Probabilistic Ontology and Knowledge Fusion for Procurement Fraud Detection in Brazil

Rommel Carvalho, Kathryn Laskey, and Paulo Costa
George Mason University
Marcelo Ladeira, Laécio Santos, and Shou Matsumoto
Universidade de Brasília

Paper - Uncertainty Reasoning for the Semantic Web
URSW - ISWC

Agenda

- Introduction
- MEBN and PR-OWL
- Procurement Fraud Detection
- Conclusion

Introduction

Introduction

- Brazilian Office of the Comptroller General (CGU) primary mission
  - Prevent and detect irregularities (corruption)
    - Gather information from a variety of sources
    - Combine the information
    - Then evaluate whether further action is necessary
- Problem
  - Information explosion
    - Growing Acceleration Program (PAC)
      - 250 billion dollars - 1,000+ projects only in SP

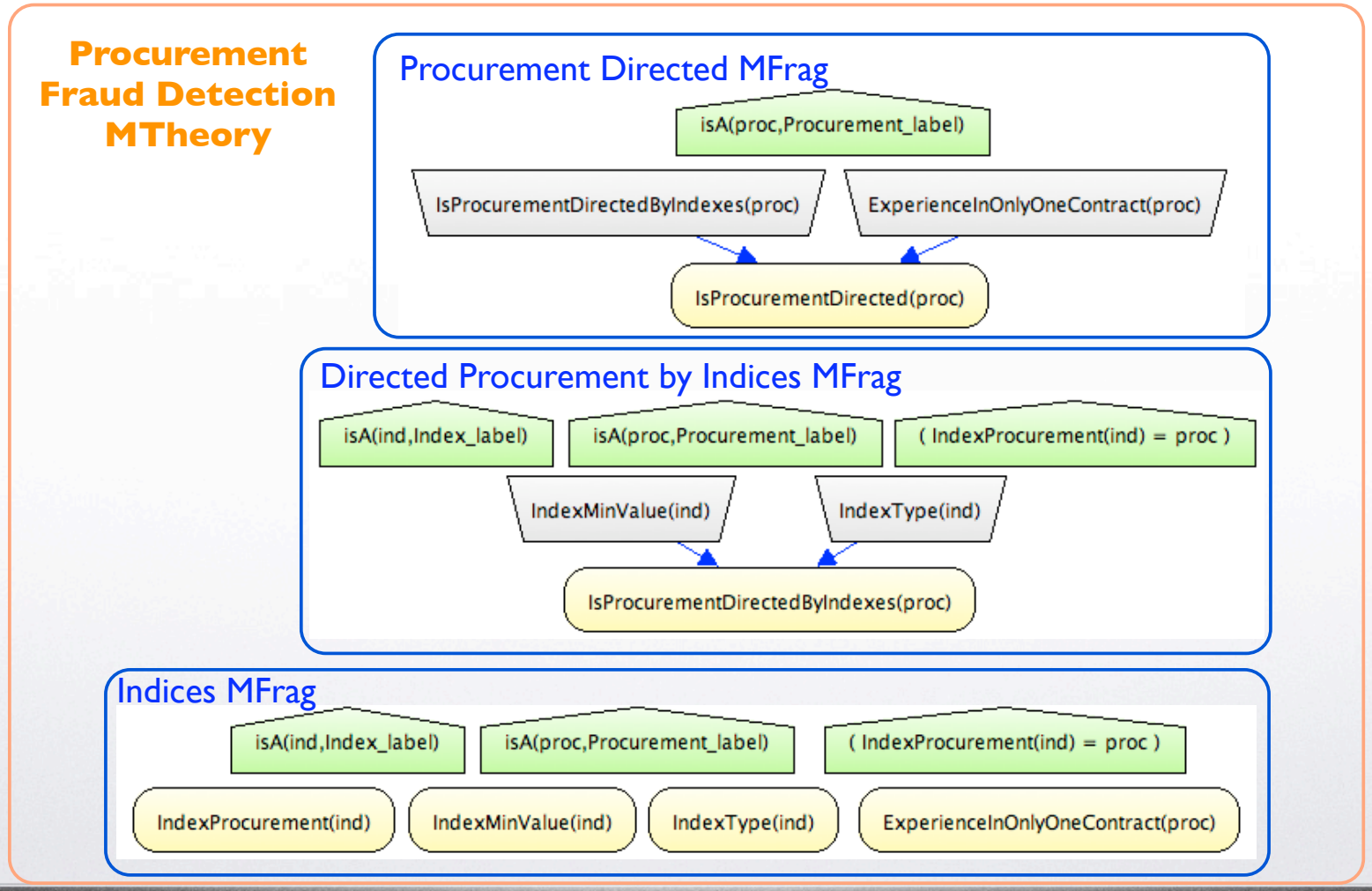
Introduction

- MEBN
  - Represent and reason with uncertainty about any propositions that can be expressed in first-order logic
- PR-OWL
  - Uses MEBN logic to provide a framework for building probabilistic ontologies
- Fraud Detection and Prevention Model
  - Uses MEBN and PR-OWL
  - Proof of concept

MEBN and PR-OWL

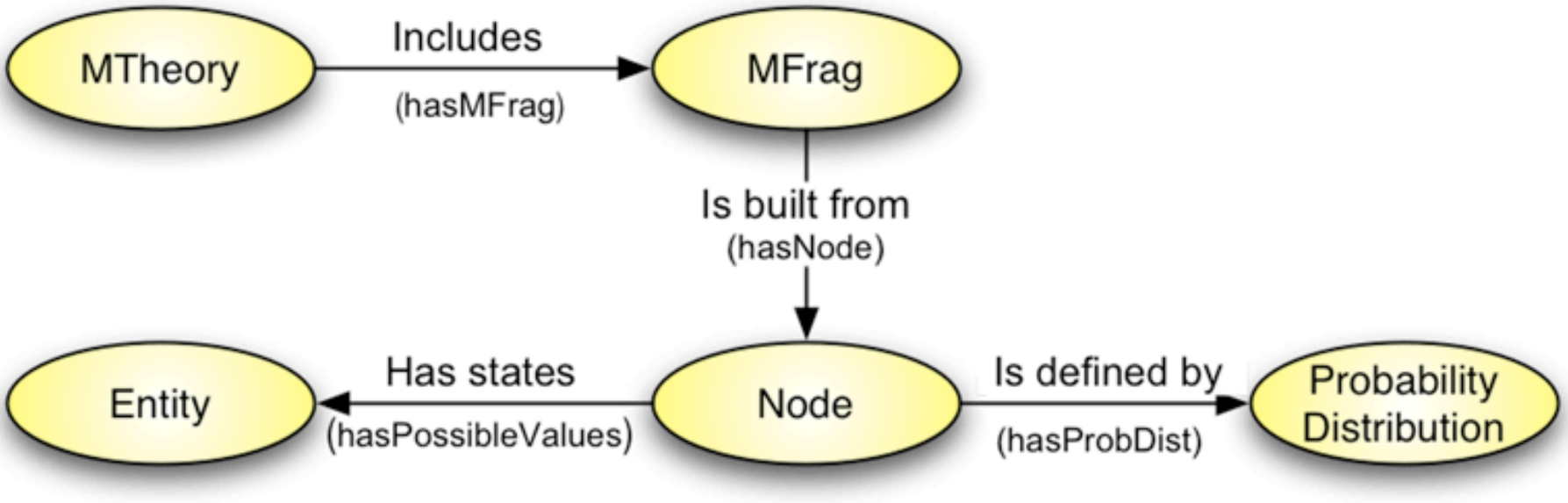
MEBN

- BN + FOL

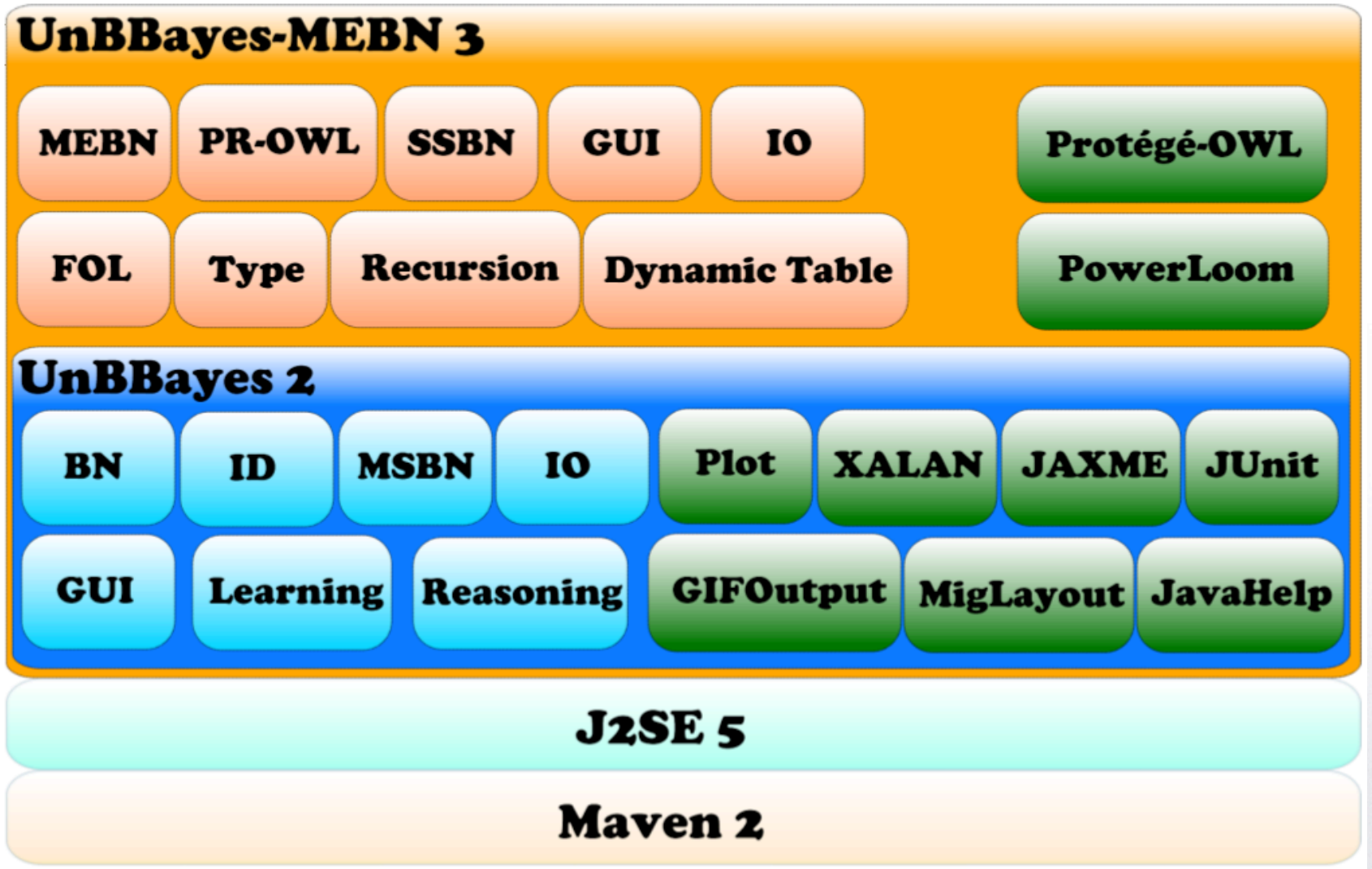

Procurement Fraud Detection MTheory
Procurement Directed MFrag
isA(proc,Procurement_label)
IsProcurementDirectedByIndexes(proc)
ExperienceInOnlyOneContract(proc)
IsProcurementDirected(proc)
Directed Procurement by Indices MFrag
isA(ind,Index_label)
isA(proc,Procurement_label)
(IndexProcurement(ind) = proc)
IndexMinValue(ind)
IndexType(ind)
IsProcurementDirectedByIndexes(proc)
Indices MFrag
isA(ind,Index_label)
isA(proc,Procurement_label)
(IndexProcurement(ind) = proc)
IndexProcurement(ind)
IndexMinValue(ind)
IndexType(ind)
ExperienceInOnlyOneContract(proc)

PR-OWL

- MEBN + OWL



UnBBayes


UnBBayes-MEBN 3
MEBN
PR-OWL
SSBN
GUI
IO
Protégé-OWL
FOL
Type
Recursion
Dynamic Table
PowerLoom
UnBBayes 2
BN
ID
MSBN
IO
Plot
XALAN
JAXME
JUnit
GUI
Learning
Reasoning
GIFOutput
MigLayout
JavaHelp
J2SE 5
Maven 2

Procurement Fraud Detection

Procurement Fraud Detection

- A major source of corruption is the procurement process
  - Laws attempt to ensure a competitive and fair process
  - Perpetrators find ways to turn the process to their advantage while appearing to be legitimate
- Specialist from CGU (Mário Spinelli)
  - Structured the different kinds of procurement frauds found in the past years

Procurement Fraud Detection

- Types of fraud
  - Characterized by criteria
    - Principle of competition is violated when we have
      - Owners who work as a front (usually someone with little or no education)
      - Use of accounting indices that are not common
- Ultimate goal
  - Structure the specialist knowledge in a way that an automated system can reason with the evidence in a manner similar to the specialist
    - Support current specialists
    - Train new ones

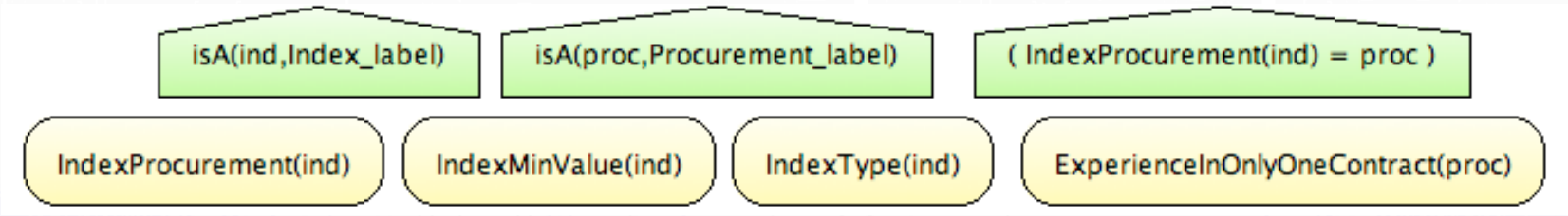
Procurement Fraud Detection

- Realistic goal (this paper)
  - Proof of concept
    - Selected just a few criteria
  - Why Semantic Web?
  - Propose an overall architecture for collecting data, reasoning with uncertainty (model designed), and reporting alerts
  - Ask specialists to analyze results (subjective)
    - No massive data used
  - Show that new criteria can be easily incorporated

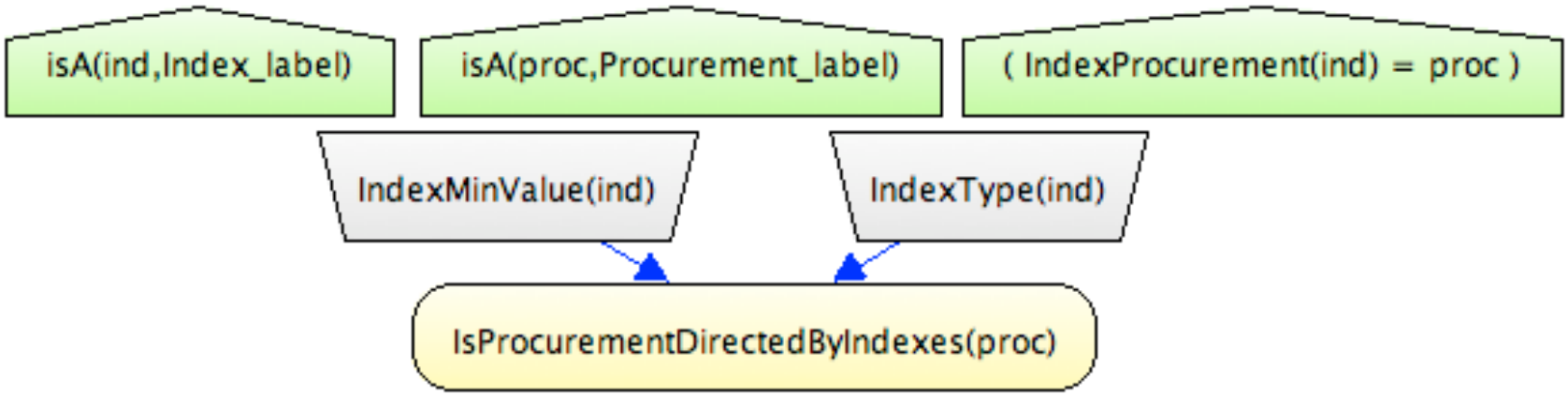
Why Semantic Web?

Architecture

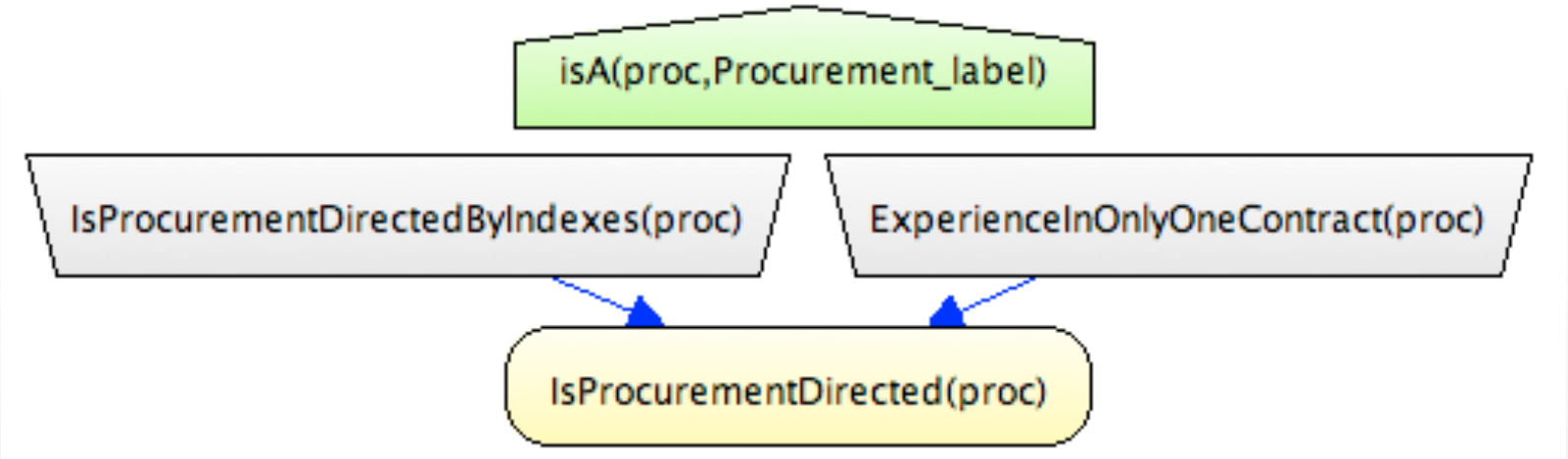
Used Criteria



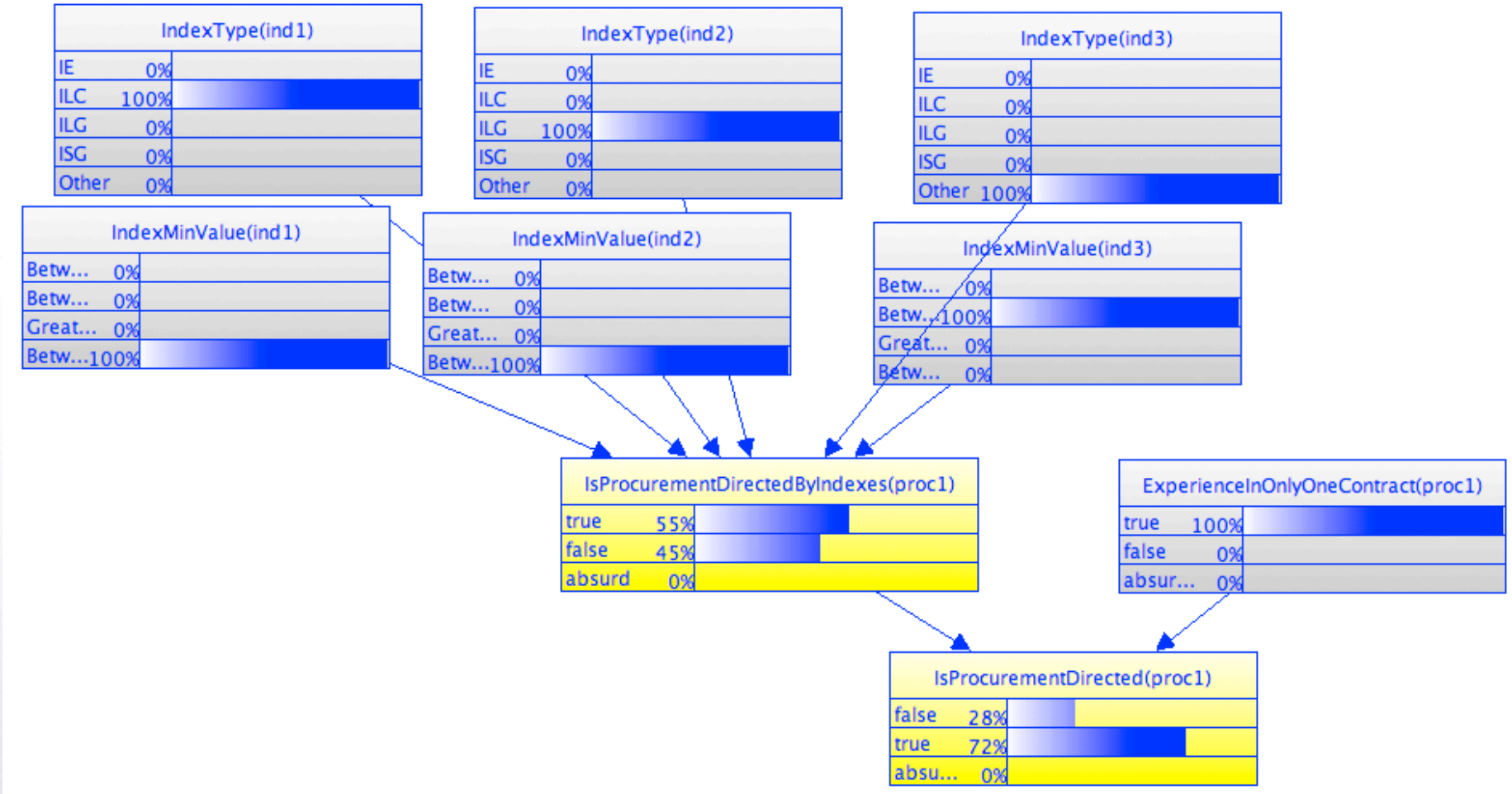
Used Criteria



Used Criteria



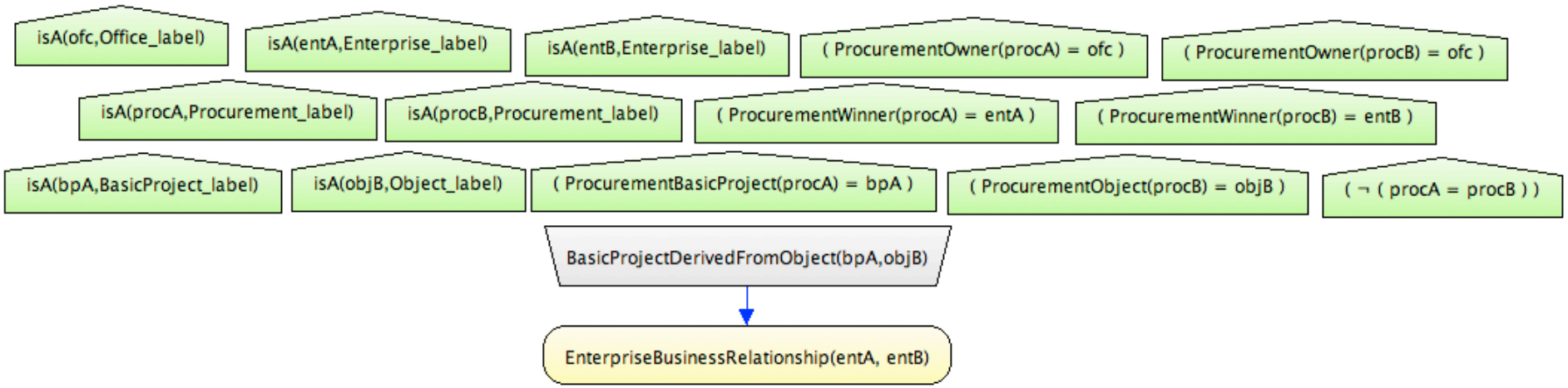
Results


IndexType(ind1)
IE 0%
ILC 100%
ILG 0%
ISG 0%
Other 0%
IndexType(ind2)
IE 0%
ILC 0%
ILG 100%
ISG 0%
Other 0%
IndexType(ind3)
IE 0%
ILC 0%
ILG 0%
ISG 0%
Other 100%
IndexMinValue(ind1)
Betw... 0%
Betw... 0%
Great... 0%
Betw...100%
IndexMinValue(ind2)
Betw... 0%
Betw... 0%
Great... 0%
Betw...100%
IndexMinValue(ind3)
Betw... 0%
Betw...100%
Great... 0%
Betw... 0%
IsProcurementDirectedByIndexes(proc1)
true 55%
false 45%
absurd 0%
ExperienceInOnlyOneContract(proc1)
true 100%
false 0%
absur... 0%
IsProcurementDirected(proc1)
false 28%
true 72%
absu... 0%

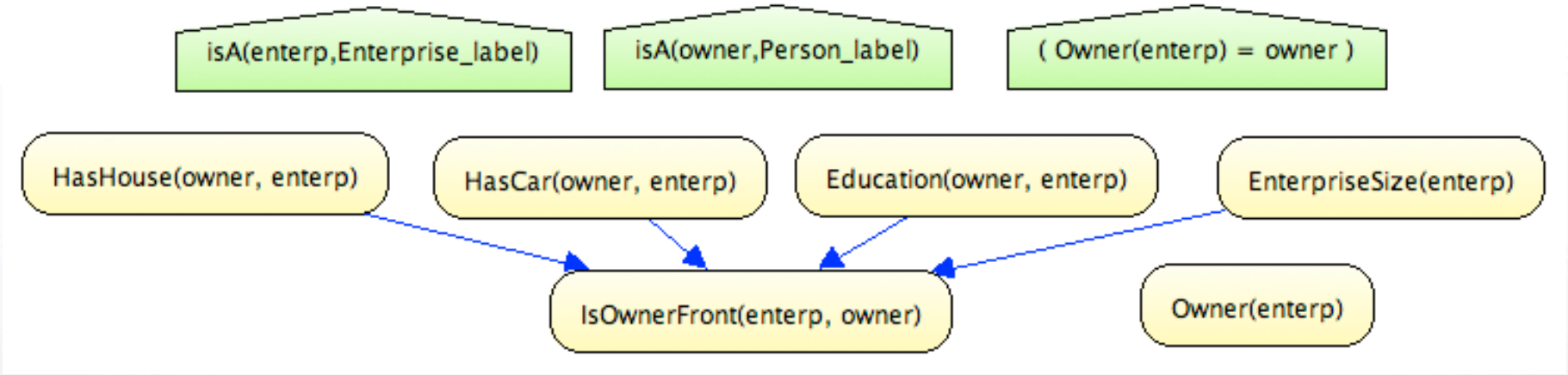
Results

- Non suspect procurement:
  - 0.01% that the procurement was directed to a specific company by using accounting indices;
  - 0.10% that the procurement was directed to a specific company.
- Suspect procurement:
  - 55.00% that the procurement was directed to a specific company by using accounting indices;
  - 29.77%, when the information about demanding experience in only one contract was omitted, and 72.00%, when it was given, that the procurement was directed to a specific company.

Adding new Criteria



Adding new Criteria



Conclusion

Conclusion

- Correct conclusion for both suspicious and non-suspicious cases
  - Results are encouraging
    - Suggesting that a fuller development of our proof of concept is promising
    - Needs more testing, especially with real data for validating the conclusions
- Advantages
  - Impartiality in the judgment
  - Scalability
  - Joint analysis of large volumes of indicators

Conclusion

- Future work
  - Choose/add new criteria
  - Collect more data for validation of the model
  - Will probably required fusion of data from different agencies
    - Good for assessing the usefulness of ontologies and the SW

Obrigado!