SHARP NLP

August, 2013  Mayo Clinic
Introduction

NLP Software
- cTAKES
- Evaluation Workbench

Manual Annotations
- Seattle Group Health clinical notes
- Mayo Clinic clinical notes
Software – cTAKES functionality

02/2013  cTAKES 3.0 released
   new Relation Extractor module, with 2 relations implemented
      location_of
      degree_of

08/2013  Apache cTAKES 3.1
   new Template Filler module
      collects data from all other modules to fill CEM template
   improvements to other modules

4Q2013  anticipate Apache cTAKES 3.2
   improved named entity attributes - e.g. history_of, negation detection
Software – cTAKES milestones

06/2012  cTAKES accepted as an Apache incubator project
02/2013  cTAKES 3.0 released
03/2013  Graduated to Top Level Project  Apache cTAKES
08/2013  Apache cTAKES 3.1 release candidate available
08/2013  anticipated release of Apache cTAKES 3.1
08/2013  integration into Data Normalization pipeline completed
10/2013  anticipated release of Data Normalization pipeline with integrated Apache cTAKES
Comparing, analyzing, and drilling down into details of differences

- Two different runs of a system
- Manual vs. system results

Video overview online
http://screencast.com/t/QzaMLwWwFe
Annotations

- Manually created by domain experts

Uses
- Training the software
- Testing and evaluating the algorithms

Levels
- UMLS entities (medications, diseases/disorders, etc.)
- UMLS relations (location_of, degree_of, etc.)
- Treebank
- (to fill in yet)
Annotations - Process

- Manually annotated
- Some annotated by pairs
- Adjudicated/reviewed by domain expert
- Used during training & testing of software
- Corrections phase (August-Sept 2013)
Annotations - Progress

- Mayo Clinic corpus – completed*
- Seattle Group Health corpus – 30% completed*

* not counting the corrections phase.
Deep dive

- MedXN
- Common type systems
- Attribution discovery
- Multi-Language layered Information Retrieval
- NLP in practice
  - Indexing
  - Phenotyping
  - Clinical decision support
Deep dive

- MedXN
- Common type systems
- Attribution discovery
- Multi-Language layered Information Retrieval
- NLP in practice
  - Indexing
  - Phenotyping
  - Clinical decision support
MedXN system

- The accurate exchange of medication information requires drug name standards
  - RxNorm addresses this issue by providing a normalized drug name
  - Today, RxNorm is becoming part of Meaningful Use to support the expanding functionality of health record technology
- Aim: to extract comprehensive medication information and normalize it to the most appropriate RxNorm CUI (RxCUI) as specifically as possible
MedXN Algorithm for medication extraction and normalization

Step 1: Medication Extraction

Step 2: Attribute Extraction

Step 3: Medication & Attribute Association

Step 4: Convert to RxNorm Standard

Step 5: Convert to RxCUI Representation

Step 6: Normalize to Specific RxCUI

"Sulfasalazine [AZULFIDINE] 500-mg
2 tabs by mouth two times a day"

Sulfasalazine [AZULFIDINE]
RxCUI="9524::IN::202770::BN"

500-mg (strength), 2 (dose),
tabs (form), mouth (route),
two times a day (frequency)

<Sulfasalazine [AZULFIDINE]>
+ <500-mg, 2, tabs, mouth,
two times a day>

sulfasalazine <in>500 mg<st>
oral tablet<df>azulfidine<bn>

9524<in>500 mg<st>
317541<df>202770<bn>

Sulfasalazine 500 MG
Oral Tablet [AZULFIDINE]
RxCUI=208437::SBD
Medication information annotations visualized in MedXN
MedXN vs. NCBO annotator vs. MedEx of RxCUI assignment

Precision
Recall
F-measure

[Bar chart showing comparison of precision, recall, and F-measure between MedXN, NCBO annotator, and MedEx]
Deep dive

- MedXN
- **Common type systems**
- Attribution discovery
- Multi-Language layered Information Retrieval
- DN/NLP in practice
  - Indexing
  - Phenotyping
  - Clinical decision support
Common NLP types in SHARPn

- Interoperability
- Clinical Element Models (CEMs)
  - Normalized semantics
  - Knowledge constraints
  - 6 generic templates
- Semantic Type system
Populating the NLP Type System

**IdentifiedAnnotation**
- polarity : int
- conditional : boolean
- uncertainty : int
- subject : String
- generic : boolean
- ontologyConceptArr : FSArray
- id : int
- typeID : int
- segmentID : String
- sentenceID : String
- discoveryTechnique : int
- confidence : Double

**Medication**
- statusChange : MedicationStatusChange
- dosage : MedicationDosage
- duration : MedicationDuration
- endDate : Date
- startDate : Date
- form : MedicationForm
- frequency : MedicationFrequency
- route : MedicationRoute
- strength : MedicationStrength
- relativeTemporalContext : TemporalRelation

**Attribute Discovery** (ctakes-assertion)
Deep dive

- MedXN
- Common type systems
- **Attribution discovery**
- Multi-Language layered Information Retrieval
- DN/NLP in practice
  - Indexing
  - Phenotyping
  - Clinical decision support
### SHARPn Attribute Discovery

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Negation</strong></td>
<td>Patient has not noticed any <em>numbness</em>.</td>
</tr>
<tr>
<td><strong>Uncertainty</strong></td>
<td>The symptoms are not inconsistent with <em>renal failure</em>.</td>
</tr>
<tr>
<td><strong>Conditional</strong></td>
<td>Pt should come back to the ED if any <em>rash</em> occurs.</td>
</tr>
<tr>
<td><strong>Subject</strong></td>
<td>Family history of <em>lupus</em>.</td>
</tr>
<tr>
<td><strong>Generic</strong></td>
<td>We discussed increased risk of <em>breast cancer</em>.</td>
</tr>
<tr>
<td><strong>HistoryOf</strong></td>
<td>PMH: <em>Hyperlipidemia</em></td>
</tr>
</tbody>
</table>

- **Context of Named Entities (NEs)**
  - 6 attributes, All CEM templates
  - For Normalization and HTP
  - For Information Retrieval
Current Polarity (Negation) methods suffer across sources

<table>
<thead>
<tr>
<th>Train set</th>
<th>Test set</th>
<th>i2b2/VA test</th>
<th>SHARP seed dev</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Prec</td>
<td>Rec</td>
</tr>
<tr>
<td>i2b2/VA assert</td>
<td>train</td>
<td>95.0</td>
<td>85.9</td>
</tr>
<tr>
<td>SHARP seed train</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline: NegEx</td>
<td></td>
<td>84.9</td>
<td>79.5</td>
</tr>
</tbody>
</table>

- SHARP seed corpus is difficult
- Overcome with advanced methods
- New results: F1 = 90.0 (vs. 62.7)
Deep dive

- MedXN
- Common type systems
- Attribution discovery
- Multi-Language layered Information Retrieval
- DN/NLP in practice
  - Indexing
  - Phenotyping
  - Clinical decision support
Why Information Retrieval (IR)?

- Clinical NLP out-of-the-box
  - Comprehensive knowledge
  - Customize? Collaborate!
- Diverse requirements
- Physician/Researcher tasks
  - Enroll patients in study
  - Define retrospective cohort
  - Consumer health search
  - Find relevant resources
- User-centric -> Information Retrieval

Somali patients (unique terms)
Drug-induced liver injury (rel’ns)
Pediatric asthma (temporal)

Text REtrieval Conference (TREC)
CLEF eHealth Task 3
MayoExpert
Attribute Discovery in IR: TREC Cohort Identification

- Attribute discovery help IR users?
- Method: Attribute → term weight
  - Ex: search="carotid endarterectomy"
    - “Pt has not had a carotid endarterectomy”
- Context of terms is important

<table>
<thead>
<tr>
<th>Run ID</th>
<th>‘11 bpref</th>
<th>‘12 bpref</th>
<th>infNDCG</th>
</tr>
</thead>
<tbody>
<tr>
<td>MayoLucene</td>
<td>0.4249</td>
<td>0.2771</td>
<td>0.3694</td>
</tr>
<tr>
<td>MayoPayload</td>
<td>0.4730</td>
<td>0.2981</td>
<td>0.4119</td>
</tr>
</tbody>
</table>
Consumer Health Search: 2013 CLEF eHealth Task 3

- Web search for health information
  - Methods: Language models
  - Excellent performance (gray=difference from median)
Semantic Textual Similarity (STS): What and Why?

Predict similarity (sentence pairs)

- Why STS? Aggregate NLP evaluation
  - Named entities
  - Attributes
  - Relations
  - Word sense disambiguation
  - Semantic role labeling
  - Co-reference resolution

- Why STS? Precision semantics in IR

S1: The right common iliac is widely patent.
S2: Both internal iliac arteries are patent.

Human: 3.60/5.00
# Results: STS at *SEM 2013

<table>
<thead>
<tr>
<th>TEAM NAME</th>
<th>headlines rank</th>
<th>OnWN rank</th>
<th>FNWN rank</th>
<th>SMT rank</th>
<th>mean rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>UMBC-EBIQUITY-ParingWords</td>
<td>0.7642</td>
<td>0.7529</td>
<td>0.5818</td>
<td>0.3804</td>
<td>0.6181</td>
</tr>
<tr>
<td>UMBC-EBIQUITY-galactus</td>
<td>0.7428</td>
<td>0.7053</td>
<td>0.5444</td>
<td>0.3705</td>
<td>0.5927</td>
</tr>
<tr>
<td>deft-baseline</td>
<td>0.6532</td>
<td>0.8431</td>
<td>0.5083</td>
<td>0.3265</td>
<td>0.5795</td>
</tr>
<tr>
<td>MayoClinicNLP-r4ALL</td>
<td>0.7275</td>
<td>0.7618</td>
<td>0.4359</td>
<td>0.3048</td>
<td>0.5707</td>
</tr>
<tr>
<td>UMBC-EBIQUITY-saiyan</td>
<td>0.7838</td>
<td>0.5593</td>
<td>0.5815</td>
<td>0.3563</td>
<td>0.5683</td>
</tr>
<tr>
<td>MayoClinicNLP-r3wtCD</td>
<td>0.6440</td>
<td>0.8295</td>
<td>0.3202</td>
<td>0.3561</td>
<td>0.5671</td>
</tr>
<tr>
<td>MayoClinicNLP-r1wtCDT</td>
<td>0.6584</td>
<td>0.7775</td>
<td>0.3735</td>
<td>0.3605</td>
<td>0.5649</td>
</tr>
<tr>
<td>CLaC-RUN2</td>
<td>0.6921</td>
<td>0.7366</td>
<td>0.3793</td>
<td>0.3375</td>
<td>0.5587</td>
</tr>
<tr>
<td>MayoClinicNLP-r2CDT</td>
<td>0.6827</td>
<td>0.6612</td>
<td>0.396</td>
<td>0.3946</td>
<td>0.5572</td>
</tr>
<tr>
<td>NTNU-RUN1</td>
<td>0.7279</td>
<td>0.5952</td>
<td>0.3215</td>
<td>0.4015</td>
<td>0.5519</td>
</tr>
<tr>
<td>CLaC-RUN1</td>
<td>0.6774</td>
<td>0.7667</td>
<td>0.3793</td>
<td>0.3068</td>
<td>0.5511</td>
</tr>
</tbody>
</table>

- 3rd of 30 teams
- 5th, 6th, 8th of 90 submissions
- Methods: Combine metrics (string, multiword, distributional)
Deep dive

- MedXN
- Common type systems
- Attribution discovery
- Multi-Language layered Information Retrieval
- DN/NLP in practice
  - Indexing
  - Phenotyping
  - Clinical decision support
Prospective and retrospective case identification for epidemiology studies, clinical trial feasibility studies and patient recruitment

Personalized decision support for preventive care

Outcome tracking and care quality monitoring

Patient status monitoring

Phenotype ("clinotype") generation

Assist in detection of adverse events

Assist in coding for billing and research

Patient and document level summarization and aggregation

Prerequisite for additional data mining and discovery
NLP for indexing

Collection Readers

- SentenceAnnotator
- Tokenizer
- Lexical Normalizer*
- SectionAnnotator
- DictionaryLookup
- CRFCMAnnotator*
- ContextAnnotator

CAS Consumers

NLP for pattern-based information extraction

Collection Readers

- SentenceAnnotator
- Tokenizer
- POS Tagger*
- ChunkAnnotator*
- SectionAnnotator
- InformationExtractor
- ContextAnnotator

CAS Consumers

OpenNLP components In cTAKES

- MedTagger (Torii et al)
- SecTag (Denny et al)
- MedTagger (Torii et al)
- MedTagger (Torii et al)
- ConText (Chapman et al)

* Indicates optional components
Normalization and Integration Logics

Visit

Structured Data
Clinical Narratives

EMR

Current Visit Information

Natural language processing

IBM Watson-based technology

Accelerated knowledge engineering utilizing three levels of intelligences

Powerful computing through IBM Watson-based technology

Drools-based inference engine

A unified architecture to provide diverse decision support needs

User Intelligence

Normalization and Integration

Decision

Guideline Logics

Triaging

Treatment

Follow-up

Big Data

Questionnaire

Clinical Data Warehouse

IBM Watson

Clinical Practice Guidelines
Population Management → CPM

- Screening
  - Colonoscopy
  - Cervical Cancer
  - Pulmonary Nodules
  - Renal lesions

- Surveillance
  - Mammograms
  - Aneurysms

- Follow-up
  - NLP DSS 98-100% accuracy
Population Management → CPM

Pap Smear

EMR
Structured Data
\( e.g. \text{Age}=35, \ \text{WBC}=4500 \)
Free Text Data
\( e.g. \text{PAPsmear report} \)

NLP
Structured Data
\( e.g. \text{HPVtest=} \text{positive} \)

Guideline Logic
USPSTF recommendation
\( e.g. \text{Repeat PAPtest in 6 months} \)
Clinical pilot for NLP-CDS system for cervical cancer screening and surveillance

12 providers have used the system for 3 months. The providers report a high accuracy and utility of the system over the current CDS system that operates only on discrete data.
Colon and Rectal Surgery improves processes, quality and outcomes as part of institutional focus on managing to reimbursement

June 28, 2013

System. This opportunity provides patients with improved medical and rehabilitation services. Mayo Clinic loses $2,000 a day for Medicare colorectal surgery patients who are hospitalized longer than 14 days. “When we discharge that patient to Lake City, for example, reimbursement covers the cost of care plus 1 percent,” he says. “It’s a huge win for patients. They are getting the care they really need in a more convenient place.”

- Natural language processing
This new technology combs the electronic medical records to find documented abnormalities, potentially before patients present with symptoms. This technology works automatically, in real time, so providers can be alerted to concerns such as internal bleeding. “We are still working on this,” says Dr. Larson. “It has huge potential to allow us to prevent problems.”
NLP-enabled postoperative complications detection

Agile NLP process

Detect complications automatically

Machine learning to learn implicit logics

Knowledge engineering: Inclusion/Exclusion rules

Contextual information analysis and annotation

Knowledge engineering: keywords

Complication definition

Surgery Workflow and Documentation Process

EMR

Demographics

Problem List

Labs

Radiology

ECG Reports

Clinical Notes

Operation Notes