NiCT/ATR in NTCIR-7 CCLQA Track

Youzheng WU, Wenliang CHEN, Hideki KASHIOKA

NiCT/ATR, Japan

NTCIR CCLQA

- Complex Cross-Lingual Question Answering.

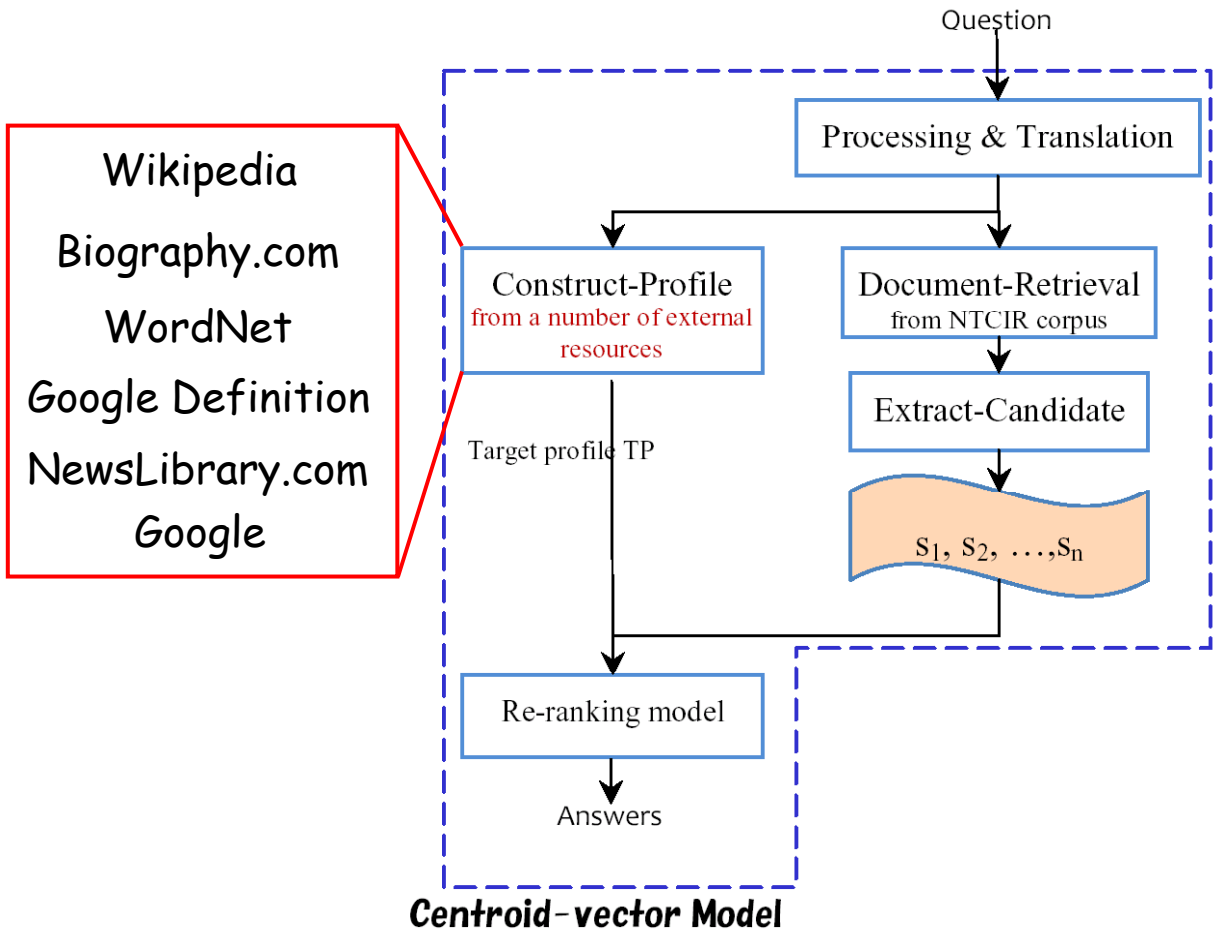
List/Event questions
List major events in formation of European Union.
Relationship questions
Does Iraq possess uranium, and if so, where did it come from?
Biography questions:
Who is Howard Dean?
Definition questions:
What are stem cells?

Questions in **English** and getting answers from Chinese (**Simplified**, Traditional) or Japanese corpus

Related studies

- Pattern-matching-based [Xu, et al. 2005] [Harabagiu, et al. 2004] [Cui, et al. 2004]
  - Basic syntactic/semantic structures like appositives, copulas; predicates and relations.
- Centroid-vector-based [Xu, et al. 2003] [Chen, et al. 2006] [Kor, et al. 2007]
  - Build a target profile for each question, and then compute the similarities between candidates and the target profile.
- Others [Biadsy, et al. 2008]
  - Unsupervised classification model to Biography production using Wikipedia.

Centroid-vector-based



Centroid-vector-based cont.

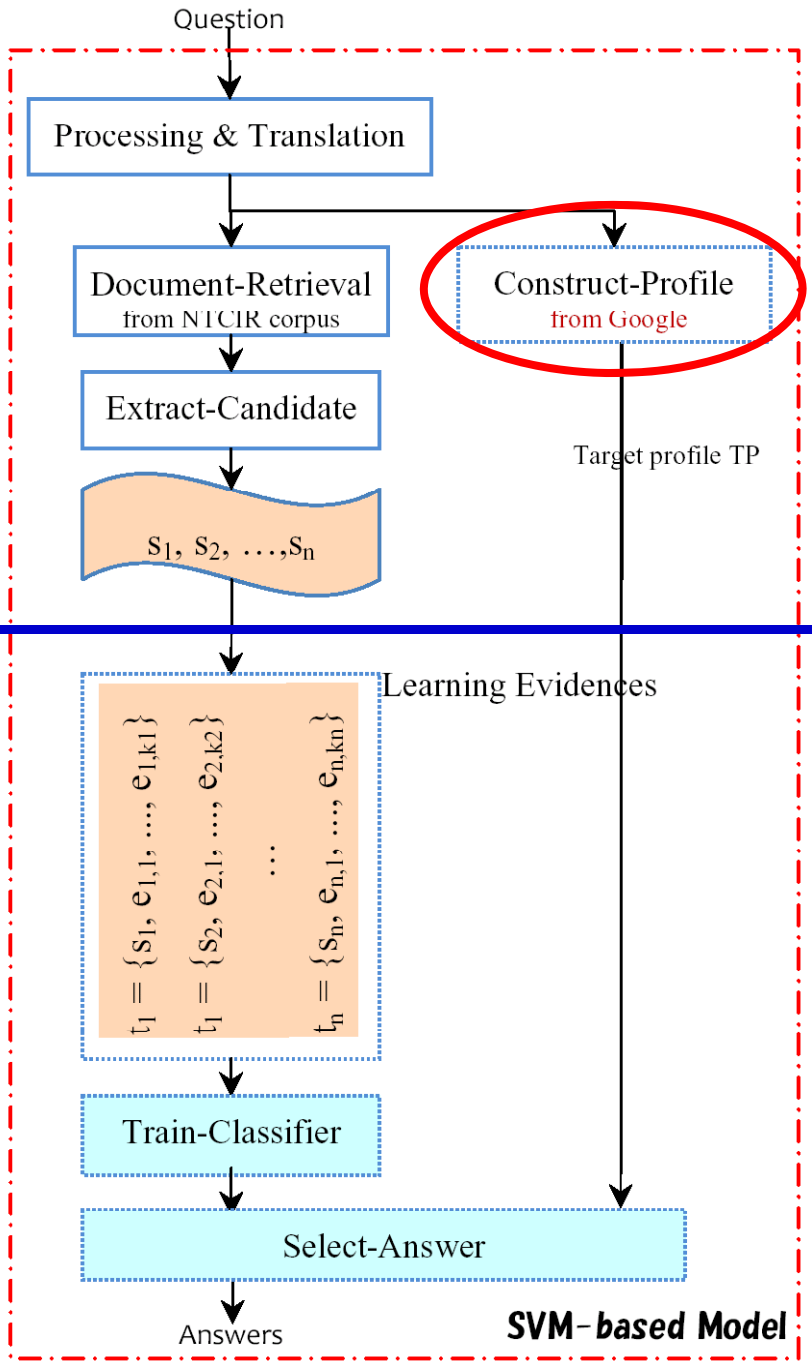
- Easy to implement, and fast in speed
- In essence, a type of **question-side expansion.**
  - Hard to obtain sometimes
    - Wikipedia, WordNet, and Biography.com contain only 82.0%, 40.4%, and 24.6% of TREC05 questions, respectively.
  - Not always contribute positively
    - Wikipedia negatively impacts Biography questions [Kor, et al. 2007].

Our solution

SVM-based model	Centroid-vector-based model
Regarding complex QA as a SVM-based classification	as a retrieval process
Applying **sentence-side expansion**	question-side expansion
Requiring no specific resources, except a general search engine (Google)	A number of external resources such as Wikipedia, Biography.com, WordNet, etc.
Incorporating multiple features	TF-IDF-similarity score

SVM-based

Same to centroid-vector model



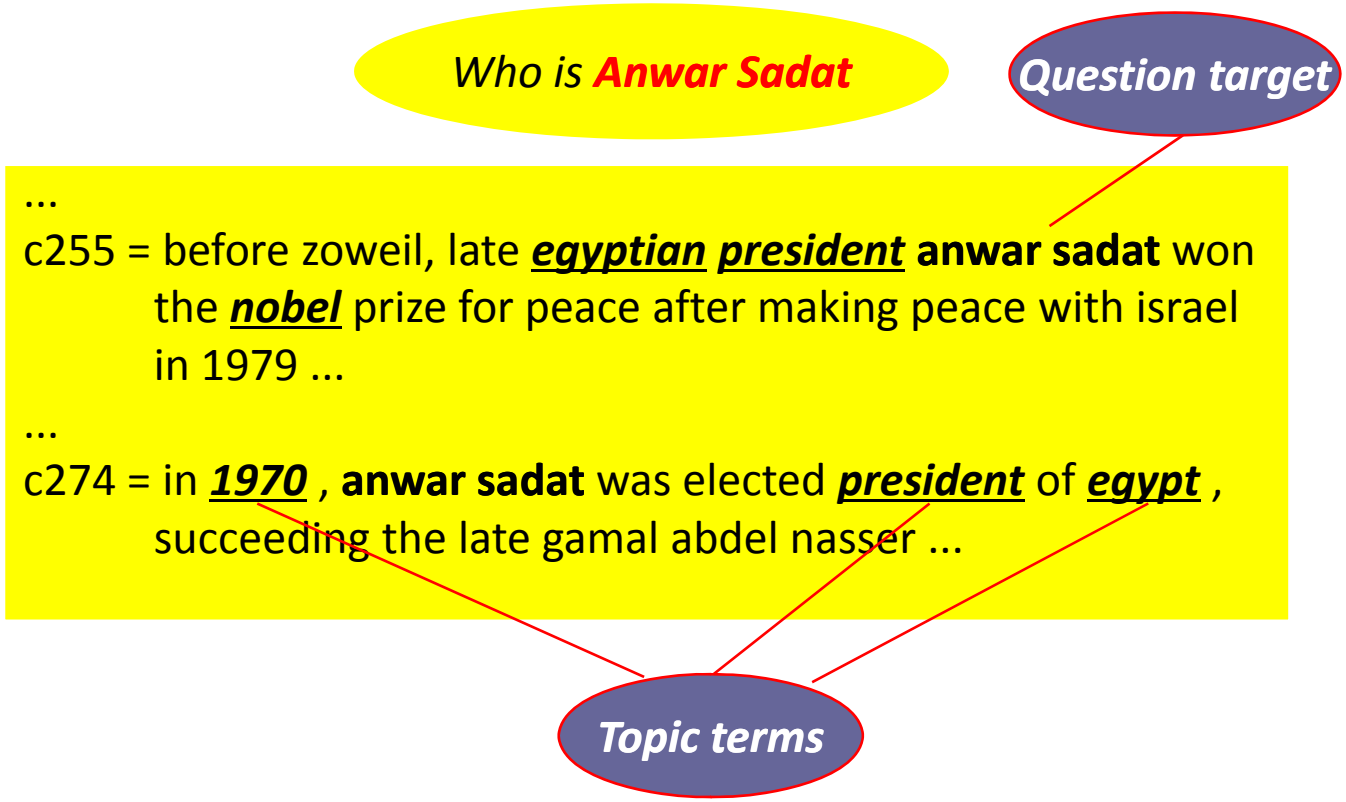
Learning Evidences by Sentence-side Expansion

For each s_i in S

1. Extract **2 or 3** nouns nearest to question target from candidate $\mathbf{s_i}$ as topic terms of $\mathbf{s_i}$, labeled as **R**.
2. Combine topic terms **R** and question target to compose a web query and submit it to Google.
3. Download the top 100 Google snippets.
4. Retain those snippets {$\mathbf{e_{i,i1}}$, …$\mathbf{e_{i,ik}}$} that contain words in question target and **R** as Web evidences for candidate S_i.

end

An Example


*Who is **Anwar Sadat***
Question target
...
c255 = before zoweil, late ***egyptian*** ***president*** **anwar sadat** won the ***nobel*** prize for peace after making peace with israel in 1979 ...
...
c274 = in ***1970*** , **anwar sadat** was elected ***president*** of ***egypt*** , succeeding the late gamal abdel nasser ...
Topic terms

An Example

Results **1** - **10** of about **15,300** for **"Anwar Sadat" 1970 president egypt**.

Google "Anwar Sadat" 1970 president egypt Search Advanced Search Preferences

Web

Anwar El Sadat - Wikipedia, the free encyclopedia
In 1983, Sadat, a miniseries based on the life of **Anwar Sadat**, ... served as acting **President** of **Egypt** for eight days following Sadat's assasination ...
en.wikipedia.org/wiki/Anwar_El_Sadat - 164k - Cached - Similar pages

CNN Cold War - Profile: **Anwar Sadat**
In 1964, Sadat became **Egypt's** vice **president**. When Nasser died suddenly of a heart attack in **1970**, Sadat succeeded him. As **president**, Sadat inherited a ...
www.cnn.com/SPECIALS/cold.war/kbank/profiles/sadat/ - 11k - Cached - Similar pages

Anwar as-Sadat
(1918-1981) **President** of **Egypt** from **1970** to 1981. Sadat's background was military education. He was active in the dissident officer's actions of freeing ...
lexicorient.com/e.o/sadat.htm - 25k - Cached - Similar pages

Anwar el-Sadat -- Britannica Online Encyclopedia
Egyptian army officer and politician who was **president** of **Egypt** from **1970** until his assassination in 1981. He initiated serious peace negotiations with ...
www.britannica.com/EBchecked/topic/515786/Anwar-el-Sadat - 56k - Cach

The Pharaonic Village - **Anwar Sadat** - Hero of War
President Nasser shaking hands with **Anwar Sadat** (then Vice **Presid**
death, elected **President** of **Egypt** October 15, **1970**. ...
www.tour**egypt**.net/Village/exhibits_**anwarsadat**.htm - 16k - Cached

Political Friendster - **Anwar Sadat** - **President Egypt** - Connections
After **Egyptian President** Gamal Abddul Nasser dies in October **1970**, he is succe d as **president** of **Egypt** by his former Vice **President**, **Anwar Sadat**. ...
www.politicalfriendster.com/showPerson.php?id=5293&name=**Anwar-Sadat---President-Egypt** - 22k - Cached - Similar pages

Anwar Sadat - MSN Encarta
Sadat, Anwar al- 1918-1981, **Egyptian** military leader and **president** **1970**-1981, ... **President** of **Egypt**, serving ... **Anwar Sadat** Biography. Anwar al-Sadat. ...
encarta.msn.com/encyclopedia_761561908/**anwar_sadat**.html - 32k - Cached - Similar pages

Anwar Sadat - Mahalo
Political and military leader **Anwar Sadat** was **president** of **Egypt** from **1970**-1981. He

Results **1** - **10** of about **42,400** for **"Anwar Sadat" egyptian president Nobel**.

Google "Anwar Sadat" egyptian president Nobel Search Advanced Search Preferences

Web Video

Tip: Save time by hitting the return key instead of clicking on "search"

Peace 1978
Nobel Prize® medal - registered trademark of the **Nobel** Foundation ... **President** of the Arab Republic of **Egypt**, Prime Minister of Israel ...
nobelprize.org/**nobel**_prizes/peace/laureates/1978/ - 16k - Cached - Similar pages

Anwar al-Sadat - Curriculum Vitae
Vice-**president** of **Egypt**, 1964-66, 1969-70. **President** of **Egypt** ... Man of Defiance: A Political Biography of **Anwar Sadat**. Totowa, N.J.: Barnes & **Nobel**, 1985. ...
nobelprize .g/**nobel**_prizes/peace/laureates/1978/al-sadat-cv.html - 18k - Cached ilar pages

gyptia resident, 1970-81, **Nobel** 1978 December 25 ...
. Born: **Anwar Sadat**, **Egyptian President**, 1970-81, **Nobel** 1978, tesman Related Year: ...
ecember_25_1918_389286.html - 7k -

dia, the free encyclopedia
ty was signed by **Anwar Sadat** and Israelli Prime ... Both re awarded the **bel** Peace Prize for creating the ...
edia.org/Anwar_El_Sadat - 164k - Cached - Similar pages

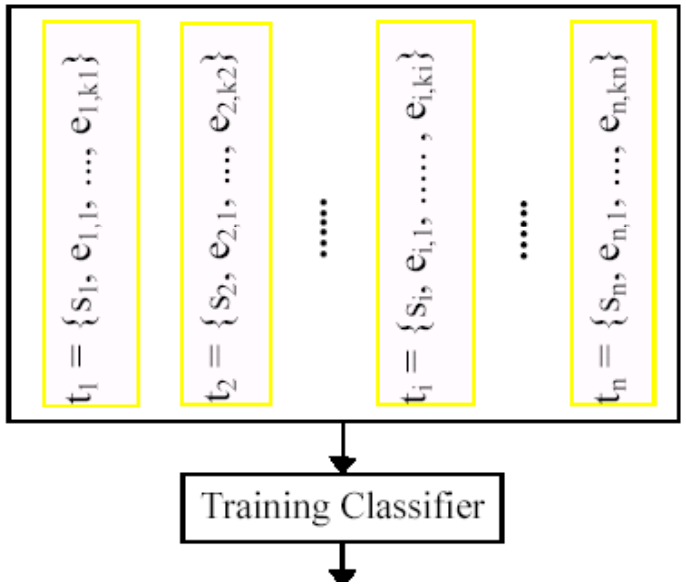
Anwar Sadat - Mahalo
Political and military leader **Anwar Sadat** was **president** of **Egypt** from 1970-1981. He negotiated a peace treaty with Israel for which he shared the 1978 **Nobel** ...
www.mahalo.com/**Anwar_Sadat** - 39k - Cached - Similar pages

YouTube - **Egyptian** Presedint **Anwar Sadat** speech ...
Nobel price is a man made award, given by groupe of ppl ,who ...
6 min - ★★★★★
www.youtube.com/watch?v=fNOxfP85TZA

CNN Cold War - Profile: **Anwar Sadat**
In 1964, Sadat became **Egypt's** vice **president**. When Nasser died suddenly of a ... Prime Minister Menachem Begin, jointly received the **Nobel** Peace Prize. ...

Bridge lexical gap between candidates and profile

SVM-based model



Train-Classifier

Features	Description
Bfull	If question target occurs in the exact form.
Bbegin	If question target occurs at the beginning of an instance or not
Bpattern	If one of predefined patterns occurs or not
Btime	If time expression occurs or not. Candidates with time expression tend to capture important events involving target
Unigram-overlap	Overlap of unigrams between an instance and target profile
Bigram-overlap	Overlap of Bigrams between an instance and target profile
TF-IDF similarity	TF-IDF-based similarity between an instance and target profile
Freq	The number of relevant pages returned by Google

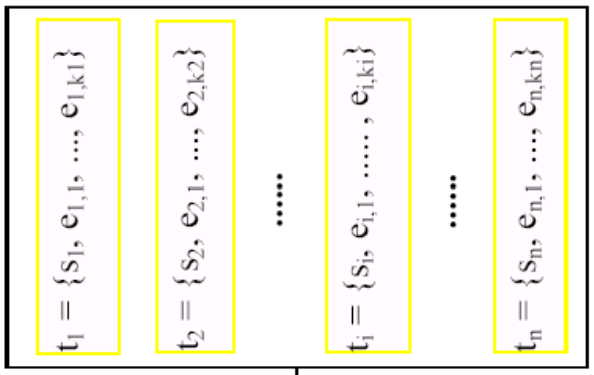
SVM Classifier

Rules for the Train-Classifier

- Manually generate from the abstract of Wikipedia.
- Useful to Biography and Definition questions.

Regular expression of patterns
DEF (*) [是|为] [一 q|一个]
DEF (*) [又 称|或 称|或 称为|或 简称|或 译|简称|旧 称|全 称为|全 名|学 名|也 称|也 称为|也 称作|也 叫|亦 称|亦 称为|亦 作|英文 名称|又 称为|又 称作|又 叫|又 名|又 译|原 称|原 为|原名|正式 名称|全 名 为|又 名|又 被 称为]
DEF (*) [是 ns 一个|位于 ns|是 位于|一个 位于|在 ns|坐落 于]
DEF (*) [是 指|指 的 是|是 由]
DEF (*) [发生|成立] [于|在] DEF (*) [是 ns|女|字]
DEF (*) [ns *家|现任 ns|本 名|生 卒]
DEF (*) [籍贯|出生于|生于|出生于|生于|祖籍] ns
DEF (*) ns [出生|人]
Legend: () is optional field; [] is obligation field; * denotes that any word can appear; | separates the elements within the obligation field, *q* and *ns* represent the part-of-speeches of quantifier and location name, respectively.

Select-Answer



Ideal Answer

assuming that all the features fire

Training Classifier

Selecting Definition

Answers

Which topic is the nearest topic to ideal answer

descending order

$<t_{p1}, p_1>$
$<t_{p2}, p_2>$
...
$<t_{pj}, p_j>$
...
$<t_{pn}, p_n>$

j=15 if the t_{pj} > average probability (1/n)
j=10 else;

Experiments

- Three runs for the EN-CS and CS-CS tasks
  - RUN-3: The Centroid-vector model (**5 external resources**).
    - Wikipedia (0.2 million Chinese entries);
    - Baidu Baike (1 million Chinese entries, http://baike.baidu.com);
    - Google Definition (e.g., define: Nobel prize);
    - Google news (1000 news sources updated continuously);
    - Google.
  - **RUN-1: The SVM-based model (Google).**
  - **RUN-2: The SVM-based model (5 external resources).**
    - **To compare with the RUN-1**

Official Result

$$\text{precision} = \begin{cases} 1 - \dfrac{L - C \times a}{L} & \text{if } L < C \times a \\ 1 & \text{else} \end{cases}$$

- Three findings: 1. the precision is low; 2. the ranking of difficulties of answering questions; 3. Cross-lingual vs. monolingual (only 10% include at most 1 error).

Table 3. The official scores of the RUN-1

	Recall (%)	Precision (%)	F-score (%)
the EN-CS task			
Event	20.70	4.8	14.54
Definition	50.13	4.2	22.16
Biography	67.04	7.2	31.58
Relation	44.78	5.6	23.35
all	43.08	5.4	**22.11**
the CS-CS task			
Event	20.87	4.3	14.30
Definition	55.38	4.9	24.15
Biography	69.40	7.9	33.76
Relation	48.15	5.6	24.29
all	45.66	5.5	**23.16**

The best (the second is 19.30%)

The second (the best is 43.29%)

Comparison of the Three Runs

EN-CS task			
	RUN-1	RUN-2	RUN-3
Event	14.54	14.08	8.09
Definition	22.16	23.37	12.57
Biography	31.58	30.27	20.77
Relation	23.35	22.80	12.10
all	22.11	21.79	12.73
CS-CS task			
Event	14.30	14.07	10.86
Definition	24.15	25.65	16.18
Biography	33.76	32.53	18.06
Relation	24.29	23.76	16.50
all	23.16	22.98	15.06

The conclusion:

- The proposed SVM-based models are much better than the baseline by comparing the RUN-1 and RUN-2 with RUN-3.
- Target profile does not play an important role in the proposed SVM-based model by comparing the RUN-1 with RUN-2.

Automatic Scores

Table 5. The F-scores (%) using automatic evaluation

	RUN-1	RUN-2	RUN-3
the EN-CS task			
Event	19.52	19.12	18.01
Definition	20.95	22.99	19.14
Biography	35.81	36.40	34.32
Relation	32.18	32.07	28.19
all	26.86	27.24	24.55
the CS-CS task			
Event	17.89	18.49	16.42
Definition	24.24	26.14	19.86
Biography	37.22	36.86	32.75
Relation	34.00	33.84	31.10
all	27.85	28.30	24.78

The best (the second is 22.90%)

The second (the best is 37.75%)

IR4QA + CCLQA

- Three retrieval results are CMUJAV1-EN-CS-**01**-T-limit50, CMUJAV1-EN-CS-**02**-T-limit50, and MITEL-EN-CS-**01**-T-limit50.

Table 6. The IR4QA results of the CJAV1, CJAV2, and MITEL

	Mean AP	Mean Q	Mean nDCG
MITEL	36.82	35.18	52.06
CJAV1	34.52 ∨	32.58	47.89
CJAV2	33.87 ∨	31.99	47.25

Table 7. The F-scores of our runs based on the CJAV1, CJAV2, and MITEL

	CJAV1	CJAV2	MITEL
RUN-1	27.63	27.30	27.50
RUN-3	24.91	24.44	24.47

IR4QA + CCLQA

Table 8. The Mean-AP scores of the CJAV1 and the MITEL over types of questions

	CJAV1		MITEL
Event	19.53	<	26.57
Definition	48.65	>	37.10
Biography	46.65	>	45.15
Relation	32.00	<	41.37

Table 9. The F-scores of the RUN-1 based on the CJAV1 and MITEL over types of questions

	CJAV1		MITEL
Event	17.74	<	18.92
Definition	23.08	>	17.23
Biography	38.37	>	37.87
Relation	33.40	<	36.01

The conclusion:

- The impacts of the IR4QA system on the CCLQA system are roughly consistent.
- However, the extent of the impacts are not the same.

Discussion

- Hard to directly evaluate the quality of web evidences learned by sentence-side expansion
  - Has the same underlying logic as that of the ROUGE metric [Lin, et al. 2003] and the nugget-pyramid metric [Lin, et al. 2006]: **use unigram overlap to match semantically**
- Speed problem
  - Have to train an SVM-classifier for each question.

Summary

- Propose an SVM-based classification model for NTCIR complex QA system:
  - Each candidate represents a topic
  - Learn training data for each topic by sentence-side expansion
  - Assume an ideal answer, and classify this ideal answer into topics to find real answers
- The SVM-based model achieves competitive performances, and relies on no specific external resources other than Google.

Thanks!