Creating an ontology driven rule base for an expert system in medical diagnosis

Valérie BERTAUD GOUNOT\textsuperscript{a} (Presenter) DMD PhD
Valéry DONFACK\textsuperscript{a}, Jérémy LASBLEIZ\textsuperscript{a} MD,
Annabel BOURDE\textsuperscript{a}, Régis DUVAUFERRIER\textsuperscript{a} MD PhD

\textsuperscript{a} INSERM unit U936, Biomedical Knowledge Modeling, IFR 140, Faculty of Medicine, University of Rennes 1, France

Monday, August 29th 2011, OSLO
INTRODUCTION

• Expert systems of the 1980s have failed on the difficulties of maintenance of large rule bases

• Ontologies:
  – Graphical display of the ontology (Protégé)
  – Easy modification
  – Integrated tools for consistency checking (Pellet)

• Ontology ≈ knowledge base

• Aim:
  – show how semantic web tools can help
  – propose a method to achieve and maintain rule bases grounded on ontologies
MATERIAL

• NCI-T (v10.7): 80,000 classes, 187 properties (relations), 57,000 restrictions
• Protege 3.4.2
• OWL1 and SWRL languages
• OWL and Jess API’s
• Pellet reasonner
METHODS

(1) Extracting a subontology (plasma cell disease) from de NCI-T

(2) Dealing with « may have » and « exclude » relationships

(3) Adding a SWRL rule layer for abductive reasonning

(4) Automatic creation of instances and assertions

(5) Driving a user interface from the ontology to capture facts (signs observed in one patient)
METHODS

• Abductive reasoning

  – Ontology classifiers (Pellet, FACT ...): deductive reasoning.
    • IF (a→b) and (a is true), THEN (b is true)

  – In medicine, even if information are missing we need to get diagnostic hypotheses from the system. Diagnostic process needs abductive reasoning
    • IF (a→b) and (b is true), THEN (a is possibly true)

→ SWRL rules to introduce abductive reasoning
RESULTS

(1) Extracting a sub-ontology
281 classes, 17 relations and 25 restrictions.
taxonomy of PLASMA_CELL_NEOPLASMs

(2) Dealing with « may-have » relationships

→ Re-organizing the relationships in the NCIT: the relationship "Disease_may_have_finding" subsumes "Disease_has_finding".

N.NOY, S de Coronado, H. Solbrig, G. Fragoso, F. Hartel et M. Musen - Representing the NCIT in OWL DL
RESULTS

(3) Creation of the « prototypal cases »

– Tbox (Terminology Box) – Indolent_Myeloma

– Abox (Assertional Box) – Indolent_Myeloma

OWL API
RESULTS

(4) Adding a SWRL rule layer for abductive reasoning

- OWL Ontology
- Jess engine
- 305 Rules (0 order logic)
- 9 SWRL rules (1st order logic)

(5) Automatic and comprehensive generation of

<table>
<thead>
<tr>
<th>Relation</th>
<th>Sign</th>
<th>Maladie</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finding_Has_Diagnosis</td>
<td></td>
<td>Recurrent_Plasma_Cell_Myeloma_0</td>
</tr>
<tr>
<td>Finding_Exclude_Diagnosis</td>
<td></td>
<td>Plasma_Cell_Leukemia_0</td>
</tr>
<tr>
<td>Finding_Absence_Has_Diagnosis</td>
<td></td>
<td>Refractory_Plasma_Cell_Neoplasm_0</td>
</tr>
<tr>
<td>Finding_Absence_Exclude_Diagnosis</td>
<td></td>
<td>Primary_Amyloidosis_0</td>
</tr>
<tr>
<td>Finding(?y) ∧ Disease_May_Have_Finding(?x, ?y) → Finding_Has_Diagnosis(?y, ?x)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finding(?y) ∧ Disease_Excludes_Finding(?x, ?y) → Finding_Excludes_Diagnosis(?y, ?x)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finding(?y) ∧ Disease_Has_Finding(?x, ?y) → Finding_Absence_Excludes_Diagnosis(?y, ?x)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finding(?y) ∧ Disease_Excludes_Finding(?x, ?y) → Finding_Absence_Has_Diagnosis(?y, ?x)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
RESULTS

(6) Driving a user interface from the ontology to capture facts
RESULTS available at: http://www.med.univ-rennes1.fr/OntoDiag/
<table>
<thead>
<tr>
<th>Signs</th>
<th>Present</th>
<th>Absent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical Signs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aggressive_Clinical_Course</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bone_Pain</td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>Impotence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stable_Disease</td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>Systemic_Disease</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unfavorable_Clinical_Outcome</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gynecomastia</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

List of diagnoses:
- Solitary_Osseous_Plasmacytoma_0 (signs number: 2/6)
- Extramedullary_Plasmacytoma_0 (signs number: 1/4)
- Solitary_Plasmacytoma_0 (signs number: 1/3)
- Plasmacytoma_0 (signs number: 1/2)
- Monoclonal_Gammopathy_of_Undetermined_Significance_0 (signs number: 1/4)
- Stage_III_Multiple_Myeloma_0 (signs number: 1/6)
- Chest_Wall_Solitary_Plasmacytoma_0 (signs number: 1/2)

List of excluded diagnoses:
- Smoldering_Myeloma_0 (signs number: 2)
- Indolent_Myeloma_0 (signs number: 2)
CONCLUSION

• **This process based on a formal ontology allows**
  – Possible re-use of existing ontologies
  – Easy generating a large number of production rules (from the SWRL rules)
  – Easy generating the user entry form
  – Easy knowledge update
  – Availability of tools for consistency checking (like Pellet)

• **SWRL rules allow**
  – Reasoning with “exclude” relationships and absent findings (negative findings)
  – Abductive reasoning

• **Improvements are needed**
  – Better classification of possible diagnoses
  – Entry form developing on demand
  – Stronger evaluation
Than you for your attention


Valerie BERTAUD-GOUNOT, University of Rennes1, FRANCE
e-mail: valerie.bertaud@univ-rennes1.fr
The sub-ontology extractor

(1) Extracting a subontology (plasma cell disease) from de NCI-T

Valerie BERTAUD - MIE 2011
RESULTS

(3) Creation of the « prototypal cases »

– Tbox (Terminology Box) – Indolent Myeloma

– Abox (Assertionnal Box) – Indolent Myeloma
RESULTS

(4) Adding a SWRL rule layer for abductive reasoning

<table>
<thead>
<tr>
<th>OWL Ontology</th>
</tr>
</thead>
<tbody>
<tr>
<td>9 SWRL rules</td>
</tr>
<tr>
<td>(1st order logic)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rule</th>
</tr>
</thead>
</table>
SWRL rules (1st order logic)

OWL Ontology

Jess engine

Rules (0 order logic)

Knowledge base (Database)

User entry form

John SMITH’s possible diagnoses

John SMITH’s findings

METHODS