A Comparative Evaluation of an Ontological Medical Decision Support System (OMeD)

John A. Doucette, Atif Khan and Robin Cohen

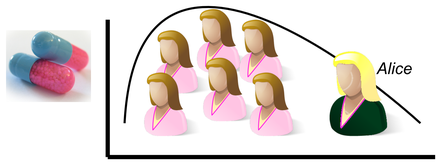
WATERLOO
CHERITON SCHOOL OF COMPUTER SCIENCE

Introduction

OMeD is a **knowledge-based medical decision support** framework that **utilizes semantic web techniques** for knowledge representation and reasoning.

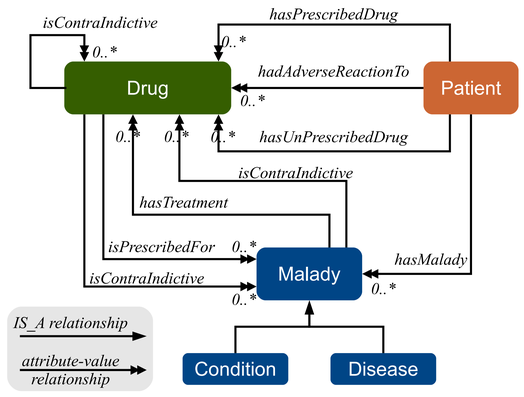
We present a **proof-of-concept implementation** of OMeD and **compare** it to a set of standard **machine learning techniques** across a series of benchmarks based on simulated patient data.

Line of inquiry: *"Should **Alice** be treated with drug X?"*



Semantic Model, Inference Rules & Semantic Proof

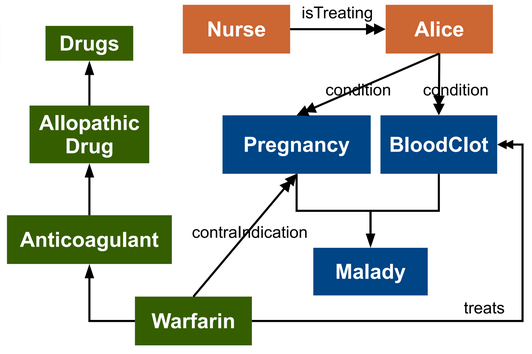
Semantic Model
(for Drug Prescription Line of Inquiry)



Query

*Can we administer **Warfarin** to **Alice**?*

Knowledge Graph



Inference Rule

If a drug is contra-indictive of a condition, and some entity has that condition, then that entity cannot be given the drug.

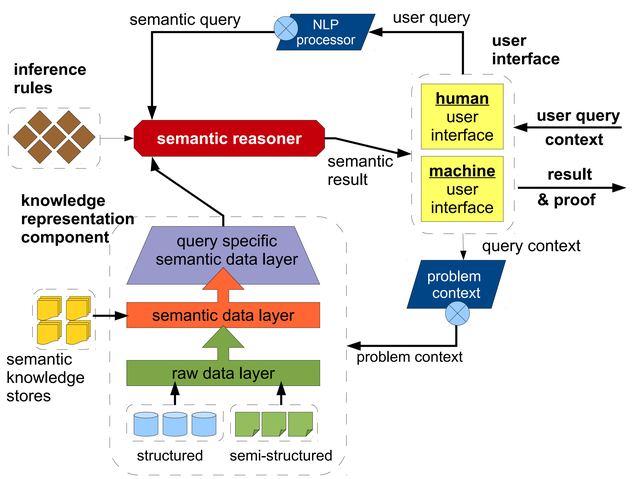
```
{?ANY  :condition ?COND.
 ?DRUG :contraIndication ?COND.
}=>{?ANY :canNotBeGiven ?DRUG }
```

Semantic Proof

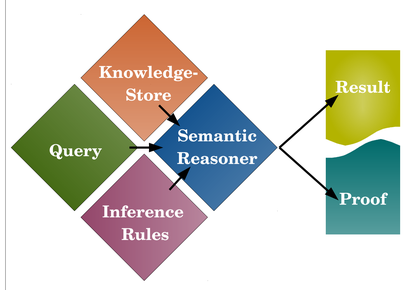
```
{{:Alice :condition :Pregnancy} e:evidence <kb#_27>.
{:Warfarin :contraIndication :Pregnancy} e:evidence <kb#_22>

}=>{
 {:Alice :canNotBeGiven :Warfarin} e:evidence <rules#_9>
}.

# Proof found in 3 steps (2970 steps/sec) using 1 engine
(18 triples) }.
```

Framework Design

Architecture & Components



Semantic Knowledge-based Decision Making



Result can be easily explained to end users as it is based on domain based concepts and terminology

Proof is a First Order Logic proof that can be independently verified

Framework Evaluation — OMeD Vs. Machine Learning Techniques

Simulated Data

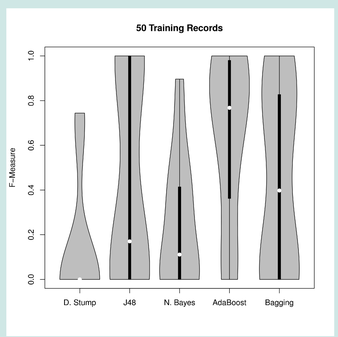
Abnormally dense patient records
Gamma-distributed attributes
Class assigned by OMeD

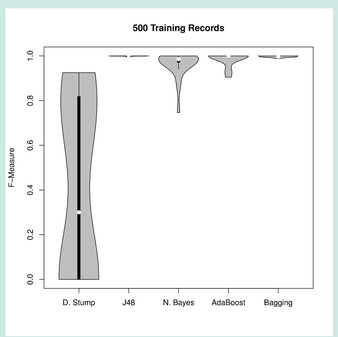
Classifiers

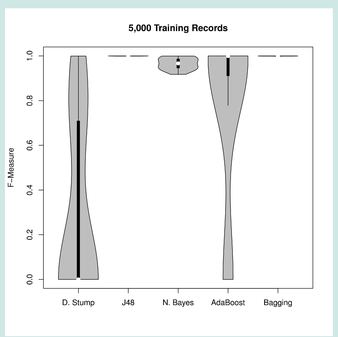
A single layer decision tree, **Decision Stump**
An IG-based decision tree, **C4.5 R8 (J48)**
A bootstrapping meta-classifier, **Bagging**
An iterative boosting meta-classifier, **AdaBoost**

Experiment 1

Goal: find the best classifier
20 drugs, 20 conditions
5, 50, 500, and 5,000 training records
F1 measure (balanced precision and recall)
J48 least affected by data poverty

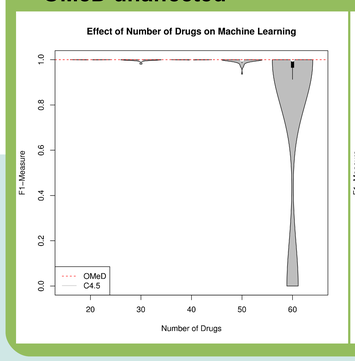


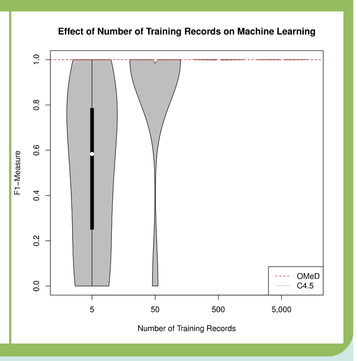




Experiment 2

Goal: measure effect of increasing sparsity on J48
Add more drugs (more features)
500 training records, 20 conditions, 20, 30, 40, 50, and 60 drugs
F1 measure (balanced precision & recall)
J48 strongly affected by increase in sparsity
OMeD unaffected





No Free Lunch (NFL) Theorem

*"in the **absence of prior knowledge** about the properties of the function, **all possible strategies** for optimization **must perform** precisely the **same on average**"* **[1]**

ML & Decision Support Systems

- Decision support systems should be reliable, interpretable and verifiable.
- ML suffers from fundamental unreliability due to NFL.
- OMeD does not rely on empirical optimization and is not subject to NFL.
- OMeD is also easy to interpret and produces verifiable results.

[1] Wolpert, D.H., and Macready, W.G. ***No Free Lunch Theorems for Optimization***, IEEE Transactions on Evolutionary Computation, 1997

Conclusion

Machine learning techniques perform poorly on simulated patient data, even when it is comparatively dense (each patient took 25% of all drugs).

OMeD's prototype was constructed and verified, demonstrating a working realization of the system.

Recent work supports our findings with real world data and more complex ontologies.

Future work examines combining OMeD with machine learning techniques to produce a ontological DSS which is resistant to noise.