The Patent Machine Translation Task
— Summary of NTCIR-9 and Plans for NTCIR-10 —

Isao Goto (NICT)

Bin Lu (City Univ. of Hong Kong / Hong Kong Institute of Education)

Ka Po Chow (Hong Kong Institute of Education)

Eiichiro Sumita (NICT)

Benjamin K. Tsou (Hong Kong Institute of Education / City Univ. of Hong Kong)

Summary of NTCIR-9

Motivation

- There is a significant **practical need** for patent translation.
  - to understand patent information written in foreign languages.
  - to apply for patents in foreign countries.
- Patents constitute one of the **challenging domains**.
  - Patent sentences can be quite **long** and contain **complex structures**.
  - Translation between **languages with largely different word order** is difficult for **long** sentences.

Goals of PatentMT

- To develop **challenging** and **significant practical** research into patent machine translation.

- To **investigate** the **performance** of state-of-the-art machine translation systems in terms of patent translations involving Japanese, English, and Chinese.

- To **compare** the effects of **different methods** of patent translation by applying them to the same test data.

- To **create** publicly-available **parallel corpora of patent documents** and human evaluations of MT results for patent information processing research.

- To **drive machine translation research**, which is an important technology for cross-lingual access of information written in unknown languages.

- The ultimate goal is **fostering scientific cooperation**.

Findings of Previous Patent Translation Tasks

NTCIR-7	Human evaluation	**RBMT** was **better** than **SMT** for **JE** and **EJ**.
	CLIR evaluation	SMT was better than RBMT for EJ. ■ The translations were used as bag-of-words. ■ This means that **word selection** by SMT was better than that by RBMT.
NTCIR-8	Automatic evaluation	A hybrid system (RBMT with statistical post edit) achieved the best score for JE.

Comparison of NTCIR-7, 8, and 9

	NTCIR-7	NTCIR-8	**NTCIR-9** New
Language	**Japanese to English English to Japanese**	**Japanese to English English to Japanese**	**Chinese to English Japanese to English English to Japanese**
Human evaluation	**Adequacy Fluency**	No human evaluation	**Adequacy Acceptability** New
Extrinsic evaluation	CLIR	CLIR	No extrinsic evaluation
Number of participants	15	8	**21**

At NTCIR-9, participants can choose subtasks from three language directions, including **Chinese to English**.

Notable Findings at NTCIR-9

- **SMT** was the **best** system for **Chinese to English** and **English to Japanese** patent translation.
    - This is the **first time** for **SMT** to be **demonstrated equal or better** quality than that of the top-level RBMT for **English to Japanese** patent translation.
    - The **pre-ordering** method of NTT-UT for SMT is very effective for English to Japanese patent translation.

- **80%** of patent sentences could be understood in the best system for **Chinese to English** patent translation.

- **RBMT** was the best system for **Japanese to English** patent translation.

Remaining Issues of NTCIR-9

- **Practical evaluation**
  - The quality of translated sentences was evaluated at NTCIR-9.
  - More practical evaluations are also expected.

Plans for NTCIR-10

Outline of the Plans for NTCIR-10

- Three subtasks:

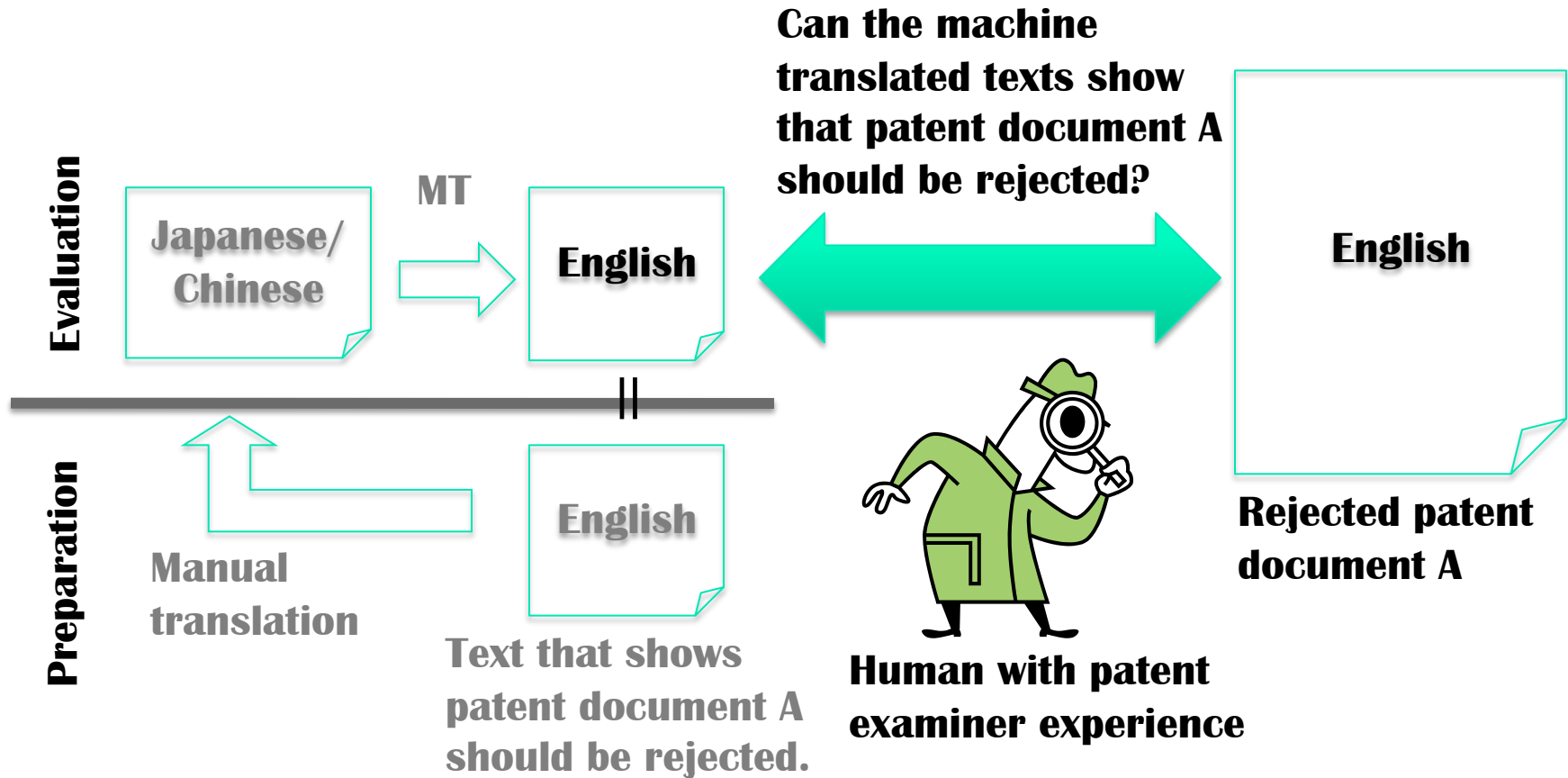| Subtasks	Training data
Chinese to English	**1 million** sentence pairs
Japanese to English	Approximately **3.2 million** sentence pairs
English to Japanese	

(Subtasks and training data are the same as at NTCIR-9)

- Participants select subtasks in which they wish to participate.

- Large scale **parallel corpora** and **new test sets** will be **provided**.

- **Practical evaluation** will be added (under consideration).

- **Human evaluation** will be carried out.

Differences from NTCIR-9

Practical Evaluation (under consideration)	**New**: To explore **practical** MT performance in appropriate fields for patent machine translation.	
Intrinsic Evaluation	**Similar to the NTCIR-9 evaluation**. Quality of translated sentences will be evaluated. Additions:	
	Chronological evaluation	Comparison between NTCIR-10 and NTCIR-9 to **measure progress**.
	Multilingual evaluation	Comparison of CE and JE translations using the **same English reference** will be added.

Possible Approach to Practical Evaluation

Can the machine translated texts show that patent document A should be rejected?

Evaluation

Japanese/ Chinese → MT → English

English

Rejected patent document A

Preparation

Manual translation

English

Text that shows patent document A should be rejected.

Human with patent examiner experience

(The feasibility of this is under investigation.
We are working hard to make necessary arrangements.)

Why is it so exciting to participate in?

- **Patents** are one of the **challenging domains** for MT.
  - Patent sentences could be quite **long** and contain **complex structures.**
  - Translation between **languages with largely different word order** is difficult for **long** sentences.
- Participants will receive **evaluation results** for their MT quality.
- Participants can use **large-scale** patent **parallel** and **monolingual corpora**.
- Participants can choose subtasks from three language directions, including **the language direction of Chinese to English**.

- We look forward to many groups participating in PatentMT at NTCIR-10!