

## Rule-based OWL Modeling with ROWLTab Protégé Plugin

- Md Kamruzzaman Sarker (1)
  - Adila Krisnadhi (1,2)
    - **David Carral (3)**
    - Pascal Hitzler (1)
- (1) Data Semantics Laboratory (DaSe Lab)
  Data Science and Security Cluster (DSSC)
  Wright State University, Dayton, OH, USA
  http://www.daselab.org/
- (2) Universitas Indonesia, Depok, Indonesia
- (3) Center for Advancing Electronics in Dresden (cfaed), TU Dresden, Germany



May 2017 – ESWC 2017 – Pascal Hitzler

Problem: directly modeling in OWL (in any syntax, including Manchester syntax) is error-prone and cumbersome.

It appears that rules are much simpler to use for expressing schema information.

Ru3:  $\operatorname{Person}(x) \wedge \operatorname{hasMother}(x, y) \rightarrow \operatorname{Parent}(y)$ 

Ax3:  $\exists$ hasMother<sup>-</sup>.Person  $\sqsubseteq$  Parent

Hence, we developed a Protégé plug-in which affords the modeling of OWL using rules (to the extent to which rules can be converted into OWL).

Non-convertible rules are stored as SWRL-Rules (with a warning to the user).



## **ROWL Protégé plug-in**

• We re-used the SWRLTab source code building our user interface, i.e. the user interaction and rule syntax used is essentially the same.

Person(?x) ^ hasChild(?x,?y) ^ Female(?y) -> hasDaughter(?x,?y)

- However, users can declare new classes directly from the ROWLTab.
- Rules are converted into OWL if possible and added to the ontology; annotation properties are used to store the rules from which they are generated.
- As usual under Protégé, there is no automatic check on RBox regularity, but this can be done (as usual) by calling a reasoner.



## **ROWL Protégé plug-in**

	g/example1) Search
ctive Ontology × Entities × Individuals by class × DL Query × SWRLTab	× OWLAx × OntoGraf × ROWLTab ×
OWL SWRL	
lame	Colort Aviant
2	Generated Axioms
omment	Select All
tabua -	(attends some Course) and (worksFor some Dept) SubClassOf StudentWorker
latus Jk	
	Select axioms which you want to integrate. Integrate Cancel
	Clear Convert to OWL Axiom
Name Body S1 Magazara A Elaphant@d ->, regularThand	Comment Comment
woosetra, ~ crephantry ~ sinalier i nanc	
A HIGGSECK A A Elephanicity - A sinalier many	Edit
SA mouseux) ~ Erephanuty ~ A sinaler i nano	Edit De No Reasoner set. Select a reasoner from the Reasoner menu Show Inferen

May 2017 – ESWC 2017 – Pascal Hitzler

## **User Evaluation**

- Subjects: 12 graduate students from Wright State University with some basic knowledge of OWL and at least minimal exposure to Protégé.
- Participants were given 12 natural language sentences to model in Protégé, half with the standard interface, half with ROWL.
  - Easy sentences: atomic subclass inclusions
  - Medium sentences: Required some role restrictions.
  - Hard sentences: Required rolifications.

Ru5:  $Person(x) \wedge hasBrother(x, y) \wedge hasSon(y, z) \rightarrow hasNephew(x, z)$ Ax5:  $Person \sqsubseteq \exists R_1.Self, R_1 \circ hasBrother \circ hasSon \sqsubseteq hasNephew$ 



Group A	Group B		Difficulty
1. Every father is a parent.	7.	Every parent is a human.	
2. Every university is an educational	8.	Every educational institution is an	easy
institution.		organization.	
3. If a person has a mother then that	9.	If a person has a parent who is fe-	
mother is a parent.		male, then this parent is a mother.	modium
4. Any educational institution that	10.	Any university that is funded by a	meqrum
awards a medical degree is a medi-		state government is a public uni-	
cal school.		versity.	
5. If a person's brother has a son,	11.	If a person has a female child, then	
then that son is the first person's		that person would have that fe-	
nephew.		male child as her daughter.	$\operatorname{hard}$
6. All forests are more biodiverse than	12.	All teenagers are younger than all	
any desert.		twens.	



## **Time used**

#### Hypothesis:



# On medium and hard sentences, participants would be able to model quicker with the ROWLTab than without it.

Sentence	Time (i	in secs)	<b>#</b> c	licks	Corre	ctness
Category	Protégé	ROWL	Protégé	ROWL	Protégé	ROWL
	avg/std	avg/std	avg/std	avg/std	avg/std	avg/std
easy	79/ $41$	47/9	44/ 38	59/ 19	2.9/0.3	2.9/0.3
$\operatorname{medium}$	312/181	116/61	216/131	141/91	2.2/0.5	2.5/0.8
hard	346/218	160/66	351/318	228/168	0.9/0.7	2.5/0.7

#### **Paired t-test:**

easy:	p = 0.002 < 0.01
medium:	p = 0.020 < 0.05
hard:	p = 0.020 < 0.05

#### Hypothesis:



On medium and hard sentences, participants would provide more correct answers with the ROWLTab than without it.

Time (i	n secs)	# cl	icks	Corre	ctness
Protégé	ROWL	Protégé	ROWL	Protégé	ROWL
avg/std	avg/std	avg/std	avg/std	avg/std	avg/std
79/ $41$	47/9	44/ 38	59/ 19	2.9/0.3	2.9/0.3
312/181	116/61	216/131	141/91	2.2/0.5	2.5/0.8
346/218	160/66	351/318	228/168	0.9/0.7	2.5/0.7
	Time (i Protégé avg/std 79/41 312/181 346/218	Time (in secs)ProtégéROWLavg/stdavg/std79/4147/9312/181116/61346/218160/66	Time (in secs) $\#$ clProtégéROWLProtégéavg/stdavg/stdavg/std79/4147/9312/181116/61216/131346/218160/66351/318	Time (in secs) $\#$ clicksProtégéROWLProtégéROWLavg/stdavg/stdavg/stdavg/std79/4147/944/3859/19312/181116/61216/131141/91346/218160/66351/318228/168	Time (in secs)    # clicks    Correction      Protégé    ROWL    Protégé    ROWL    Protégé      avg/std    avg/std    avg/std    avg/std    avg/std      79/41    47/9    44/38    59/19    2.9/0.3      312/181    116/61    216/131    141/91    2.2/0.5      346/218    160/66    351/318    228/168    0.9/0.7

#### **Paired t-test:**

easy:	p = 1.0000 > 0.05
medium:	p = 0.180 > 0.05
hard:	p = 0.0001 < 0.01

## Clicks

#### Hypothesis:

# DaSe Lab

#### None (this was for information only)

Sentence	Time (i	in secs)	# c	licks	Corre	ctness
Category	Protégé	ROWL	Protégé	ROWL	Protégé	ROWL
	avg/std	avg/std	avg/std	avg/std	avg/std	avg/std
easy	79/ $41$	47/9	44/ 38	59/ 19	2.9/0.3	2.9/0.3
$\operatorname{medium}$	312/181	116/61	216/131	141/91	2.2/0.5	2.5/0.8
hard	346/218	160/66	351/318	228/168	0.9/0.7	2.5/0.7
	1					

#### **Paired t-test:**

easy:	p = 0.092 > 0.05	
medium:	p = 0.030 < 0.05	(significant time difference)
hard:	p = 0.173 > 0.05	(significant correctness
		difference)



May 2017 – ESWC 2017 – Pascal Hitzler

### Assessment

 The hypotheses for time and for correctness (hard questions) were confirmed. For correctness (medium questions) the hypothesis was rejected.

category	$\operatorname{time}$	clicks	correctness
easy	significant $(p < 0.05)$	not significant	not significant
medium	significant $(p < 0.01)$	significant $(p < 0.05)$	not significant
hard	significant $(p < 0.05)$	not significant	significant $(p < 0.01)$

It appears that medium modeling problems (with some role restrictions) can be done correctly with the standard Protégé interface by this type of user, although more time is needed than when using ROWLTab.

It appears that hard problems (requiring rolification) cannot really be solved using the standard Protégé interface, and the unsuccessful solution attempts in addition require more time.



There is a lot of scope for improving the core functionality.

However we first want to see if there is uptake, before we put more work into it.

All feedback (and feature requests) are most welcome.

- Naming of fresh properties (from rolification)
- Automatic regularity checks.
- Use of nominal schemas if rule cannot be rendered in OWL.
- Add right-hand side disjunctions and existentials, and left-hand side universals, plus perhaps other syntax extensions.
   Goal: Make it possible to express all OWL axioms in some rule-type syntax



## Conclusions

- It is clearly easier to model axioms using the ROWLTab interface, than using the standard Protégé interface.
- We hope for feedback, it will encourage us to improve the interface.

Acknowledgements:

- Support by NSF on the TROn project.
- Support by the DFG within the Cluster of Excellence "Center for Advancing Electronics Dresden" (dfaed)
- Tanvi Banerjee and Derek Doran for advice on statistics





## Thanks!

#### http://daselab.org/content/modeling-owl-rules



May 2017 – ESWC 2017 – Pascal Hitzler