that defines **tree structure**



<title>
Use
</title>

<p>
<uicontrol>Delivery Address**</uicontrol>**
facet is available of the type
<uicontrol>Standard Item**</uicontrol>**.
</p>
...

<section>
<sectiontitle>
Creating a sourcing project
</sectiontitle>
<p>
The value help is available for the
<uicontrol>Name**</uicontrol>** field.
The data that has previously been
defined for the selected entry is then
copied into the **<uicontrol>**Delivery
Address**</uicontrol>** section.
</p>
</section>



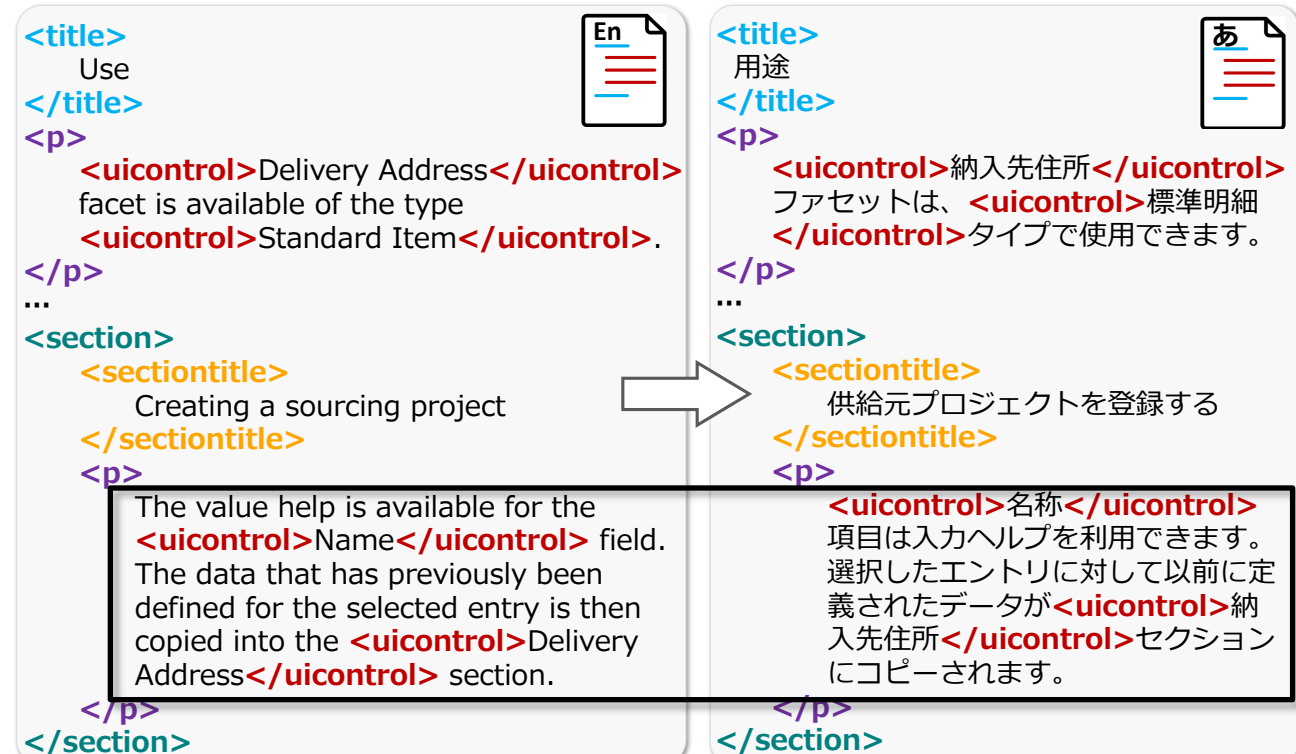
Structured Document Translation

- is just putting the translated sentence in the correct position?

<pre><title> Use </title> <p> <uicontrol>Delivery Address</uicontrol> facet is available of the type <uicontrol>Standard Item</uicontrol>. </p> ... <section> <sectiontitle> Creating a sourcing project </sectiontitle> <p> The value help is available for the <uicontrol>Name</uicontrol> field. The data that has previously been defined for the selected entry is then copied into the <uicontrol>Delivery Address</uicontrol> section. </p> </section></pre>		<pre><title> 用途 </title> <p> <uicontrol>納入先住所</uicontrol> ファセットは、<uicontrol>標準明細 </uicontrol>タイプで使用できます。 </p> ... <section> <sectiontitle> 供給元プロジェクトを登録する </sectiontitle> <p> <uicontrol>名称</uicontrol> 項目は入力ヘルプを利用できます。 選択したエントリに対して以前に定 義されたデータが<uicontrol>納 入先住所</uicontrol>セクション にコピーされます。 </p> </section></pre>
---	--	--

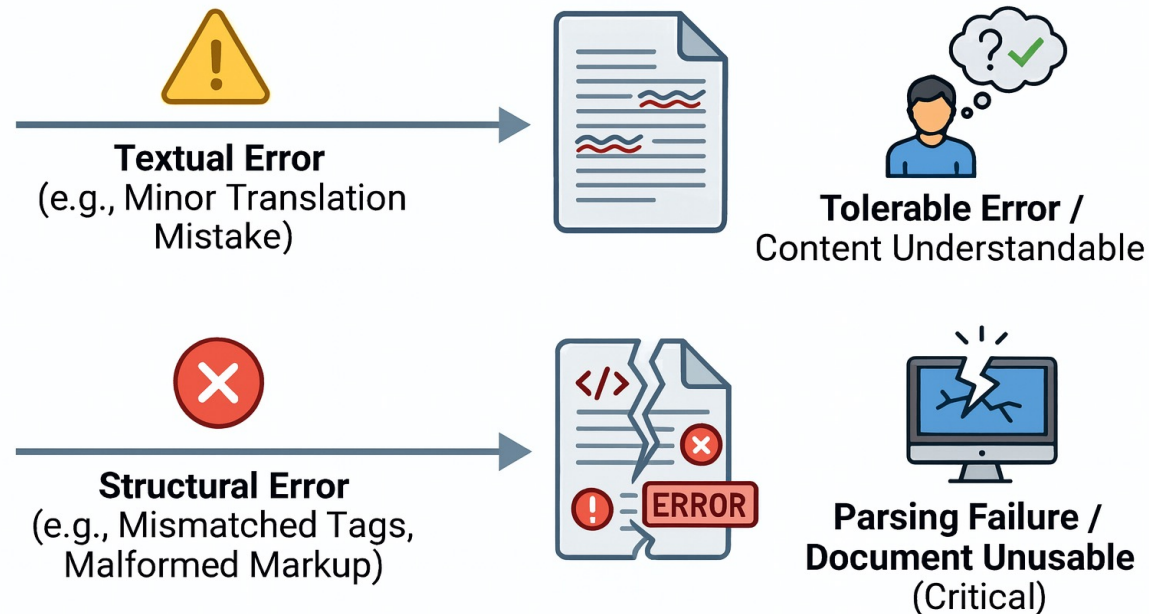
Structured Document Translation

- is **more than** just putting the translated sentence in the correct position!
 - Languages have their own **word order**
 - Doc Translation relies on context
- LLM can handle this task in an end-to-end way!



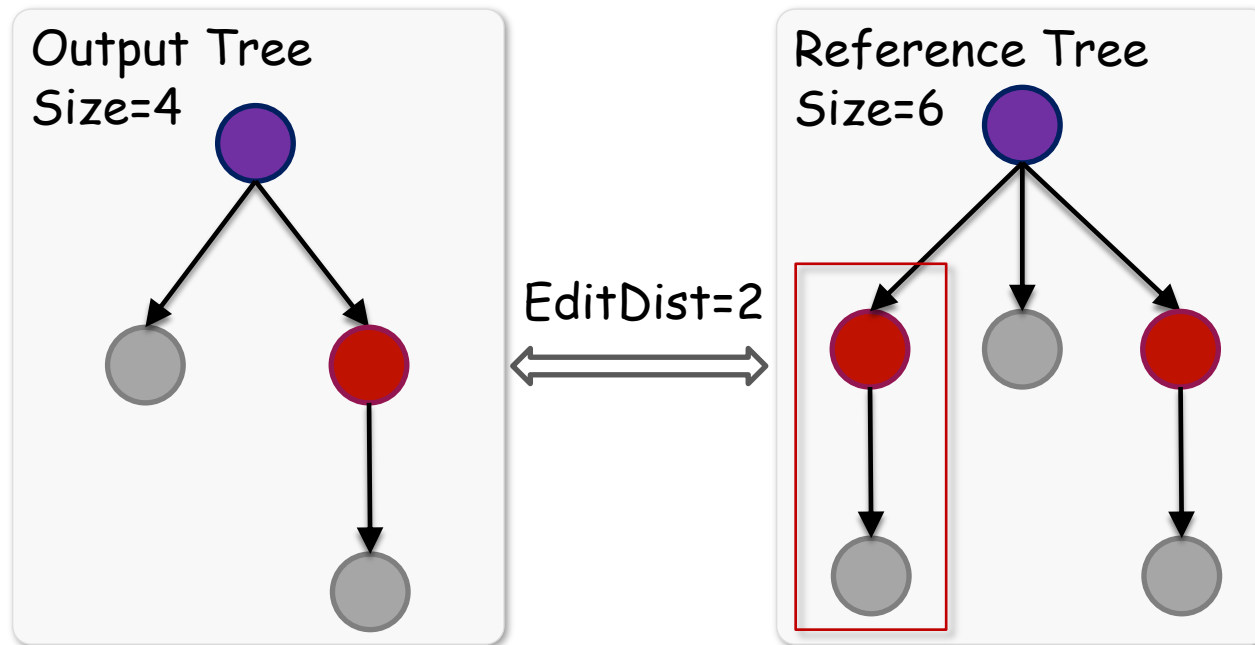
The Challenge

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



Key Idea

- **Reinforce structure preservation** ability through rewards
 - **TreeSim**: Tree similarity between output tree and reference tree
 - **NodeChrF**: ChrF score for each node in tree, compared to reference

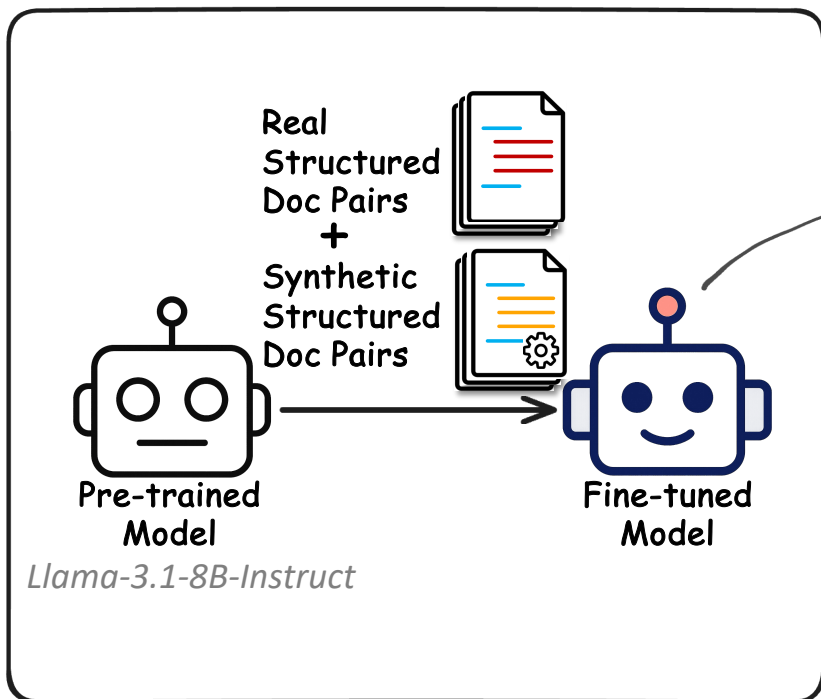


$$\begin{aligned} TreeSim &= 1 - \frac{EditDist}{Max(Size_1, Size_2)} \\ &= 1 - \frac{2}{6} \\ &= \frac{2}{3} \end{aligned}$$

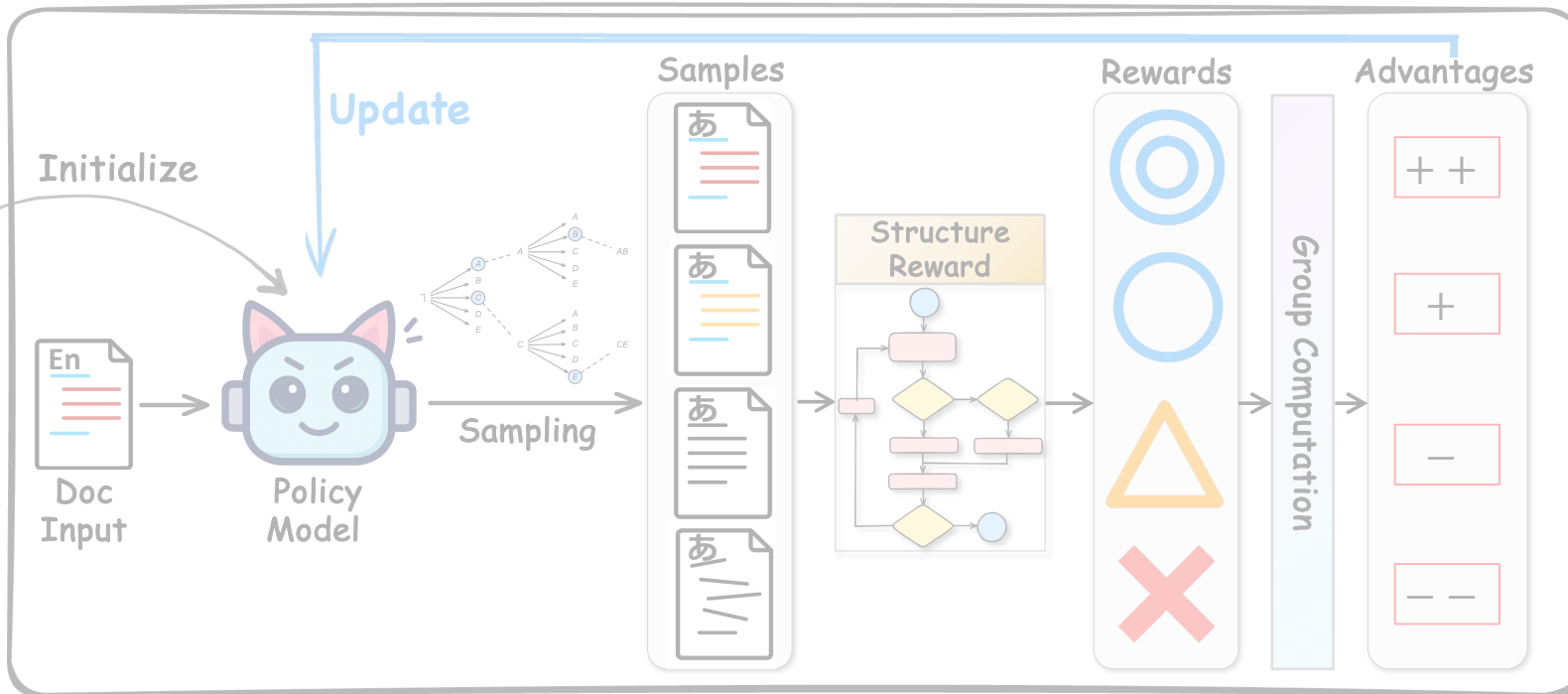
Two-stage Training

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

Supervised Fine-Tuning



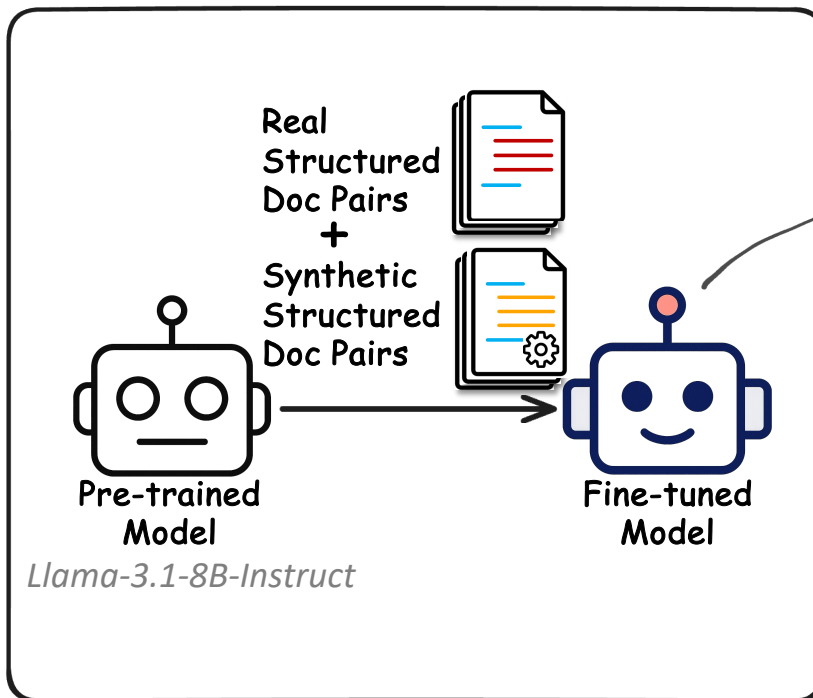
Format Reinforcement Learning



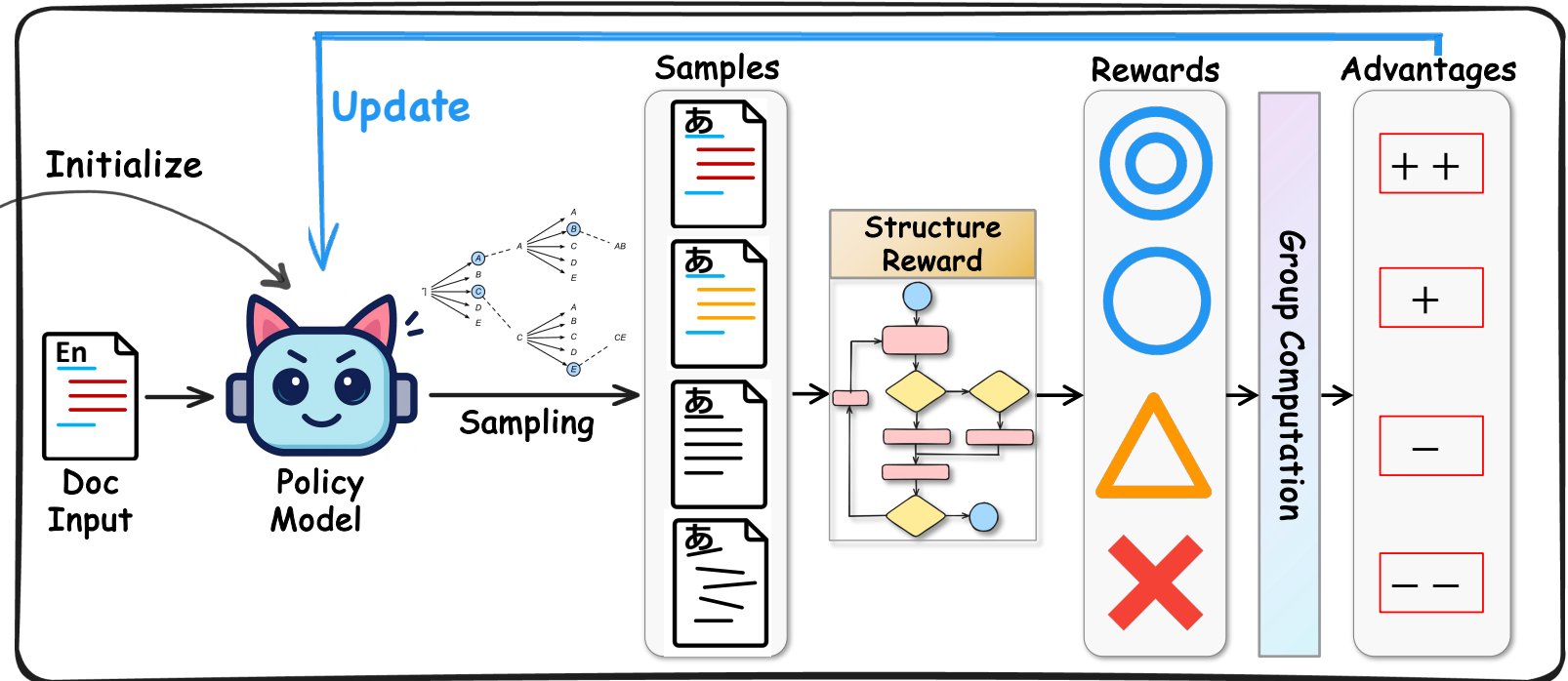
Two-stage Training

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

Supervised Fine-Tuning



Format Reinforcement Learning



Main Results

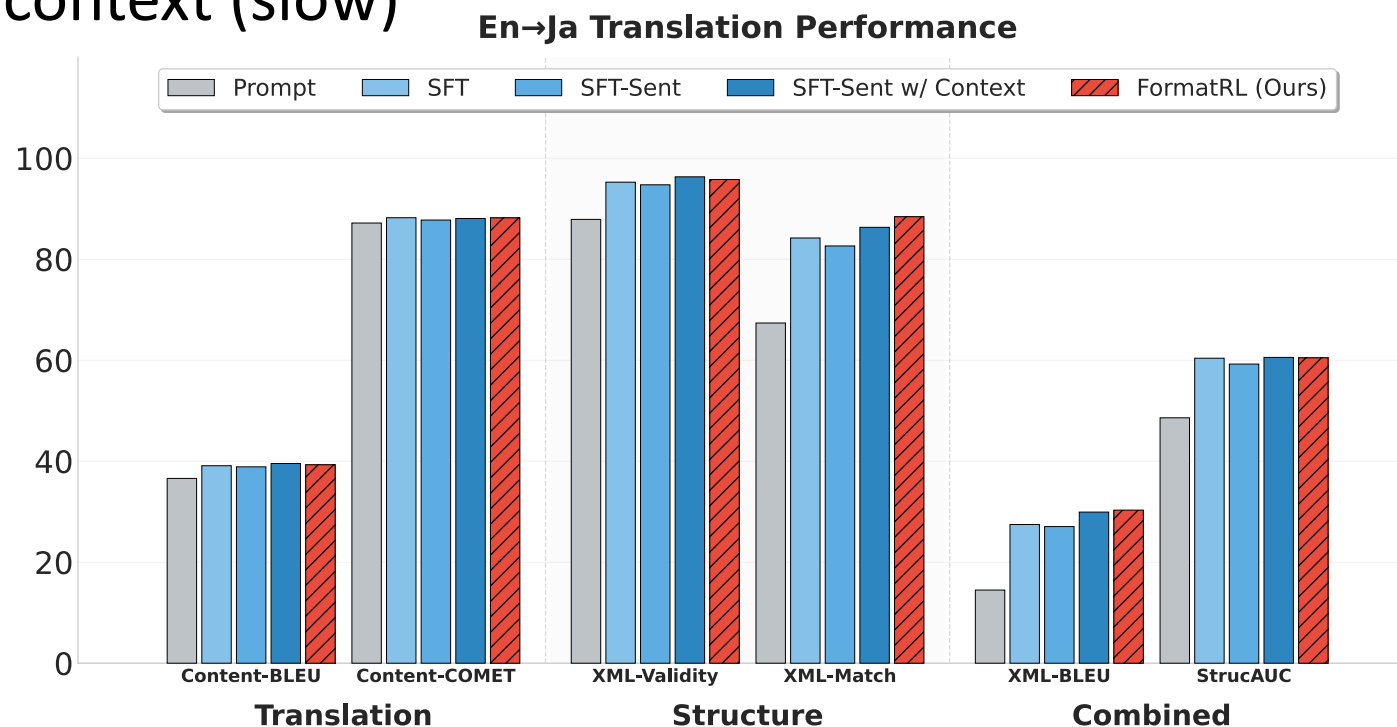
- **Structural accuracy improved.**
- **Combined metrics improved.** They take account both translation and structure accuracy

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

■ Baselines:

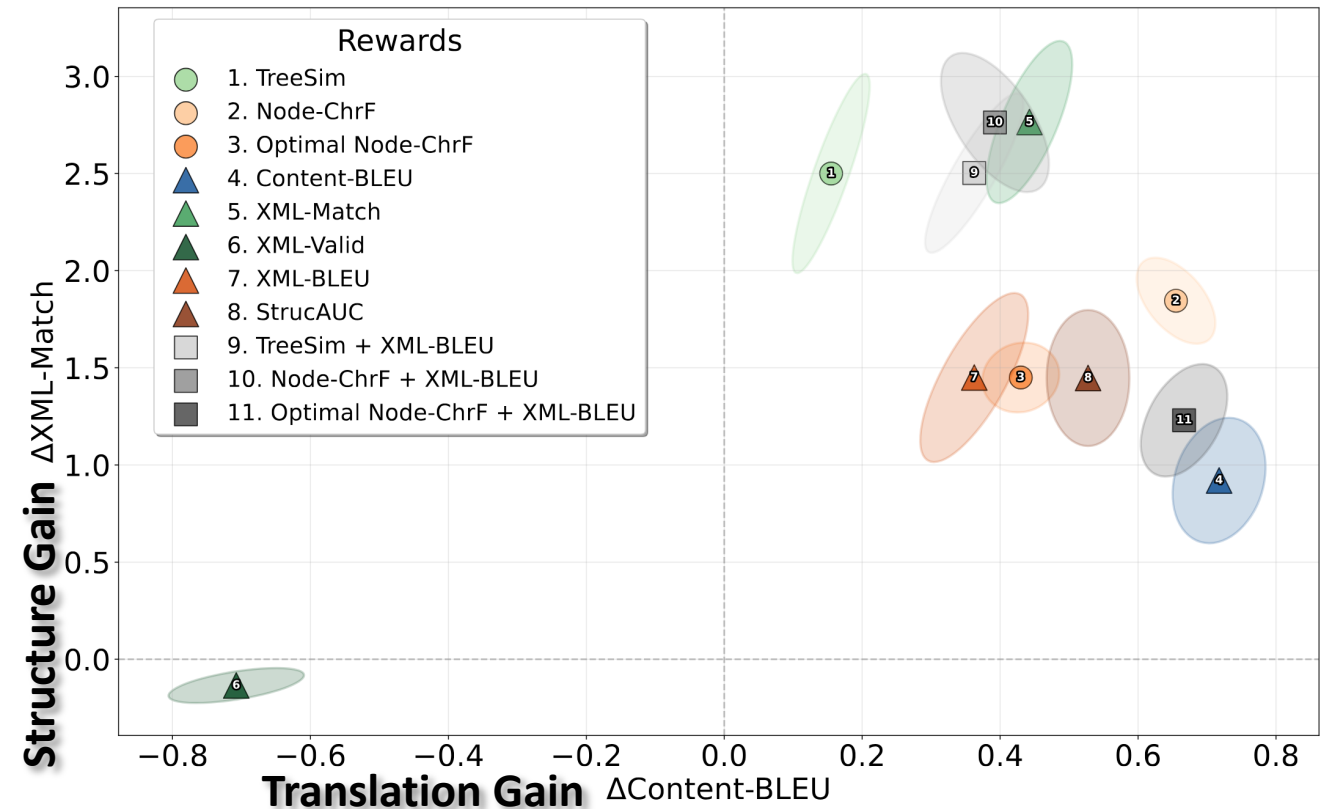
- SFT-Sent: parse-and-assemble, sentence-level translation w/o context
- SFT-Sent w/ Context: parse-and-assemble, sentence-level translation w/ Doc as context (slow)



Diverse Rewards

- Structural rewards (1. TreeSim) improve structural acc.
- Content translation rewards (4. BLEU) improve translation acc.

→ Achieve different goals by simply changing rewards!



Takeaways

- **Reinforcement learning** is a **powerful** tool
 - We can design rewards **for fine-grained needs**
 - It's simple & efficient
- Format reinforcement learning with TreeSim reward can boost structural accuracy in structured document translation