Comparing Drools and ontology reasoning approaches for telecardiology decision support

Pascal VAN HILLE a, Julie JACQUES b, Julien TAILLARD b, Arnaud ROSIER a,
David DELERUE b, Anita BURGUN a, Olivier DAMERON a

a Inserm UMR936, Université de Rennes 1, 35043 Rennes, France
b ALICANTE SA, 50 rue Philippe de Girard, 59113 Seclin, France
Introduction: Telecardiology

• More and more patients receive cardiac rhythm management device (+27% between 2005 and 2011 in Europe)
• Devices have now remote monitoring capabilities
• Transmission of pre-defined alerts to the physician has the potential to offer improved patient safety and quality of care
Introduction: Problems

• Automatic triage of alerts according to emergency level is necessary to keep up with this overwhelming flow of alerts efficiently.

• This is a difficult task because the risk associated with an alert depends on multiple interdependent factors such as:
  – the patient’s medical history
  – his current pathologies
  – his current treatment
Introduction: The CHA2DS2VASc score

- Most of the alerts are related to episodes of atrial fibrillation (AF)
- Stroke risk is evaluated using CHA2DS2VASc score

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Congestive heart failure/left ventricular dysfunction</td>
<td>1</td>
</tr>
<tr>
<td>Hypertension</td>
<td>1</td>
</tr>
<tr>
<td>Age&gt;=75</td>
<td>2</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>1</td>
</tr>
<tr>
<td>Stroke/transient ischemic attack/thromboembolism</td>
<td>2</td>
</tr>
<tr>
<td>Vascular disease (prior myocardial infarction, peripheral artery disease, aortic plaque)</td>
<td>1</td>
</tr>
<tr>
<td>65&lt;=Age&lt;75</td>
<td>1</td>
</tr>
<tr>
<td>Sex category (ie, female gender)</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total</th>
<th>0=&lt; score =&lt;9</th>
</tr>
</thead>
</table>

Introduction: Akenaton project

• The Akenaton project* [1] developed an automatic classification of AF alerts according to their severity
• It hinges on the CHA2DS2VASc score
• The score evaluation requires domain knowledge for reasoning about the patient's clinical context


*The AKENATON project is funded by the French Agence Nationale pour la Recherche (ANR -07-TecSan- 001).
Introduction: The Akenaton prototype

1. reads AF alerts issued by the patient's device
2. extracts relevant data from patient's medical record with natural language processing [2]
3. integrates alert data and patient clinical context
4. reasons on these data in order to determine the patient's CHA2DS2VASc score
5. computes the alert severity level according to AF duration, the CHA2DS2VASc score and the Vitamin K antagonists status

Introduction: Objectives

We compared two approaches for implementing the reasoning module:

1. based on Drools, a business rule management system

2. based on the Web Ontology Language (OWL) and the Semantic Web Rule Language (SWRL)
Methods : Drools-based reasoning module

• Business rule management system with a forward chaining inference based rules engine
• 176 concepts identified with the help of a cardiologist expert
• 14 Drools Expert rules created with a domain-specific language (DSL) composed of 52 instructions
• DSL rules are automatically translated into Java
Methods : Drools-based reasoning module

• One main rule per criterion
• Example: For diabetes criterion

```java
rule "CHA2DS2VASc Diabetes"
  when Patient is diagnosed with diabetes
  Then Patient's CHADS2 criteria "D" is checked
end
```

• Granularity is handled by additional rules

```java
rule "insulin dependent diabetes mellitus"
  when Patient is diagnosed with type 1 diabetes
  Then Patient is diagnosed with diabetes
end
```

```java
rule "non-insulin dependent diabetes mellitus"
  when Patient is diagnosed with type 2 diabetes
  Then Patient is diagnosed with diabetes
end
```
Methods: Ontology-based reasoning module

- Created in java with OWL API
- Reasoning based on a domain ontology in description logic with SWRL rules
- The domain ontology has 840 classes describing diseases, medicines and equipment from the field of telecardiology [4]
- Data from each patient record were represented as instances of the ontology

Methods: Ontology-based reasoning module

- CHA2DS2VASC score was computed using the best of 10 modeling strategies [5]
- Criteria D, C, H, S2, and V which potentially require reference to domain knowledge to be reconciled with patient data were evaluated with the ontology
- Age and sex criteria were evaluated directly in the Java program
- The sum of points is also computed directly in Java

Methods: Ontology-based reasoning module

• Five SWRL rules (one per criterion) were created
• The Pellet reasoner was used to perform the inference
Methods: Ontology-based reasoning module

- SWRL rules are of the form of an implication between an antecedent and consequent

Example: For diabetes criterion

\[
\text{dia-criterion-score}(\text{?region}),
\text{diabetes-mellitus}(\text{?d}),
\text{patient}(\text{?p}),
\text{is-affected-by}(\text{?p, ?d}),
\text{quality-quale}(\text{?p, ?region})
\]\n
\[\rightarrow \text{has-for-integer-value}(\text{?region, 1})\]
Methods: Ontology-based reasoning module

- SWRL takes into account subsumption
- No additional rules are necessary to handle granularity
Methods : Evaluation set of patients

- Correctness of the score and performances were evaluated on a set of 62 patients with a Paradym ICD from the Department of Cardiology at Rennes University Hospital
## Results

<table>
<thead>
<tr>
<th></th>
<th>Drools</th>
<th>OWL+SWRL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correctness of the score</td>
<td>ok</td>
<td>ok</td>
</tr>
</tbody>
</table>

Primitives:
- 14 rules with 52 DSL instructions.
- 5 rules with 11 specifics.

Expressiveness of the rules language:
- Good
- Medium

Editing tools:
- Good

Maintainability:
- Difficult
- Easy

29/08/2012

MIE 2012 - Pascal van Hille - Inserm
UMR936, University of Rennes 1, France
Discussion : performance

• Reasoning time is short enough in both approaches for not being a discriminating factor

• This aspect could be reconsidered if a large number of alerts should be classified in real time

• The performance impact of adding a large number of rules has not been measured and should also be evaluated
Discussion : granularity

• With a finer level of granularity, the number of rules can be much more important using Drools than using the ontology

• This is because SWRL uses the OWL hierarchy to directly take subsumption into account

• In our example, 3 rules are necessary with Drools to take into account "diabetes mellitus" and its subconcepts whereas only one SWRL rule is necessary

• An important number of rules can result in a maintenance issue
Discussion: broader domain

• Whether the results obtained in this study are applicable to a broader domain, or one with a finer granularity remains to be studied
  – The Drools engine performances would probably scale up, but the number and the complexity of the resulting Drools rules would make them difficult to develop and to maintain
  – The ontology-based approach would benefit from modularity and from possible reuse of existing resources, but the performances would probably decrease
• A direction worth exploring would consist in using the ontology for automatically generating some Drools rules
Discussion: ontology advanced reasoning capabilities

• The reasoning underlying the CHA2DS2VASC score determination involves both subsumption and classification
  – Subsumption reconciles the patient's data which are usually precise with the criteria which are more general
  – Classification is necessary to infer whether a pathology is a peripheral artery disease

• These characteristics require advanced reasoning capabilities that favor the OWL and SWRL combination over DROOLS
Conclusion

• Drools and Ontology are two valid approaches with comparable performances for determining the CHA2DS2VASc score using domain knowledge

• Reasoning based on ontology has the advantage of being simpler to implement and to maintain since:
  – the granularity of information is managed in the ontology and not by rules
  – domain knowledge (in the ontology) is separated from operational knowledge (in the java code)

• Reasoning based on Drools provides greater expressiveness than OWL + SWRL
Thank you for your attention.

pascal.vanhille@univ-rennes1.fr