Prediction of Class and Property Assertions on OWL Ontologies through Evidence Combination

Giuseppe **Rizzo**
Claudia **d'Amato**

Nicola **Fanizzi**
Floriana **Esposito**

LACAM - Computer Science Dept.
University of Bari, Italy

WIMS'11

Motivation

Semantic Web knowledge bases characterized by **uncertainty**

- incompleteness / inconsistency
- Purely dedutcive methods may fall short

Exploiting alternative (approximate / inductive) approaches to perform data mining tasks

Proposed Approach

In particular: task of **prediction** of assertions

  - class-membership
  - object and data-type props filler

Proposal

- Nearest Neighbors approach
- ***Dempster-Shafer*** Evidence Theory (DST)
  - BBA, Belief, Plausibility, Confirmation
- Evidence **combination**
  - DS, Yager, other combination rules

DL Knowledge Bases

Knowledge Base $\mathcal{K} = <\mathcal{T}, \mathcal{A}>$

- TBox $\mathcal{T}$: set of axioms
  defining *concepts* and *properties*
- ABox $\mathcal{A}$: set of assertions
  concerning the world-state
  - Facts that involve the individuals (resources)
    using concepts and properties
- Reasoning services
  - open-world semantics

Dissimilarity Measures/1

- Given a **context** of concepts

$$C = \{ C_1, C_2, \ldots, C_m \}$$

- *Projection* function:

$$\forall a \in \mathsf{Ind}(\mathcal{A}) \qquad \pi_i(a) = \begin{cases} 1 & \mathcal{K} \models C_i(a) \\ 0 & \mathcal{K} \models \neg C_i(a) \\ \pi_i & \text{otherwise} \end{cases}$$

- *Discernibility* function for C_i:

$$\delta_i(a,b) = |\pi_i(a) - \pi_i(b)|$$

Dissimilarity Measures/2

- Given a context C , $p \in \mathbf{R}$ and $w \in \mathbf{R}^n$ family of **dissimilarity** measures:

$$d_p^{\mathsf{C}}(a,b) = \left[\sum_{C_i \in \mathsf{C}} w_i \delta_i(a,b)^p \right]^{\frac{1}{p}}$$

Evidence Theory

Frame of discernment Ω

- set of hypotheses for a certain domain

Basic belief assignment (BBA) $m : 2^{\Omega} \rightarrow [0,1]$

- $\sum_{A} m(A) = 1$
- $m(A)$ belief committed <u>exactly</u> to A
    - *no additional claims about its subsets*
- $m(A) > 0 \Rightarrow A$ is a **focal** element

Belief and Plausibility

- **Belief** function:

$$\forall A \in 2^{\Omega} \qquad Bel(A) = \sum_{\emptyset \neq B \subseteq A} m(B) \in [0,1]$$

- **Plausibility** function:

$$\forall A \in 2^{\Omega} \qquad Pl(A) = \sum_{B \cap A \neq \emptyset} m(B) \in [0,1]$$

Rules of Combination

Given BBAs m_1 and m_2

- DS rule

$$m_{12}(A) = (m_1 \oplus m_2)(A) = \frac{\sum_{B \cap C = A} m_1(B)\, m_2(C)}{1 - \sum_{B \cap C = \emptyset} m_1(B)\, m_2(C)}$$

normalized version:

- $1 - c$ hides the **contrast** between the BBAs

Rules of Combination/2

- Yager's rule

$$m_{12}(A) = \begin{cases} \sum_{B \cap C = A} m_1(B)\, m_2(C) & A \neq \Omega \wedge A \neq \emptyset \\ m_1(\Omega)\, m_2(\Omega) + c & A = \Omega \\ 0 & A = \emptyset \end{cases}$$

  - more *epistemologically* sound:
    contrast attributed to the case A = Ω
    (total ignorance)

- Other rules used in the experiments:
  Dubois-Pradé, Mixing

Evidential Nearest-Neighbors

- Given
  - A ***set of values*** V (to be predicted)
  - a ***training set*** of labeled individuals

    TrSet = $\{(x_1, v_1), \ldots, (x_M, v_M)\} \subseteq \mathrm{Ind}(\mathcal{A}) \times V$
  - a ***query individual*** x_q
- Select the set of k nearest neighbors $N_k(x_q)$ according to a (dis)similarity measure

Evidential Nearest-Neighbors

- Each (x_i, v_i) in $N_k(x_q)$ induces a BBA m_i regarding the value to be predicted for x_q

$$m_i(A) = \begin{cases} \lambda\sigma(d(x_q, x_i)) & A = \{v_i\} \\ 1 - \lambda\sigma(d(x_q, x_i)) & A = V \\ 0 & \text{otherwise} \end{cases}$$

- Combine the induced BBAs:

$$\bar{m} = \bigoplus_{j=1}^{k} m_j = m_1 \oplus \cdots \oplus m_k$$

Evidential Nearest-Neighbors

- Predict based on belief / plausibility values:

$$v_q = \operatorname*{argmax}_{(x_i, v_i) \in N_k(x_q)} \overline{Bel}(\{v_i\})$$

$$v_q = \operatorname*{argmax}_{(x_i, v_i) \in N_k(x_q)} \overline{Pl}(\{v_i\})$$

Evidential Nearest-Neighbors

- Alternatively, use a ***confirmation*** function

$$C(A) = Bel(A) + Pl(A) - 1$$

then:

$$v_q = \operatorname*{argmax}_{(x_i, v_i) \in N_k(x_q)} \overline{C}(\{v_i\})$$

Prediction Tasks

- Class-membership w.r.t. Q :

$$V_Q = \{-1,+1\} \quad \text{or} \quad V_Q = \{-1,0,+1\}$$

- Object property R filler:

$$V_R = \mathrm{Ind}(\mathcal{A})$$

- Datatype property P value:

$$V_P = \{ v \mid \exists\, P(a, v) \in \mathcal{A} \}$$

Experiments

- Ontologies from standard repositories

Ontology	DL language	#concepts	#object properties	#datatype properties	#individuals
FSM	$\mathcal{SOF}(\mathcal{D})$	20	10	7	37
BCO	$\mathcal{ALCROF}(\mathcal{D})$	196	22	3	112
IMDB	$\mathcal{ALIN}(\mathcal{D})$	7	5	13	302
BioPax	$\mathcal{ALCIF}(\mathcal{D})$	74	70	40	323
HDis	$\mathcal{ALCIF}(\mathcal{D})$	1498	10	15	639

- 10 fold cross validation
- $k = \log|\text{TSet}|$
- *4 combination rules*
- Random classes created with $\mathcal{ALC}$ ops
- 5 built-in *functional* properties

Indices

Using a reasoner to decide the ground truth:

- **Match** rate (M%)
- **Omission** error rate (O%)
- **Commission** error rate (C%)
- **Induction** rate (I%)

Outcomes
Class Membership

Ontology		Dempster	Dubois-Prade	Mixing	Yager
FSM	M%	86.60 ± 04.42	84.75 ± 04.49	85.80 ± 03.90	89.00 ± 04.65
	C%	04.69 ± 03.05	06.65 ± 03.06	05.49 ± 02.33	02.29 ± 02.76
	O%	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00
	I%	08.71 ± 00.29	08.71 ± 00.29	08.71 ± 00.29	08.71 ± 00.29
BioPax	M%	94.93 ± 00.32	94.76 ± 00.32	94.93 ± 00.32	94.93 ± 00.32
	C%	00.15 ± 00.00	00.32 ± 00.00	00.15 ± 00.00	00.15 ± 00.00
	O%	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00
	I%	04.91 ± 00.29	04.91 ± 00.29	04.91 ± 00.29	04.91 ± 00.29
BCO	M%	85.21 ± 04.04	84.54 ± 04.83	85.21 ± 04.04	85.45 ± 04.18
	C%	00.81 ± 00.56	01.47 ± 01.54	00.81 ± 00.56	00.57 ± 00.70
	O%	00.05 ± 00.14	00.14 ± 00.23	00.05 ± 00.14	00.05 ± 00.14
	I%	13.93 ± 03.72	13.95 ± 03.64	13.93 ± 03.72	13.93 ± 03.72

Outcomes
Object Property Values

Ontology		Dempster	Dubois-Prade	Mixing	Yager
FSM	M%	99.64 ± 00.33	99.64 ± 00.33	99.98 ± 00.07	99.64 ± 00.33
	C%	00.02 ± 00.07	00.36 ± 00.33	00.02 ± 00.07	00.36 ± 00.33
	O%	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00
	I%	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00
BioPax	M%	100.00 ± 00.00	100.00 ± 00.00	100.00 ± 00.00	100.00 ± 00.00
	C%	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00
	O%	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00
	I%	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00
BCO	M%	100.00 ± 00.00	100.00 ± 00.00	100.00 ± 00.00	100.00 ± 00.00
	C%	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00
	O%	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00
	I%	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00

Outcomes
Data Property Values

Ontology		Dempster	Dubois-Prade	Mixing	Yager
BCO	M%	64.15 ± 13.53	33.79 ± 11.64	63.52 ± 15.08	71.14 ± 10.00
	C%	35.85 ± 13.53	13.61 ± 10.52	36.48 ± 15.08	28.86 ± 10.00
	O%	00.00 ± 00.00	52.60 ± 15.95	00.00 ± 00.00	00.00 ± 00.00
	I%	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00
IMDB	M%	65.60 ± 06.38	39.73 ± 14.19	66.25 ± 05.94	61.34 ± 08.28
	C%	30.74 ± 06.57	13.62 ± 10.52	30.09 ± 06.13	35.00 ± 09.78
	O%	03.66 ± 03.74	43.01 ± 19.99	03.66 ± 03.74	03.66 ± 03.74
	I%	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00
HDIS	M%	61.00 ± 19.15	61.00 ± 19.15	61.00 ± 19.15	61.00 ± 19.15
	C%	35.62 ± 17.32	35.62 ± 17.32	35.62 ± 17.32	35.62 ± 17.32
	O%	03.38 ± 04.94	03.38 ± 04.94	03.38 ± 04.94	03.38 ± 04.94
	I%	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00	00.00 ± 00.00

Conclusions

Contribution

- Evidential NN procedure based on
  - DST
  - Dissim. measure
- Prediction of
  - class-membership
  - (functional) role fillers

Outlook

- Tackle prediction of non-functional properties vals
- Regression/Ranking
  - based on non-explicit criteria
- Integration with Rough DL

The End

Thank you

Questions ?

<u>Offline</u>

Find us at: http://lacam.di.uniba.it:8000/