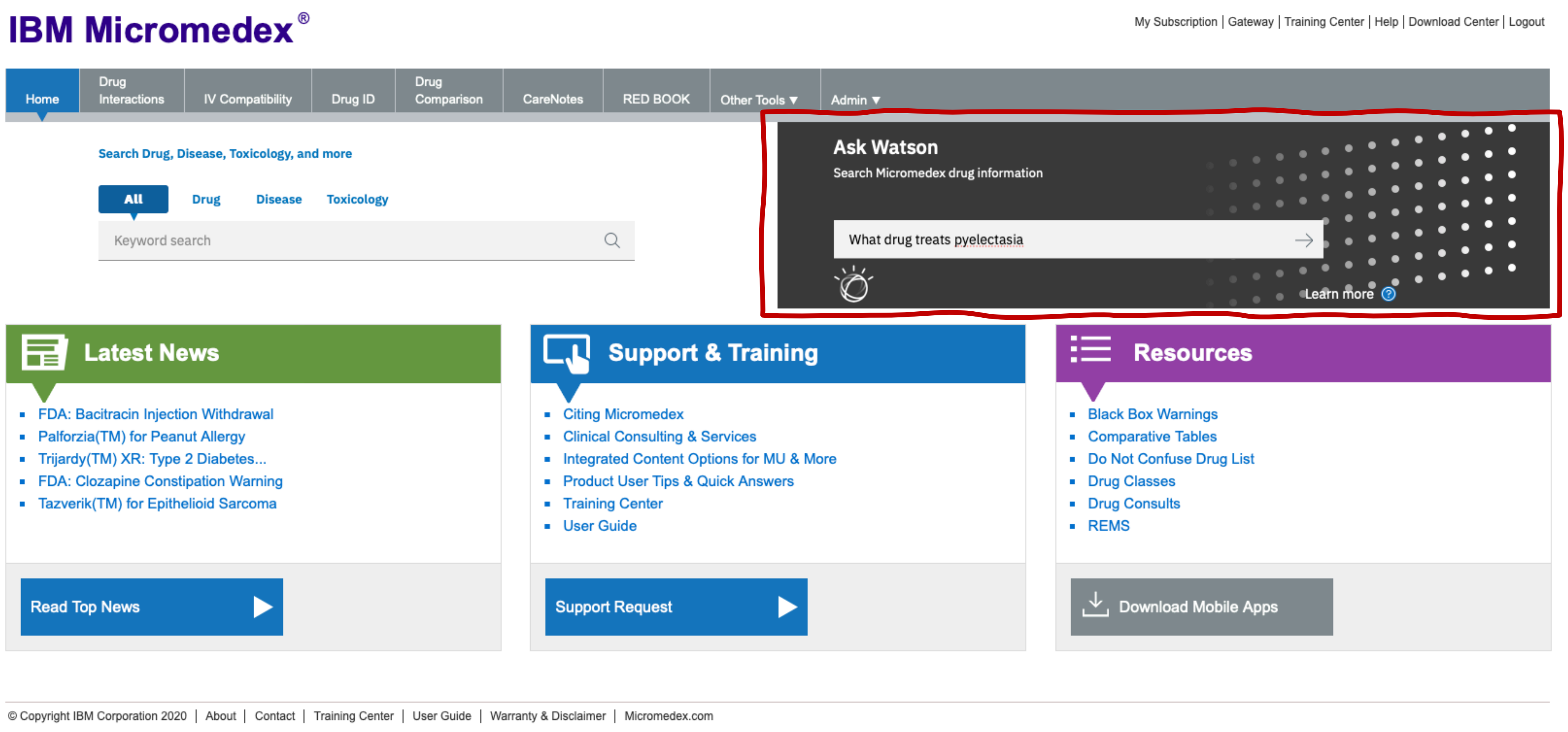
IBM Research

Expanding Query Answers on Medical Knowledge Bases

Chuan Lei
Vasilis Efthymiou
Rebecca Geis
Fatma Özcan

Querying medical knowledge bases



Query relaxation

<u>Problem:</u>
Users do not always formulate their queries precisely to match the terms in the KB

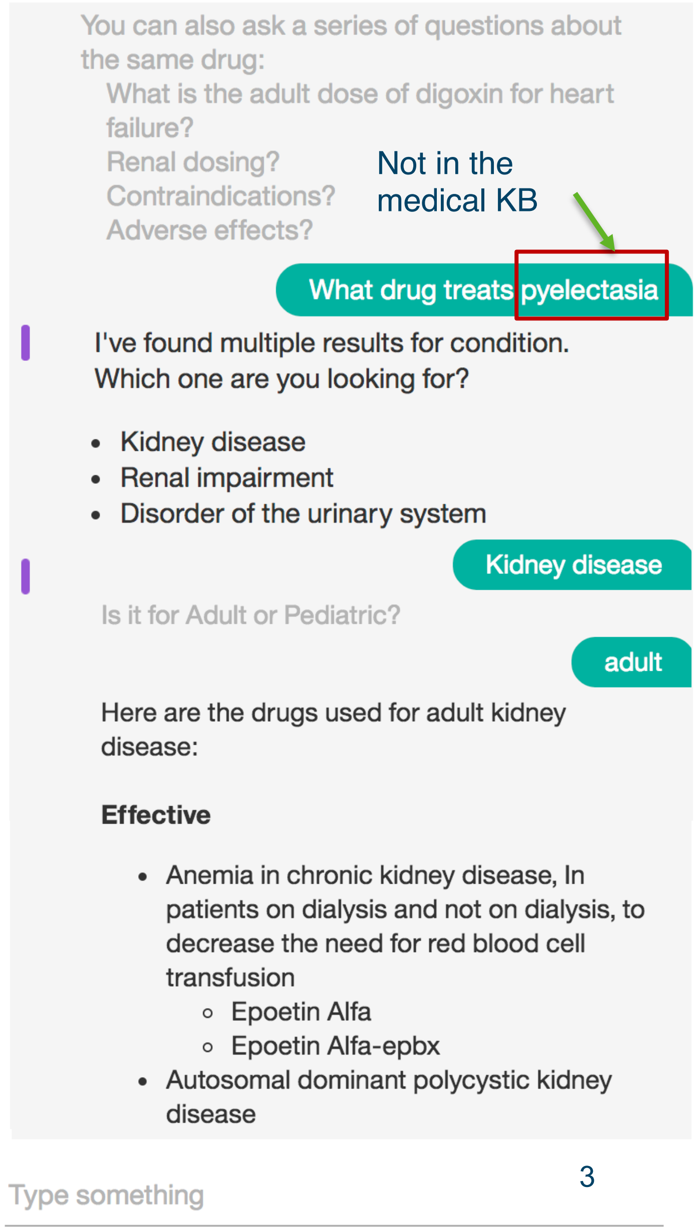
- ➢ No answer or incomplete answers returned

<u>Goal:</u>
Query relaxation (QR) transforms the query in a way that the user's intent is better represented

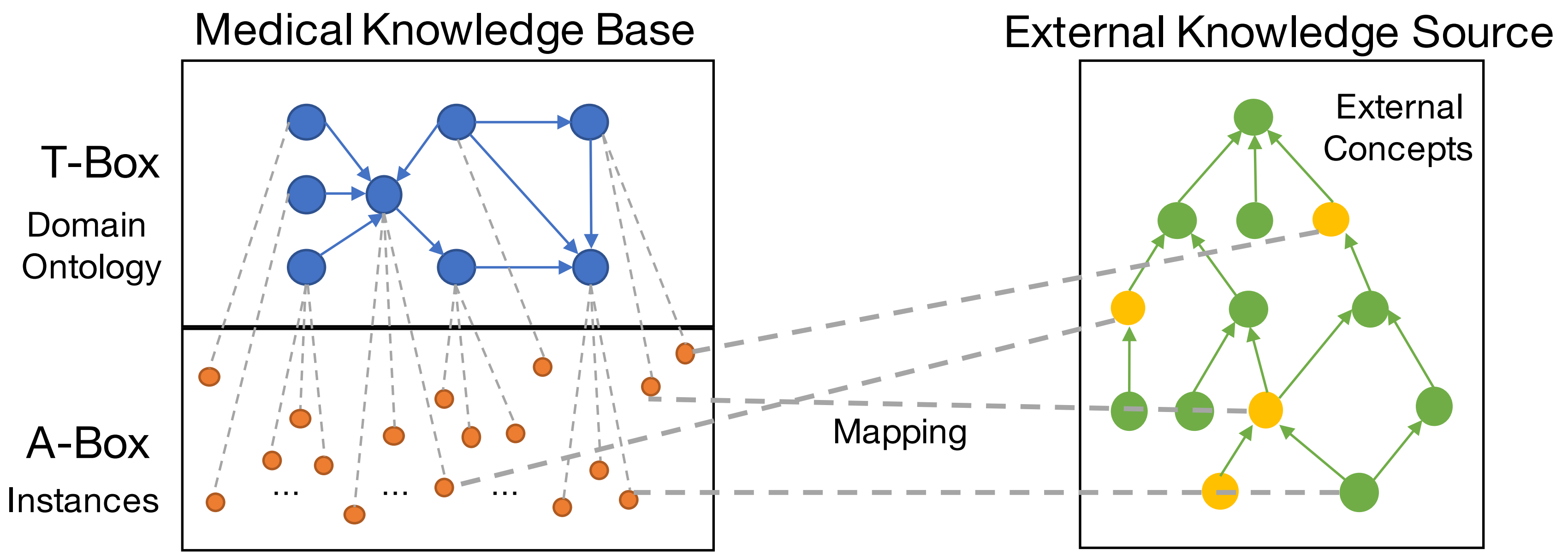
- ➢ greatly improving the flexibility and usability of a medical KB

<u>Contributions</u>:

- an effective offline external knowledge source incorporation
- a novel similarity metric to identify semantically related concepts
- a programmatic way to incorporate our QR into existing systems
- experimental evaluation shows our QR outperforms existing methods



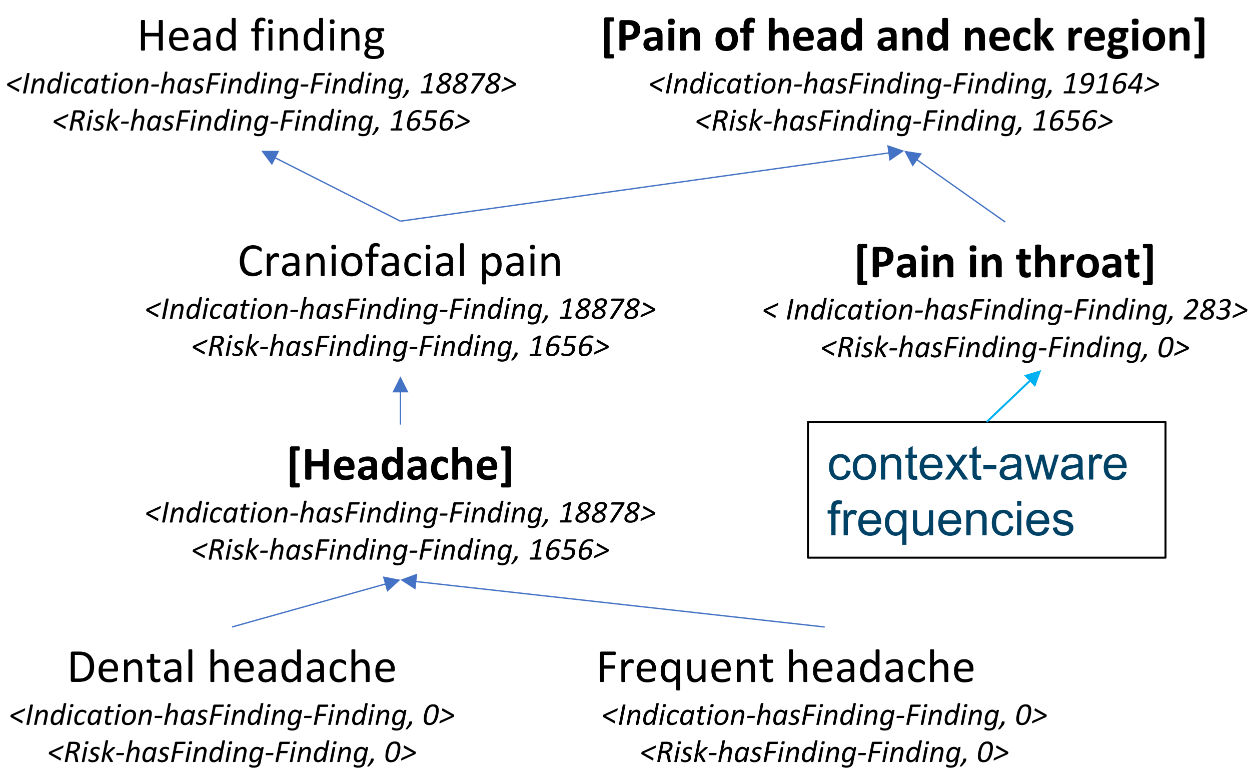
Two-phase approach (overview)



<u>Offline phase</u> (*aka* external knowledge source incorporation):
(i) Initialize the set of contexts, *(ii)* compute concept frequencies, *(iii)* generate mappings

<u>Online phase</u> (*aka* online query relaxation):
(i) map query term to external concept, *(ii)* return top-k external concepts

External knowledge source incorporation



Mapping medical KB to external knowledge source

- exact match / fuzzy match / embeddings / …

The *context* of a query term can be represented by a relationship and its associated concepts from the domain ontology

Concept frequency

$$freq(A) = |A| + \sum_{A_i \sqsubseteq A} freq(A_i)$$

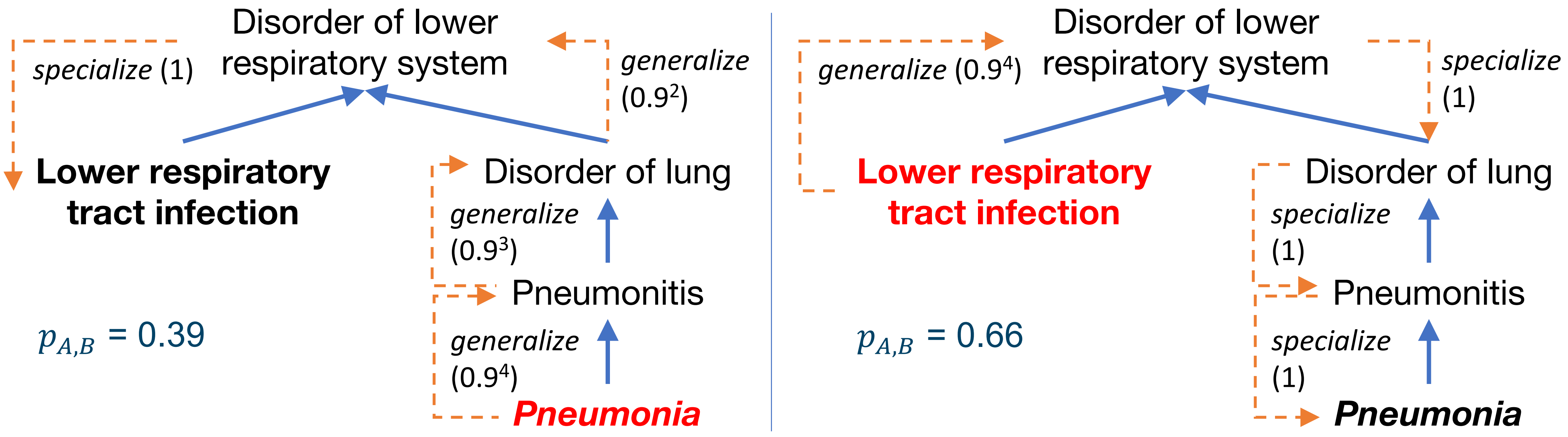
Information content-based similarity

$$IC(A) = -\log(freq(A))$$

$$sim_{IC}(A, B) = \frac{2 \times IC(lcs(A, B))}{IC(A) + IC(B)}$$

Online query relaxation

Generalization vs specialization



The weight of a path connecting two external concepts A and B: $p_{A,B} = \prod_{i}^{|D|} w_i^{D-i}$

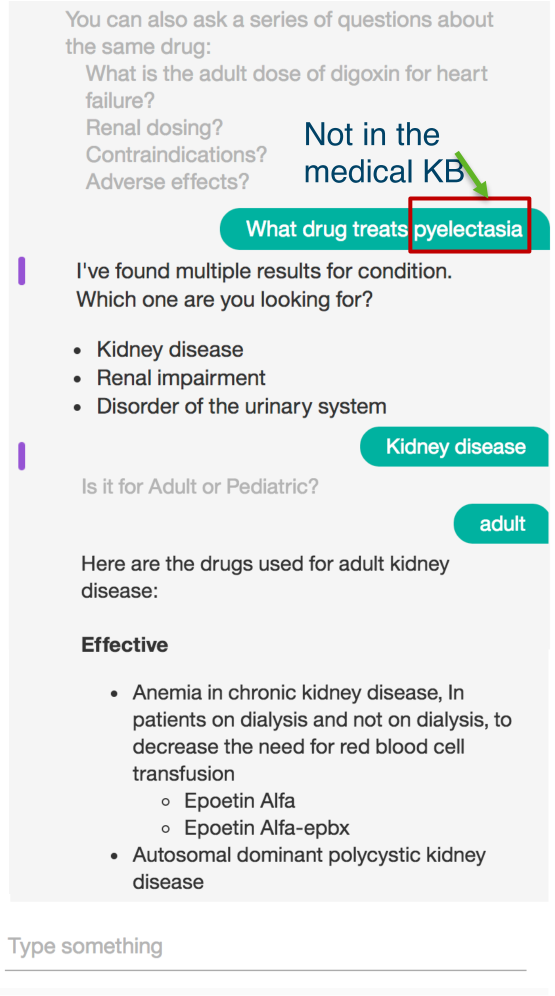
Overall concept similarity: $sim(A,B) = p_{A,B} \times sim_{IC}(A,B)$

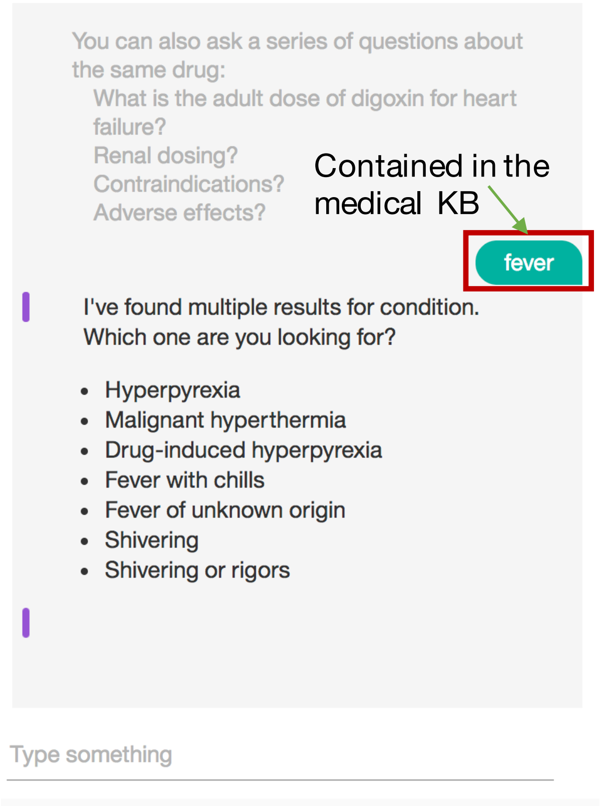
Putting it all together

- Given a query term *q*, the query relaxation method
  1. finds an external concept *A* that matches *q*
  2. searches for the external concepts within ***r*** distance from *A*
  3. retrieves the top-k pre-computed similarity between *A* and each external concept in its neighborhood. Top-***k*** relaxed results are returned based on their overall similarity scores

- ***r*** can be:
  – set as a fixed value by empirical studies, or
  – dynamically decided if a fixed *r* cannot provide *k* results
- ***k*** can be application-specific or defined by users

Integration with IBM Watson Assistant





A. Quamar, C. Lei, D. Miller, F. Özcan, J. Kreulen, R. Moore, V. Efthymiou. An Ontology-Based Conversation System for Knowledge Bases. *SIGMOD 2020*

Experimental evaluation

Accuracy of mapping methods

Methods	Precision	Recall	F1
EXACT	**100**	83.33	90.01
EDIT	96.36	88.33	92.17
EMBEDDING	96.49	**91.67**	**94.02**

Overall effectiveness of query relaxation (QR)

Methods	P@10	R@10	F1
QR	**90.51**	**82.64**	**86.40**
QR-no-context	85.45	77.27	81.15
QR-no-corpus	78.23	70.91	74.39
IC	75.55	68.18	71.68
Embedding-pre-trained	66.14	60.13	62.99
Embedding-trained	79.37	71.81	75.40

Setup

- KB: IBM Micromedex
- External knowledge source: SNOMED CT
- Corpus: a few thousand in-depth documents describing drugs, findings, adverse effects

Results

- IC baseline is not as good as QR even the variations without context or corpus information
- QR without contextual information is reasonable
- QR without corpus is much worse
- pre-trained* is off-the-shelf, but worst results
- trained: using glove and fasttext

* http://bio.nlplab.org

Experimental evaluation – user study

User study with 20 medical SMEs:
Watson Assistant with and without query relaxation (QR)

	QR		*no QR*	
Score	***T1***	***T2***	***T1***	***T2***
1 (Very dissatisfied)	2.1%	10.55%	13.06%	11.11%
2 (Dissatisfied)	10.35%	11.07%	16.87%	38.26%
3 (Okay)	25.59%	29.33%	36.29%	30.85%
4 (Satisfied)	35.21%	33.37%	18.25%	12.47%
5 (Very satisfied)	26.85%	15.68%	15.53%	7.31%
AVG	3.73	3.31	3.06	2.67

T1: for 20 fixed concepts, SMEs pick 20 questions
T2: SMEs are free to ask 10 questions about anything

Observations

- QR improved the user experience in both tasks on average by 20% compared to no QR
- T1 results better than T2
- User feedback for not satisfying answers:
  - expected answers are not contained in the given KB
  - not ideal conversational flow (irrespective of QR results)
  - the amount of information returned is overwhelming

Summary

- A novel two-phase query relaxation method
  - leverages external knowledge sources
  - empowers semantically related concepts with a novel similarity metric

- Integration with two exemplary systems
  - a conversational system
  - a natural language query system

- Our method outperforms state-of-the-art ones in precision and recall
- User study shows our method
  - expands the query results
  - improves their quality for medical KBs

Thank you!


thank you
danke
gracias
merci
obrigado
teşekkür ederim
go raibh maith agat
ngiyabonga
謝謝
спасибо
bedankt
dziękuję
sukriya
kop khun krap
mochchakkeram
terima kasih
arigatô
dakujem
grazie
tapadh leat
dank je
감사합니다
ありがとう
мерси
хвала
рахмат
Баярлалаа
ευχαριστώ