

MedEx-UIMA

MedEx – a rule-based tool to identify drug information from clinical text

- Semantic-based parsing (Drug names and signatures)
- Developed at Vanderbilt, evaluated on text at other sites
- Implemented in Python

Findings	Prec	Rec	F-Score
DrugName	95.0	91.5	93.2
Strength	98.8	90.5	94.5
Route	98.8	89.6	93.9
Frequency	98.9	93.2	96.0

Table 1. Evaluation on **discharge summaries** from Vanderbilt.

Findings	Prec	Rec	F-Score
DrugName	96.7	88.0	92.1
Strength	94.7	94.7	94.7
Route	96.0	87.0	91.3
Frequency	96.8	89.2	92.9

Table 2. Evaluation on **clinic visit notes** from Vanderbilt.

Xu et al. JAMIA 2010; 17:19-24

Findings	Prec	Rec	F-Score
DrugName	84.2	87.1	85.6
Dose	89.5	81.8	85.5
Route	91.8	85.8	88.7
Frequency	87.9	85.8	86.8
Reason	45.9	29.6	36.0
Duration	36.4	35.8	36.1
All	83.9	80.3	82.1

Table 3. Evaluation on 2009 i2b2 data set: **discharge summaries and progress notes. Ranked 2nd out of 20 participating teams!**

Doan et al. JAMIA 2010; 17: 528-31

MedEx-UIMA

- Implemented in Java
- Based on the UIMA framework
- Similar performance as MedEx
- Open Source, available at:
<https://code.google.com/p/medex-uima/>
- Extended functions
 - Drug name encoding (RxNORM CUIs: specific vs. generic)
 - Frequency normalization (TIMEX3)
 - Support multiple Operating Systems (Windows, Linux, Unix, Mac etc.)
 - Two versions for downloading: User vs. Developer
 - Drool rule engine for tag disambiguation
- Evaluation on 125 notes shows that it performs slightly better than the python MedEx

User Version

- Prerequisite - Java 1.7
- Install MedEx-UIMA
 - Download and unzip the file to a local folder
- Run MedEx-UIMA
 - Java application mode: outputs delimited textual file
 - GUI mode: outputs XML file that can be visualized at the UIMA interface

Java Application Mode

- Enter into MedEx-UIMA folder and run
 - Windows: `java -cp lib/*;bin org.apache.medex.Main arg1 arg2 ...`
 - Linux/Unix/Mac: `java -cp lib/*:bin org.apache.medex.Main arg1 arg2 ...`
- Outputs: one drug per line, fields separated by “|”

```
Sentence index (start from 1)
Sentence text
Drug name      (e.g. 'simvastatin[0, 11]')
Brand name     (e.g. 'zocor[12, 17]')
Drug form      (e.g. 'tablet[19, 25]')
Strength       (e.g. '10mg[20, 24]')
Dose amount    (e.g. '2 tablets[2, 11]')
Route          (e.g. 'by mouth[10, 18]')
Frequency      (normalized frequency) (e.g. 'b.i.d.(R1P12H)[10, 16]',
                'R1P12H' is the TIMEX3 format of 'b.i.d.')
Duration       (e.g. 'for 10 days[10, 21]')
Necessity      (e.g. 'prn[10, 13]')
UMLS CUI
RXNORM RxCUI
RXNORM RxCUI for generic name
Generic name   (associated with RxCUI code)|
```

Java Application Mode

- Enter into MedEx-UIMA folder and run
 - Windows: `java -cp lib/*;bin org.apache.medex.Main arg1 arg2 ...`
 - Linux/Unix/Mac: `java -cp lib/*:bin org.apache.medex.Main arg1 arg2 ...`
- Outputs: one drug per line, fields separated by “|”

Zinc Sulfate 220 mg p.o. q day

Drug Name	Dose Amount	Route	Frequency	UMLS CUI	RxNORM CUI - specific	RxNORM CUI – generic	Generic Name
Zinc sulfate	220 mg	p.o.	q day	C0991474	316931 (Zinc sulfate 220 mg)	39954 (Zinc sulfate)	Zinc sulfate

GUI Mode

- Install UIMA
- Modify the input and output folders for UIMA CPE application (XML file)
- Run UIMA CPE program of MedEx:
 - Windows: "java -cp lib/*;bin org.apache.UIMA.CPE.medex.MedexUIMACPE"
 - Linux/Unix/Mac: "java -cp lib/*:bin org.apache.UIMA.CPE.medex.MedexUIMACPE"
- Run UIMA GUI to visualize the results
 - Windows user: Run "[UIMA_HOME]\bin\annotationViewer.bat"
 - Linux/Unix/Mac user: Run "[UIMA_HOME]/bin/annotationViewer.sh"

Annotation Results for test.txt.xmi in C:\medex\MedEx_UIMA\uima-output

POTENTIALLY SERIOUS INTERACTION: LEVOFLOXACIN; Lipitor was initially helped for an initial transaminitis presumed to be secondary to shock liver , that at METOLAZONE 2.5 MG PO DAILY Starting Today August PRN Other:Weight gain including amitriptyline , Flexeril , and Valium for reported history of sciatica and low back pain.

MEDICATIONS ON ADMISSION :

1. Digoxin .125 mg qod , .25 mg qod .
2. Colace 100 mg po tid .
3. Levoxyil 100 po qd .
4. Prinivil 20 mg po qd .

Click In Text to See Annotation Detail

- Annotations
 - Drug
 - Drug ("METOLAZONE")
 - begin = 157
 - end = 167
 - semantic_type =

Legend

<input type="checkbox"/> DocumentAnnota...	<input checked="" type="checkbox"/> Drug	<input checked="" type="checkbox"/> Frequency	<input checked="" type="checkbox"/> Neccessity	<input checked="" type="checkbox"/> Route
<input checked="" type="checkbox"/> Strength				

Select All Deselect All Hide Unselected

System Input Arguments

- Five input arguments
 - i (input directory)
 - o (output directory)
 - b [yn]
 - use built-in sentence detector (default) or not
 - f [yn]
 - normalize frequency to TIMEX3 format or not
 - n=default
 - e.g. 'b.i.d.' -> 'R1P12H'
 - d [yn]
 - use drool engine or not
 - The default setting is to use the built-in rules for disambiguation (faster)

Developer Version

- Prerequisite: Java 1.7
- Set up development environment
 - Set the JAVA_HOME environment variable as your root directory of JAVA
 - Set up UIMA
 - Install Eclipse
 - Install UIMA Eclipse plug-ins
 - Set up a class path variable named "UIMA_HOME"
- Open Eclipse and import the "MedEx_UIMA" project
- Develop and test
- For more details, please refer to "Readme.txt"

Contact Information

- For collaboration/support
 - Hua Xu – hua.xu@uth.tmc.edu
 - Josh Denny – Josh.Denny@Vanderbilt.edu
- For technical support
 - Min Jiang (Programmer) – min.jiang@uth.tmc.edu