

**OWL model of
Clinical Trial Eligibility Criteria
Compatible
with Partially-Known Information**

Olivier Dameron

IRISA équipe Dyliss
Université de Rennes 1, France

<http://www.irisa.fr/dyliss/olivier.dameron>

Outline

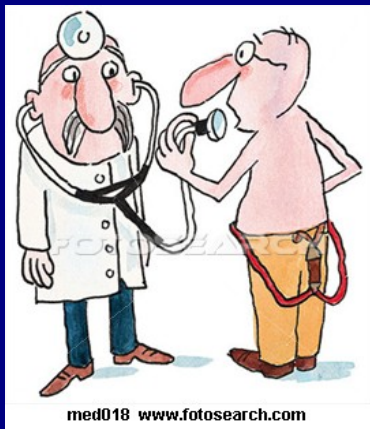
- Context
 - Clinical trial eligibility
 - Missing information leads to over-rejection
- Eligibility criteria design pattern
 - Compatible with missing information
 - Based on open world assumption
- Validation on simulated data
- Evaluation on real CT and patients

Context

Patient recruitment in clinical trials

- Clinical trial are important
 - For patients
 - For research
- Patient recruitment rate in cancer CT is low
 - Currently around 3% (France, all cancers)
 - Goal = 10% by 2013

Clinical trial recruitment workflow



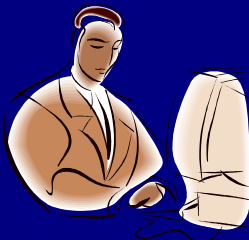
MDM Secretary



MDM Report



SCREENING



Check up



Inclusion (Y/N)



PRE-SCREENING

Potential inclusion (Y/N)



MDM report is displayed during MDM

Patient recruitment in clinical trials

- **Bottlenecks**

- Number of clinical trials
 - www.clinicaltrials.gov: 136.447 CT
- Number and complexity of eligibility criteria
- Therapy decision during multidisciplinary meeting
 - Time is a major constraint
 - Inclusion decision is a supplementary workload for physicians

- **Requirements**

- Automation
 - Handling lots of data
 - Handling complex data and complex criteria
- Ontologies
 - Reconciling data with abstract eligibility criteria
- Logics
 - Combining criteria

Patient recruitment in clinical trials

Current situation:

- EHR are gaining acceptance
 - Brittany: ~15,000 records
 - Structured and standardized information ++

Assumption

- Data contained in EHR is valuable for facilitating clinical trial recruitment

ASTEC project (2009 – 2012)

- Patients with prostate cancer
 - Enough patients
 - Enough data for each patient
 - Recruitment was already > national average
- Showed that ontologies are useful for
 - Data integration (treatments, antecedents...)
 - Data - criteria reconciliation
- What is left:
 - Combine criteria...
 - ... even if some required information is missing

Modeling eligibility criteria

Eligibility criteria

Patient eligible iff:

- All inclusion criteria (I_i) are met
- None of the exclusion criteria (E_j) are met

$$\text{patient eligible} \Leftrightarrow \left(\bigwedge_{i=0}^n I_i(p) \right) \wedge \left(\bigwedge_{j=0}^m \neg E_j(p) \right)$$

E.g: patient eligible $\Leftrightarrow I_0(p) \wedge I_1(p) \wedge (\neg E_0(p)) \wedge (\neg E_1(p))$

Inclusion

- I_0 : prostate adenocarcinoma
- I_1 : absence of metastasis

Exclusion

- E_0 : patient > 70 y.o.
- E_1 : diabetes

The problem of partially-known information

Missing information => failed eligibility

- When some of the criteria are not known, the conjunction can never be true

$$\text{patient eligible} \Leftrightarrow I_0(p) \wedge I_1(p) \wedge (\neg E_0(p)) \wedge (\neg E_1(p))$$

- The inability to assert or reject a criterion should not necessarily result in ineligibility
- Distinction between:
 - Formal inclusion
 - Formal rejection
 - Proposal (pending further information)

Patient	I_0	I_1	E_0	E_1	$I_0 \wedge I_1 \wedge \neg E_0 \wedge \neg E_1$	Decision
p_0	T	T	F	T	F	Exclude (E_1)
p_1	T	T	F	F	T	Include
p_2	T	T	F	?	F cannot assert $\neg E_1$	Propose (assume $\neg E_1$)
p_3	T	F	F	T	F	Exclude (both $\neg I_1$ and E_1)
p_4	T	F	F	F	F	Exclude ($\neg I_1$)
p_5	T	F	F	?	F	Exclude ($\neg I_1$)
p_6	T	?	F	T	F	Exclude (E_1)
p_7	T	?	F	F	F cannot assert I_1	Propose (assume I_1)
p_8	T	?	F	?	F cannot assert I_1 cannot assert $\neg E_1$	Propose (assume both I_1 and $\neg E_1$)



Missing information

Out of 9 combinations of True, False and Unknown:

- 4: all information available (p0, p1, p3, p4)
- 2: some missing information but what is know is enough to infer exclusion (p5, p6)
 - Failure to meet an inclusion criterion (p5)
 - An exclusion criterion is met (p6)
- 3: some missing information but no conclusion can be inferred from the available information (p2, p7, p8)

How much information is missing?

- 286 cases examined in Rennes' urology MDM october 2008 → march 2009
- MDM file: 59 fields supporting distinction btw known and unknown values
- 58.6% of values were unknown
- Average for each case: 34.6 fields unknown
 - “best” case: 19 fields unknown
- 54 fields (91.5%) were unknown in at least 1 case
 - Systematically known: patient ID, MDM date, gender, tumor anatomic site, primary histol. site

Does missing information really matter?

- For a particular CT, some of the 59 fields are not relevant
- For a patient with some fields unknown, maybe what we know is enough to infer exclusion correctly

Does missing information really matter?

	CT1	CT2	CT3	CT4
Nb inclusion fields	15	19	17	10
Nb exclusion fields	10	9	13	11
Unkn. Values proportion	50%	62%	53%	43%
Nb patients with all values known	0	0	0	0
Nb eligible patients (orig)	17	0	7	2
Nb eligible patients (review)	30	23	7	2

Assuming the value of missing items is tricky

- CT1:
 - I0 = presence of X
 - I1 = absence of Y
- If X and Y values are unknown, why assume
 - That X is true...
 - ... but that Y is false?
 - X and Y assumed values depend on the kind of criterion, not even on X and Y probability
- Even worse: CT2
 - I0 = absence of X
 - Now we should assume that X is false whereas we assumed that it was true for CT1

Closed / open world reasoning

- Closed world reasoning
 - Assumes exhaustive description
 - Negation as failure: if you cannot prove that it is true (or if you cannot find it in the table) then it is false
 - “Classical” reasoning, typically used in databases,...
- Open world reasoning
 - Assumes (potentially) incomplete description
 - A statement's truth value is independent of whether or not it is known to be true
 - Typically used with OWL

Open vs closed world assumption (1/3)

- Is Albert Einstein a Nobel prize winner?
- Is Mozart?

Open vs closed world assumption (2/3)

- Is Albert Einstein a Nobel prize winner?
- Is Mozart?
- Is Pierre Chasin-Dupuis?

Open vs closed world assumption (3/3)

- Is Albert Einstein a Nobel prize winner?
- Is Mozart?
- Is Pierre Chasin-Dupuis?

(exact, but sheer luck)

Closed world assumption => yes, no, no

Open world assumption => yes, no, ?

(requires additional knowledge :-)

Result: EC Design Pattern

Modeling criteria

- Each criterion (inclusion and exclusion alike) represented by class C_i
 - inclusion and exclusion alike
 - class of the patients meeting the criterion
 - how one actual patient may become an instance of C_i is not important at this point (asserted, defined class, SWRL rule,...)
- The negation of each criterion is represented explicitly by class $\text{Not_}C_i$
 - defined class: $\text{Not_}C_i = (\text{Criterion and not } C_i)$
 - Class of the patients known not to meet C_i
 - shortcut that makes modeling CT easier

clinicalTrial Protégé 3.4.4 (file:/home/olivier/recherche/astec/unknownInformation/clinicalTrial.pprj, OWL / RDF Files)

File Edit Project OWL Reasoning Code Tools Window Collaboration Help

Metadata(clinicalTrial.owl) OWLClasses Properties Individuals Forms

SUBCLASS EXPLORER

For Project: clinicalTrial-paolo

Asserted Hierarchy

- owl:Thing
 - ClinicalTrial
 - Criterion
 - E_0
 - E_1
 - L_0
 - L_1
 - Not_E_0
 - Not_E_1
 - Not_I_0
 - Not_I_1
 - Patient

CLASS EDITOR for Not_I_0 (instance of owl:Class)

For Class: http://www.u936.univ-rennes1.fr/dameron/ontologies/clinicalTrial.owl#Not_I_0 Inferred View

Property	Value	Lang
rdfs:comment		

Annotations

Asserted Conditions

- Criterion — NECESSARY & SUFFICIENT
- not I_0 — NECESSARY

Disjoints

Logic View Properties View

Modeling clinical trials: inclusion

- Each clinical trial is represented by a CT_k class
- CT_k_include subclass of CT_k
 - Defined class:
C_0 and C_1 and Not_C_2 and Not_C_3
 - Class of the patients who are known to be formally eligible

Modeling clinical trials: exclusion

- CT_k_exclude subclass of CT_k
 - Placeholder: class of the patients who are known to be not eligible
- CT_k_exclude_at_least_one_exclusion_criterion
 - Subclass of CT_k_exclude
 - Class of the patients who meet at least 1 excl. crit.
 - Defined class: C_2 or C_3
- CT_k_exclude_at_least_one_failed_inclusion_cr.
 - Subclass of CT_k_exclude
 - Class of the patients who fail to meet ≥ 1 incl. Crit
 - Defined class: Not_C_1 or Not_C_2

Modeling clinical trials: reasoning

- Patients classified as instances of CT_k include are known to be eligible for CT_k
- Patients classified as instances of CT_k_exclude are known to be not eligible for CT_k
 - And we even know which criteria caused rejection
- Patients classified neither as instances of CT_k_include nor as instances of CT_k_exclude
 - May be eligible: what is known
 - Is compatible with the inclusion requirements
 - Does not meet any exclusion condition
 - And we even know which criteria are unknown ²⁸

Validation

Validation

- Fictitious clinical trial
 - 2 inclusion criteria: I0 and I1
 - 2 exclusion criteria: E0 and E1
- Data set: 9 fictitious patients
 - I0 always true
 - E0 always false
 - I1 and E0 cover all combinations of True, False, ?
- Results identical to table slide #13

clinicalTrial Protégé 3.4.4 (file:/home/olivier/recherche/astec/unknownInfo)

File Edit Project OWL Reasoning Code Tools Window Collaboration Help

Metadata(clinicalTrial.owl) OWLClasses Properties Individuals Forms

SUBCLASS EXPLORER
For Project: clinicalTrial-paolo

Asserted Hierarchy

- owl:Thing
 - ClinicalTrial
 - Ct_DL
 - Ct_DL_exclude
 - Ct_DL_include**
 - Ct_DL_propose
 - Criterion
 - E_0
 - E_1
 - L_0
 - L_1
 - Not_E_0
 - Not_E_1
 - Not_L_0
 - Not_L_1
 - Patient

CLASS EDITOR for Ct_DL_include (instance of owl:Class)

For Class: www.u936.univ-rennes1.fr/dameron/ontologies/clinicalTrial.owl#Ct_DL_include Inferred View

Property	Value
rdfs:comment	

L_0
 L_1
 not E_0
 not E_1

Ct_DL
 Ct_DL_include or Ct_DL_exclude or Ct_DL_propose [fr]

Ct_DL_exclude

Logic View Properties View

Class Hierarchy

- owl:Thing (0 / 9)
 - ClinicalTrial (0 / 9)
 - Ct_DL (0 / 9)
 - Ct_DL_exclude (0 / 5)
 - Ct_DL_include (0 / 1)**
 - Ct_DL_propose

Asserted **Inferred**

Inferred Instances

- p1

clinicalTrial Protégé 3.4.4 (file:/home/olivier/recherche/astec/)

File Edit Project OWL Reasoning Code Tools Window Collaboration

Metadata(clinicalTrial.owl) OWLClasses Properties Individuals

SUBCLASS EXPLORER

For Project: clinicalTrial-paolo

Asserted Hierarchy

- owl:Thing
 - ClinicalTrial
 - Ct_DL
 - Ct_DL_exclude
 - At_least_one_exclusion_criterion
 - At_least_one_failed_inclusion_criterion
 - Ct_DL_include
 - Ct_DL_propose
 - Criterion
 - E_0
 - E_1
 - L_0
 - L_1
 - Not_E_0
 - Not_E_1
 - Not_L_0
 - Not_L_1
 - Patient

CLASS HIERARCHY

- owl:Thing (0 / 9)
 - ClinicalTrial (0 / 9)
 - Ct_DL (0 / 9)
 - Ct_DL_exclude (0 / 5)
 - At_least_one_exclusion_criterion (0 / 3)
 - At_least_one_failed_inclusion_criterion (0 / 3)
 - Ct_DL_include (0 / 1)
 - Ct_DL_propose

CLASS EDITOR for At_least_one_exclusion_criterion (instance of owl:Class)

For Class: fr/dameron/ontologies/clinicalTrial.owl#At_least_one_exclusion_criterion Inferred View

Property	Value
rdfs:comment	

E_0 or E_1

Ct_DL_exclude

Ct_DL_include or Ct_DL_exclude or Ct_DL_propose

Logic View Properties View

clinicalTrial Protégé 3.4.4 (file:/home/olivier/recherche/astec/)

File Edit Project OWL Reasoning Code Tools Window Collaboration

Metadata(clinicalTrial.owl) OWLClasses Properties Individuals

SUBCLASS EXPLORER

For Project: clinicalTrial-paolo

Asserted Hierarchy

- owl:Thing
 - ClinicalTrial
 - Ct_DL
 - Ct_DL_exclude
 - At_least_one_exclusion_criterion
 - At_least_one_failed_inclusion_criterion
 - Ct_DL_include
 - Ct_DL_propose
 - Criterion
 - E_0
 - E_1
 - L_0
 - L_1
 - Not_E_0
 - Not_E_1
 - Not_L_0
 - Not_L_1
 - Patient

Class Hierarchy

- owl:Thing (0 / 9)
 - ClinicalTrial (0 / 9)
 - Ct_DL (0 / 9)
 - Ct_DL_exclude (0 / 5)
 - At_least_one_exclusion_criterion (0 / 3)
 - At_least_one_failed_inclusion_criterion (0 / 3)
 - Ct_DL_include (0 / 1)
 - Ct_DL_propose

Asserted Inferred

Inferred Instances

- p3
- p4
- p5

CLASS EDITOR for At_least_one_failed_inclusion_criterion (instance of owl:Class)

For Class: dameron/ontologies/clinicalTrial.owl#At_least_one_failed_inclusion_criterion Inferred View

Property	Value
rdfs:comment	

Not_I_0 or Not_I_1
 Ct_DL_exclude
 Ct_DL_include or Ct_DL_exclude or Ct_DL_propose

Logic View Properties View

Evaluation

Evaluation

- We evaluated our design pattern on:
 - The 286 prostate cancer cases
 - Clinical trial CT1 (getug14);
CT2, CT3 and CT4 pending
- Results:
 - 0 patients formally eligible
 - 170 patients excluded
 - 116 patients potentially eligible

TP	FP
30	86
FN	TN
0	170

Precision: 0.26
Recall: 1.0

Discussion

Discussion

- Among the 86 false positive cases:
 - The expert agreed that they were all “theoretically” eligible
 - Non-official criteria (eg rejection because Gleason score too low)
- Additional inclusion criteria Gleason > 6
 - 54 patients potentially eligible
 - 25 TP
 - 5 FN
 - 29 FP
 - 8: new results indicated that they did not have cancer
 - 3: rejected because too much information was missing
 - 4: surgery preferred treatment

Discussion

- Missing information could be handled by negation as failure with ad-hoc conversion btw inclusion and exclusion criteria
 - Inclusion “absence of ischemic heart disease” → exclusion “presence of ischemic heart d.”
 - But what if “absence of ischemic heart disease” had been an exclusion criterion?
 - Converted to inclusion “presence of IHD”
 - All the patients with no mention of IHD would be rejected
- Use negation as failure to model exclusion rather than inclusion

Conclusion

- Automatic assistance needed for increasing patient enrollment into clinical trials
- Ignoring missing information leads to over-rejection
- We proposed an OWL design pattern for eligibility criteria capable of determining
 - Formal inclusion
 - Formal exclusion
 - Potential inclusion
- Our system correctly predicts exclusion and leaves more time to examine the other cases