

IBM Research

Expanding Query Answers on Medical Knowledge Bases

Chuan Lei

Vasilis Efthymiou

Rebecca Geis

Fatma Özcan

IBM



Querying medical knowledge bases

IBM Micromedex®

[My Subscription](#) | [Gateway](#) | [Training Center](#) | [Help](#) | [Download Center](#) | [Logout](#)

Home | Drug Interactions | IV Compatibility | Drug ID | Drug Comparison | CareNotes | RED BOOK | Other Tools | Admin

Search Drug, Disease, Toxicology, and more

All | Drug | Disease | Toxicology

Keyword search

Ask Watson

Search Micromedex drug information

What drug treats pyelectasia

Learn more

Latest News

- FDA: Bacitracin Injection Withdrawal
- Palforzia(TM) for Peanut Allergy
- Trijardy(TM) XR: Type 2 Diabetes...
- FDA: Clozapine Constipation Warning
- Tazverik(TM) for Epithelioid Sarcoma

Read Top News

Support & Training

- Citing Micromedex
- Clinical Consulting & Services
- Integrated Content Options for MU & More
- Product User Tips & Quick Answers
- Training Center
- User Guide

Support Request

Resources

- Black Box Warnings
- Comparative Tables
- Do Not Confuse Drug List
- Drug Classes
- Drug Consults
- REMS

Download Mobile Apps

Query relaxation

Problem:

Users do not always formulate their queries precisely to match the terms in the KB

- No answer or incomplete answers returned

Goal:

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

Contributions:

- 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

You can also ask a series of questions about the same drug:

What is the adult dose of digoxin for heart failure?

Renal dosing?

Contraindications?

Adverse effects?

Not in the
medical KB

What drug treats pyelectasia

I've found multiple results for condition. Which one are you looking for?

- Kidney disease
- Renal impairment
- Disorder of the urinary system

Kidney disease

Is it for Adult or Pediatric?

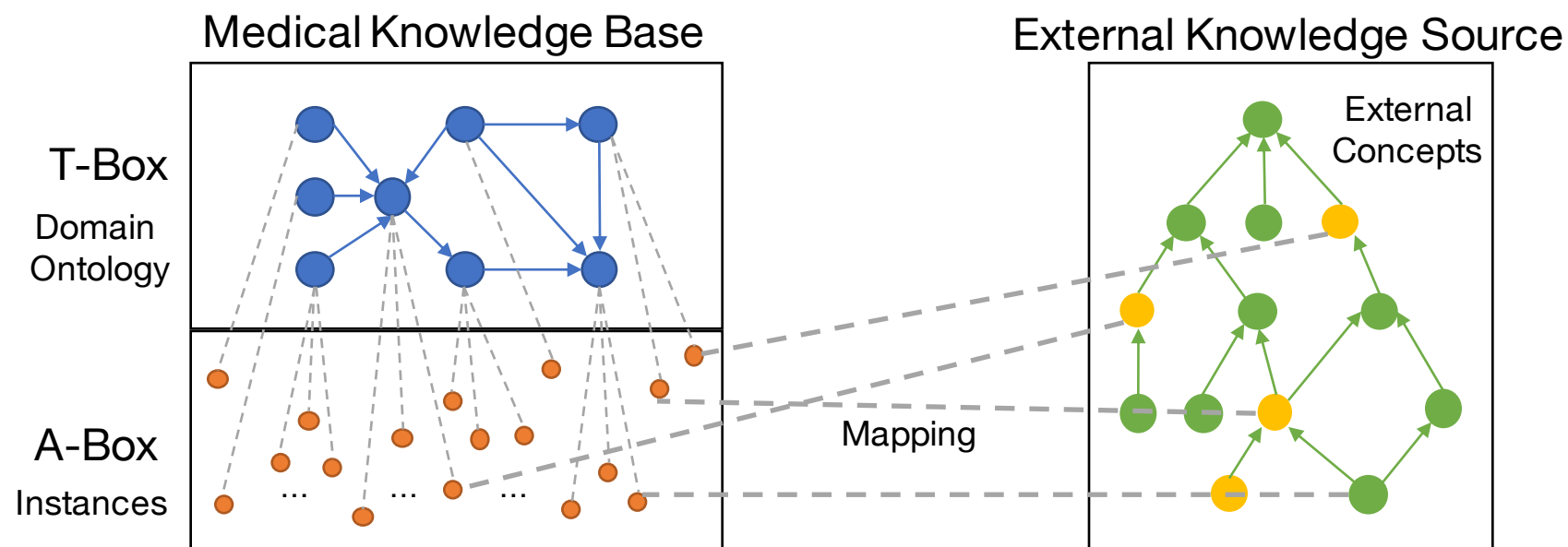
adult

Here are the drugs used for adult kidney disease:

Effective

- Anemia in chronic kidney disease, In patients on dialysis and not on dialysis, to decrease the need for red blood cell transfusion
 - Epoetin Alfa
 - Epoetin Alfa-epbx
- Autosomal dominant polycystic kidney disease

Two-phase approach (overview)



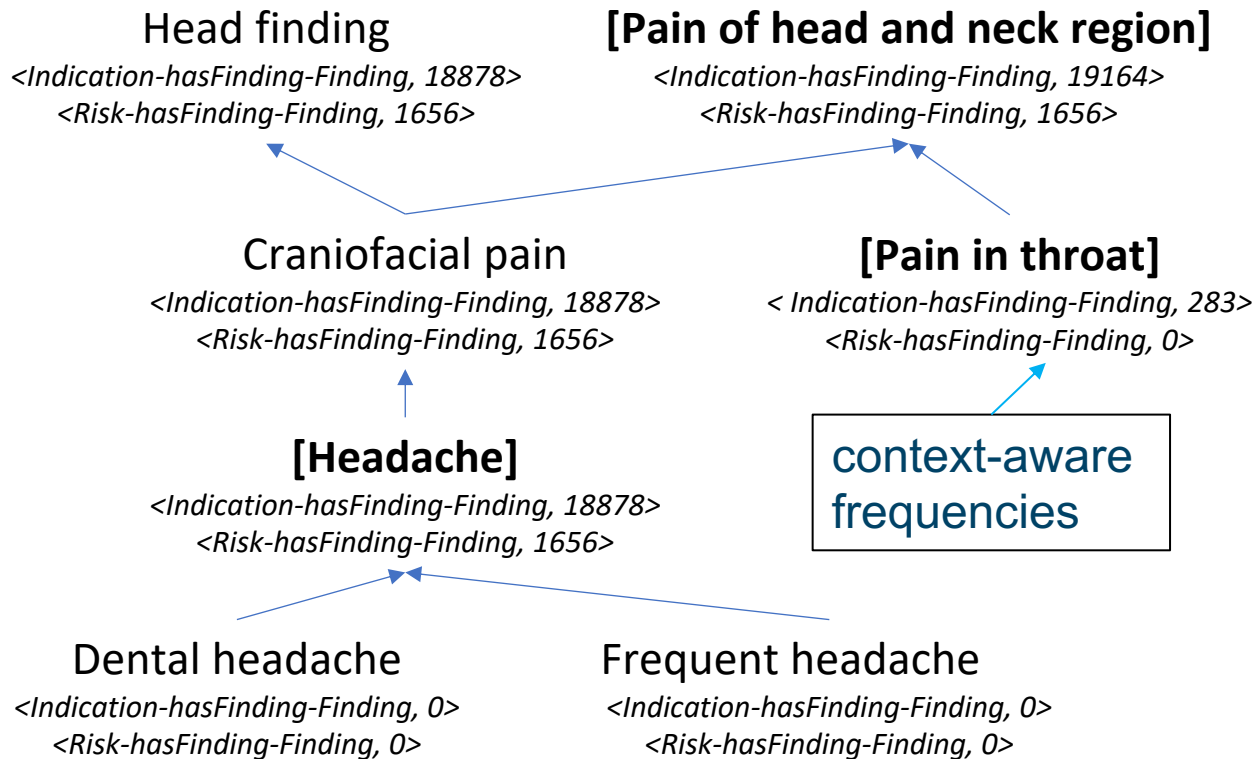
Offline phase (aka external knowledge source incorporation):

(i) Initialize the set of contexts, (ii) compute concept frequencies, (iii) generate mappings

Online phase (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

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

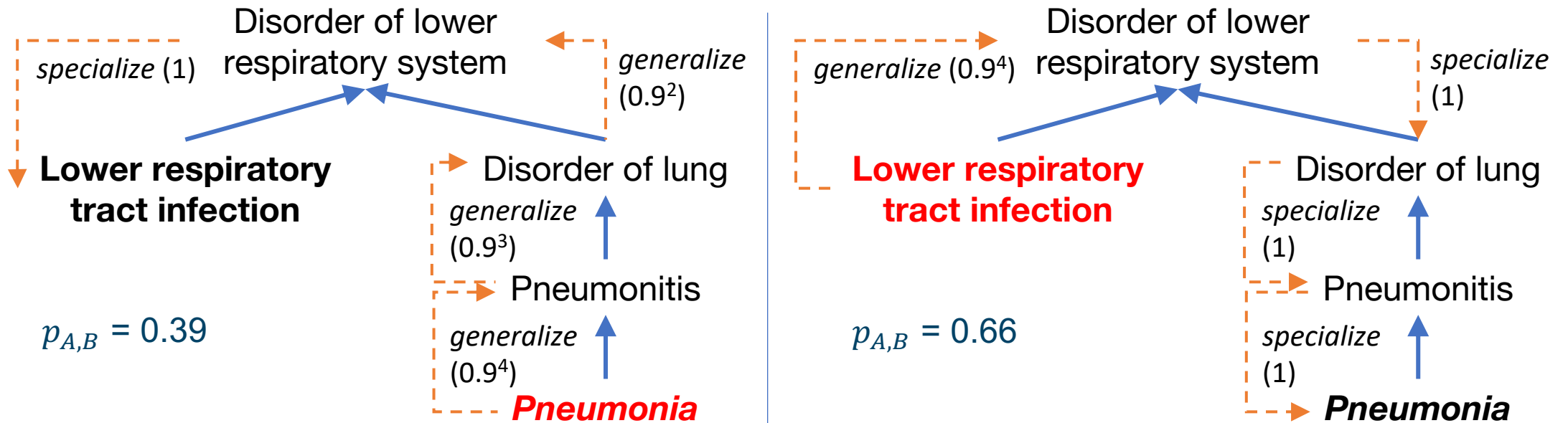
Information content-based similarity

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

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

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

You can also ask a series of questions about the same drug:

- What is the adult dose of digoxin for heart failure?
- Renal dosing?
- Contraindications?
- Adverse effects?

Not in the medical KB

What drug treats pyelectasia

I've found multiple results for condition. Which one are you looking for?

- Kidney disease
- Renal impairment
- Disorder of the urinary system

Kidney disease

Is it for Adult or Pediatric?

adult

Here are the drugs used for adult kidney disease:

Effective

- Anemia in chronic kidney disease, In patients on dialysis and not on dialysis, to decrease the need for red blood cell transfusion
 - Epoetin Alfa
 - Epoetin Alfa-epbx
- Autosomal dominant polycystic kidney disease

Type something

You can also ask a series of questions about the same drug:

- What is the adult dose of digoxin for heart failure?
- Renal dosing?
- Contraindications?
- Adverse effects?

Contained in the medical KB

fever

I've found multiple results for condition. Which one are you looking for?

- Hyperpyrexia
- Malignant hyperthermia
- Drug-induced hyperpyrexia
- Fever with chills
- Fever of unknown origin
- Shivering
- Shivering or rigors

Type something

Experimental evaluation

Accuracy of mapping methods

Methods	Precision	Recall	F1
<i>EXACT</i>	100	83.33	90.01
<i>EDIT</i>	96.36	88.33	92.17
<i>EMBEDDING</i>	96.49	91.67	94.02

Overall effectiveness of query relaxation (QR)

Methods	P@10	R@10	F1
<i>QR</i>	90.51	82.64	86.40
<i>QR-no-context</i>	85.45	77.27	81.15
<i>QR-no-corpus</i>	78.23	70.91	74.39
<i>IC</i>	75.55	68.18	71.68
<i>Embedding-pre-trained</i>	66.14	60.13	62.99
<i>Embedding-trained</i>	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)

	<i>QR</i>		<i>no QR</i>	
Score	<i>T1</i>	<i>T2</i>	<i>T1</i>	<i>T2</i>
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

