Formal Ontology in Information Systems
T. Prince Sales et al. (Eds.)

doi:10.3233/FAIA250488

Measuring Similarities of Literary Characters

Daniele PORELLO [a], Emilio M. SANFILIPPO [b,c,1], and Alessandro MOSCA [b]
[a] *DAFIST, University of Genoa, Genoa, Italy*
[b] *CNR ISTC Laboratory for Applied Ontology, Trento-Catania, Italy*
[c] *CESR University of Tours, Tours, France*

Abstract. One could argue that fictional characters in literary texts, films, television series, and other artistic works become part of our lives. Although they do not possess spatio-temporal existence like flesh-and-blood individuals, their presence can be traced in our cities through museums—such as the Sherlock Holmes Museum in London or Juliet's house in Verona—or through statues like the Rocky Balboa statue in Philadelphia. In this paper, we explore ontology-based approaches to representing and comparing interpretations of literary characters. In particular, we introduce and discuss several similarity measures between interpreted literary characters, and discuss conditions for the identification of characters across heterogeneous interpretations.

Keywords. literary characters, fiction, similarity measures, literature

1. Introduction

In research contexts relative to literary studies, philosophy, or art criticism, literary characters play a prominent role, e.g., because they serve as key elements for exploring the theoretical frameworks or cultural assumptions within which they were conceived.[2]

In particular, both analytic philosophy and formal models in the Digital Humanities (DH) often adopt an "objectivist" approach to tell which properties a character satisfies (for philosophy, see [2]; for research in DH, see [3,4]). For instance, it is often assumed that, given a literary character, it is the text featuring the character to characterize it through certain properties [5]. This view is challenged by a strong tradition in hermeneutics and literary theory, with some cases in analytic philosophy, too, that questions the objective status of a text's content, emphasizing the interplay between texts and interpreters [6,7]. An important distinction is made between *ideal* and *real-world* inter-

[1]Corresponding Author: Emilio M. Sanfilippo, emilio.sanfilippo@cnr.it
[2]Prominent examples of literary characters include Humpty Dumpty, Sherlock Holmes, Anna Karenina and Harry Potter. One might ask whether the London in which the stories about Holmes occur or Harry Potter's cloak of invisibility are also literary characters. For the purposes of this paper, we will not address questions relating to how to distinguish strictly speaking between 'characters' and other entities that occur in fictional stories (see [1] for further reading). Although we will primarily consider characters like those mentioned above, the considerations set out in this paper can apply to other types of fictional entities, as well as real-world individuals.

preters (though terminology may vary) [8]. Ideal interpreters are abstract models, often conceived as capable of grasping all possible nuances of meaning in a text. By contrast, real-world interpreters are flesh-and-blood individuals shaped by their historical and cultural contexts, common sense, and individual perspectives. As a result, their interpretations are subjective and contingent on their background. This latter understanding of interpreters aligns with the view that a single text can be interpreted in multiple, and even conflicting, ways. Furthermore, the interpretations provided by real-world interpreters are generally considered partial with respect to the texts they interpret. For example, they may approach texts from particular angles while leaving others unexplored.

The purpose of this paper is to contribute to the foundations of an ontology-based modeling framework to analyze interpretations of texts and their characters as they are provided by real-world interpreters. The framework is intended to support experts in exploring alternative interpretations of texts, analyzing their similarities and departing points. In particular, given that a text is open to multiple interpretations, it could be in principle the case that interpreters do not talk of the same character even if interpreting the same text. From a formal perspective, to express and compare interpretations, we rely on research in knowledge representation and reasoning, where an ontology is a set of formulas in a logical language which are used to describe a particular domain of interest (here, we restrict to description logics). We will introduce and discuss different measures to compute similarity between alternative knowledge bases, standing, in our scenario for specific literary interpretations. Specifically, we will tailor several similarity measures–extensively discussed in the literature on cognitive science [9], belief merging [10], entity linking [11], and judgment aggregation [12]–to account for the specificities of quantifying proximity and differences between literary characters.

The paper is structured as follows. Section 2 gives some hints concerning how we conceive literary characters in our approach in the light of the current philosophical debate. Section 3 introduces the formal framework we rely on and discusses different measures to evaluate the similarity between sets of interpreters' assertions, by distinguishing the scenario in which the interpreters share the same ontology (Section 3.1) from the one in which multiple ontologies are in place (Section 3.2). Section 4 concludes the paper.

2. Literary Characters: A Stipulative Perspective

There is an ongoing and wide-ranging debate in philosophy concerning the ontology of literary characters within the broader discussion on fictional entities (*ficta*) [2]. Philosophers commonly ask whether literary characters exist and, if so, what they are. The philosophical agenda on these topics is broad, with a high-level division between *realists*, who accept the existence of characters, and *anti-realists*, who deny it. Additionally, there are various competing accounts of how to conceive of characters, whether as individuals or as types (for an overview, see [13, pp. 3-64]).

In contrast, literary critics and scholars discuss characters without necessarily engaging with the ontological questions that concern philosophers. For the former, literary characters are simply objects of study, much like other aspects of literature like themes, motifs, or narrative structures (for an example, see [14]). However, as pragmatic as this approach can be, it cannot escape questions relative to the relations between characters, especially when they are characterized in different manners. One significant challenge is the re-identification of characters across different texts. For example, Saint-Gelais [15] discusses cases where literary characters appear in texts by different authors, raising the question of whether they remain the same entity or merely share common features.

For our purposes, particularly in representing and comparing interpretations in the form of *interpretive data* [16], we adopt a *stipulative* approach to the existence of characters, which is meant to bypass any further metaphysical concern about their nature. This is a common methodological choice for formal modeling and computational processing, especially in the DH community (see, for instance, [3,4,17]).

Stipulative approaches are found in philosophy, too. Thomasson [18], for instance, draws an analogy between literature and social practices, such as sports, which are governed by public rules accepted by participants. In these contexts, new terms are introduced along with rules that determine their usage and reference. As Thomasson explains:"[...] the rules for baseball are more explicitly stipulated than the 'rules' implicit in discussions of literature, but the idea is the same: both are sets of public practices, where those competent in the practice stipulatively introduce new terms with certain rules of use embodying at least tacit existence, survival, and identity conditions [...]" [18, pp.145-46]. This analogy supports the idea that literary characters can be treated as objects of study with sorts of operational definitions, even if their metaphysical status remains contested.

Given this general consideration, we will assume that literary characters are domain entities upon which we can predicate. Also, following research in both philosophy and literary studies [13,19], we assume that a character can be characterized in terms of various properties. For example, in Doyle's *A Study in Scarlet*, the character of Sherlock Holmes has the properties of *being a detective*, *living in Baker Street*, and *collaborating with Watson*, among others. It is fundamental to stress that, following our focus on interpretation, the way in which characters are ascribed properties is *interpreter-dependent*. This is a crucial difference with respect to state of the art ontological theories in philosophy [2]. As a consequence, the resulting approach bears an inherent weakness: if the relations between characters can be given only through the analysis of interpretations, one can barely claim that two characters are *identical*, considering that interpretations are only partial and contextual. In this sense, through a focus on interpretation, we can only provide hints to tell when characters should be distinguished or identified *according to what interpreters claim*. In our understanding, this aligns with interpretation practices in literary contexts, where any sort of claim relative to characters is based on the information made explicitly available by interpreters.

Finally, following [19,20], we assume that interpreters use *diagnostic properties* to characterize characters. These are properties that interpreters consider particularly relevant. From this perspective, diagnostic properties can be considered as sort of essential properties. However, as mentioned above regarding the interpretational basis of our study, "essentiality" is only stipulative and cannot be understood in strict metaphysical terms. The attribution of diagnostic properties to characters is subjective; therefore, different scholars may emphasize different properties for a character. As we will discuss further below, this has consequences for characters' identification.

3. Formal Framework

We assume that a set $\mathcal{N}$ of n interpreters is given and that ontologies are written in some description logics, $\mathcal{DL}$. Moreover, we assume that each ontology is partitioned into an upper-level part, dealing with the most general classes and relations (such as 'object', 'event', 'parthood'), and a domain level (dealing with domain concepts such as 'detective', 'tobacco', 'doctor', etc.). As a simplification, we illustrate domain ontologies in $\mathcal{ALC}$, so for instance, we do not use role hierarchies.

We distinguish between two cases: *(i)* multiple interpreters share the same ontology (see Section 3.1), which defines the properties ascribed to characters, and interpreters might disagree on the ABox, i.e. about the properties ascribed to characters; *(ii)* multiple interpreters possibly have different ontologies (see Section 3.2), thus, they might disagree on the meaning of the properties as well as on the ABox. Both cases (*i*) and (*ii*) cover the situation where multiple texts are interpreted by different interpreters, as well as the case where many interpreters provide different interpretations of the same text. Notice that a description logic ABox can be viewed as an instance of a relational database with only unary or binary relations. However, the semantics of ABoxes differs from the usual semantics of database instances: contrary to the *closed-world assumption* semantics, which is typical of relational databases, the semantics of ABoxes is based on the *open-world assumption* (OWA). OWA is a feature that makes description logics particularly suitable for dealing with interpreters' viewpoints and critical arguments. Contrary to what happens in traditional database, where lack of knowledge is interpreted as negative information, in an ABox it only indicates lack of knowledge. For instance, the absence of the statement $\mathsf{Person}(\mathsf{Holmes})$ in a interpreter's ABox does not imply the inclusion of its negation $\neg\mathsf{Person}(\mathsf{Holmes})$. As said, interpreters are supposed to provide only *partial* and *incomplete* interpretations over artworks that would be erroneous to be treated as a closed set of formulas.

Even in the case of multiple interpreters with different TBoxes, we will assume that they all agree on the top-level ontology, to which their ontologies are attached as domain modules. For the sake of example, we rely on DOLCEbasic$_{\mathrm{OWL}}$ (here abbireviated $\mathcal{D}$), the OWL2 version of DOLCE, as top-level part [21]. Clearly, other top-level ontologies, such as BFO [22] or UFO [23], could be used as well. The agreement on a top-level ontology is a technical move to provide a minimal common ground for the possibly distinct interpreters' TBoxes. For instance, interpreters' domain ontologies will contain axioms that connect their domain concepts to DOLCE, such as $\mathsf{Detective} \sqsubseteq \mathsf{AgentivePhysicalObject}$, $\mathsf{Violin} \sqsubseteq \mathsf{NonAgentivePhysicalObject}$, $\mathsf{Tobacco} \sqsubseteq \mathsf{AmountOfMatter}$.[3]

We assume that a knowledge base KB is given by a TBox ($\mathcal{T}$) and an ABox ($\mathcal{A}$), i.e. $KB = \mathcal{T} \cup \mathcal{A}$. The set of *subconcepts* of KB is given by the subconcepts of the concepts included in $\mathcal{T}$ and in $\mathcal{A}$ [24]. By $\mathsf{sub}(C)$ we denote the set of subconcepts of C, inductively defined over the structure of C. The set of subconcepts in an axiom $C \sqsubseteq D$ is thus $\mathsf{sub}(C \sqsubseteq D) = \mathsf{sub}(C) \cup \mathsf{sub}(D)$ and $\mathsf{sub}(C(a)) = \mathsf{sub}(C)$.

Given an ABox $\mathcal{A}$, we focus on a subset of $\mathcal{A}$ that contains statements about a literary character ch, written $\mathcal{A}^{\mathsf{ch}} \subseteq \mathcal{A}$. We denote by ϕ a concept $C \in \mathsf{sub}(KB)$ or a role name R occurring in KB. Thus, $\phi(\mathsf{ch})$ may indicate a statement of the form $C(\mathsf{ch})$, $R(\mathsf{ch}, c)$ or $R(c, \mathsf{ch})$, for some individual name c in KB. Therefore, $\mathcal{A}^{\mathsf{ch}} = \{\phi(\mathsf{ch}) \mid \phi(\mathsf{ch}) \in \mathcal{A}\}$.

In both scenarios (*i*) and (*ii*), we assume that interpreters have possibly conflicting ABoxes $\mathcal{A}_i$ including different statements about the character ch_i, $\mathcal{A}_i^{\mathsf{ch}_i} \subseteq \mathcal{A}_i$ (so they have possibly distinct knowledge bases KB_i). We also assume that each individual name

[3]By restricting to $\mathcal{ALC}$ domain ontologies, we do not enable role hierarchies. So roles are not connected to the top-level. We leave the semantics of relations for a dedicated work.

ch_i can occur only in the corresponding $\mathcal{A}_i$ and that ch is a fresh individual name that does not occur in any $\mathcal{A}_i$ nor in $\mathcal{T}_i$.

We denote by $\mathcal{A}_i^{\mathsf{ch}_i/\mathsf{ch}}$ the uniform substitution of all occurrences of ch_i with ch in each statement of $\mathcal{A}_i^{\mathsf{ch}_i}$. This is a purely technical move that is useful to compare interpretations by computing similarities between the ch_i. Finally, we assume that each TBox $\mathcal{T}_i$ (or the shared TBox $\mathcal{T}$in case (*i*)) is consistent, and that each $\mathcal{A}_i$ is consistent with $\mathcal{T}_i$ (or with $\mathcal{T}$ in scenario (*i*)). It follows that, if $\mathcal{A}_i^{\mathsf{ch}_i}$ is consistent with $\mathcal{T}$, then $\mathcal{A}_i^{\mathsf{ch}_i/\mathsf{ch}}$ is consistent with $\mathcal{T}$, too.

In general, ABoxes are not deductively closed with respect to their TBoxes. We define the *deductive closure* of $\mathcal{A}^{\mathsf{ch}}$ w.r.t the TBox $\mathcal{T}_i$, the subconcepts of the ontology, and ch, as $cl_{\mathcal{T}_i}(\mathcal{A}) = \{\phi(\mathsf{ch}) \mid \mathcal{T}_i, \mathcal{A}^{\mathsf{ch}} \models \phi(\mathsf{ch})\}$.

3.1. Multiple Interpreters, Single Ontology

In this case, *n* interpreters have the same TBox $\mathcal{T}$. According to what said above, this implies that they share a common upper-level ontology and fully agree on the specific ontology describing the characters' properties. The features ascribed to characters by interpreters are represented by concepts or roles ϕ occurring in their $\mathcal{A}_i^{\mathsf{ch}_i}$.

In this scenario, interpreters agree on the meanings of the properties in the shared ontology $\mathcal{T}$, while they might disagree on which of these properties apply to the interpreted character.

Claim 1 (Character identification) *Given N interpreters equipped with a single TBox $\mathcal{T}$, and ABoxes $A_1^{ch_1}, \dots, A_n^{ch_n}$, if $\bigcup_{i \in \mathcal{N}} \mathcal{A}_i^{ch_i/ch}$ is inconsistent with $\mathcal{T}$, then we cannot identify a single character ch.*

Claim 1 provides a *necessary* condition for possible identification, while it gives room for distinguishing multiple characters on the basis of more restrictive grounds than the sole compatibility of information.[4]

A more severe restriction would be to claim that consistency is also *sufficient* for identifying characters: if $\bigcup_{i \in \mathcal{N}} \mathcal{A}_i^{\mathsf{ch}_i/\mathsf{ch}}$ is consistent, then we can identify a single ch. This stronger conditions amounts to assuming that interpreters, in order to distinguish characters, have to explicitly offer conflicting character properties, whereas this is not always the case, as shown in Example 1.

Example 1 *Suppose $\mathcal{T} = \{C \sqsubseteq D, D \sqsubseteq E\}$ and interpreters ABoxes are as follows.*

$\mathcal{A}_1^{ch_1/ch}$	$\{C(ch),$	$D(ch)\}$	
$\mathcal{A}_2^{ch_2/ch}$		$\{D(ch),$	$E(ch)\}$

In this case, interpreters 1 and 2's ABoxes are different. Once we substitute ch_i with ch, their union $\{C(ch), D(ch), E(ch)\}$ is consistent with $\mathcal{T}$. In this example, interpreters provide no explicit information to differentiate the characters. However, notice that interpreter 2's ABox is consistent with both $C(ch_2)$ and $\neg C(ch_2)$, and we do not know

[4]We are thus excluding here a Meinongian view of characters by which there can be characters exhibiting contradictory properties. We also reject the comprehension principle, hence, it is not the case that all sets of properties return a character; see [2] for some discussion on Meinongian approaches.

which is 2's view. For this reason, identifying characters only on ground of consistency might be too restrictive.[5] *In a nutshell, according to Claim 1, in this case it is possible to identify the characters. However, we might have no sufficient reasons to do it.*

The interesting case is therefore when the union of the properties ascribed to the character by interpreters is plainly inconsistent with respect to $\mathcal{T}$. This might happen even in simple cases, as in Example 2. According to Claim 1, in this case we cannot assume that a single character among different interpreters exists.

Example 2 *Suppose* $\mathcal{T} = \{C \sqsubseteq \neg D\}$ *and interpreters 1 and 2 have the following ABoxes.*

$\mathcal{A}_1^{ch_1/ch}$	$\{C(ch)\}$	
$\mathcal{A}_2^{ch_2/ch}$		$\{D(ch)\}$

Each interpreter has an ABox consistent with $\mathcal{T}$. *However, their union* $\{C(ch), D(ch)\}$ *is inconsistent with* $\mathcal{T}$ *(ch is both D and* $\neg$*D). Therefore, in this case, we have explicit reasons to claim that the two interpreters are intending distinct characters, ch-according-to-1 (call it* ch_1*), and ch-according-to-2 (*ch_2*).*

We now turn to discuss a number of similarity measures with values in $[0,1]$, which can be used either to formulate sufficient conditions for character identification (e.g., when the similarity between ch_1 and ch_2 is sufficiently high) or to quantify differences between characters.

3.1.1. Similarity Measures

In general, when (possibly) distinct characters ch_1 and ch_2 are described by ABoxes $\mathcal{A}_1^{\mathsf{ch}_1}$ and $\mathcal{A}_2^{\mathsf{ch}_2}$ (respectively), we can ask how similar the two are. In our view, the problem of defining a relation of similarity between characters leverages the similarity between sets of ABox statements.

We start by introducing a simple notion of syntactic similarity counting the properties that are *explicitly* asserted about characters.

$$\mathsf{ch}_1 \sim_{syn} \mathsf{ch}_2 = \frac{|\{\phi \mid \phi(\mathsf{ch}) \in \mathcal{A}_1^{\mathsf{ch}_1/\mathsf{ch}} \cap \mathcal{A}_2^{\mathsf{ch}_2/\mathsf{ch}}\}|}{|\{\phi | \phi(\mathsf{ch}) \in \mathcal{A}_1^{\mathsf{ch}_1/\mathsf{ch}} \cup \mathcal{A}_2^{\mathsf{ch}_2/\mathsf{ch}}\}|} \tag{1}$$

This measure simply counts the properties that explicitly belong to both $\mathcal{A}_1$ and $\mathcal{A}_2$, over the set of all properties that occur in $\mathcal{A}_1$ or in $\mathcal{A}_2$.

Example 3 *Consider the TBox* $\mathcal{T} = \{C \sqsubseteq R, S \sqsubseteq R\}$, $\mathcal{A}_1^{ch_1} = \{C(ch_1)\}$ *and* $\mathcal{A}_2^{ch_2} = \{S(ch_2)\}$. *Thus,* $\mathcal{A}_1^{ch_1/ch} = \{C(ch)\}$ *and* $\mathcal{A}_2^{ch_2/ch} = \{S(ch)\}$. *The similarity between* ch_1 *and* ch_2*, according to the syntactic similarity, is then* $ch_1 \sim_{syn} ch_2 = \frac{0}{2} = 0$. *I.e., the characters have no properties in common. If* $\mathcal{A}_2^{ch_2/ch} = \{C(ch)\}$*, then* $ch_1 \sim_{syn} ch_2 = 1$.

Consider Example 1. In this case, the syntactic similarity is $ch_1 \sim_{syn} ch_2 = \frac{1}{3}$.

[5] Although multiple approaches to formally deal with the evolution of ontologies in both their extensional and intensional parts (see, for instance, [25,26]) exist, in the present work we do not consider the scenario in which an interpreter wants to keep track of, and possibly reason on the changes in time of its character interpretations. We leave this topic to future research relative to our work.

Equation 1 only counts explicit information contained in the interpreters' ABox, regardless the information that can be *inferred* from the ABoxes by means of $\mathcal{T}$. Differently, Equation 2 takes into account characters' properties which can be inferred from their ABox statements. Notice, however, that computing deductive closures is usually computationally costly.

$$\mathsf{ch}_1 \sim_{sem} \mathsf{ch}_2 = \frac{|\{\phi \mid \phi(\mathsf{ch}) \in cl(\mathcal{A}_1^{\mathsf{ch}_1/\mathsf{ch}}) \cap cl(\mathcal{A}_2^{\mathsf{ch}_2/\mathsf{ch}})\}|}{|\{\phi \mid \phi(\mathsf{ch}) \in cl(\mathcal{A}_1^{\mathsf{ch}_1/\mathsf{ch}} \cup \mathcal{A}_2^{\mathsf{ch}_2/\mathsf{ch}})\}|} \tag{2}$$

The difference between syntactic and semantic similarities is evident when considering a TBox $\mathcal{T} = \{C \equiv D\}$ and ABoxes $\mathcal{A}_1^{\mathsf{ch}_1/\mathsf{ch}} = \{C(\mathsf{ch})\}$ and $\mathcal{A}_2^{\mathsf{ch}_2/\mathsf{ch}} = \{D(\mathsf{ch})\}$. In this case the syntactic similarity is 0, whereas the semantic similarity is 1. Further differences are illustrated by the following two examples.

Example 4 *Consider again the case of Example 3. The set* $cl(\mathcal{A}_1^{ch_1/ch} \cup \mathcal{A}_2^{ch_2/ch})$ *is* $\{C(ch), R(ch), S(ch)\}$. *Thus, the semantic similarity is* $ch_1 \sim_{sem} ch_2 = \frac{1}{3}$, *while* $ch_1 \sim_{syn} ch_2 = 0$. *Therefore, the similarity increases, in this case, by taking into account inferred properties.*

Example 5 *Assume that the shared TBox is* $\mathcal{T} = \{C \sqsubseteq \exists R.D, \exists R.D \sqsubseteq E, E \sqsubseteq H \sqcap I, G \sqsubseteq F\}$, *and the interpreters' ABoxes are:*

$\mathcal{A}_1^{ch_1/ch}$	$=$	$\{C(ch), I(ch), R(ch,b)\}$
$\mathcal{A}_2^{ch_2/ch}$	$=$	$\{G(ch), I(ch), R(ch,b)\}$

In this case, $ch_1 \sim_{syn} ch_2 = \frac{2}{4} = 0.5$. *To compute the semantic similarity, we look at the deductive closures:*

$cl(\mathcal{A}_1^{ch_1/ch})$	$=$	$\{C(ch), \exists R.D(ch), E(ch), H \sqcap I(ch), H(ch), I(ch), R(ch,b)\}$
$cl(\mathcal{A}_2^{ch_2/ch})$	$=$	$\{G(ch), I(ch), R(ch,b), F(ch)\}$

So $ch_1 \sim_{sem} ch_2 = \frac{2}{9} \approx 0.22$. *Thus, in this case, by considering also inferred information, the semantic similarity is lower than the syntactic one.*

Syntactic and semantic similarities can be used to quantify differences between distinct characters, notice that they also differ significantly in the way they cope with inconsistent ABoxes.

Example 6 *Consider the TBox* $\mathcal{T} = \{C \sqsubseteq \exists R.D, \exists R.D \sqsubseteq E, E \sqsubseteq H \sqcap I, G \equiv \neg C, I \sqsubseteq L, L \sqsubseteq \neg M\}$ *and interpreters' ABoxes as in Example 5. Now, the union of the previous ABoxes is inconsistent w.r.t.* $\mathcal{T}$, *i.e., the subset* $\{C(ch), G(ch)\}$ *is inconsistent with the TBox axiom* $G \equiv \neg C$.

The syntactic similarity in this case is the same as in Example 5, i.e, $ch_1 \sim_{syn} ch_2 = 0.5$. *By contrast, the semantic similarity significantly decreases. Since the union of the ABoxes is inconsistent, by 'ex falso quodlibet', the denominator will contain all the possible statements* $\phi(ch)$, *for any* $\phi \in \mathsf{sub}(KB)$. *I.e.,* $ch_1 \sim_{sem} ch_2 = \frac{2}{12} \approx 0.16$, *as there are 12 subconcpts and roles in* $\mathsf{sub}(KB)$.

Therefore, the semantic similarity of two characters which are associated to inconsistent ascriptions of properties will be very small. In particular, it can never be 1 —given that each $\mathcal{A}_i$ is assumed to be consistent— and it significantly decreases as the number of concepts used in $\mathsf{sub}(KB)$ *increases.*

A quite standard choice to measure distances (and also similarity) between sets is by means of the Hamming distance [27]. Here, the Hamming distance counts the differences between two sets of formulas. Let $\mathcal{A}(\phi(\mathsf{ch})) = 1$ if $\phi(\mathsf{ch}) \in \mathcal{A}$ and $\mathcal{A}(\phi(\mathsf{ch})) = 0$, if $\phi(\mathsf{ch}) \notin \mathcal{A}$. To define the Hamming distance between $\mathcal{A}_1^{\mathsf{ch}_1}$ and $\mathcal{A}_2^{\mathsf{ch}_2}$ we replace ch_i with ch. Accordingly, we can define a similarity based on h, $\sim_H$.

$$h(\mathcal{A}_1^{\mathsf{ch}_1/\mathsf{ch}}, \mathcal{A}_2^{\mathsf{ch}_2/\mathsf{ch}}) = \sum_{\phi(\mathsf{ch}) \in \mathcal{A}_1 \cup \mathcal{A}_2} |\mathcal{A}_1(\phi(\mathsf{ch})) - \mathcal{A}_2(\phi(\mathsf{ch}))| \tag{3}$$

$$\mathsf{ch}_1 \sim_H \mathsf{ch}_2 = 1 - \frac{h(\mathcal{A}_1^{\mathsf{ch}_1/\mathsf{ch}}, \mathcal{A}_2^{\mathsf{ch}_2/\mathsf{ch}})}{|\mathcal{A}_1^{\mathsf{ch}_1/\mathsf{ch}} \cup \mathcal{A}_2^{\mathsf{ch}_2/\mathsf{ch}}|} \tag{4}$$

As noted in [12], the Hamming distance may not be a good choice when discussing KBs based on $\mathcal{DL}$, which endorse OWA. In particular, the Hamming distance counts the difference between two sets S and S' by giving the same weight (i.e., 1) to the case when a statement belongs to S and not S', and to the case when a statement belongs to S' and not S. However, due to the asymmetry of positive and negative information in $\mathcal{DL}$, it is interesting to introduce in our framework an *asymmetric* difference measure (strictly speaking, due to the lack of symmetry, it does not return a *distance*), denoted by aH (for asymmetric Hamming difference), cf. [28,12]. The asymmetric similarity reflects the fact that interpreter 1 might be more keen to consider disagreement when 1 is asserting something and 2 is not, rather than when 1 is not asserting something that 2 does. Accordingly, we define an asymmetric similarity measure as follows.

$$aH(\mathcal{A}_1, \mathcal{A}_2) = |\{\phi(\mathsf{ch}) \in \mathcal{A}_1 \cup \mathcal{A}_2 \mid \mathcal{A}_1(\phi(\mathsf{ch})) = 1 \text{ and } \mathcal{A}_2(\phi(\mathsf{ch})) = 0\}| \tag{5}$$

$$\mathsf{ch}_1 \sim_{aH} \mathsf{ch}_2 = 1 - \frac{aH(\mathcal{A}_1^{\mathsf{ch}_1/\mathsf{ch}}, \mathcal{A}_2^{\mathsf{ch}_2/\mathsf{ch}})}{|\mathcal{A}_1^{\mathsf{ch}_1/\mathsf{ch}} \cup \mathcal{A}_2^{\mathsf{ch}_2/\mathsf{ch}}|} \tag{6}$$

We can also introduce a semantic version of the Hamming similarities. E.g., for the asymmetric case, we are counting the asymmetric difference of inferred properties.

$$\mathsf{ch}_1 \sim_{saH} \mathsf{ch}_2 = 1 - \frac{aH(cl(\mathcal{A}_1^{\mathsf{ch}_1/\mathsf{ch}}), cl(\mathcal{A}_2^{\mathsf{ch}_2/\mathsf{ch}}))}{|cl(\mathcal{A}_1^{\mathsf{ch}_1/\mathsf{ch}} \cup \mathcal{A}_2^{\mathsf{ch}_2/\mathsf{ch}})|} \tag{7}$$

Example 7 *Consider the TBox $\mathcal{T}$ of Example 5. Interpreters' statements are depicted as follows.*

	$C(ch)$	$\exists R.D(ch)$	$E(ch)$	$H(ch)$	$I(ch)$	$G(ch)$	$F(ch)$	$R(ch,b)$
$\mathcal{A}_1^{ch_1/ch}(\phi(ch))$	1	1	1	0	0	0	1	1
$\mathcal{A}_2^{ch_2/ch}(\phi(ch))$	0	0	1	1	1	1	1	1

The Hamming distance between the Aboxes is 5, so $ch_1 \sim_H ch_2$ *is then* $1-5/8 \approx 0.37$. *The asymmetric Hamming distances, instead, are 2 and 3 (respectively). Therefore,* $ch_1 \sim_{aH} ch_2 = 1-2/8 = 0.75$ *and* $ch_2 \sim_{aH} ch_1 = 1-3/8 = 0.62$. *Accordingly,* ch_1 *is more similar to* ch_2 *than* ch_2 *is to* ch_1*, since more assertions of* 2 *are not accepted by* 1 *than the converse.*

The semantic Hamming distances consider the deductive closures of the ABoxes.

	$C(ch)$	$\exists R.D(ch)$	$E(ch)$	$H(ch)$	$I(ch)$	$G(ch)$	$F(ch)$	$R(ch,b)$
$cl(\mathcal{A}_1^{ch_1/ch}(\phi(ch)))$	1	1	1	1	1	0	1	1
$cl(\mathcal{A}_2^{ch_2/ch}(\phi(ch)))$	0	0	1	1	1	1	1	1

So the (semantic) Hamming similarity would be $1-3/8 = 0.62$. *The semantic asymmetric similarities are:* $ch_1 \sim_{saH} ch_2 = 1-\frac{2}{8} = 0.75$ *and* $ch_2 \sim_{saH} ch_1 = 1-\frac{1}{8} = 0.87$. *Thus, including inferred properties,* ch_2 *is more similar to* ch_1 *than the converse.*

We conclude our overview of similarity measures by discussing an important metric in cognitive science, namely the Tversky's contrast rule [9,29].[6] Tversky's similarity was originally used to define the distance between the prototype of a concept and an instance in order to provide cognitively adequate conditions of classification. A common aspect with our setting is that similarity is computed in terms of the attributes. A significant difference is that, while Tversky's rule introduces weights for the attributes (i.e., diagnosticity values), here we do not do that. Namely, we cannot assume that all interpreters agree on the values representing the relevance of a property for a character.[7]

Tversky's contrast rules, without diagnosticity values, define the similarity between two entities a and b as the ratio between the number of attributes that a and b have in common, and the attributes that a has and b does not have plus the attributes that b has and a does not have (cf. [29]). Given these restrictions, Tversky's similarity can be defined by means of the asymmetric (or even symmetric) Hamming distance. As as an example, we offer a semantic version of the Tversky similarity that uses inferred information, where we abbreviate by $\mathcal{A}_i$ the notation $\mathcal{A}_i^{\mathsf{ch}_i/\mathsf{ch}}$.

$$\mathsf{ch}_1 \sim_{sTv} \mathsf{ch}_2 = \frac{|cl(\mathcal{A}_1)\cap cl(\mathcal{A}_2)|}{|cl(\mathcal{A}_1)\cap cl(\mathcal{A}_2)| \, + \, aH(cl(\mathcal{A}_1),cl(\mathcal{A}_2)) \, + \, aH(cl(\mathcal{A}_2),cl(\mathcal{A}_1))} \tag{8}$$

Example 8 *Consider again the TBox and the ABoxes of Example 7. In that case, the semantic Tversky similarity is computed as follows:* $ch_1 \sim_{sTv} ch_2 = \frac{5}{5+2+1} = 0.62$

3.2. Multiple Interpreters, Multiple Ontologies

In this section, we discuss the viable definitions of similarity when interpreters might have different domain ontologies. Following the discussion of Claim 1, we make explicit the condition of possible character identification also in this setting. In principle, we have two choices: *(i)* we find a common reference ontology $\mathcal{R}$, such that $\mathcal{R} \subseteq \bigcup_i \mathcal{T}_i$, with respect to which we assess the consistency of the ABoxes, or *(ii)* we let interpreters

[6]For a version of Tversky measure to compute similarities between individual names in ontologies, see [30].

[7]Interpreters might have different views of the diagnostic property or of the importance of a property; we will briefly discuss this point in Section 3.3.

maintain their semantic views, so that each interpreter assesses an ABox w.r.t. its $\mathcal{T}_i$. The second choice is at odds with Claim 1 for the following reason.

Example 9 *Suppose* $\mathcal{T}_1 = \{C \sqsubseteq D\}$ *and* $\mathcal{T}_2 = \{C \sqsubseteq \neg D\}$, $\mathcal{A}_1^{ch_1/ch} = \{C(ch)\}$, $\mathcal{A}_2^{ch_1/ch} = \{C(ch)\}$. *Each ABox is consistent with its TBox, however* $\mathcal{T}_1 \cup \mathcal{T}_2$ *is inconsistent with the ABoxes. So we cannot use the union of the TBoxes to assess the consistency of the properties ascribed to the character. Notice also that the syntactic similarity between* ch_1 *and* ch_2 *is 1. It is puzzling to consider this situation as a possible identification of a single character* ch*, not knowing with respect to which ontology its consistency is assessed.*

For these reasons, we then rephrase the idea of Claim 1 as follows.

Claim 2 (Character identification (multiple ontologies)) *Given N interpreters equipped with TBoxes* $\mathcal{T}_i$*, and ABoxes* $A_1^{ch_1}, \ldots, A_n^{ch_n}$*, if* $\bigcup_{i \in \mathcal{N}} \mathcal{A}_i^{ch_i/ch}$ *is inconsistent with a certain reference ontology* $\mathcal{R}$*, then we cannot identify a single character* ch *w.r.t* $\mathcal{R}$.

Therefore, in this scenario the problem becomes the selection of a reference ontology $\mathcal{R}$ to assess the consistency of the assertions of interpreters. Firstly, notice that the union of the $\mathcal{T}_i$ is likely to result in an inconsistent TBox (as in Example 9). However, if $\bigcup_i \mathcal{T}_i$ is consistent with each $\mathcal{A}_i$, then we can reduce this scenario to the cases of Section 3.1 without losing information from interpreters. On the other hand, when $\bigcup_i \mathcal{T}_i$ is inconsistent, we have to select an appropriate consistent subset. According to our assumptions, interpreters' disagreement about the TBoxes concerns the domain aspects, whereas they all agree on a single top-level ontology. That means that there exists a non-empty *consistent* reference ontology that is unanimously accepted among all interpreters, i.e. $\mathcal{R} = \bigcap_i \mathcal{T}_i$. In the worst case, $\mathcal{R}$ coincides with the top-level ontology.[8] Selections of the common $\mathcal{R}$ that are less restrictive than the intersection can be studied as a problem of ontology debugging, integration, or aggregation and can be approached with techniques from [12,24,31,32], among others. We do not enter here into the details relative to the selection of a common ontology; we only exhibit a few possibilities by means of the following example.

Example 10 *Consider the following two TBoxes, where* $\mathcal{T}_1$ *complies with Conan Doyles' traditional view of Sherlock Holmes, whereas* $\mathcal{T}_2$ *corresponds to the anthropomorphized animal view of the animated version "Sherlock Hound", where most characters are in fact dogs. Here* APO *abbreviates the class 'agentive physical object' of* $\mathcal{D}$.[9]

$$\begin{aligned} \mathcal{T}_1 = \quad & \mathcal{D} \cup \{Detective \sqsubseteq Rational, Rational \sqsubseteq APO, Detective \sqsubseteq APO, \\ & Rational \sqsubseteq \neg Dog,\ \ Chihuahua \sqsubseteq Dog\} \\ \mathcal{T}_2 = \quad & \mathcal{D} \cup \{Detective \sqsubseteq Dog, Dog \sqsubseteq APO, Detective \sqsubseteq APO\} \end{aligned}$$

The union $\mathcal{T}_1 \cup \mathcal{T}_2$ *is inconsistent with any ABox statement about detectives. The most restrictive choice is to consider the intersection* $\mathcal{T}_1 \cap \mathcal{T}_2 = \mathcal{D} \cup \{Detective \sqsubseteq APO\}$.

[8]If each $\mathcal{T}_i$ is consistent with $\mathcal{A}_i$, then their intersection is consistent with $\mathcal{A}_i$. Notice that, if interpreters disagree also on the axioms connecting the domain concepts to the top-level, then a concept C occurring in A_i is not in the common reference ontology. So the deductive closure in this case just returns the ABox $\mathcal{A}_i$.

[9]In principle, we should consider the intersection of the deductive closure of the $\mathcal{T}_i$. However, including $\mathcal{D}$ results in very large sets of axioms; so, for the sake of example, we illustrate the case of the intersection of $\mathcal{T}_i$.

The ontology is consistent, but it does not infer the classification of detectives as dogs or rational beings, only as agentive physical objects.

A debugging solution (as discussed, e.g., in [31]) searches for a maximal consistent subset of $\mathcal{T}_1 \cup \mathcal{T}_2$, e.g., by removing $\mathsf{Rational} \sqsubseteq \neg\mathsf{Dog}$. In this case, classifications of detectives as dogs and as rational beings are preserved, consenting rational dogs.

A solution based on axiom weakening (see [24]) can replace $\mathsf{Rational} \sqsubseteq \neg\mathsf{Dog}$ with a (logically) weaker version of it, e.g., by replacing it with $\mathsf{Rational} \sqsubseteq \neg\mathsf{Chihuahua}$, thereby weakening the meaning of the concept 'rational'.

3.2.1. Similarity Measures

The definitions of syntactic similarities ($\sim_{syn}$, $\sim_H$, $\sim_{aH}$) can be directly imported in this scenario even in the absence of a common reference ontology. Their significance is however questionable (as shown in Example 9). By contrast, the definition of semantic similarities ($\sim_{sem}$, $\sim_{saH}$, and $\sim_{sTv}$) depends on the reference ontology $\mathcal{R}$.

Once $\mathcal{R}$ is selected, we can measure its compatibility with an interpreter's $\mathcal{T}_i$ by means of a number of measures (e.g, the *happiness* measures in [32]). As an example, we consider the following one, computing the ratio of inferences of $\mathcal{T}_i$ preserved by $\mathcal{R}$.

$$c_i(\mathcal{T}_i, \mathcal{R}) = \frac{|\{\phi \mid \phi(\mathsf{ch}) \in cl_{\mathcal{T}_i}(\mathcal{A}_i^{\mathsf{ch}_i/\mathsf{ch}}) \cap cl_{\mathcal{R}}(\mathcal{A}_i^{\mathsf{ch}_i/\mathsf{ch}})\}|}{|\{\phi \mid \phi(\mathsf{ch}) \in cl_{\mathcal{T}_i}(\mathcal{A}_i^{\mathsf{ch}_i/\mathsf{ch}})\}|} \tag{9}$$

Notice that in this context the values of the semantic similarities can vary across interpreters, depending on how close is $\mathcal{R}$ to their $\mathcal{T}_i$. For this reason, we define a notion of a semantic similarity for interpreter i. Denote by $\sim_S$, for $S \in \{sem, saH, sTv\}$, any of the semantic similarities. The semantic similarity for interpreter i w.r.t. $\mathcal{R}$ is then: $\mathsf{ch}_1 \sim_S^i \mathsf{ch}_2 = c_i \times \mathsf{ch}_1 \sim_S \mathsf{ch}_2$. For example, $\mathsf{ch}_1 \sim_S^i \mathsf{ch}_2 = \mathsf{ch}_1 \sim_S \mathsf{ch}_2$, when $c_i = 1$, namely when $\mathcal{T}_i$ is closest to $\mathcal{R}$.

To define a global semantic similarity relation, not depending on a single interpreter and common to all interpreters in $\mathcal{N}$, several choices are possible.

As an example, we offer the following definition. As we work with similarity with values in $[0,1]$, its maximal value is always $N = |\mathcal{N}|$.[10]

$$\mathsf{ch}_1 \sim_S^{\mathcal{N}} \mathsf{ch}_2 = \frac{\sum_{i \in \mathcal{N}} (\mathsf{ch}_1 \sim_S^i \mathsf{ch}_2)}{N} \tag{10}$$

Example 11 *Suppose $\mathcal{A}_1^{ch_1/ch} = \{\mathsf{Human}(ch)\}$, $\mathcal{A}_2^{ch_2/ch} = \{\mathsf{Dog}(ch)\}$, and*

[10]This definition is inspired by the classical *utilitarian* measure of multiagent welfare. Other choices are possible, e.g., the minimum of the values of c_i (*egalitarian* welfare) or the product of the c_i (*Nash* welfare). Each of these choices entails assumptions to assess the interpreters' contributions, cf. [33]. We leave a discussion of their features for future work.

$$\begin{array}{lcr} \mathcal{T}_1 & = & \{Human \sqsubseteq Rational, Human \sqsubseteq APO, Human \sqsubseteq \neg Dog\} \\ \mathcal{T}_2 & = & \{Dog \sqsubseteq Rational, Dog \sqsubseteq APO, Rational \sqsubseteq \neg Human\} \\ \mathcal{R} & = & \{Dog \sqsubseteq Rational, Dog \sqsubseteq APO, Human \sqsubseteq APO, \\ & & Rational \sqsubseteq \neg Human, Human \sqsubseteq \neg Dog\} \\ cl_{\mathcal{T}_1}(\mathcal{A}_1^{ch_1/ch}) & = & \{Human(ch), Rational(ch), APO(ch), \neg Dog(ch)\} \\ cl_{\mathcal{T}_2}(\mathcal{A}_2^{ch_2/ch}) & = & \{Dog(ch), Rational(ch), APO(ch), \neg Human(ch)\} \\ cl_{\mathcal{R}}(\mathcal{A}_1^{ch_1/ch}) & = & \{Human(ch), APO(ch), \neg Rational(ch), \neg Dog(ch)\} \\ cl_{\mathcal{R}}(\mathcal{A}_2^{ch_2/ch}) & = & \{Dog(ch), Rational(ch), APO(ch), \neg Human(ch)\} \end{array}$$

We have: $c_1(\mathcal{T}_1, \mathcal{R}) = 3/4 = 0.75$ *and* $c_2(\mathcal{T}_2, \mathcal{R}) = 4/4 = 1$. *The semantic similarity between* $cl_{\mathcal{R}}(\mathcal{A}_1^{ch_1/ch})$ *and* $cl_{\mathcal{R}}(\mathcal{A}_2^{ch_2/ch})$ *is* $1/7 \approx 0.14$. *Therefore,* $ch_1 \sim^1_{sem} ch_2 = 0.75 \times 0.14 \approx 0.1$ *and* $ch_1 \sim^2_{sem} ch_2 = 1 \times 0.14 = 0.14$. *A semantic similarity common to interpreters 1 and 2 can be obtained by means of Equation 10:* $ch_1 \sim^{\mathcal{N}}_{S} ch_2 = (0.75 \times 0.14) + (1 \times 0.14)/2 = 0.1 + 0.14/2 \approx 0.12$.

3.3. Types of Properties

So far, we discussed similarity measures that do not take into account the different types of properties that can occur in the interpreters' $\mathcal{A}_i$. Specifically, all the axioms and properties represented in the KBs have been interpreted as equally important. In this section, we will briefly comment on this aspect, leaving to future work a more extended treatment of the topic. In particular, among the various properties that interpreters may use in their interpretations, we focus here on *diagnostic* properties. It is important to stress that, as discussed in Section 2, the fact that certain properties are relevant to identify characters is interpretation-based rather than a strong metaphysical fact. Hence, it is perfectly admissible for interpreters to select different diagnostic properties to interpret the character at stake, in which case one wonders whether they mean the same character.[11]

Accordingly, Claim 1 and 2 could be revised, by demanding the consistency of the TBox only with respect to the diagnostic properties of the character.

To represent interpreters' ABoxes containing properties of heterogeneous importance, we could use, for each ϕ and interpreter i, a weight $w^i_\phi \in [0,1]$. The value might represent the diagnostic contribution of the property to the character identification (when w^i_ϕ is 1, we might view ϕ as a very relevant diagnostic property of the character, according to interpreter i). In a general setting, we cannot assume that interpreters agree on the values of each attribute, which is what causes the similarity measures based on interpreters' weighted properties to be asymmetric by definition. Thus, we assume that each interpreter is equipped with a weight $w^i_\phi > 0$, for each ϕ such that $\phi(\mathsf{ch}_i) \in \mathcal{A}_i$, while for ϕ not occurring in $\mathcal{A}_i$, we may set, as a simplification, $w^i_\phi = 0$.

We can sketch how to define a similarity measure that takes into account the weights. We only illustrate the case of a Hamming-inspired measure, leaving a dedicated analysis to future work.

[11] We also leave the case of measuring different levels of importance for the axioms occurring in the TBoxes to future work. Methods allowing this have been proposed, e.g., in [34,32].

$$h_w^i(\mathcal{A}_i^{\mathsf{ch}_i/\mathsf{ch}}, \mathcal{A}_j^{\mathsf{ch}_j/\mathsf{ch}}) = \sum_{\phi(\mathsf{ch}) \in \mathcal{A}_i \cup \mathcal{A}_j} w_\phi^i \times |\mathcal{A}_i(\phi(\mathsf{ch})) - \mathcal{A}_j(\phi(\mathsf{ch}))| \tag{11}$$

$$\mathsf{ch}_i \sim_{w^i H} \mathsf{ch}_j = 1 - \frac{h_w^i(\mathcal{A}_i^{\mathsf{ch}_i/\mathsf{ch}}, \mathcal{A}_j^{\mathsf{ch}_j/\mathsf{ch}})}{\sum_{\{\phi | \phi(\mathsf{ch}) \in \mathcal{A}_i^{\mathsf{ch}_i/\mathsf{ch}} \cup \mathcal{A}_j^{\mathsf{ch}_j/\mathsf{ch}}\}} w_\phi^i} \tag{12}$$

This similarity is defined from interpreter i's perspective on the importance of the properties. The weighted Hamming distance h_w^i assesses the discriminating properties between the two ABoxes, by multiplying them by their weight. Finally, the weighted similarity $\sim_{w^i H}$ is 1 minus the ratio of weighted differences over the sum of weights.

Example 12 *Suppose* $\mathcal{A}_1^{ch_1/ch} = \{C(ch), D(ch)\}$*,* $\mathcal{A}_1^{ch_2/ch} = \{D(ch), E(ch)\}$*. The Hamming similarity is 1/3. Suppose the weights are as follows:* $w_C^1 = 0.8$*,* $w_D^1 = 0.2$*,* $w_E^1 = 0$*, and* $w_C^2 = 0$*,* $w_D^2 = 0.5$*,* $w_E^2 = 0.9$*. The weighted distance for interpreter 1 is* $0.8 \times 1 + 0.2 \times 0 + 0 \times 1 = 0.8$*, while for interpreter 2,* $0 \times 1 + 0.5 \times 0 + 0.9 \times 1 = 0.9$*. So* $\sim_{w^1 H}$ *is* 0.2*, while* $\sim_{w^2 H}$ *is* 0.36*. The shared D is more diagnostic for 2 than for 1.*

4. Discussion and Conclusion

We introduced and discussed in the paper a number of formal means, based on description logics and automated reasoning, that support the analysis of multiple, possibly heterogeneous, interpretations of literary characters.

The core idea is that interpreters provide their interpretations of characters by ascribing them properties which are formally represented in ontologies. The attribution of incompatible properties, leading to logical inconsistencies, is taken as an evidence that interpreters talk of different characters (see Claims 1 and 2). On the other hand, that interpretations are compatible is a hint that they might refer to the same character. Coreference is not forced, because of the inherent incompleteness of interpretations. Similarity measures help in analyzing how similar (different) interpretations are.

We distinguished between two main cases depending on whether multiple interpreters use or not the same ontology. From an application perspective, an advantage of the first approach is that one can compare interpretations in a straightforward way. In addition, the first approach aligns well with tasks relative to data integration, which can be useful when merging data of different parties. The design of a common ontology remains challenging, especially when it comes to literary characters which, as said, can be understood in different manners. The second approach provides interpreters with more flexibility since they are free to use the ontology they prefer. However, as inconsistencies are likely to emerge when comparing interpretations, our proposal is to rely on a consistent subset of the ontologies to identify similarities and differences.

Concerning the presented similarity measures, further work on case studies is needed to assess more precisely their ideal contexts of use. For the sake of the discussion, there are at least four key aspects to be considered: *(i)* explicit vs. implicit information: measures based on *syntactic similarity* (Equation 1) can be useful in applications to compare what interpretations explicitly say, whereas measures based on *semantic similarity* (Equation2) are useful to compute metrics on inferences on the knowledge captured in the ontology at use. As said, semantic similarity has some costs in terms of computability; *(ii)* symmetry vs. asymmetry: metrics on the Hamming distance can be helpful to stress what are the differences between interpretations in terms of what is assessed and not assessed by either one or the other interpretation; *(iii)* quantitative vs. qualitative values: the presented measures can *quantitatively* tell how similar (different) certain interpretations are. One should however make sense of them from a *qualitative* perspective and especially from a literary perspective. One could also think in terms of qualitative ranges of similarity to tell when two interpretations are more or less similar; *(iv)* property types: similarity measures can be tuned by taking into account different types of properties that interpreters may use for their interpretations. We started considering diagnostic properties as an interesting example but other property types should be considered as well, including, e.g., historical properties relative to the creation of characters.

As a concluding remark, it is important to stress that aspects like the inherent complexity of the literary criticism discourse, the difficulty in identifying shared, codified practices and theories in the community of reference, as well as the impact of the usage of metaphorical language as an expression of creativity in literary criticism argumentation are extremely difficult to be captured in their entirety in a formal setting. Any formal approach that aims at staying computationally tractable will resolve into an approximation of the intricacies and nuances of the literacy criticism discourse. On the other hand, as we elaborated in the present work, the formal approach provides a fundamental support for: *(i)* reducing ambiguities in the arguments by the various interpreters, *(ii)* making explicit the argumentative assumptions of the various positions on a given subject, *(iii)* qualifying the proximity or distance between the advocated theses, *(iv)* comparing the contents of the various positions with each other, and *(v)* identifying the discursive elements that generate convergence or divergence between them.

Acknowledgments

Research leading to this work was (partially) supported by the following projects: *MITE – Make it explicit: Documenting interpretations of literary fictions with conceptual formal models* (CUP B53D23028830001) funded by the European Union - Next Generation EU; the Italian PNRR MUR project PE0000013-FAIR, *Future Artificial Intelligence Research*.

References

[1] Margolin U. Characterization in narrative: Some theoretical prolegomena. Neophilologus. 1983;67(1):1-14.
[2] Kroon F, Voltolini A. Fictional Entities. In: Zalta EN, Nodelman U, editors. The Stanford Encyclopedia of Philosophy. Fall 2023 ed. Metaphysics Research Lab, Stanford University; 2023. .
[3] Damiano R, Lombardo V, Pizzo A. The ontology of drama. Applied Ontology. 2019;14(1):79-118.
[4] Pannach F, Yang X, Solissa NV, Yu Z, Van Cranenburgh A, Van Der Ree M, et al. The GOLEM Triple Store: A Graph-based Representation of Narrative and Fiction. In: Joint of the ESWC 2024 Workshops and Tutorials, ESWC-JP 2024. CEUR (vol. 3749); 2024. .
[5] Reicher ME. The ontology of fictional characters. Characters in Fictional Worlds. 2010:111-33.
[6] Hempfer KW. Fundamentals of Literary Theory. Springer; 2024.
[7] Paganini E. Vague fictional objects. Inquiry. 2020;63(2):158-84.
[8] Willis I. Reception. Routledge; 2017.
[9] Tversky A. Features of similarity. Psychological review. 1977;84(4):327.

[10] Konieczny S, Pino Pérez R. Logic based merging. Journal of Philosophical Logic. 2011;40:239-70.
[11] Shen W, Wang J, Han J. Entity linking with a knowledge base: Issues, techniques, and solutions. IEEE Transactions on Knowledge and Data Engineering. 2014;27(2):443-60.
[12] Porello D, Endriss U. Ontology merging as social choice: Judgment aggregation under the open world assumption. Journal of Logic and Computation. 2014;24(6):1229-49.
[13] Voltolini A. How ficta follow fiction: A syncretistic account of fictional entities. Springer Science & Business Media; 2006.
[14] Barolini T. Dante and Francesca da Rimini: Realpolitik, Romance, Gender. Speculum. 2000;75(1):1-28.
[15] Saint-Gelais R. Fictions transfuges: la transfictionnalité et ses enjeux. Seuil; 2011.
[16] Sanfilippo EM, Ferrario R. D3.1 - Observations Modeling: State of the Art. MITE – Make it explicit: Documenting interpretations of literary fictions with conceptual formalism; 2024. Funded by European Union – Next Generation EU. Available from: https://www.loa.istc.cnr.it/mite/wp-content/uploads/2024/09/MITE_D3_1.pdf.
[17] Ciotti F. Toward a formal ontology for narrative. MATLIT. 2016;4(1):29-44.
[18] Thomasson AL. Fictional characters and literary practices. The British Journal of Aesthetics. 2003;43(2):138-57.
[19] Eco U. On the ontology of fictional characters: A semiotic approach. Σημειωτκή-Sign Systems Studies. 2009;37(1-2):82-98.
[20] Richardson B. Transtextual Characters. In: Eder J, Jannidis F, Schneider R, editors. Characters in Fictional Worlds: Understanding Imaginary Beings in Literature, Film, and Other Media. Berlin: De Gruyter; 2011. p. 527-41.
[21] Porello D, Vieu L, Terkaj W, Borgo S, Compagno F, Sanfilippo EM. DOLCE in OWL: The Core Theory. In: Proceedings of the Joint Ontology Workshops 2024, EPISODE X: The Tukker Zomer of Ontology. CEUR (vol. 3882); 2024. p. 1-14.
[22] Otte JN, Beverley J, Ruttenberg A. BFO: Basic Formal Ontology. Applied Ontology. 2022;17(1):17–43.
[23] Guizzardi G, Benevides AB, Fonseca CM, Almeida JP, Sales TP, Porello D. UFO: Unified Foundational Ontology. Applied Ontology. 2022;1:167-210.
[24] Troquard N, Confalonieri R, Galliani P, Peñaloza R, Porello D, Kutz O. Repairing Ontologies via Axiom Weakening. In: Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2-7, 2018. AAAI Press; 2018. p. 1981-8.
[25] Smith B, Ceusters W. Ontological realism: A methodology for coordinated evolution of scientific ontologies. Applied Ontology. 2010;5(3-4):139-88.
[26] De Leenheer P, Mens T. In: Hepp M, De Leenheer P, De Moor A, Sure Y, editors. Ontology Evolution. Boston, MA: Springer US; 2008. p. 131-76.
[27] Hamming RW. Error detecting and error correcting codes. The Bell System Technical Journal. 1950;29(2):147-60.
[28] Albertoni R, De Martino M. Asymmetric and context-dependent semantic similarity among ontology instances. In: Journal on data semantics X. Springer; 2008. p. 1-30.
[29] Smith EE, Osherson DN, Rips LJ, Keane M. Combining prototypes: A selective modification model. Cognitive science. 1988;12(4):485-527.
[30] Porello D, Righetti G, Troquard N, Confalonieri R, Kutz O. An Ontological Modelling of Prototype Theories. In: Proceedings of the 9th Workshop on Formal and Cognitive Reasoning co-located with the 46th German Conference on Artificial Intelligence (KI 2023), Berlin, Germany, September 26, 2023. CEUR (vol. 3500); 2023. p. 20-31.
[31] Ji Q, Gao Z, Huang Z, Zhu M. Measuring effectiveness of ontology debugging systems. Knowledge-Based Systems. 2014;71:169-86.
[32] Porello D, Troquard N, Peñaloza R, Confalonieri R, Galliani P, Kutz O, et al. Two Approaches to Ontology Aggregation Based on Axiom Weakening. In: IJCAI. vol. 2018. AAAI Press; 2018. p. 1942-8.
[33] Endriss U, Maudet N, Sadri F, Toni F. Negotiating socially optimal allocations of resources. Journal of artificial intelligence research. 2006;25:315-48.
[34] Righetti G, Porello D, Troquard N, Kutz O, Hedblom MM, Galliani P. Asymmetric Hybrids: Dialogues for Computational Concept Combination. In: Formal Ontology in Information Systems - Proceedings of the Twelfth International Conference, FOIS 2021, Bozen-Bolzano, Italy, September 11-18, 2021. vol. 344 of Frontiers in Artificial Intelligence and Applications. IOS Press; 2021. p. 81-96.