

nukl's QA System at the NTCIR-16 QA Lab-PoliInfo-3

Yasuhiro Ogawa, Yugo Kato, Katsuhiko Toyama

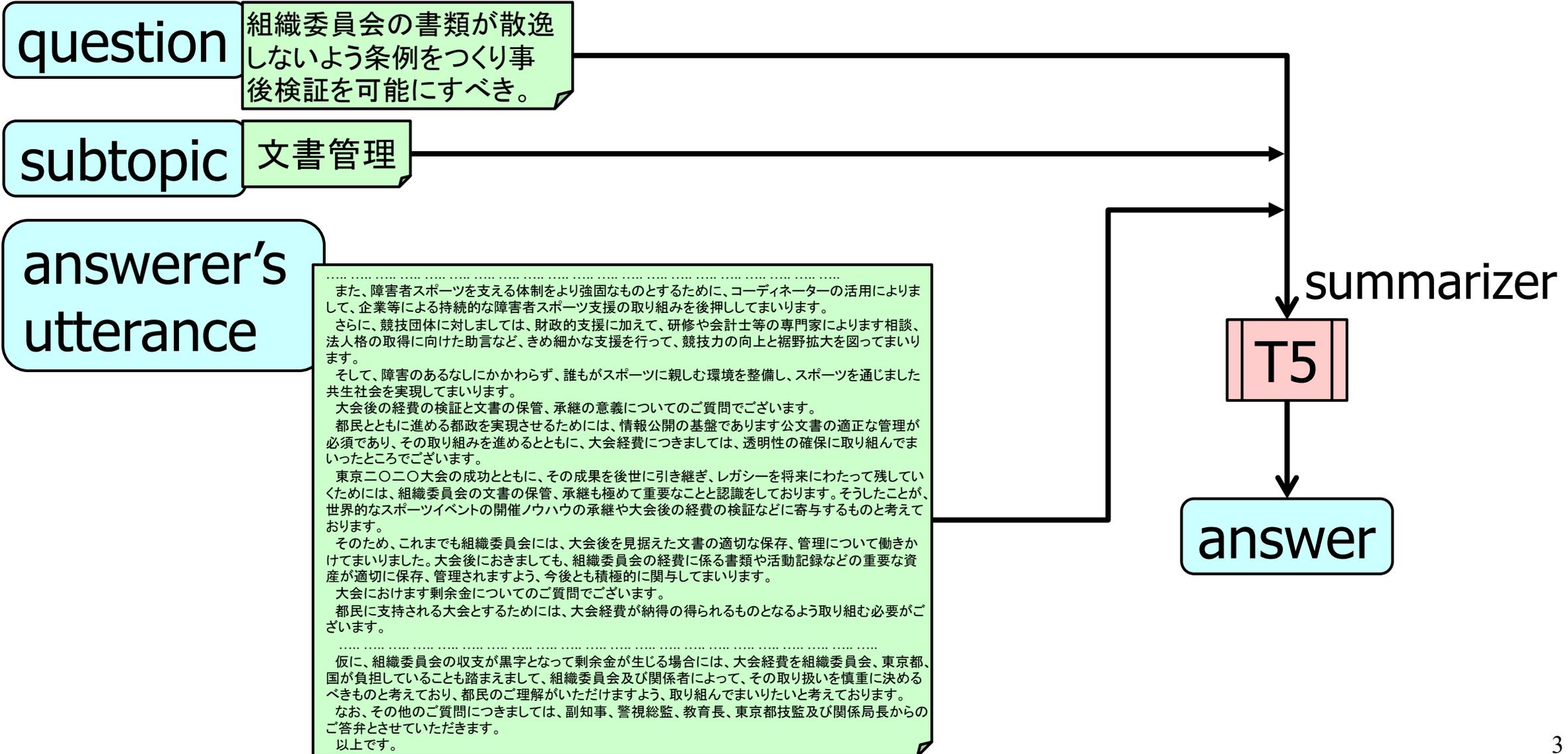
Nagoya University, JAPAN

Outline

Question Answering subtask

- Using T5 as a summarizer 
- Two input types
 - Answerer's entire utterance
 - Answer text corresponding to the input question
- **Achieved the highest score**

Method 1: Input the Entire Utterance



Method 1: Input the Entire Utterance

question 組織委員会の書類が散逸しないよう条例をつくり事後検証を可能にすべき。

subtopic 文書管理

answerer's utterance

too long

また、障害者スポーツを支える体制をより強固なものとするために、コーディネーターの活用によりまして、企業等による持続的な障害者スポーツ支援の取り組みを後押ししてまいります。

さらに、競技団体に対しましては、財政的支援に加えて、研修や会計士等の専門家によります相談、法人格の取得に向けた助言など、きめ細かな支援を行って、競技力の向上と裾野拡大を図ってまいります。

そして、障害のあるなしにかかわらず、誰もがスポーツに親しむ環境を整備し、スポーツを通じた共生社会を実現してまいります。

大会後の経費の検証と文書の保管、承継の意義についてのご質問でございます。

都民とともに進める都政を実現させるためには、情報公開の基盤であります公文書の適正な管理が必須であり、その取り組みを進めるとともに、大会経費につきましては、透明性の確保に取り組んでまいったところでございます。

東京二〇二〇大会の成功とともに、その成果を後世に引き継ぎ、レガシーを将来にわたって残していくためには、組織委員会の文書の保管、承継も極めて重要なことと認識しております。そうしたことが、世界的なスポーツイベントの開催ノウハウの承継や大会後の経費の検証などに寄与するものと考えております。

そのため、これまでも組織委員会には、大会後を見据えた文書の適切な保存、管理について働きかけてまいりました。大会後におきましても、組織委員会の経費に係る書類や活動記録などの重要な資産が適切に保存、管理されますよう、今後とも積極的に関与してまいります。

大会におけます剰余金についてのご質問でございます。

都民に支持される大会とするためには、大会経費が納得の得られるものとなるよう取り組む必要がございます。

仮に、組織委員会の収支が黒字となって剰余金が生じる場合には、大会経費を組織委員会、東京都、国が負担していることも踏まえまして、組織委員会及び関係者によって、その取り扱いを慎重に決めるべきものと考えており、都民のご理解がいただけますよう、取り組んでまいりたいと考えております。

なお、その他のご質問につきましては、副知事、警視総監、教育長、東京都技監及び関係局長からのご答弁とさせていただきます。

以上です。

last sentences <1024 tokens

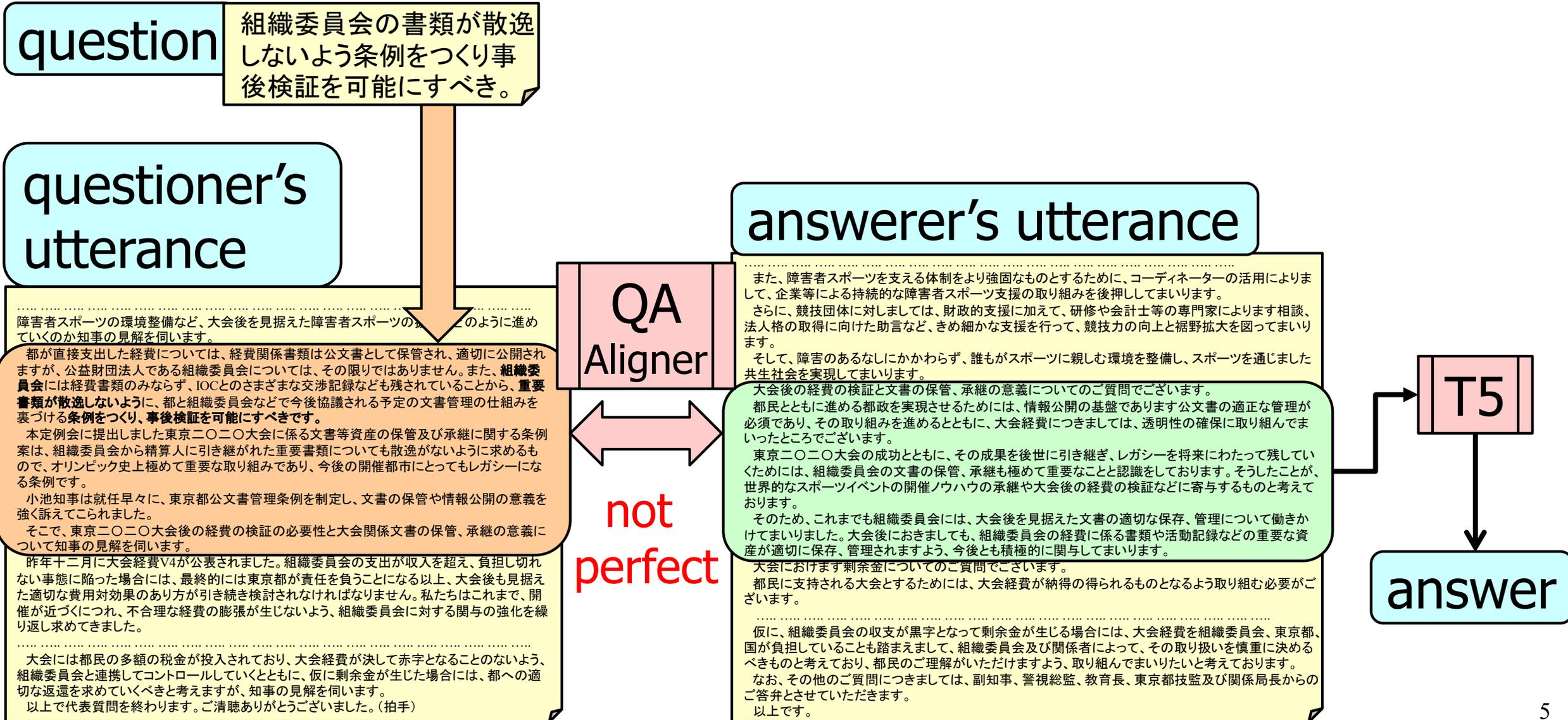
summarizer

T5

answer

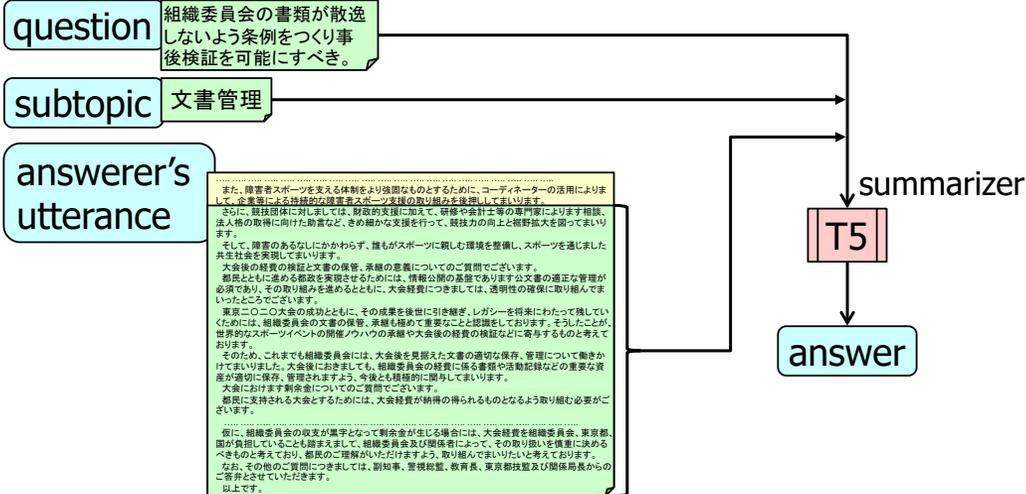
target part may be missing

Method 2: Input Aligned Text

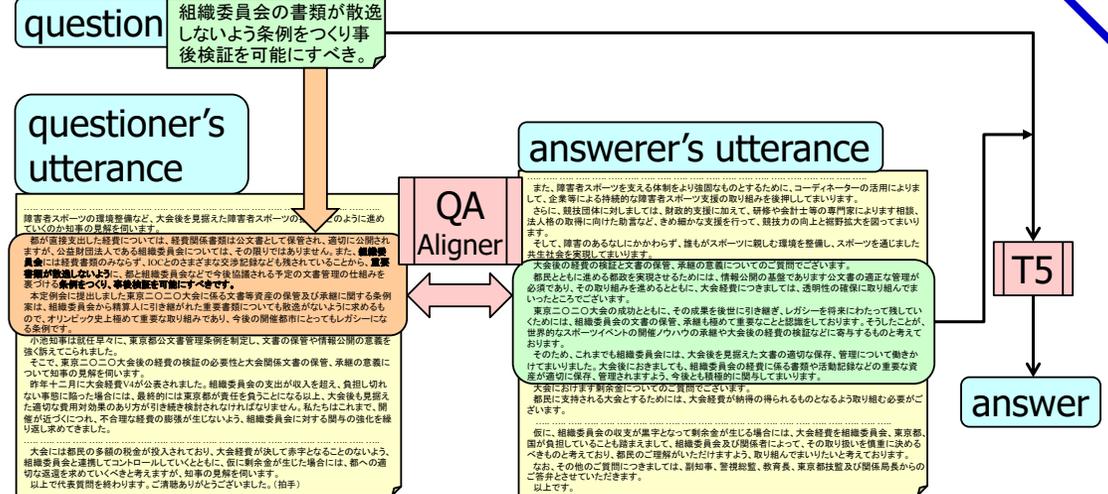


Proposed Method: Mixed

Method 1



Method 2



Select the answer by the utterance's length (threshold θ)

answer

Experimental Setting

Pre-trained T5 model	https://huggingface.co/sonoisa/t5-base-japanese
Tokenizer	SentencePiece
Maximum input length	1,024 tokens
Maximum output length	64 tokens
Number of epochs	4
Batch size	2
Threshold θ	1,000, 2,000, 2,500 characters
Training data (Method 1)	7,627 tuples (2001 – 2019)
Training data (Method 2)	2,171 tuples (2011 – 2016)
QA Aligner	AKBL team's system (ID 235)

Thanks AKBL team

Experimental Results (Automatic Evaluation)

ID	System	ROUGE-1-F
310	Proposed ($\theta = 2,000$)	0.3132
313	Proposed ($\theta = 2,500$)	0.3129
311	Proposed ($\theta = 1,000$)	0.3051
266	Method 1	0.2823
316	Method 2	0.2787
288	ditlab	0.3031
190	AKBL	0.2306
166	TO	0.0767

Experimental Results (Human Evaluation)

ID	Team	Correspondence				Content				Well-formed				Overall			
		A	B	C	Score	A	B	C	Score	A	B	C	Score	A	B	C	Score
Gold		377	20	3	774	208	170	22	586	391	8	1	790	217	164	19	598
310	nukl	363	25	12	751	138	211	51	487	381	19	0	<u>781</u>	148	203	49	499
288	ditlab	348	33	19	<u>729</u>	138	200	62	<u>476</u>	379	17	4	775	142	200	58	<u>484</u>
269	ditlab	346	31	23	723	129	209	62	467	384	16	0	784	136	207	57	479
190	AKBL	320	42	38	682	104	196	100	404	381	6	13	768	103	203	94	409
166	TO	83	77	240	243	4	58	338	66	99	33	268	231	4	36	360	44

Our system achieved the highest score in both automatic and human evaluations

Additional Experiment

- QA alignment with Gold Data (GD)

ID	System	ROUGE-1-F
	Proposed ($\theta = 2,000$) with GD	0.3333
310	Proposed ($\theta = 2,000$)	0.3132
	Method 2 with GD	0.3049
266	Method 1	0.2823
316	Method 2	0.2787

Better QA alignment produces better answers

Conclusion

Question Answering subtask

- Using T5 as a summarizer
- Two input types (entire or corresponding part)
- Achieved the highest score

Future Works

- Increasing training data for Method 2

More discussions at the poster session