Report Ontologies and Knowledge Bases

Yasar Mahomed Abbas, Franziska Pannach,
Danielle Russell, Yuvika Singh

May 28, 2018

1 Objective and Motivation

Folk and Fairy tales are a substantial part of oral history. They play an important role in the cultural heritage of regions, nations or cultural minorities. In European context, fairy tales have been collected and editored by the Grimm brother's in the beginning of the 19th century.[1] In the African context, the oral tradition of folk tales existed way longer. Presumably, African folk tales are therefore different than European fairy tales in terms of structure and motifs. This project aims to construct an ontology of African folk tales following the approach of Russian folklorist Vladimir Propp. We hope to not only collect and structure African folk tales, but also investigate how they follow Propp's formalism and how motifs and agents are verbalised. Hence, the ontology is going to contain:

- The Proppian functions and entities encoded within.

- Specific folk tale motifs according to the Aarne-Thompson-Uther Index (ATU).

- The representation of the functions and motifs in selected African Folktales.

2 Domain

2.1 Motifs Indices

Folk tales motifs are usually classified by two motif indices. The Aarne-Thompson-Uther index (ATU)[1] is used to classify tales into one category. The categories are relatively wide, describing the main story line of the tale. Therefore, each tale can only have one ATU type. In contrast, the Thompson-Motif-Index (TMI) is more fine grained, describing single motifs, i.e. repeated elements, e.g. characters. The TMI motifs are organized in a hierarchical structure. A tale can be described with more than one TMI motif.[2]

---

[1] http://www.mftd.org/
[2] https://sites.ualberta.ca/urban/Projects/English/Motif_Index.htm

2.2 Propp Functions

Russian folklorist Vladimir Propp introduced 31 invariant functions describing the morphology of the Russian magic folk tale. In his ground breaking 1928 work 'Morphology of the Folk tale', he argues that the narrative of folk tales always follows the same pattern. The narrative functions and the *Dramatis Personae* (agents in the story) he introduced are strictly defined and specify recurrent units from which the tales are constructed. Propp [2] set four axioms:

1. Functions of characters serve as stable, constant elements in a tale, independent of how and by whom they are fulfilled. They constitute the fundamental components of a tale.

2. The number of functions known to the fairy tale is limited.

3. The sequence of functions is always identical.

4. All fairy tales are of one type in regard to their structure.

The high formality of this structuralism allows something as complex and highly emotional as the folk tale to be pressed in a strict pattern. Thus, they can be further used in automatically processing or when generating new tales. In Computational Linguistics, Propp's functions are used in various ways, such as for automatic markup, classification and annotation ([3]), or as a foundation of an independent XML dialect ([3], [4]). His approach is still widely used not only in folk tale research but also applied to contemporary work such as the Star Wars Trilogy[3]. Proppian functions can appear in the tale as they are, with modifiers or they can be inverted, e.g. *Hero leaves home → Hero does not leave home* (explicitely).

3 Conceptualization

3.1 Existing Work

Declerck et al. 2016/2017 [5] have introduced an integrated ontology that combined the ATU and the TMI motifs in a complex way. They suggested extending the ontology by including

- Adding more fairy tales that fall into one of the ATU classes

- Adding more tales from specific collections

- Add Proppian functions

The aim of this project is to fulfil these three aspects. For the time being, we will concentrate on Fairy Tales anthologies from the Southern African context. Our ontology will be independent but can be easily added to the existing work once it's published by Declerck et al.[5]

---

[3]http://jaced.com/2013/02/06/vladimir-propp-science-of-the-fairy-tale/

3.2 Definition of the Ontology

We describe our ontology by the following properties $< C,I,A,R >$

- $c_i \in C$ set of Classes: Dramatis Personae according to Propp, elements in Proppian functions, e.g. *the hero*, *the claim*, Proppian functions and subfunctions, classes that describe the publications of the tale (tale, anthology, editor/author)

- $i_i \in I$ set of Instances: The representation of c_i in the fairy tales from the anthologies such as [6], e.g. *The Girl Who Lived In A Cave*, the functions appearing in those fairytales, the Dramatis Personae, motifs according to ATU[4]

- $a_i \in A$ set of Axioms: Proppian functions, e.g. *Acquisition of Magical Agent*, and axioms describing the publication

- $r_i \in R$ set of Relations: Relationships between classes that model the functions, sequential relations of functions, e.g. *FollowedBy*, *PrecededBy* or *applies* to link the function to a tale, *appearsIn* to link a Dramatis Personae to the tales, relationships linking the Dramatis Personae together, such as *causesHarm*, *combats* or *relatedTo*

We are using the ATU index for the classification of our motifs, ignoring the TMI motifs for now, since the classification of tales in TMI motifs requires a vast amount of knowledge in literary studies. Since Declerck et al.'s ontology will cover the TMI motifs, this is not considered a drawback of our work.

3.3 Compentency Questions

Our final ontology will be able to answer the following questions.

1. Which folk tales fall into a given motif class, e.g. ATU 70-99 Other Wild Animals?

2. Which Dramatis Personae appear in a given tale?

3. Which Proppian functions appear in African folk tales?

4. How are Dramatis Personae interacting in the African folk tales, e.g. which figures use the "interdiction" relation?

5. Is there a dominating interaction between certain classes of Dramatis Personae, e.g. the villain causes harm to the hero's parents?

6. Which sequences of Proppian functions appear in the given tale?

7. Which Proppian functions follow a given function predominantely, i.e. are there patterns withing the Proppian sequences?

8. Who is the editor of an anthology with folk tales from a given origin?

9. How are Proppian functions verbalised, i.e. which words are used to describe events that fall into a given function class?

---

[4]ATU motifs in contrast to TMI motifs are not always single concepts, they can also be a description of content such as *Ogre Frightened by Man* (ATU 1145-1154), nevertheless they will be classes within the scope of this project in contrast to axioms

3.4 Axioms

We define some axioms for the publication and the classification of the fairy tales.

- Each tale is published in an anthology.
- Each anthology has at least one editor, a title, a publisher, and a date of publication.
- Each tale has a title.
- A tale can have an author and an origin.
- Each tale has a set of Dramatis Personae.
- Each person is represented by one or many verbalisations.[5]
- Each tale falls into one of the ATU classes.
- Each ATU class has an ATU number and a description.
- If a Proppian function applies for a tale, there is some verbalisation in the text.
- Proppian functions follow a specific order (see below), this order is represented by a sequence.

Furthermore, following Propp's approach, we derive our axioms for the possible description of the narrative as follows [2]:

- α The initial situation. A text may only have a single Initial Situation function.

1. A member of a family leaves home (the hero is introduced);
2. An interdiction is addressed to the hero ('don't go there', 'go to this place');
3. The interdiction is violated (villain enters the tale);
4. The villain makes an attempt at reconnaissance (either villain tries to find the children/jewels etc; or intended victim questions the villain);
5. The villain gains information about the victim;
6. The villain attempts to deceive the victim to take possession of victim or victim's belongings (trickery; villain disguised, tries to win confidence of victim);
7. Victim taken in by deception, unwittingly helping the enemy;
8. Villain causes harm/injury to family member (by abduction, theft of magical agent, spoiling crops, plunders in other forms, causes a disappearance, expels someone, casts spell on someone, substitutes child etc, commits murder, imprisons/detains someone, threatens forced marriage, provides nightly torments); Alternatively, a member of family lacks something or desires something (magical potion etc);
9. Misfortune or lack is made known, (hero is dispatched, hears call for help etc/ alternative is that victimized hero is sent away, freed from imprisonment);
10. Seeker agrees to, or decides upon counter-action;
11. Hero leaves home;
12. Hero is tested, interrogated, attacked etc, preparing the way for his/her receiving magical agent or helper (donor);
13. Hero reacts to actions of future donor (withstands/fails the test, frees captive, reconciles disputants, performs service, uses adversary's powers against them);

---

[5]e.g. in Snow White 'the stepmother' and 'the evil queen' describe the same individual

14. Hero acquires use of a magical agent (directly transferred, located, purchased, prepared, spontaneously appears, eaten/drunk, help offered by other characters);

15. Hero is transferred, delivered or led to whereabouts of an object of the search;

16. Hero and villain join in direct combat;

17. Hero is branded (wounded/marked, receives ring or scarf);

18. Villain is defeated (killed in combat, defeated in contest, killed while asleep, banished);

19. Initial misfortune or lack is resolved (object of search distributed, spell broken, slain person revived, captive freed);

20. Hero returns;

21. Hero is pursued (pursuer tries to kill, eat, undermine the hero);

22. Hero is rescued from pursuit (obstacles delay pursuer, hero hides or is hidden, hero transforms unrecognizably, hero saved from attempt on his/her life);

23. Hero unrecognized, arrives home or in another country;

24. False hero presents unfounded claims;

25. Difficult task proposed to the hero (trial by ordeal, riddles, test of strength/endurance, other tasks);

26. Task is resolved;

27. Hero is recognized (by mark, brand, or thing given to him/her);

28. False hero or villain is exposed;

29. Hero is given a new appearance (is made whole, handsome, new garments etc);

30. Villain is punished;

31. Hero marries and ascends the throne (is rewarded/promoted).

If a function applies for a tale, the axiom holds. Not all of the functions/axioms need to be fulfilled by every tale.

3.5 Classes

Tale, Anthology, Editor, Author, Title, Publisher, Date, Fictional Person, Person, Object, ATU Class, ATU Number, Description, Proppian Function, Verbalisation, Symbol, Family Member, Hero, Villain, Victim, Seeker, Helper, Donor, Dispatcher, Princess, Princess's Father, False Hero, Magical Agent, Desired Object, Task, Reward

3.6 Relations

We define the relations based on the axioms in 3.4 .

publishedIn(tale, anthology), editedBy(anthology, editor), appearsIn(fictional Person, tale), hasClass(tale, ATU), applies(tale, function), FollowedBy(function, function), PrecededBy(function, function), isInversed(function), interdicts(fictional person, hero), reconnaissance(villain, object), reconnaissance(villain, victim), gainsInformation(villain, victim), deceives(villain, victim), isDeceivedBy(victim, villain), takesPossessionFrom(villain, victim), helpTheEnemy(victim, villain), related(hero, family member), causesHarm(villain, related(hero, family member)), lacks(hero, object), testedBy(hero, donor), interrogated(hero, donor), attackedBy(hero, donor), reactsTo(hero, donor), acquires(hero, magical agent), acquiresFrom(hero, donor), ledTo(hero, desired object), combat(hero, villain), brandedBy(hero, object), defeats(hero,villain), pursuedBy(hero, fictional person), rescuedBy(hero, fictional person), proposedTo(task,

hero), proposedBy(task, fictional person), resolves(hero, task), recognizedBy(hero, object), marries(hero, fictional person), isRewarded(hero, object)

Each of the relations describing the relationship between two fictional persons, is tied to one of the functions we described earlier. For example, if and only if the function 31. *W Marriage* applies, we can use the function *marries(hero, fictional person)*. That means, the relationship is tied to the appearance of the function and cannot be used in a different context, e.g. at a beginning of a tale. However, the application of the function *W Marriage* does not guarantee that *marries(hero, fictional person)* applies, e.g. when in the tale *Guinea Fowl Child* the heroine is courted, the function *W Wedding* still applies although there is no fictional person instance to fulfill the relationship *marries*. Therefore, we cannot connect the two concepts of the class of the function and it's respective function although they belong together.

We could formulate a relationship between the class of the Proppian function and the relation for the given example as:

$$\text{marries(A,B)} \cap \text{precededBy}(\bigcup_{i=1}^{30} \text{ProppianFunction}_i) \cap \neg\text{followedBy}(\bigcup_{i=1}^{30} \text{ProppianFunction}_i)$$
$$\cap \exists\ \text{Hero(A)} \cap \exists\ \text{FictionalPerson(B)} \rightarrow (\ \exists\ \text{W\_ Marriage} \subseteq \text{applies.Tale}\ )^6$$

For the scope of this project, it is unrealistic to implement these relationssships for all 31 classes of functions and their subfunctions, especially since it is not always given that the subfunctions inherit the relationship, e.g. the marries(A,B) function applies not in case the subclass *w1 Promise of marriage* applies.

With respect to the unary relations like *leavesHome(hero)*, *isPunished(villain)*, we believe they are sufficiently represented by the function itself and its verbalisation, i.e. the instance of the class. The binary functions, however, allow us to model the relationships between fictional persons in a sense that would otherwise likely be lost.

---

[6]The function *W Wedding* is function number 31 and therefore the last function is the sequence.

3.7 UML Diagram
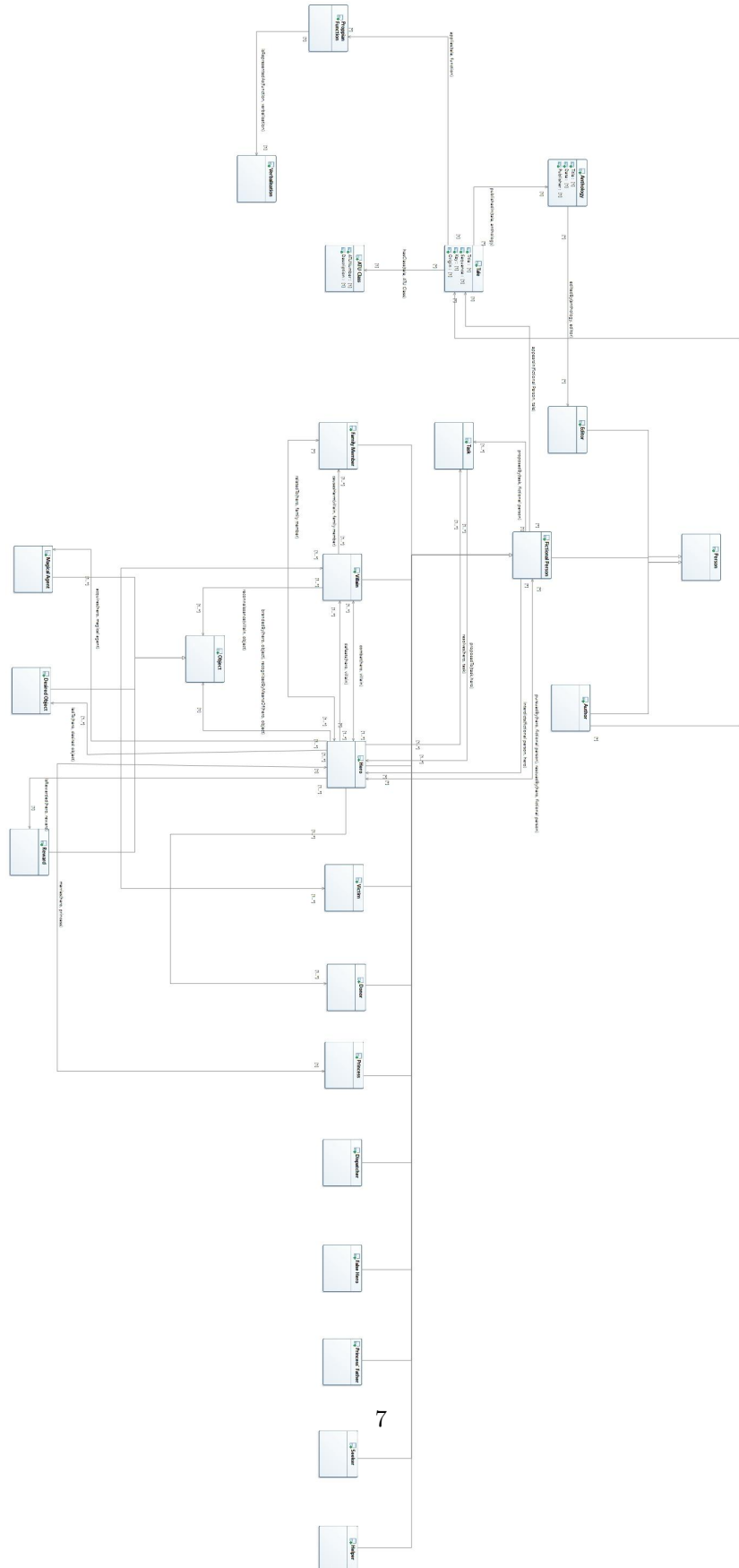
For the scope of this report, we refrain from putting all 31 functions and their subfunction in the UML diagram, since they are merely subfunctions of the class *Proppian Function* and inherit the property *Verbalisation* as well as the *applies* property that connects them to the *Tale* class.[7]

3.8 Description Logic

The following tables describes the most important concepts of the folk tale ontology in Description Logic.

Axioms of Class Hierarchy	Important Concepts
FictionalPerson $\subseteq \exists$ isA.Person	FictionalPerson $\subseteq \exists$ appearsIn.Tale
Family Member $\subseteq \exists$ isA.FictionalPerson	Tale $\subseteq \exists$ hasClass $= 1$ ATU_Class
Princess $\subseteq \exists$ isA.FictionalPerson	ProppianFunction $\subseteq \exists$ applies.Tale
Princess' Father $\subseteq \exists$ isA.FictionalPerson	Tale $\subseteq \exists$ publishedIn.Anthology
Hero $\subseteq \exists$ isA.FictionalPerson	Anthology $\subseteq \exists$ editedBy.Editor
False Hero $\subseteq \exists$ isA.FictionalPerson	False Hero $\subseteq \neg$ Hero
Victim $\subseteq \exists$ Hero $\cup$ FamilyMember $\cup$ Princess	FamilyMember $\subseteq \exists$ relatedTo.Hero
Dispatcher $\subseteq \exists$ isA.FictionalPerson	Villain $\subseteq \neg$ (Victim $\cap$ Hero)
Villain $\subseteq \exists$ isA.FictionalPerson	Editor $\subseteq \neg$ FictionalPerson
Donor $\subseteq \exists$ isA.FictionalPerson	Author $\subseteq \neg$ FictionalPerson
Seeker $\subseteq \exists$ isA.FictionalPerson	Reward $\subseteq \neg$ MagicalAgent
Helper $\subseteq \exists$ isA.FictionalPerson	Princess $\subseteq \neg$ Princess'Father
Editor $\subseteq \exists$ isA.Person	$\text{SubFunction}_i \subseteq \neg (\text{ProppianFunction} \setminus \text{SubFunction}_i)$
Author $\subseteq \exists$ isA.Person	
Desired Object $\subseteq \exists$ isA.Object	
Magical Agent $\subseteq \exists$ isA.Object	

4 Implementation

In the following section, we describe the implementation of the ontology as it stands May 28, 2018. We constructed 208 subclasses of OWL:Thing, 37 of which are disjoint. The following three pictures show a part of the classes with their respective subclasses. We followed a naming convention for the Proppian functions, trying to stay as close to the source [2] as possible. Propp numbered all the functions using roman letters, gave each of them a unique symbol and a short description. Therefore, our class names for the Proppian functions follow the scheme *IX_B_Mediation*.

---

[7]In the same manner, we use Subfunction in the Description Logic representative for the 31 Proppian functions and their child functions, instead of declaring them as disjoint for each subclass.

Figure 1: Upper level of the subclasses

Figure 2: Level of Proppian functions and their subfunctions
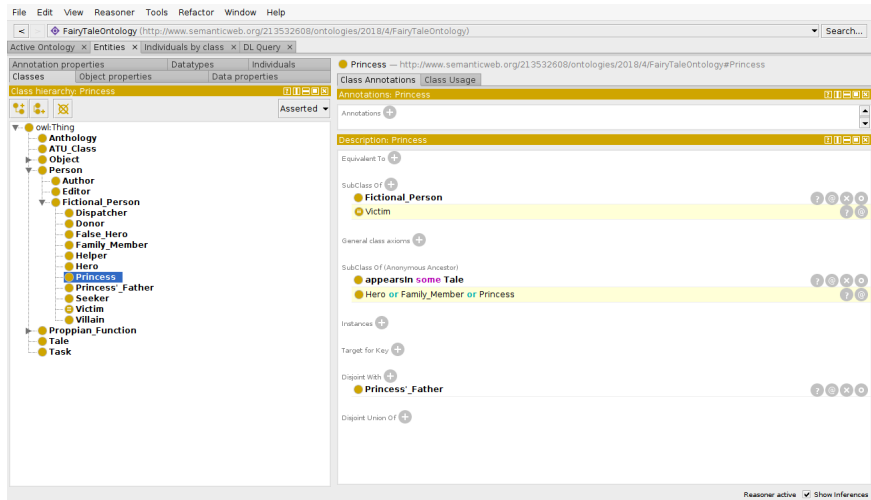
Figure 3: Level of Dramatis Personae

The object properties are mainly the relations between the Dramatis Personae. They do not inherently follow a hierarchical structure. Therefore, the hierarchy of the object properties is wide but relatively flat. Where necessary, an rdf:comment describes the use of the property.
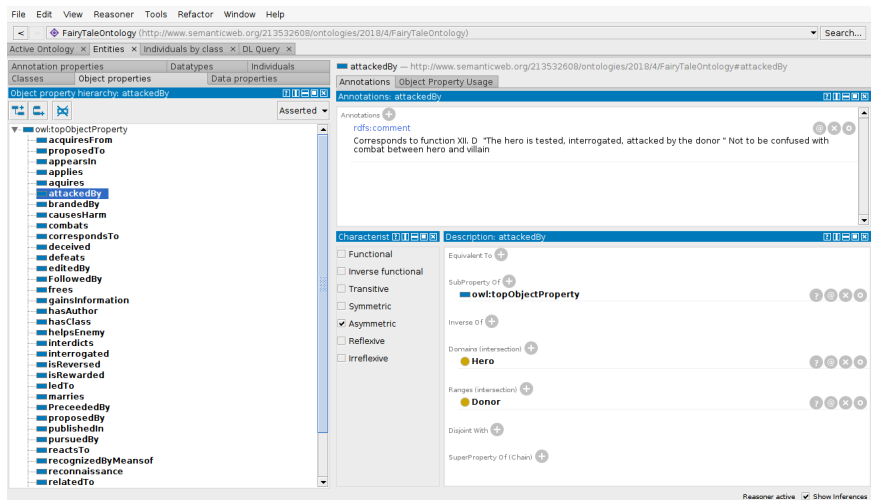

Figure 4: Object Properties

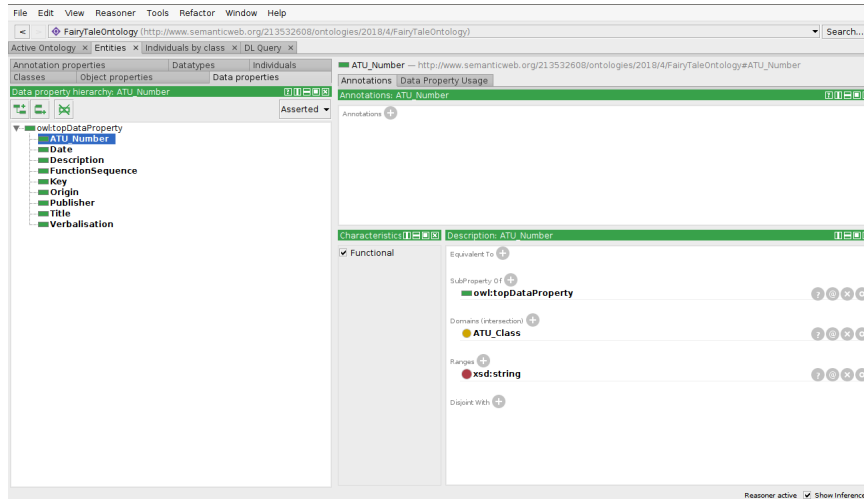Data properties are kept relatively simple. They are used to describe the publication of the anthology, the ATU Class, the tale, or the proppian function.

Figure 5: Data Properties

We addeed the anthologies, folk tales, their ATU classes, their functions as well as the Dramatis Personae as individuals to our database. Currently, May 28, 2018, we included XXXXX fairy tales, including XXXX functions and XXXX dramatis personae.
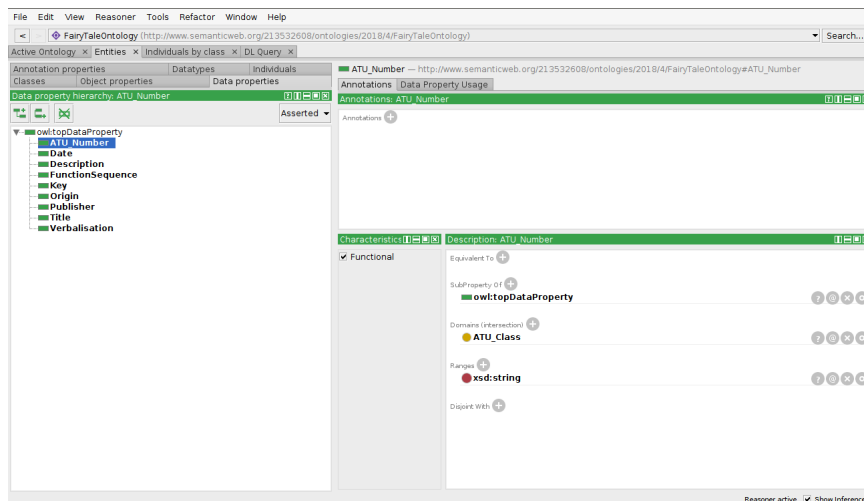

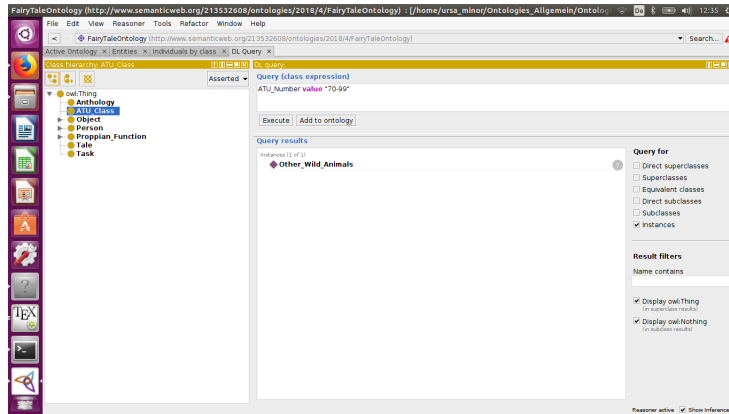Figure 6: Instances HERE INSTANCES EINFUEGEN, Wrong bild

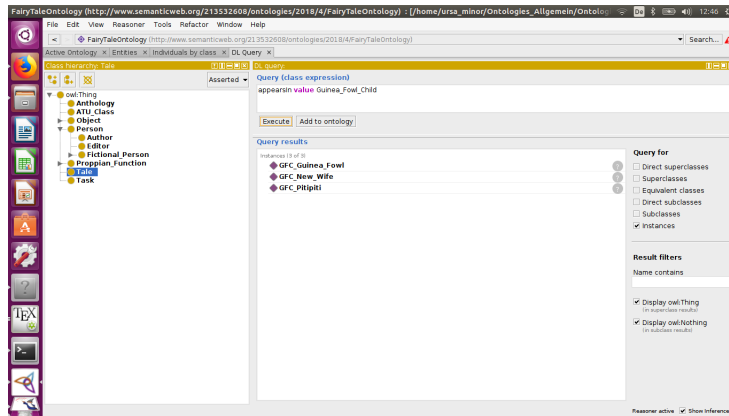Figure 7: Instances HERE INSTANCES EINFUEGEN, Wrong bild

Figure 8: Instances HERE INSTANCES EINFUEGEN, Wrong bild

4.1 Answering the Competency Questions

We can answer question number 1, "Which folk tales fall into a given motif class, e.g. ATU 70-99 Other Wild Animals?" by using the DL Tab of the protege software and employing the Hermit Reasoner while querying the instances of our ontology.

In the same manner, we can answer the second question "Which Dramatis Personae appear in a given tale?". The answer to question 3 could be constructed as follows:

5 Conclusion and Outview

We built a sophisticated ontology that represents the Proppian theory about the Morphology of the Folk Tale.[2] It also allows users to comfortably add more stories. The ontology can be alonestanding, or it can be connected to Declerck et.al.'s [5] work once their fairy tale motif ontology is published. In the long term, modelling the relationship between the functions, and the Dramatis Personae, like in the example in section 3.6 could be useful. Also, adding more annotations with comments about the usage of classes of Proppian functions and their connection to the relations as described in section 3.6 would certainly increase the usability of the system. Furthermore, Propp also made statements about how characters make their appearance into the course of
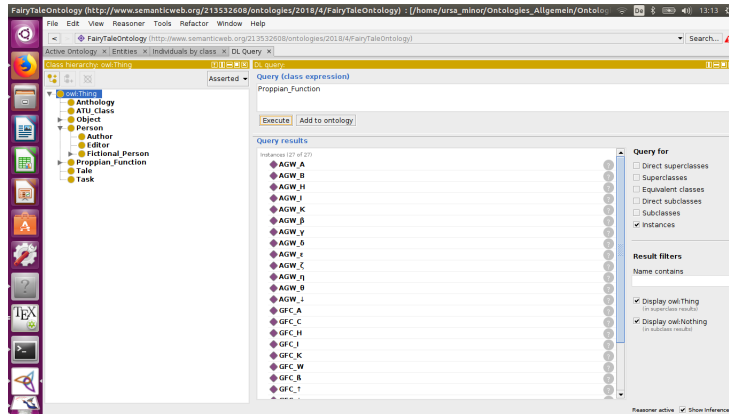
Figure 9: Instances HERE INSTANCES EINFUEGEN, Wrong bild

action, a feature that could be added to the ontology in the future. The more fairy tales of different origin are added, the more and complexer the deductions about the structure and verbalisation of African folk tales can be made. Certainly, these deductions can only be made with the help of scholars from folklorist or literary studies. Furthermore, it would be interesting to compare these findings about African folk tales to studies of folk tales of different origin, such as Russian, Central European or Native American tales.

6   Bibliography

References

[1] J. Grimm and W. Grimm, eds., *Kinder- und Hausmärchen.* Verlag der Dieterich-schen Buchhandlung, 7. ed., 1857.

[2] V. Propp, *Morphology of the Folktale*, vol. 10. University of Texas Press, 1968.

[3] S. Malec, "Autopropp: Toward the automatic markup, classification, and annotation of russian magic tales.," in *Proceedings of the First International AMICUS Workshop on Automated Motif Discovery in Cultural Heritage and Scientific Communication Texts* (P. Lendvai, ed.), pp. 112–115, 2010.

[4] P. Lendvai, T. Declerck, S. Darányi, and S. Malec, "Propp revisited: Integration of linguistic markup into structured content descriptors of tales," in *Digital Humanities 2010*, Oxford University Press, 7 2010.

[5] T. Declerck, L. Schäfer, and N. Koleva, "Integrated ontologies for the classification of folk tales." online, multiple presentations combined, 2017.

[6] A. M. Smith, ed., *Children of Wax: African Folk Tales.* Canongate, 1989.