clinithink

add meaning:get knowledge

Clinical Natural Language Processing tools for SNOMED CT

SNOMED CT Implementation Showcase 2014

Dr Richard Gain
Senior Clinical Terminologist, Clinithink

Aims of this presentation

In this presentation we will explore:

- some issues around making sense of unstructured clinical data

- some of the difficulties of using SNOMED CT

- how Clinical NLP and SNOMED CT can be integrated

- abstraction of information from coded records using queries

clinithink
add meaning:get knowledge

CliX
PATENT PENDING

Why are we doing this?

Structured clinical data is challenging enough

- Written clinical records often contain large amounts of narrative which are rich in information content

- This information is hard for computers to interpret and understand

- Turning free text into structured data makes the information accessible

Unstructured clinical data is….

- "Messy"
- Dispersed & disconnected
- Complex
- Non-uniform & non-standardized
- Varied
- Inconsistent

clinithink
add meaning:get knowledge

CliX
PATENT PENDING

Adding "Clinical" to NLP = CNLP

Unlock the meaning and value in healthcare data

- SNOMED CT is a compositional terminology which works best when used in post-coordinated expressions

- These SNOMED CT expressions can be difficult to construct and analyse

- Clinical Natural Language Processing (CNLP) can be used to unleash the power of the SNOMED CT terminology

clinithink
add meaning:get knowledge

CliX
PATENT PENDING

Essential NLP Concepts

NLP is different from Clinical NLP (CNLP).

Voice Recognition with NLP: **APPLE SIRI**

HX: She presented with occipital headache for last six hours. She has been intermittently nauseated but has not vomited. Mild fever. She has no past history of headaches but has a past history of asthma. Mother has a severe allergy to penicillin but she has no known penicillin allergy. Currently taking paracetamol prn.

OE: BP is 102/60, pulse 70, respiratory rate 20, temperature 37.2 C. Nasophayrynx normal, both tympanic membranes normal.

Clinical Natural Language Processing: **CLINITHINK CLiX**

I didn't find anything for 'if it'll headache for last six hours she has been intermittently nauseated but has not vomited mild fever she is.'

History
- Occipital headache
  - six hours
- Nauseated
- **not** Vomiting
- Fever
  - Mild

Past History
- **no** History of headache
- History of - asthma

Family History
- Allergy to penicillin
  - Severe
  - Mother

Allergies, Risks, Warnings
- **no known** Penicillin allergy

Medications
- Paracetamol
  - As required

Observations and Findings
- Blood pressure
  - 102/60 units: none
- Pulse finding
  - 70 units: none
- Respiratory rate
  - 20 units: none
- Body temperature
  - 37.2 units: none
- Nasopharynx normal
- Tympanic membrane normal
  - Right and left

clinithink
add meaning:get knowledge

CLiX
PATENT PENDING

How can CNLP aid data entry?

Use of CNLP tools in preference to pre-defined data entry forms:

- Allows clinicians to write notes in their own preferred style.

- Simplifies the learning curve for interacting with electronic medical record systems

- Allows for easy integration with voice recognition software

- Permits existing sources of narrative data to be batch processed and analysed to unlock the clinical information it contains

clinithink
add meaning:get knowledge

CliX
PATENT PENDING

Benefits of using SNOMED CT

- Rich clinical terminology containing many thousands of clinical concepts

- Defining attributes based on anatomy and pathology

- Concept model controls how concepts can be combined

- Subsumption testing allows for analysis of post-coordinated expressions

But this comes at a cost

- SNOMED CT is most powerful when used as a compositional terminology

- Writing SNOMED CT expressions is difficult and not possible for clinical users

clinithink
add meaning:get knowledge

CliX
PATENT PENDING

Benefits of using CNLP

Correct interpretation of meaning of records containing:

- Misspellings, word derivations and inflections

- Absence / negation / exclusion expressed in many different ways

- Acronyms and abbreviations and other synonyms

- Severity, certainty, temporality and subject specification

- Finding site and laterality

- Contextual variation defined by record section headings

- Complex sentences composed of lists of items

- Courses and episodes of illnesses

- Procedure contexts, medication doses, dates

clinithink
add meaning:get knowledge

CliX
PATENT PENDING

Here's how we do it

- SNOMED CT is most powerful when used in post-coordinated expressions

- Writing post-coordinated expressions is too hard

- Template based data entry is unnatural and too restrictive

So -

- Allow clinicians to write or dictate patient stories as narrative

- Headings provide important context and are fully supported

- Use CNLP to extract meaning from clinical narrative

- Use SNOMED CT terminology and grammar to construct expressions

- Use complex SNOMED CT based queries to abstract information from data

How does Clinithink CLiX CNLP work?

- Breaks sentence into chunks

- Corrects spelling errors

- Replaces acronyms and abbreviations with expanded text

- Looks for matches to SNOMED CT concepts

- Looks for matching attributes and values for these concepts

- Looks for possible post-coordination opportunities

- Evaluates which possibilities are most likely to be correct

- Constructs valid SNOMED CT expression as output

- Tests subsumption against standard SNOMED CT expressions

clinithink
add meaning:get knowledge

CLiX
PATENT PENDING

What does the output look like?

- Clinical text
  - "Endoscopy revealed an <u>acute gastric ulcer</u> but no evidence of <u>gastric bleeding</u> or <u>perforation of the stomach</u>"

- SNOMED output
  - 243796009 | Situation with explicit context | : { 408731000 | Temporal context | = 410512000 | Current or specified | , 246090004 | Associated finding | = 95529005 | Acute gastric ulcer | , 408732007 | Subject relationship context | = 410604004 | Subject of record | , 408729009 | Finding context | = 410515003 | Known present | }
  - 243796009 | Situation with explicit context | : { 408729009 | Finding context | = 410516002 | Known absent | , 246090004 | Associated finding | = 61401005 | Gastric bleeding | , 408731000 | Temporal context | = 410512000 | Current or specified | , 408732007 | Subject relationship context | = 410604004 | Subject of record | }
  - 243796009 | Situation with explicit context | : { 408729009 | Finding context | = 410516002 | Known absent | , 246090004 | Associated finding | = 235674005 | Perforation of stomach | , 408731000 | Temporal context | = 410512000 | Current or specified | , 408732007 | Subject relationship context | = 410604004 | Subject of record | }

clinithink
add meaning:get knowledge

ClIX
PATENT PENDING

Analysis of SNOMED encoded data - Principles

Abstraction of information from encoded data using queries

- **Queries**
  - A query is a pre-defined post-coordinated SNOMED CT expression
  - Each query describes a discrete clinical concept with any number of attributes

- **Criteria**
  - Complex combinations of queries using Boolean logic
  - Allows conjunctions such as "with", "without", "due to" to be tested separately

- **Abstractions**
  - Links each criterion to a specific output category
  - Abstractions can be standard coding schemes or completely customer specific groups

clinithink
add meaning:get knowledge

CliX
PATENT PENDING

Analysis of SNOMED encoded data - Example

- Queries
  - acute_ulcer_query

    243796009|Situation with explicit context|:{246090004|Associated finding|=(64572001|Disease|:{116676008|Associated morphology|=26317001|Acute ulcer|,363698007|Finding site|=69695003|Stomach structure|}),408729009|Finding context|=410515003|Known present|,408731000|Temporal context| =410512000|Current or specified|,408732007|Subject relationship context|=410604004|Subject of record|}

  - hemorrhage_query

    243796009|Situation with explicit context|:{246090004|Associated finding|=(64572001|Disease|:{116676008|Associated morphology|=50960005| Haemorrhage|,363698007|Finding site|=122865005|Gastrointestinal tract structure|}),408729009|Finding context|=410515003|Known present|, 408731000|Temporal context|=410512000|Current or specified|,408732007|Subject relationship context|=410604004|Subject of record|}

  - perforation_query

    243796009|Situation with explicit context|:{246090004|Associated finding|=(64572001|Disease|:{116676008|Associated morphology|=36191001| Perforation|,363698007|Finding site|=122865005|Gastrointestinal tract structure|}),408729009|Finding context|=410515003|Known present|,408731000| Temporal context|=410512000|Current or specified|,408732007|Subject relationship context|=410604004|Subject of record|}

- Criteria definition
  - acute_ulcer_query AND NOT (hemorrhage_query OR perforation_query)

- Classification
  - K253 Acute gastric ulcer without hemorrhage or perforation

clinithink
add meaning:get knowledge

CliX
PATENT PENDING

Demonstrations

Using CLiX Notes for data entry

Browsing SNOMED CT and creating expressions

Query-based abstraction of coded information

clinithink
add meaning:get knowledge

CLiX
PATENT PENDING

What have we learned?

Large volumes of useful data can be freed from narrative records

Query-based analytics on narrative text opens up new possibilities

CNLP is not the same as NLP

SNOMED CT and CNLP work well together

clinithink

add meaning:get knowledge

CliX
PATENT PENDING