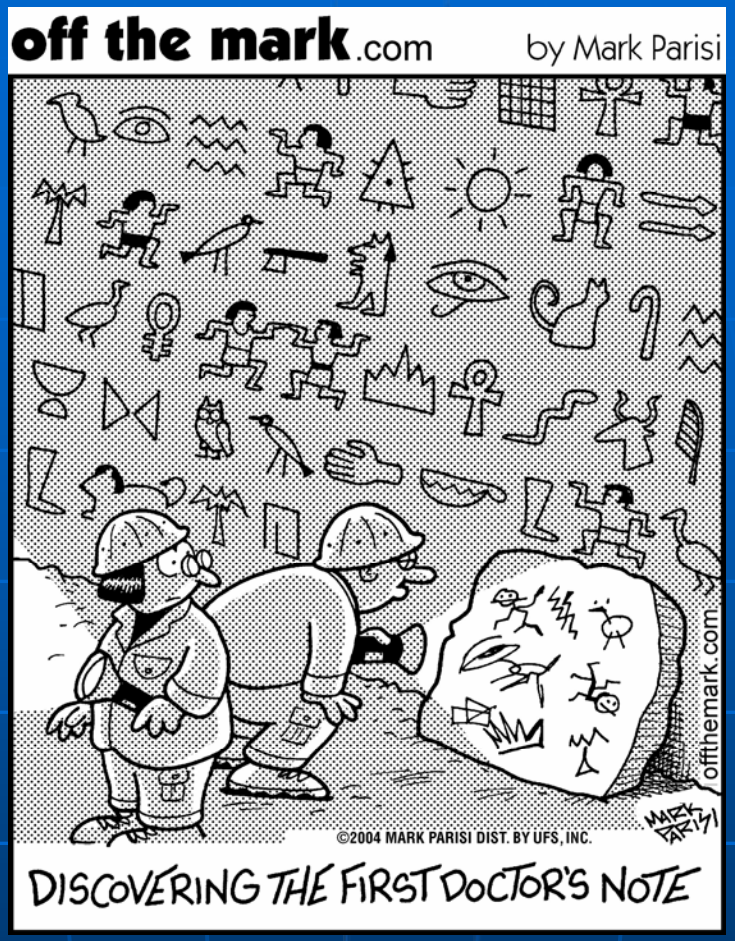
A framework to extract, interpret and structure clinical information from free-texts

Gisele Azambuja Tatjana Zrimec

Centre for Health Informatics
University of New South Wales
Sydney
AUSTRALIA

Andrew Hopkins

Cardiology Department
The Liverpool Hospital
Sydney
AUSTRALIA



"For better or worse, most clinical data accessible to computer processing is still in the form of unstructured natural language."

(Peter Szolovits 2006 http://groups.csail.mit.edu/medg/projects/text/)

Reasons for having clinical data as free text

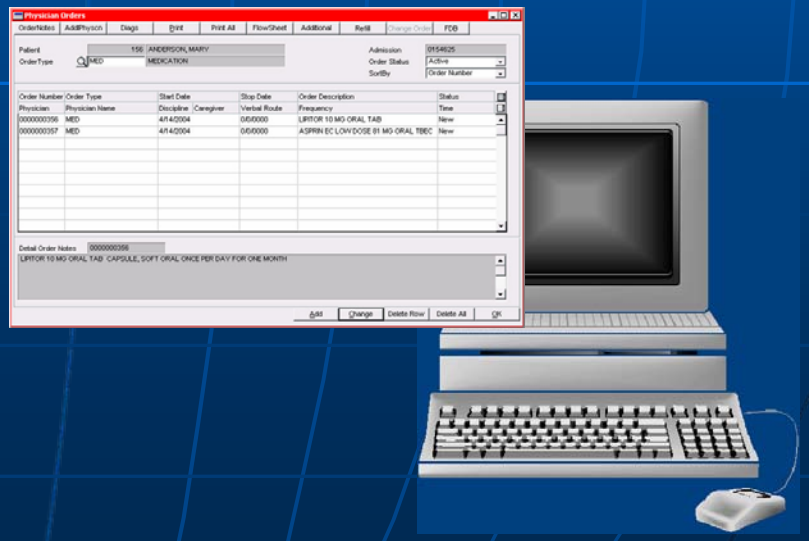
- Original paper-based medical records have been transferred to electronic systems;
- Lack of formalization and standardization;
- Textual Literature – main source of knowledge;
- Computerized Clinical Systems failed to provide structures for clinical information with same degree of expressiveness than free-text.

Some facts about the problem

Conventional medical narratives are commonly stored in existing electronic medical records as free-text clinical notes.

Searching for information in free-text is labour-intensive and demanding task

Computer applications require electronic data to be available in a structured format



To provide a framework and devise automatic methods for extracting relevant clinical information from clinical free-text reports.

Requirements

The system should be applicable to different medical domains

Extracted information should be represented in a form that can be easily accessed by

- medical professionals and
- other computer applications

Approach

We assume that physicians have 5 major questions in mind when reading free-text clinical notes:

What?

Where?

When?

Why?

Outcomes?

Approach

In our schema for representing clinical information, these questions are related to general categories of information.

What => Event

Where=>Place

When=>Time

Why=>Reason

Outcomes=>Outcome

We translated the categories into semantic types from Unified Medical Language System (UMLS).

Event=>"Health Care Activity"

Place=>"Health Care Related Organization"

Time=> "Temporal Concept"

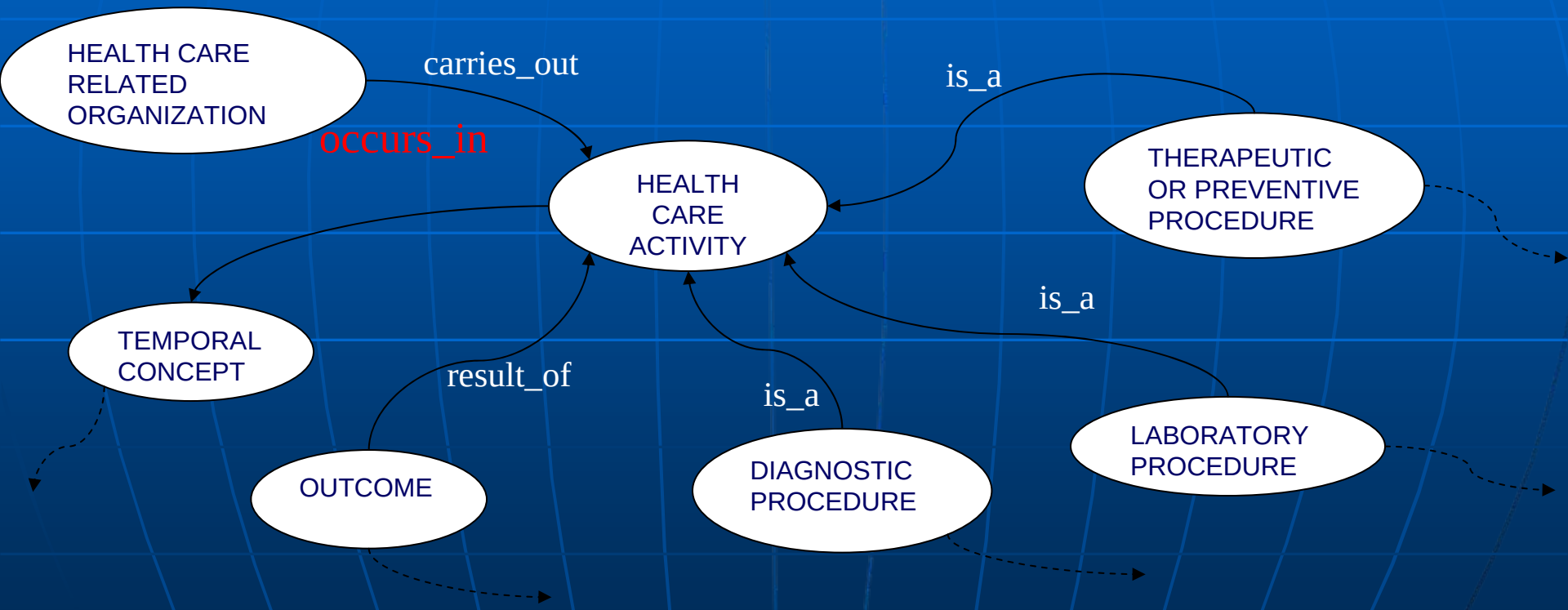
Reason=> " Disease or Syndrome " and "Sign or Symptom"

Outcome=> …..

Some classes like "Outcome" are not related to UMLS semantic types.

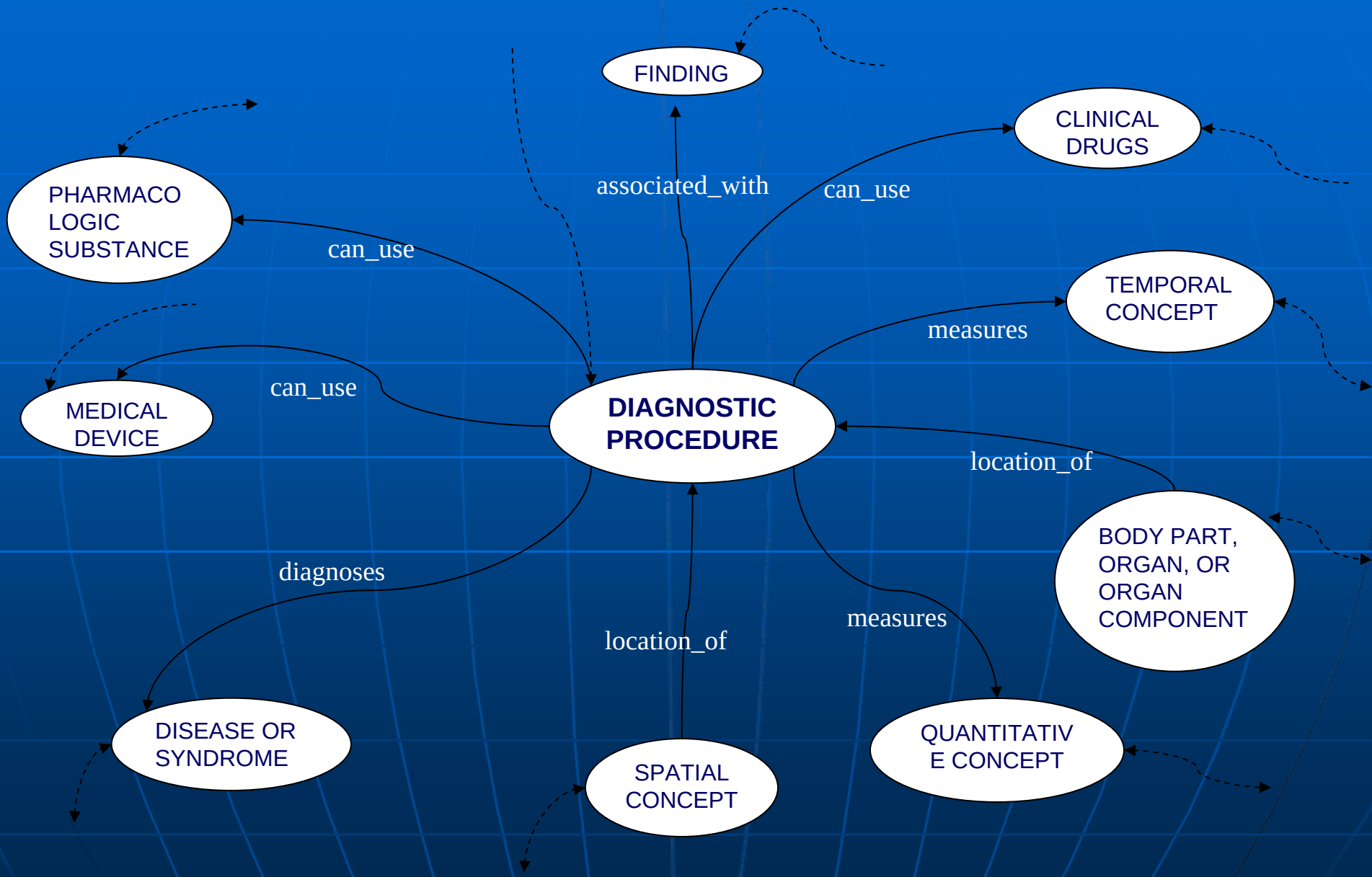
A structure for representing information from free-text clinical notes

The schema inherited predicates existing in the UMLS Semantic Network for classes related to UMLS semantic types.



We also added some relationships between classes that are well known as existing in free-text medical narratives, even though they do not exist in the UMLS Semantic Network.

Diagnostic Procedure - component



Results

Created a schema for representing information from clinical notes using an expandable structure which consists of

- classes of information,
- relationships among them and
- a set of attributes for each class of information.

[HEALTH CARE ACTIVITY]
(*occurs_in*) [TEMPORAL CONCEPT]
[HEALTH CARE RELATED ORGANIZATION] (*carries_out*)
[LABORATORY PROC] | [DIAGNOSTIC PROC] | [THERAPEUTIC OR PREVENTIVE PROC] (*is_a*)
[OUTCOME] (*result_of*)

[DIAGNOSTIC PROCEDURE]
[BODY PART, ORGAN, OR ORGAN COMPONENT] (*location_of*)
[SPATIAL CONCEPT] (*location_of*)
(*diagnoses*) [DISEASE OR SYNDROME]
(*measures*) [TEMPORAL CONCEPT]
(*measures*) [QUANTITATIVE CONCEPT]
(*can_uses*) [CLINICAL DRUGS]
(*can_uses*) [PHARMACOLOGIC SUBSTANCE]
(*can_uses*) [MEDICAL DEVICE]
[FINDING] (*associated_with*)

[BODY PART, ORGAN, ORGAN COMPONENT]
[SPATIAL CONCEPT] (*location_of*)
[QUALITATIVE CONCEPT] (*associated_with*)

[DISEASE OR SYNDROME]
[BODY PART, ORGAN, ORGAN COMPONENT]
(*location_of*)
[SPATIAL CONCEPT] (*location_of*)
[QUANTITATIVE CONCEPT] (*associated_with*)
[FINDING] [FIND] (*associated_with*)

Examples of the schema components

Current status

- Implementation - the schema represents 19 different entities and 34 relationships among them
- Representation - evaluating the schema
- Information extraction - applying automated techniques to discover rules for the processing task

Contact details:

Gisele Azambuja
gisele@student.unsw.edu.au
Tatjana Zrimec
tatjana@cse.unsw.edu.au

Andrew Hopkins
Andrew.Hopkins@sswahs.nsw.gov.au