Applied Ontology 0 (2025) 1–0
IOS Press


# An Observational Approach to Representing Interpretation

Claudio Masolo [a], Emilio M. Sanfilippo [a,b,*], Emanuele Bottazzi [a], Roberta Ferrario [a], Alessandro Mosca [a] and Marta M. Vilardo [a]
[a] *ISTC Laboratory for Applied Ontology, National Research Council (CNR), Italy*
*E-mails: claudio.masolo@loa.istc.cnr.it, emilio.sanfilippo@cnr.it, emanuele.bottazzi@cnr.it, roberta.ferrario@cnr.it, alessandro.mosca@cnr.it, marta.vilardo@istc.cnr.it*
[b] *Centre d'études supérieures de la Renaissance, University of Tours, France*

**Abstract.** Research in the Digital Humanities requires computer systems that can document and analyze interpretations of creative works, such as literary texts. These systems are designed to provide access to existing interpretations and identify their similarities and differences. To this end, we propose an approach to formally representing interpretations. In particular, we develop an ontology of observations aimed at capturing, through what we call observational vocabularies, what interpreters claim about the texts that they interpret. We distinguish between different types of observations, in particular, basic observations, which represent specific and domain-dependent claims (e.g., regarding the analysis of literary characters' traits), and observations such as assertion, denial, support, and defeat to document more nuanced claims. We also introduce formal mechanisms that can be used to analyze particular (sets of) observations, texts, etc.Throughout the paper, we illustrate the discussion with examples from literary studies.

Keywords: ontology, computational literary studies, digital humanities, interpretation, observations.

## 1. Introduction

There is an increasing attention in the Digital Humanities toward computer systems that are able to manage how scholars, critics, and even laypeople interpret artworks such as literary texts, musical compositions, theater plays, and other forms of creative expression (Barabucci et al., 2021; Pianzola et al., 2025; Sanfilippo et al., 2023; Sartini et al., 2023; Daquino et al., 2020). In the context of digital platforms, it is easy to envision a future where the metadata of artworks are enriched with data reflecting the critical discourse surrounding these works (Sanfilippo and Ferrario, 2024).

For an example relevant to this paper, which focuses on literary studies and criticism, consider the abundance of texts about classic Italian authors like Dante and Boccaccio. It would be highly valuable for research and teaching purposes to be able to access and analyze the interpretations presented in these texts using semi-automated computational methods. These methods could help to identify commonalities and differences in interpretation. Scholars and students would benefit from two key resources: a knowledge base of data on literary text interpretations and analytic transparent tools that enable the examination of these data in a manner consistent with expert knowledge.[1]

*Corresponding author. E-mail: emilio.sanfilippo@cnr.it.

[1]Viewed in a broader context of knowledge representation, this objective is comparable to practices in industrial engineering, where experts document and reason over data produced through measurement procedures (Janowicz et al., 2019).

Building on previous research (Masolo et al., 2018, 2021; Sanfilippo et al., 2023, 2024a), this paper aims to present an ontology to support the documentation of what we term *observational data*, i.e. data pertaining to *observations* (*statements*, *claims*) made by observers about subjects of their interest. While we focus on claims related to the interpretation of the contents of (literary) texts, the proposal can be applied more broadly. This is also because we do not draw a cut-off distinction between literary and critical texts. If literature is a "strange institution which allows one to say everything" (Derrida and Attridge, 1991, p.36), such a distinction is not only theoretically problematic, but it has also been challenged by artistic practice itself (consider, for example, Nabokov's *Pale Fire*, where criticism, poetry, and fiction intertwine and blur into one another). Before we dig into the proposal, for the sake of clarity, we will outline some of the assumptions that underpin our approach.

First, texts play a prominent role in the scenarios we have in mind. For example, in the case of literary criticism, one might be interested in documenting and analyzing what scholarly texts claim about Dante's poems; in fine art, in texts on Rodin's sculptures; in musicology, in texts on Mozart's compositions, to make just a few examples. The focus on texts has at least two motivations: $(i)$ we remain close to interpretive practices in which scholars continuously interpret texts and produce their own to share their interpretations with others, and $(ii)$ we depart from mentalistic approaches that focus on what experts have in mind. An example of this latter approach is discussed by Doerr et al. (2011). Instead, we consider what experts make public via texts that can be intersubjectively accessed.

Second, we assume that a literary text does not have a single prescribed content but that it can be interpreted in various and possibly incompatible manners. This does not mean that interpretations are arbitrary: the text itself provides the evidential basis and constraints against which interpretations must be justified. Regarding the debate in analytic philosophy on interpretation (Krausz, 2010), our position is compatible with both *pluralistic* and *monistic* perspectives. The main difference between the two is that only the former recognizes the possibility of interpreting a text in various *valid* ways. As we will see, our theory allows for documenting various interpretations of a text. If there are criteria to judge the *validity* of interpretations, they can be taken into account to determine *one* valid interpretation or to acknowledge *multiple* equally valid interpretations. Similarly, considering the dispute on the role of authors' intentionality when interpreting texts (Krausz, 2010), our theory takes a more abstract position. The interpretation of a text provided by its author can be taken into account to *support* or even *determine* the interpretation of the text, if available. From a methodological perspective, the ontology presented in the paper is a general theory for representing interpretations. It can be therefore tuned and refined to align with specific philosophical views, if desired.

Third, further developing the previous point, we are interested in what we call *real-life* interpreters – following the terminology of Willis (2017). Each interpreter so understood interprets a text in their own way, according to their literary competences, background knowledge, cognitive abilities, and other dimensions. In addition, we assume that interpreters articulate their interpretations only to a certain degree of completeness. In the words of Horstmann et al. (2020): "[...] the interpretation process itself [in literary contexts] is in principle incomplete: Literary Studies are a discursive disciplines."

Fourth, our proposal remains neutral with respect to the ontological status of the things featured in literary texts, such as literary characters or, more generally, fictional entities (*ficta*). As is well known, there is an ongoing debate about their nature (Kroon and Voltolini, 2023) and our approach is compatible with both realist and anti-realist perspectives. Specifically, we only consider the statements and potential agreements of interpreters without detailing ontological considerations about ficta. This approach seems legitimate in the context of literary interpretation, where experts rarely consider ontological issues when discussing the existence of literary characters. Some philosophers converge with the idea of literary

theorists that in narrating a story there is something akin to mathematics; in a way, they both are *stipulative* and *exploratory*: "let us see where and how far a given assumption or basic situation can lead us" (Margolin, 2012). According to Galván (2017), this can be done without the need to resort to the concept of "fictional truth." In any case, it is not, we insist, a pressing problem at the level of interpretation that of *referring to* some sort of fictional reality.

With these considerations in mind, we can further clarify the concept of interpretation as it will be used throughout the paper. We adopt the attitude of Gius and Jacke (2017), who understand interpretation as "the formulation of hypotheses about aspects of meaning in literary texts. These hypotheses regarding meaning are generated by *reasoning processes* that apply *inference rules* based on *consistency*, *coherence*, and other aspects in order to give the best explanation for the (portion of) text in question" (emphasis is ours). In line with this, as we will detail in the next sections, we assume that interpreting a text involves formulating hypotheses about the text based on inferences that interpreters make through their beliefs.

One can compare the activity of interpretation that scholars pursuit to that of a *game* in the sense of Wittgenstein's *Philosophical Investigations* (Wittgenstein, 2009). In this game, interpreters are the players, the statements they produce in their interpretations are their moves (the actions taken by the players), and the texts they adhere to in their interpretations are the rules, i.e., the constraints to which they have to adhere for the interpretation of a text to be an interpretation of *that* text. As Wittgenstein observed, a rule cannot uniquely determine a specific course of action. Similarly, in line with the communitarian solution of Kripke (1982), understanding how a rule is followed emerges through observing its practical applications within a community. In the same way, the scholarly humanistic endeavour sees that there is a better understanding of a text through the interplay of multiple interpretations, each grounded in careful methodological analysis. Our formal theory supports this approach by enabling comparison of different interpretative perspectives. A crucial step in this direction is to view interpretation as involving a form of *justification* toward what is asserted. This commitment is expressed through a scientific or analytic approach to the text, namely, one that seeks to provide evidence based on methods and instruments, much like measurement apparatuses. If communication requires agreement in both definitions and judgments (Wittgenstein, 2009, sect. 242), a deeper understanding of a text requires both shared methods (as we will see, committing to the same text and *observational vocabulary*) and consistent results (sharing judgments about the text through that vocabulary). This dual requirement grounds both the empirical nature of interpretative practice and the possibility of genuine communication among interpreters, while fostering productive scholarly dialogue through well-supported differences in perspective.

The remaining of the paper is structured as follows. Section 2 presents an illustrative example which we use throughout the paper to exemplify the ontology. Section 3 introduces the grounding notions of our ontology of observations. Section 4 deepens the basics of the ontology as follows: it introduces the notions of text and report (Section 4.1), two types of (complex) observations to model observations' sources (Sections 4.2–4.3), some criteria to compare reports and texts (Section 4.4), as well as relations between reports (Section 4.5). Section 5 adds two further kinds of observations to model the arguments put forward by scholars in their discourses to defend or attack a certain claim. In Section 6 we compare our approach with the state of the art. Finally, Section 7 concludes the paper.

## 2. A guiding example

We introduce here the key aspects of an example that we will use throughout the paper to show how the proposed theory can be used to represent interpretations in literary contexts. The examples do not provide an exhaustive representation of a real case study, which remains to be addressed as part of future work.

We are interested in documenting the types of observations one can make about a text, whether it is a literary text or a scholarly essay about a literary text. From an application perspective, once these observations are represented, one can consider data management solutions to make them available for further use, as well as computational methods to analyze them, e.g., similarly to what done in the context of Computational Literary Studies (Hatzel et al., 2023; Schöch et al., 2022).

For the sake of presentation, we focus on observations related to a literary text as well as criticism relative to the text. Specifically, we consider Boccaccio's *Decameron* (14th century) and, in particular, the novella X.10, recounting the story of the characters of Griselda and Gualtieri. The tale follows Gualtieri, a feudal lord, who marries Griselda, a poor peasant, and subjects her to a series of increasingly cruel trials to test her love and fidelity. These trials include making her believe that their children must be killed. Despite the ordeals, Griselda endures everything without complaint. In the end, Gualtieri reveals the truth and they supposedly live happily ever after. Due to its provocative and controversial plot, the tale continues to spark critical debate. For the example, we will consider an application context where interpreters wish to express observations relative to some of the (relational) features characterizing Gualtieri and Griselda such as their moral or social conditions, or their parental bond.

Concerning critical texts, we will consider some aspects of the interpretations provided by Branca (1956) and Picone (2008). As part of their work, both scholars explored connections between the *Decameron* and other literary texts, following the methodology, sometimes called *hypertextuality*, of studying how texts are interrelated. We do not mean here to fully report on Branca's and Picone's critical works but to show how some of their observations can be represented in our theory. In particular, we will focus on some aspects relative to Branca's and Picone's claims on the *similarities* between literary characters taken from *Decameron*'s tale X.10 and those of the other texts that they analyzed.

In particular, Branca's analysis relates Boccaccio's work to medieval culture. When commenting on tale X.10, Branca argues that the character of Griselda is inspired by the Virgin Mary as described in hagiographic texts. Branca supports this claim by attributing similar traits to both characters. On the other hand, Picone emphasizes the relationship between tale X.10 and feudal society, tracing its origins to chivalric and courtly literature. Along these lines, Picone identifies the *Lai de Fresne* (second half of the 12th century) as the main source of the tale, supporting his argument by referencing several passages from Boccaccio's text that closely echo the *Lai de Fresne*. The comparison between Branca and Picone is particularly interesting because Picone explicitly rejects Branca's thesis about the relationship between the tale and biblical and hagiographic texts. This includes rejecting Branca's claim that Griselda is similar to the Virgin Mary.

## 3. Theory of observations: The basics

In this work, *observations* are understood as the classification of domain entities through empirical measurements (e.g., carbon-14 measurement), cognitive tasks, computational analyses, or scholarly argumentation (e.g., a scholar attributing a moral value to a literary character), among other cases.

Observations abstract from both the processes that generate them and the observers thereby involved. However observations $(i)$ always have at least a *source observer*[2] and $(ii)$ are expressed using a controlled vocabulary, called *observational vocabulary*, agreed upon by all observers. The idea is that domain experts use shared observational vocabularies to express their observations in a controlled manner, making them

[2]Several notions are required to make precise the sense in which observations have a source, see (a19) in Sect. 4.3.

uniformly comparable.[3] Observations can be understood as *declarative sentences* (aka *statements*) to which at least an observer *commits*. The set of observations is then a subset of the set of sentences of the observational language; it is the set of *entertained* (declarative) sentences resulting, as said, from different sorts of processes (measurement, perception, thought, etc.), regardless of whether they are true or false.[4] For example, suppose that the observational vocabulary contains the concepts *being just* and *being unjust* as well as the individual *Gualtieri*. If no literary scholar commits to *Gualtieri being just*, the corresponding observation does not exist. Conversely, if several scholars commit to *Gualtieri being unjust*, even in different ways (e.g., by asserting, denying or supposing), then there exists a single observation entertained by all of them. Hypothetical or contested information, as well as information produced by (physical) sensors or other sources, including (speculative) scholarly research are all accepted. Furthermore, even observations directly asserted by observers are not necessarily veridical. With respect to our field of application here, some scholars even question whether concepts such as truth or falsehood apply (Feagin, 1982; Hempfer, 2024, chp. 1). This is because there is not a backbone reality against which, e.g., a literary interpretation can be mapped to tell its truth or falsity. Following this line of thought, an interpretation can be more or less justified, e.g., because it complies with criteria which scholars accept as valid. In the words of (Hempfer, 2024, pp. 22-23): "[...] there is no ultimately valid, 'true' interpretation, because both the data and the inferential processes can be challenged – what does exist, however, are more persuasive and less persuasive interpretations."

In addition to tracking observations – i.e., representing the observers who commit to a particular classification of entities in terms of the observational vocabulary – two requirements are fundamental to our scenario: $(i)$ to represent *chains* of observations – i.e., observations about observations – which are especially common in the humanities (e.g., Aristotle said that Plato said that Socrate said...); $(ii)$ to collect conflicting observations – since different observers may disagree on the classification of certain entities – while maintaining logical consistency.

From a technical perspective, we adopt a classical first-order language. This is motivated by, first, the fact that first-order logic is widely used in knowledge representation; second, by the intention of approximating the theory in the Web Ontology Language (OWL)[5] to implement a system capable of automatically reasoning about the collected observations (see the work of Sanfilippo et al. (2024b) for an example). Given the previous requirements and this constraint on the adopted logic, we reify observations into the domain of quantification, in line with state-of-the-art works such as Masolo et al. (2018), Janowicz et al. (2019), and W3C recommendations.[6] For example, rather than representing that the literary character of Gualtieri is a feudal lord using the predication `FeudalLord(gua)`, we reify this sentence in the domain of quantification, representing it as an individual constant. We then link this constant to Gualtieri using an appropriate binary relation (we say that Gualtieri is an "argument" of the observation) and we state its kind (see Sect. 3.1 for the formal details).

Observational vocabularies are primarily sets of *unary* predicates representing different kinds of observations. These predicates are organized into a taxonomy according to their level of generality.[7] Fig. 1 shows the taxonomy of the main kinds of observations that we will use to formalize the guiding example.

[3] Observational vocabularies can be understood as specific types of controlled vocabularies like those used for text annotation in NLP tasks. See Sect. 6 for more discussion on this topic.

[4] For more on the theoretical foundations of our work, the reader can refer to Masolo et al. (2018).

[5] https://www.w3.org/TR/owl2-primer/.

[6] https://www.w3.org/TR/swbp-n-aryRelations/.

[7] As it can be seen from Fig. 1, observational vocabularies can also model *relational* observational kinds, like *Similar*, involving more than one entity.

In Sect. 3.1 we will show how the number and type of the arguments of the instances of the observation kinds can be constrained and how incoherence and correlations constraints can be introduced between them. It is important to emphasize that it is up to the users of our theory to decide which observational vocabulary to adopt for a given domain of interest, and which declarative sentences to model using observations instead of standard predication. However, our theory commits to some kinds (not included in Fig. 1) that are part of all observational vocabularies: $(i)$ Sect. 4.2 and Sect. 4.3 introduce the observation kinds *assertion* and *interpretative assertion* which are used to represent the source of an observation, i.e., who is committing to it (dually for *denial* and *interpretative denial*); $(ii)$ Sect. 5 introduces the observation kinds *support* and *defeat* which are used to represent the arguments that observers consider in favor of or against given observations.

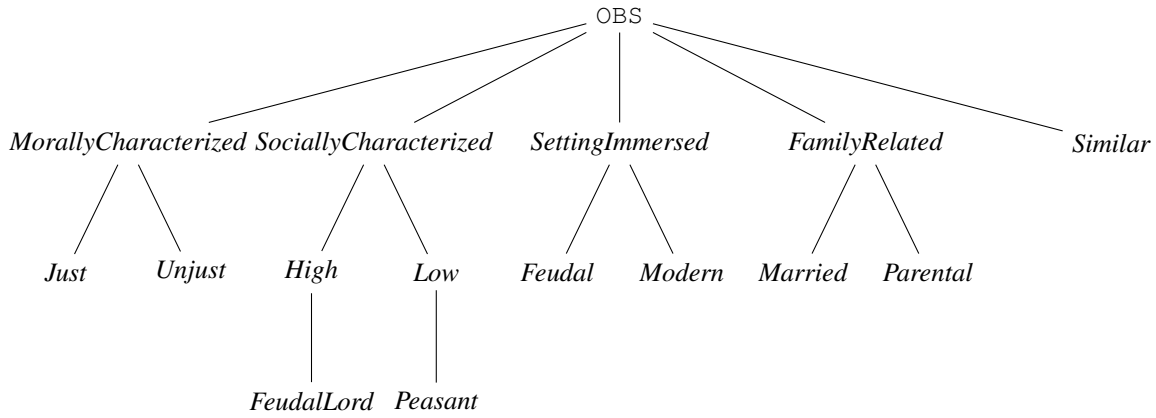


Fig. 1. Exemplification of an observational vocabulary

### *3.1. Specification of Observational Vocabularies*

Observational vocabularies must be designed through collaboration with domain experts to agree on the meaning of the adopted concepts. As with any formal language, their meaning is formally characterized by axioms constraining the primitives of the observational vocabulary at stake. The decision to reify observations in the domain of quantification requires a specific methodology, which we will explain throughout the paper.

The formal characterization of the primitives of an observational vocabulary can be divided into three main steps:

**Step 1** Identify and taxonomically structure a *finite* set $\mathcal{P}$ of (first-order) *unary* predicates, representing specific kinds of observations according to the level of generality of such kinds;

**Step 2** For each observation-kind $P \in \mathcal{P}$, determine the number and type of the entities involved in $P$'s instances, i.e., the arguments of $P$;

**Step 3** Introduce general incoherence relations, as well as correlations between the instances of different kinds in $\mathcal{P}$.[8]

[8]Strictly speaking, the observational vocabulary also contains individuals constants, e.g., the constant representing Gualtieri. The three steps discussed here concern the intensional characterization of the primitives that can be used for specific observations.

We will now explain these three steps and illustrate them using the guiding example previously introduced.

### 3.2. Step 1: The Taxonomy of Observation Kinds

The starting point is the choice of the set $\mathcal{P}$ of observation kinds and their taxonomic structure. To avoid confusion, we note the predicates in $\mathcal{P}$ in *Italics*, while the other primitives of the theory are noted in `Typewriter`. Observation kinds collect observations corresponding to the classification of an entity (or several entities) under a (the *same*) property (or relation). We call *basic observations* those that do not concern other observations.

The set $\mathcal{P}$ is mainly chosen by users according to their needs and domain of interest. However, in Sect. 4 and Sect. 5 we will introduce some kinds of observations that are fundamental for the representation of interpretations of texts and their argumentation. These kinds – which are not depicted in Fig. 1 for simplification purposes – are always supposed to be in $\mathcal{P}$. In particular, we will see that the so-called *illocutionary observations* (and their interpretative counterparts) guarantee that observations have always sources that determine their provenance (see (a19) in the below sections).

The general category *observation* (`OBS`) collects all the observations and it is covered by the kinds in $\mathcal{P}$, i.e., every observation in `OBS` must be an instance of at least one of the kinds in $\mathcal{P}$. The taxonomic structure of $\mathcal{P}$ is formalized via axioms with the form $P(x) \rightarrow Q(x)$ (with $P, Q \in \mathcal{P}$). To minimize these axioms one can focus only on *direct* subsumptions.

**Example cont'd. (observational vocabulary taxonomy)**. Fig. 1 depicts the taxonomy of the observational-kinds used in our guiding example (excluding the general ones that we will introduce in Sect. 4 and Sect. 5). The proposed observational vocabulary is not meant to exhaust all the possible categories that can be used to characterize literary characters; for a real case of interpretation, the vocabulary should be extended with other observation-kinds reflecting specific theories in literary studies, as well as specific desiderata of interpreters. The taxonomy in Fig. 1 has five main branches for observations of moral conditions (*MorallyCharacterized*), social conditions (*SociallyCharacterized*), settings (*SettingImmersed*), family relations (*FamilyRelated*), similarities (*Similar*). *MorallyCharacterized* refers to moral traits like being just (*Just*) or unjust (*Unjust*). *SociallyCharacterized* refers to social statuses like high- and low- social conditions (*High*, *Low*); being a feudal lord (*FeudalLord*) is an example of high social condition; being peasant (*Peasant*) is an example of low social condition. An observation of setting provides information about the context in which the character of a text is immersed (and such context is inherited by the text itself). The vocabulary distinguishes between feudal (*Feudal*) and modern (*Modern*) settings. Observations of family bonds represent family relationships like being married (*Married*) or standing in parental bonds (*Parental*). Finally, *Similar* accounts for a general notion of similarity.

We assume the disjointedness of all the observation kinds which are at the same depth in the taxonomy in Fig. 1. For example, the left branch of the taxonomy is formally captured by formulas (f1)-(f8). Similarly for the other branches. Note that $\mathit{MorallyCharacterized}(o) \rightarrow (\mathit{Just}(o) \vee \mathit{Unjust}(o))$ does not hold because it is possible to have an observation with only the level of detail of *MorallyCharacterized*, i.e., one knows that an agent has a moral connotation without knowing whether they are just or unjust.

**f1** $\mathit{Just}(o) \rightarrow \mathit{MorallyCharacterized}(o)$

**f2** $\mathit{Unjust}(o) \rightarrow \mathit{MorallyCharacterized}(o)$

**f3** $\mathit{Just}(o) \rightarrow \neg \mathit{Unjust}(o)$

**f4** $\mathit{MorallyCharacterized}(o) \rightarrow \mathtt{OBS}(o)$

**f5** $\mathit{MorallyCharacterized}(o) \rightarrow \neg \mathit{SociallyCharacterized}(o)$

**f6** $\mathit{MorallyCharacterized}(o) \rightarrow \neg \mathit{SettingImmersed}(o)$

**f7** $\mathit{MorallyCharacterized}(o) \rightarrow \neg \mathit{FamilyRelated}(o)$

**f8** $\mathit{MorallyCharacterized}(o) \rightarrow \neg \mathit{Similar}(o)$ □

### *3.3. Step 2: Arities and Argument Restrictions for Observation Kinds*

Once $\mathcal{P}$ and the taxonomy are introduced, one can constrain the number and the types of entities involved in the observations of a given type. To recall the previous example with the observation *Gualtieri is unjust*, this is an observation of kind *Unjust* involving one single entity: the fictional character Gualtieri. We assume that all instances of an observation kind $P \in \mathcal{P}$ have the same number of entities (abusing the terminology, we say the *arity* of $P$) which are subject to similar restrictions (abusing the terminology again, we say the *argument restrictions* for $P$).

To formally capture arities and argument restrictions, we introduce $\alpha$ binary predicates $\texttt{ARG}_i$ – where $\alpha$ is the maximal arity of the kinds in $\mathcal{P}$, i.e., observations involve at most $\alpha$-entities – to identify the $i^{th}$ entity involved in an observation, i.e., the $i^{th}$ *argument*, see axioms (a1) and (a2).[9] By axiom (a2), the second argument of $\texttt{ARG}_i$ is always an observation whereas it does not constrain the range of the first argument. As we will see in the examples throughout the paper, this allows us to model chains of observations, namely, observations about observations; e.g., when a scholar criticizes an observation done by others (or the observations introduced in Sect. 4 and Sect. 5).

**a1** $\bigwedge_{i=1}^{\alpha}(\texttt{ARG}_i(x,o) \wedge \texttt{ARG}_i(y,o) \rightarrow x = y)$

**a2** $\bigwedge_{i=1}^{\alpha}(\texttt{ARG}_i(x,o) \rightarrow \texttt{OBS}(o))$

By means of the $\alpha$ $\texttt{ARG}_i$ primitives, for each kind $P \in \mathcal{P}$ it is possible to specify its arity $n$ by an axiom with the form in (a3), where $\texttt{ARG}(x_1,\ldots,x_n,o)$ is a shortcut with the form in (d1) which identifies (in the correct order) all the entities involved in an observation $o$ of a kind with arity $n$. Argument restrictions have the form in (a4) where the $\texttt{T}_i$ are specific categories in our ontology. Note that all observations under a kind that is a root in the taxonomy of observation kinds (i.e., it has no more general observation kinds, e.g., *MorallyCharacterized* in Fig. 1) have the same arity, i.e., they necessarily have the same number of arguments but its subkinds may have more restrictive argument restrictions.

**d1** $\texttt{ARG}(x_1,\ldots,x_n,o) := \bigwedge_{i=1}^{n}(\texttt{ARG}_i(x_i,o)) \wedge \bigwedge_{i=n+1}^{\alpha} \neg\exists x(\texttt{ARG}_i(x,o))$

**a3** $P(o) \rightarrow \exists x_1,\ldots,x_n(\texttt{ARG}(x_1,\ldots,x_n,o))$ (where $n$ is the arity of $P$)

**a4** $P(o) \wedge \texttt{ARG}(x_1,\ldots,x_n,o) \rightarrow \texttt{T}_1(x) \wedge \cdots \wedge \texttt{T}_n(x)$ (where $n$ is the arity of $P$ and $\texttt{T}_i$ is a category)

**Example cnt. (arities and argument restrictions)**. We assume two additional categories in our ontology, both disjoint from observations: TXT for texts and AGN for agents.[10] TXT and AGN are not observation kinds. This means that the classification under these categories is not a matter of debate among domain

[9] Note that $\texttt{ARG}_i$-relations as well as the $\mathcal{P}$-predicates are not temporally qualified, i.e., observations cannot vary their arguments or migrate from a kind to the other. The theory can be extended to take time into account.

[10] See Sect. 4.1 for more discussion on texts. Also, we are not interested in characterizing the notion of agent, which is introduced here only for the sake of the example.

experts. On the other hand, if this were the case, one could introduce text and agent as observation kinds, adding correlations between them (see Sect. 3.4), instead of introducing axioms of the form in (a4).

Formulas (f9)-(f13) determine arities and argument restrictions for all the root observations-kinds in Fig. 1. We do not consider here any refinement of the argument restrictions for the subkinds of the root kinds.

**f9** $\mathit{MorallyCharacterized}(o) \rightarrow \exists x(\mathtt{ARG}(x,o)) \wedge \forall x(\mathtt{ARG}(x,o) \rightarrow \mathtt{AGN}(x))$

**f10** $\mathit{SociallyCharacterized}(o) \rightarrow \exists x(\mathtt{ARG}(x,o)) \wedge \forall x(\mathtt{ARG}(x,o) \rightarrow \mathtt{AGN}(x))$

**f11** $\mathit{SettingImmersed}(o) \rightarrow \exists xy(\mathtt{ARG}(x,y,o)) \wedge \forall xy(\mathtt{ARG}(x,y,o) \rightarrow \mathtt{TXT}(y))$

**f12** $\mathit{FamilyRelated}(o) \rightarrow \exists xy(\mathtt{ARG}(x,y,o)) \wedge \forall xy(\mathtt{ARG}(x,y,o) \rightarrow \mathtt{AGN}(x) \wedge \mathtt{AGN}(y))$

**f13** $\mathit{Similar}(o) \rightarrow \exists xy(\mathtt{ARG}(x,y,o))$ □

As said, we can have observations at different levels of generality. For instance, *Unjust*-observations are more specific than *MorallyCharactized*-observations. The *minimal kind* of an observation $o$ is the kind $P \in \mathcal{P}$ such that: $(i)$ $o$ is an instance of $P$, and $(ii)$ $o$ is an instance of no subkinds of $P$. We assume that the minimal kind of an observation is unique and that two observations with the same minimal kind and the same arguments are identical, see the work of Masolo et al. (2018) for the formal details. However, it is possible, for instance, to have an observation $o$ with minimal kind *Unjust* and a different observation $\bar{o}$ with minimal kind *MorallyCharacterized* (i.e., $\mathit{MorallyCharacterized}(\bar{o}) \wedge \neg \mathit{Just}(\bar{o}) \wedge \neg \mathit{Unjust}(\bar{o})$) both with Gualtieri as single argument. In this case the information that $o$ and $\bar{o}$ convey is different (given the different level of generality of *Unjust* and *MorallyCharacterized*).

To simplify the notation, we will write $\mathbf{p}(\mathtt{a}_1, \ldots, \mathtt{a}_n)$ to denote the unique observation with minimal kind $P$ (of arity $n$) and with the arguments $\mathtt{a}_1, \ldots, \mathtt{a}_n$, where the $\mathtt{a}_i$s are individual constants. For example, **morallyCharacterizated**(gua) is a conventional name of an individual constant such that:

$$\mathit{MorallyCharacterized}(\mathbf{morallyCharacterized}(\mathtt{gua})) \wedge \mathtt{ARG}(\mathtt{gua}, \mathbf{morallyCharacterized}(\mathtt{gua})) \wedge \neg \mathit{Just}(\mathbf{moralCharacterization}(\mathtt{gua})) \wedge \neg \mathit{Unjust}(\mathbf{moralCharacterization}(\mathtt{gua})).$$

We also use the notation $\mathbf{p}(x_1, \ldots, x_n)$, where the $x_i$s are variables, to indicate a general observation with minimal kind $P$ and generic arguments $x_1, \ldots, x_n$.

Finally, an observation $o_1$ (improperly) *refines* (with respect to the taxonomy of the observation-kinds in $\mathcal{P}$) an observation $o_2$ (written $\mathtt{REF}(o_1, o_2)$) when either $o_1$ and $o_2$ are identical, or they have the same arguments and the minimal-kind of $o_1$ is subsumed by the minimal-kind of $o_2$. For instance, given the taxonomy in Fig. 1, one could write $\mathtt{REF}(\mathbf{unjust}(\mathtt{gua}), \mathbf{morallyCharacterized}(\mathtt{gua}))$ between the two observations **unjust**(gua) and **morallyCharacterized**(gua), meaning that the former provides a more specific level of detail than the latter. Given the transitivity of subsumption, REF is transitive, too.

### *3.4. Step 3: Incoherences and Correlations between Observations*

The technical move of introducing observations into the domain of quantification allows us to maintain the consistency of our first-order theory even in the presence of contradictory observations. However, this also leads to the impossibility of using the logical connectives (in particular implication and negation) to introduce complex constraints between observations of different kinds. For example, in a classical setting where *being just* and *being unjust* are represented by predicates that apply to agents, $\mathtt{Just}(x) \rightarrow \neg \mathtt{Unjust}(x)$ expresses that an agent considered as just cannot be also unjust (and vice versa), i.e., it is inconsistent to say that an agent is both just and unjust. In our framework, *Just* and *Unjust* are disjoint

kinds of observations – i.e., $\mathit{Just}(o) \rightarrow \neg \mathit{Unjust}(o)$, see (f3) – but it is possible to have **just**(gua) and **unjust**(gua) (two different observations) without running into logical contradiction. This is an essential feature for collecting and comparing contrasting interpretations on a single subject. However, we also want to express that these observations are a sign of disagreement among interpreters or incoherence on the part of someone who claims both. Analogously, in a classical first-order framework it is easy to represent that feudal lords are always immersed in a feudal setting, whereas in our framework we need to introduce a mechanism to link observations of the kinds *FeudalLord* and *Feudal*.

To partially characterize these kinds of constraints, we introduce two new primitives: (*i*) *incoherence* – with $\texttt{INC}(o_1, o_2)$ standing for "$o_1$ and $o_2$ are *incoherent* observations" – and *correlation* – with $\texttt{COR}(o_1, o_2)$ standing for "the observation $o_1$ correlates with the observation $o_2$."

We assume that INC is irreflexive and symmetric, while COR is irreflexive (i.e., it is a proper and direct correlation) and transitive. The general idea is that INC represents a form of inconsistency while REF (as introduced at the end of the previous section) and COR represent different forms of inference. Specifically, REF captures a specialization (generalization) relation on the adopted observational vocabulary, while COR connects observations whose kinds belong to different branches in the taxonomy (but see the discussion in Sect. 5 for some counterexamples). We do not commit to any particular notion of inconsistency and inference, hence either a classical or a probabilistic view can be supported. We indicate with IMP (standing for *implies*) the disjunction between REF and COR, see (d2). Finally, (a5) ensures that implication is a form of coherence; (a6) ensures that to get incoherent observations we have to start from incoherent ones; and (a7) states that if an observation correlates with a second observation, then any refinement of the first observation correlates with a generalization of the second observation.

**d2** $\texttt{IMP}(o_1, o_2) :\equiv \texttt{REF}(o_1, o_2) \vee \texttt{COR}(o_1, o_2)$

**a5** $\texttt{IMP}(o_1, o_2) \rightarrow \neg\texttt{INC}(o_1, o_2)$

**a6** $\texttt{IMP}(o_1, o_2) \wedge \texttt{IMP}(o_3, o_4) \wedge \texttt{INC}(o_2, o_4) \rightarrow \texttt{INC}(o_1, o_3)$

**a7** $\texttt{REF}(o_1, o_2) \wedge \texttt{REF}(o_3, o_4) \wedge \texttt{COR}(o_2, o_3) \rightarrow \texttt{COR}(o_1, o_4)$

The idea is to use these primitives to partially capture the previous constraints. Although INC and COR are defined on specific observations, they are often used to capture some "laws" that reflect how the application domain is organized, i.e., incoherence or correlation hold for all the instances of given observation kinds. However, note that neither IMP nor COR allows one to infer new observations (i.e., extend the ABox) from the given ones. These primitives simply establish some dependencies between existing observations (observations already present in the ABox).

**Example cnt. (incoherences and correlations)**. The observational vocabulary can be further characterized by introducing incoherence (see (f14)-(f17)) and correlation (see (f18)) laws applying to all the instances of the represented observation-kinds. First, note that, given the symmetry of INC, from (f14) it also follows that $\texttt{INC}(o_2, o_1)$; similarly for formulas (f15)-(f17). Second, given that *FeudalLord*-observations refine *High*-observations and that *Peasant*-observations refine *Low*-observations, by (a6) and (f15), it follows that $\mathit{FeudalLord}(o_1) \wedge \mathit{Peasant}(o_2) \wedge \texttt{ARG}(x, o_1) \wedge \texttt{ARG}(x, o_2) \rightarrow \texttt{INC}(o_1, o_2)$. Similarly one can infer incoherences between *Peasant*- and *Low*-observations as well as *Peasant*- and *High*-observations.

**f14** $\mathit{Just}(o_1) \wedge \mathit{Unjust}(o_2) \wedge \texttt{ARG}(x, o_1) \wedge \texttt{ARG}(x, o_2) \rightarrow \texttt{INC}(o_1, o_2)$

**f15** $\mathit{High}(o_1) \wedge \mathit{Low}(o_2) \wedge \texttt{ARG}(x, o_1) \wedge \texttt{ARG}(x, o_2) \rightarrow \texttt{INC}(o_1, o_2)$

**f16** $\mathit{Feudal}(o_1) \wedge \mathit{Modern}(o_2) \wedge \texttt{ARG}(x, y, o_1) \wedge \texttt{ARG}(x, y, o_2) \rightarrow \texttt{INC}(o_1, o_2)$

**f17** $\mathit{Married}(o_1) \wedge \mathit{Parental}(o_2) \wedge \texttt{ARG}(x, y, o_1) \wedge \texttt{ARG}(x, y, o_2) \rightarrow \texttt{INC}(o_1, o_2)$[11]

**f18** $\mathit{FeudalLord}(o_1) \wedge \mathit{Feudal}(o_2) \wedge \texttt{ARG}(x, o_1) \wedge \texttt{ARG}(x, y, o_2) \rightarrow \texttt{COR}(o_1, o_2)$[12] □

## 4. Texts and Illocutionary Acts

We introduce in the following subsections texts, reports, and some observation-kinds to document texts' contents. In particular, texts and reports are introduced in Sect. 4.1. Building on the theory of speech acts in philosophy, Sect. 4.2 introduces what we call *illocutionary observations*, and Sect. 4.3 presents illocutionary observations bearing an interpretative dimension. Finally, Sect. 4.5 presents formal means to model relations of accessibility between reports.

### *4.1. Texts and Reports*

We intend *texts* (TXT) as sequences of words in one or more languages as made public by a certain (group of) author(s) at a time. Hence, a text cannot be made public by different authors at different times; also, texts differing even for a single word are not identical, although they can be similarly interpreted and can stand in various relations. A text exists in time but it is not, e.g., the individual pattern of ink on a paper; in this sense it is a text *type*.[13]

A *report* (REP) is a text[14] expressed in an observational vocabulary, i.e. it is an *observational text*. One can think of reports as representations of complex and coherent *speech acts* through which an agent makes explicit and public their observations through an observational vocabulary. Recall that speech acts are intended in philosophy as *acts* through which agents "do things with words", to recall Austin (1962). Among others, *illocutionary acts* are speech acts characterized by their specific force. For instance, as Austin notes, when I *promise* or *state* that I will be home for dinner, the content is the same but the illocutionary force differs – one has the force of a promise, the other of a statement. For our purposes, we focus on *assertions* and *denials* as basic types of illocutionary acts suitable for representing certain types of interpretative claims about texts.

Reports and texts play an important role in our application context. As said in the Introduction, the aim of interpreting a text is to make hypotheses about aspects of its content according to a (group of) interpreter(s). The observations in a report make explicit such hypotheses.

Before moving on to the next sections, we comment on some assumptions about agents and their capabilities, especially when it comes to their relations with texts.

First, agents have beliefs and reasoning capabilities. By assuming that $\beta(a)$ identifies the *belief-base* of agent $a$,[15] the inference mechanism allows $a$ to make explicit beliefs that remain implicit in $\beta(a)$. All the agents share the same observational vocabulary and constraints characterizing its primitives, i.e., they

[11] The symmetry of *being married* can be characterized in terms of COR, e.g., $\mathit{Married}(o_1) \wedge \mathit{Married}(o_2) \wedge \texttt{ARG}(x, y, o_1) \wedge \texttt{ARG}(y, x, o_2) \rightarrow \texttt{COR}(o_1, o_2) \wedge \texttt{COR}(o_2, o_1)$. For the sake of simplicity, we refrain from introducing this constraint.

[12] We refrain from introducing the existential constraint $\mathit{FeudalLord}(o_1) \wedge \texttt{ARG}(x, o_1) \rightarrow \exists y o_2(\mathit{Feudal}(o_2) \wedge \texttt{ARG}(x, y, o_2))$, because, given our documentary perspective, we only consider observations that authors explicitly make.

[13] We do not enter into the debate whether text types can exist in a pure abstract sense (Sanfilippo, 2021). The intuition we want to grasp is that a text type can have multiple materializations.

[14] Formally, we have $\texttt{REP}(x) \rightarrow \texttt{TXT}(x)$.

[15] A belief-base can be understood as a knowledge-base in the standard sense used in knowledge representation. We wish to stress here that it collects the set of *beliefs* of an agent rather than its *knowledge*. The distinction between belief and knowledge is widely discussed; see, e.g., the work of Schwitzgebel (2024).

share the intersubjectively developed theory, which is noted $\mathcal{T}_{\text{OBS}}$, through the aforementioned three steps. Agents could express their belief-bases in different languages. Here we assume that these languages can all express $\mathcal{T}_{\text{OBS}}$. Furthermore, agents can detect inconsistencies between sources of information, and they can resolve these by discarding some of the information provided by the sources. For our purposes it is not necessary to commit to a specific notion of inference or consistency; for instance, both classical deduction and probabilistic inference may be considered.

Second, agents are supposed to be competent speakers of the languages of the texts that they interpret. In particular, they are able to recognize and solve possible grammatical ambiguities and surface inconsistencies in the texts, as well as to (partially) map natural language terms into the language used in their belief-bases. The function $\tau(t,a)$ identifies the resolution of ambiguities and surface inconsistencies in text $t$ performed by agent $a$. Translating a text into the language of $\beta(a)$ may require an iterative process that considers lexical and common-sense knowledge contained in $\beta(a)$, as well as resolving possible inconsistencies between the text and $\beta(a)$. These aspects will be discussed in more detail in Sect. 4.3.

Third, the authors of reports – $a(r)$ identifies the author of report $r$[16] – are sincere, rational, and coherent, i.e., they have reasons for their observations, and a single report does not contain incoherent observations. Since the beliefs of agents, as well as how they resolve ambiguities or make translations, may change over time, the same author could produce different reports that are incoherent with each other; hence, the consistency of authors through their reports is not guaranteed.

It is important to note that the functions $a$, $\beta$, and $\tau$ are not part of our formal theory: they are used to clarify the intended meaning of illocutionary acts and their coherence (see Sect. 4.2 and Sect. 4.3).

### 4.2. Illocutionary Observations

We now present complex observations aiming at capturing claims made public via texts or reports. The idea is to introduce several types of observations to represent different kinds of illocutionary acts.

Assertive illocutionary acts seem to have a strong commitment toward the truth of what is asserted. However, several other illocutionary acts, while sharing features with assertions, lack their strong epistemic commitment or have no commitment at all.[17] These can be arranged on a scale of commitment. Conjectures, although weaker than assertions, still aim at truth while acknowledging possible counterevidence (Williamson, 2002). At the lower end, guesses require no belief in what is said (Bach and Harnish, 1979; Searle, 1985). Literary interpretation presents even more complex cases. Kafka's or Borges's paradoxes resemble liar's sentences, appearing to commit simultaneously to a proposition and its negation (Trafford, 2014). Still more challenging is poetry, where the Austinian view that poetic language is non-serious, and thus involves no genuine speech acts (Austin, 1962), remains controversial.

In what follows we focus on *assertions* (*ASR*) and *denials* (*DNY*) which, in contexts related to literary studies and criticism in a more general sense, are relevant to document scholars' views on texts. The arity and argument restrictions for *ASR* and *DNY* are specified in (a8) and (a9). Intuitively, an *ASR*- (*DNY*-) observation states that a text *asserts* (*denies*) an observation. For instance, the observation $\mathbf{asr}(\texttt{tx10}, \mathbf{married}(\texttt{gua}, \texttt{gri}))$ says that the tale X.10 of the *Decameron* (`tx10`) asserts that Gualtieri (`gua`) and Griselda (`gri`) are married.

**a8** $ASR(o) \rightarrow \exists xy(\texttt{ARG}(x,y,o)) \wedge \forall xy(\texttt{ARG}(x,y,o) \rightarrow \texttt{TXT}(x) \wedge \texttt{OBS}(y))$

[16] While the author of a report could be a group, we leave this aspect aside in this paper.

[17] Benton and Van Elswyk (2020) provide a comprehensive list of what are known in the literature as *constative* speech acts, including: insist, conjecture, assure, state, swear, guess, claim, testify, argue, admit, conclude, remind, predict, confess, report, and hypothesize.

**a9** $DNY(o) \rightarrow \exists xy(\texttt{ARG}(x,y,o)) \wedge \forall xy(\texttt{ARG}(x,y,o) \rightarrow \texttt{TXT}(x) \wedge \texttt{OBS}(y))$

Given that assertions and denials are themselves observations, one can recursively construct further assertions or denials of observations. For instance, $\mathbf{asr}(\texttt{r}, \mathbf{asr}(\texttt{tx10}, \mathbf{married}(\texttt{gua}, \texttt{gri})))$ represents that the report `r` asserts that the text `tx10` asserts that **married** applies to `gua` and `gri`. As said in Sect. 3, *chains* of assertions are particularly important in the humanities, where it is common for scholars to make arguments on what others have said.

To clarify the notion of assertion, we consider a slight variation of the position put forward by Marsili (2024). In particular, according to Marsili, an illocutionary act of *asserting* an observation $(i)$ presupposes a commitment of the asserting agent towards the truth of the observation, i.e., the observation is presented as true; $(ii)$ it *explicitly* expresses the observation, as opposed to *inferring* it; and $(iii)$ it comes with normative consequences, i.e., the assertor can be publicly criticized and challenged by others, possibly loosing their authority or reputation. Dually for denials.

In our approach, as we will see when we discuss the intended meaning of assertions and denials, one has to think in terms of *rational justifications* for observations rather than *truth*. Roughly, when an observation $o$ is asserted in a report, we are not entitled to say that it is true; we just know that the author of the report has some justification to rationally support, namely infer, $o$. On the other hand, it is also possible that others have reasons to reject $o$. Furthermore, as we will see, in case of interpretative illocutionary acts, we just know that the author of observation $o$ has some justification to support $o$ *given what is written in a text*.

In our theory, while the assertions (or denials) of an observation in natural language can be differently interpreted, the assertions made in reports $r$ are unambiguous, because they are expressed in (shared and controlled) observational vocabularies. From this perspective, they contribute to make public the beliefs of $a(r)$ on a certain subject.

To make more precise in which sense a report $r$ partially makes explicit and public the beliefs of $a(r)$, we characterize the intended meaning of the assertions and denials in $r$ by relying on the belief-base $\beta(a(r))$ of the author of $r$:[18]

- if $ASR(x) \wedge \texttt{ARG}(r,o,x) \wedge \texttt{REP}(r)$ then $\beta(a(r))$ allows to prove $o$;
- if $DNY(x) \wedge \texttt{ARG}(r,o,x) \wedge \texttt{REP}(r)$ then $\beta(a(r))$ is inconsistent with $o$.

Some comments are due. First, as said, the authors of reports are assumed to share the same observational theory $\mathcal{T}_{\text{OBS}}$ (this including observation-kinds) although their individual languages could also contain additional and not necessarily shared concepts. However, this is only a simplification hypothesis, and one could consider reports based on different observational vocabularies and theories shared only by specific communities of (and not all) agents. Second, with an abuse of notation, we say that $\beta(a(r))$ allows to prove $o$, even though $o$ is an observation rather than a formula (and $r$ and $o$ are variables of our theory). To be more precise, we should introduce the translation of $o$ in the language of $\beta(a(r))$. Furthermore, as said, $\beta(a(r))$ can change over time. Without complicating the notation, $\beta(a(r))$ should be here understood as the belief-base of $a(r)$ when the report $r$ is written and made public. Third, the

[18] Such a characterization is consistent with that provided by Carrara et al. (2024) when $\beta(a(r))$ is interpreted as the *information scenario* (*epistemic state*). We will see that in the case of interpretative illocutionary observations, the information scenario is extended by taking into account also information coming from the text, see Sect. 4.3. The characterization of the illocutionary acts with weak commitment discussed at the beginning of this section seems to require many truth-values. For conjectures, however, one could try to exploit the notion of *guaranteed possibility* introduced by Dubois et al. (2000). In this perspective, if a report $r$ conjectures $o$, then the models of $o$ must be included in those of $\beta(a(r))$. The conjecture of $o$ would then be much stronger than the mere absence of denials of $o$.

proposed characterization states the conditions for the *justification* of assertions and denials. It is not expressed in terms of necessary and sufficient conditions, because reports contain only information that their authors make public. For instance, that $\beta(a(r))$ allows to prove $o$ does not imply the existence of the observation $\mathbf{asr}(r,o)$ because $a(r)$ may decide to keep this information as private. To provide a sufficient condition, it seems fundamental to refer to the explicit choices of agents, an aspect that is not considered in the current account.[19] Similarly to Carrara et al. (2024), we do not commit to a specific inference system.[20] In particular, our framework has an intuitionistic flavor, e.g., that the assertion of $o$ is not justified does not imply that the denial of $o$ is justified (and vice versa). Fourth, and this is a relevant aspect of our theory, the characterization of the intended meaning of assertions and denials presented above does not apply to *ASR*-observations where the first argument is a text $t$ but not a report. This is because, given a text $t$ in natural language, it is not possible to unambiguously say if $t$ asserts a given observation $o$, that is, whether the author of $t$ has a commitment towards $o$. In the epistemological perspective we embrace, we limit to what is written in a text, making hypotheses on their authors' commitments only by interpreting the texts. This is the reason why assertions involving texts like $\mathbf{asr}(t,o)$ cannot exist in isolation but must always be included in observations ending in reports, e.g., $\mathbf{asr}(r,\mathbf{asr}(t,o))$. In this manner, it is formally explicit who is committing to $\mathbf{asr}(t,o)$, namely $a(r)$. However, note that it is possible to have $\mathbf{asr}(r,\mathbf{asr}(t,o))$ even though $a(r)$ did not directly interpret $t$. For instance, $r$ may "import" $\mathbf{asr}(t,o)$ from a trusted source. In this perspective, one could contemplate a report $r$ asserting both $\mathbf{asr}(t,o)$ and $\mathbf{dny}(t,o)$ since $r$ could import them from different trusted sources which are in disagreement. This would make explicit the fact that $a(r)$ is aware of a disagreement about the interpretations of $t$.[21]

As said, we assume the coherence of reports. A first constraint is to exclude, in agreement with our intended meaning, the existence of reports where both $\mathtt{A}(r,o)$ and $\mathtt{D}(r,o)$ hold (see (d3) and (d4)), i.e., reports that assert and deny the same observation. Indeed, this can be generalized by considering refinements and correlations as in (a10). A second constraint takes into account the incoherence relation between observations introduced in Sect. 3.4: (a11) excludes reports asserting incoherent observations.

**d3** $\mathtt{A}(r,o) :\equiv \mathtt{REP}(r) \wedge \exists \bar{o}(\bar{o} = \mathbf{asr}(r,o))$ (*assertion of o by r*)

**d4** $\mathtt{D}(r,o) :\equiv \mathtt{REP}(r) \wedge \exists \bar{o}(\bar{o} = \mathbf{dny}(r,o))$ (*denial of o by r*)

**a10** $\mathtt{IMP}(o_1,o_2) \rightarrow \neg\exists r(\mathtt{A}(r,o_1) \wedge \mathtt{D}(r,o_2))$

**a11** $\mathtt{INC}(o_1,o_2) \rightarrow \neg\exists r(\mathtt{A}(r,o_1) \wedge \mathtt{A}(r,o_2))$

Going back to the example of disagreements about the interpretations of texts, notice that by accepting both $\mathbf{asr}(r,\mathbf{asr}(t,o))$ and $\mathbf{asr}(r,\mathbf{dny}(t,o))$, the incoherence constraint $\mathtt{INC}(\mathbf{asr}(t,o),\mathbf{dny}(t,o))$ cannot be applied, because the situation where $r$ asserts both $\mathbf{asr}(t,o)$ and $\mathbf{dny}(t,o)$ is more a sign of incoherence or ambiguity of $t$ rather than a sign of incoherence of $r$.

[19] Note that INC and COR suffer a similar issue. One could characterize their intended meaning as: (*i*) if $\mathtt{COR}(o_1,o_2)$ then $\mathcal{T}_{\mathrm{OBS}}$ extended with $o_1$ allows to prove $o_2$; and (*ii*) if $\mathtt{INC}(o_1,o_2)$ then $\mathcal{T}_{\mathrm{OBS}}$ extended with $o_1$ is inconsistent with $o_2$. However, not all inferences are correlations and not all inconsistencies are incoherencies, i.e., sufficient conditions for INC and COR seem to require additional, and non strictly logical, information.

[20] Carrara et al. (2024) considers also probabilistic inferences.

[21] Note that also the situation in which a report $r$ asserts $\mathbf{dny}(t,o)$ and denies $\mathbf{asr}(t,o)$ (or, dually, asserts $\mathbf{asr}(t,o)$ and denies $\mathbf{dny}(t,o)$) is not necessarily incoherent since, according to our characterization of assertions and denials, the fact that $\beta(a(r))$ allows to prove $\mathbf{dny}(t,o)$ does not necessarily imply that $\beta(a(r))$ is inconsistent with $\mathbf{asr}(t,o)$.

### 4.3. Interpretative Illocutionary Observations

To represent interpretations of texts, we introduce what we call *interpretative illocutionary observations*. Following what done in Sect. 4.2, assertion- and denial-interpretative observations are particularly relevant. Accordingly, we add to the theory two kinds of interpretative illocutionary observations, *iASR* and *iDNY*, whose arity and argument restrictions are specified in (a12) and (a13), respectively.[22] Intuitively, the observation $\mathbf{iasr}(r,t,o)$ ($\mathbf{idny}(r,t,o)$) represents that, according to the interpretation of $t$ by $a(r)$, $t$ asserts (denies) the observation $o$.[23] This means that interpretative illocutionary observations have an additional contextual dimension given by the text $t$.

**a12** $iASR(o) \rightarrow \exists xyz(\texttt{ARG}(x,y,z,o)) \wedge \forall xyz(\texttt{ARG}(x,y,z,o) \rightarrow \texttt{TXT}(x) \wedge \texttt{TXT}(y) \wedge x \neq y \wedge \texttt{OBS}(z))$

**a13** $iDNY(o) \rightarrow \exists xyz(\texttt{ARG}(x,y,z,o)) \wedge \forall xyz(\texttt{ARG}(x,y,z,o) \rightarrow \texttt{TXT}(x) \wedge \texttt{TXT}(y) \wedge x \neq y \wedge \texttt{OBS}(z))$

For the sake of clarity, notice that observations $\mathbf{iasr}(r,t,o)$ and $\mathbf{asr}(r,\mathbf{asr}(t,o))$ are not the same: the former presupposes an interpretative act by $a(r)$, while the latter just presupposes that $a(r)$ believes $\mathbf{asr}(t,o)$ but this belief could be the result of a complex reasoning on information that $a(r)$ has without necessarily requiring a direct interpretation of $t$. For instance, $a(r)$ could found $\mathbf{asr}(r,\mathbf{asr}(t,o))$ on $\mathbf{iasr}(r,t,o)$, i.e., as the result of their direct interpretation of $t$, but also, as already said, on information coming from scholars who are expert interpreters of text $t$.

The annotations on texts by literary scholars or critics can be intended as reports containing only interpretative illocutionary observations. For instance, as we will see in presenting our guiding example, when one wishes to document (and analyze) what a literary scholar says in their book $t$, according to our theory, one produces a report $r$ containing observations with form $\mathbf{iasr}(r,t,o)$ or $\mathbf{idny}(r,t,o)$, i.e., explicit commitments of the reports' author on what the scholarly text $t$ asserts or denies.

Following what done in Sect. 4.2, we now make explicit the intended meaning of interpretative illocutionary acts. The idea is that the contextualization of interpretative illocutionary observations in a report $r$ does not reduce, as for illocutionary observations, to what $a(r)$ believes but it also considers what the interpreted text $t$ says according to the interpreter $a(r)$. Hence, the information scenario (epistemic state) for evaluating $o$ depends on both what $a(r)$ believes and the content of text $t$.

We slightly modify what was previously done by Sanfilippo et al. (2024a), although the general idea is compatible with alternative views. According to Sanfilippo et al. (2024a), the aim of interpreting a (literary) text is to make hypotheses about its content according to a (group of) interpreter(s). A sort of stipulative (or pretence) attitude of interpreters is here assumed (Galván, 2017), wherein interpreting a text requires, above all, accepting what is written in the text even when it conflicts with prior beliefs of the interpreter. In this view, $\mathbf{iasr}(r,t,o)$ communicates the result of $a(r)$'s inference processes: reading the text $t$, interpreter $a(r)$ dynamically builds a given body of information from which, according to additional information $a(r)$ relies on, $a(r)$ infers the information provided by $o$. In other words, by accepting what is in $t$, $a(r)$ infers $o$. We will show that the approach $(i)$ is compatible with different theories of meaning, and $(ii)$ does not presuppose a specific (shared) ontology. It is therefore possible to have $\mathbf{iasr}(r,t,o)$ and $\mathbf{iasr}(r',t,o)$ even though $a(r)$ and $a(r')$ do not share any ontology, nor the approach requires $a(r)$ to access the way in which $a(r')$ grounds their illocutionary acts, and vice versa.

[22] For the sake of clarity, considering axioms (a12) and (a13), the variable $o$ is the (interpretative) assertion (denial) observation, whereas the variable $z$ is the asserted (denied) observation.

[23] We will see in which sense a report can be interpreted.

Since $t$ may be a natural language text, to follow to the kind of characterization done for the intended meaning of assertions and denials, $t$ must be translated into a formal language, as was done for the interpreters' belief-bases. For this purpose, we consider *Discourse Representation Theory* (DRT) (Kamp and Reyle, 2013) and later extensions such as the *Segmented Discourse Representation Theory* (SDRT) (Lascarides and Asher, 2007) that widen Montague grammar to apply to sequences of sentences called *discourses*. In these approaches, discourses (which are close to our texts) are dynamically translated into *Discourse Representation Structures* (DRSs) formally representing discourses. The translation from natural language to DRSs primarily relies on syntactic and grammatical bases. However, as discussed in detail in SDRT, DRSs can also incorporate lexical or common-sense knowledge, among other factors, that are assumed to be shared by all competent speakers of a language. Importantly, DRSs (and, indirectly, discourses) can be, in their turn, translated into first order formulas. When a discourse is (syntactically and grammatically) ambiguous, and lexical or common-sense knowledge are insufficient to disambiguate it, different DRSs must be considered. The translation from discourses to DRSs is therefore a one-to-many relation, since the same discourse can be translated into alternative DRSs. In our scenario, one can think that interpreters can disambiguate texts on the basis of cognitive, cultural, psychological, etc. grounds. For this reason, we consider an *interpreter-dependent* translation from texts to formulas where interpreters can be more selective than DRT: an interpreter does not necessarily solve all the ambiguities in a text, but they select a subset of all the formulas associated to the DRSs that translate the (possibly ambiguous) text.

As said, we indicate with $\tau(a,t)$ the set of "surface formulas" (i.e., formulas that use predicates that directly correspond to the terms of the language) representing the text $t$, once the interpreter $a$ has solved possible grammatical ambiguities[24], i.e., the subset of the DRSs representing $t$ selected by $a$. A complication arises when texts are (superficially) logically inconsistent, e.g., when a text claims something and its negation. As a simplifying hypothesis, we assume that $\tau(a,t)$ is consistent, i.e., the interpreter $a$ is able to solve any possible inconsistencies. Alternatively, one may manage logical inconsistencies, e.g., through paraconsistent logic (Priest, 2002), or by considering only consistent fragments of the obtained set of formulas. The process of *mapping* the surface formulas in $\tau(a,t)$ into formulas of the language of $\beta(a)$ usually involves the interpreters' lexical knowledge, which is assumed to be intersubjectively shared by all competent speakers. However, it may also involve specific past experiences or private beliefs, introducing a subjective dimension at this level as well. We indicate this mapping with $\mu(a,t)$. Note that $\mu(a,t)$ may be inconsistent with $\beta(a)$, e.g., when the text $t$ contradicts common-sense or natural laws. Interpretative assertions and denials assume that interpreters adopt the perspective of the text, hence, they must modify their belief-bases while preserving the text's content. In these cases, therefore, only part of the interpreter's belief-base can be used for their interpretative commitments. We indicate with $\beta_\mu(a,t)$ the update of the original belief-base $\beta(a)$ after incorporating $\mu(a,t)$ and resolving any possible inconsistencies.[25]

We can now characterize the intended meaning of interpretative illocutionary observations. Similarly to the cases of assertions and denials, we provide only necessary conditions:[26]

[24] When $a$ recognizes the ambiguity of $t$, $\tau(a,t)$ contains the logical disjunctions of the formulas corresponding to all the DRSs associated to $t$ accepted by $a$.

[25] Note that $\mu(a,t)$ may be inconsistent with the (translation of the) observational theory $\mathcal{T}_{\mathrm{OBS}}$, which we assume is included in all agents' belief-base, requiring a modification of this theory. This would imply that the agent avoids using the observational concepts affected by the change. This situation is not considered in this work.

[26] Similarly to the case of illocutionary observations, we provide the necessary conditions only for interpretative assertions and denials involving reports. Furthermore, since $\tau(a(r),t)$, $\mu(a(r),t)$, and $\beta_\mu(a(r),t)$ can all change over time, they must be considered when the report $r$ is written and made public.

– if $iASR(x) \wedge \texttt{ARG}(r,t,o,x) \wedge \texttt{REP}(r)$ then either $(i)$ $\mu(a(r),t)$ allows to prove $o$; or $(ii)$ $\beta_\mu(a(r),t)$ allows to prove $o$ but $\beta(a(r))$ alone does not allow to prove $o$;

– if $iDNY(x) \wedge \texttt{ARG}(r,t,o,x) \wedge \texttt{REP}(r)$ then either $(i)$ $\mu(a(r),t)$ is incompatible with $o$; or $(ii)$ $\beta_\mu(a(r),t)$ is incompatible with $o$ but $o$ is compatible with $\beta(a(r))$ alone.

The idea is that, for $a(r)$, what is written in $t$ is necessary to assert or deny $o$. In particular, such characterization contemplates the case where $o$ follows from $\beta(a(r))$, i.e., $a(r)$ already believes $o$, but this occurs only in the case where $o$ also follows from the text (modulo the disambiguation $\tau$ and the mapping $\mu$), i.e., the text must also state $o$. Dually for the denial of $o$.[27]

It is important to stress that interpretative assertions and denials are subjective to an interpreter $a$ in three ways: $(i)$ $\tau(a,t)$ depends on how $a$ solves linguistic ambiguities and possible internal inconsistencies of $t$; $(ii)$ $\mu(a,t)$ depends on $a$'s lexical knowledge and $a$'s prior beliefs; $(iii)$ $\beta_\mu(a,t)$ depends on $a$'s prior beliefs, as well as on the way in which $a$ solves possible inconsistencies between their prior belief-base and $\mu(a,t)$. We can then see in which sense $\mathbf{iasr}(r,r',o)$ makes sense even when $r'$ is a report: the information in $r'$ must be taken into account to prove $o$, i.e., $\mathbf{iasr}(r,r',o)$ presupposes to embrace the commitments in $r'$ (see also Sect. 4.5 on this aspect).

A fourth element of subjectivity in commitments concerns the possibility for the interpreter $a$ to have a specific *ontological view*, i.e., $a$ can interpret the formal language in a restrictive way by considering a proper subset of the whole set of the structures of the language. In this way, one can accommodate both realist and anti-realist positions about the literary entities and a non-monotonic inference mechanism (see Sanfilippo et al. (2024a) for more details).

Following what done for illocutionary observations, we characterize the coherence of reports by taking into account also interpretative illocutionary observations. Axioms (a14) and (a15) are the counterparts of (a10) and (a11), respectively (iA and iD are defined in (d5) and (d6), respectively). These axioms make sense only when we assume that the texts are consistent or interpreters solved the inconsistencies of texts when present. Otherwise, the interpretation of a text could produce incoherent results due to the incoherence of the text itself. Axioms (a16) and (a17) concern the interpretations of reports. They guarantee that what is explicitly (interpretatively) asserted (denied) in a report cannot be denied (asserted) by an interpreter of the report, i.e., reports are not ambiguous with respect to their explicit commitments. Notice that axiom (a17) introduces a link between an interpretative illocutory act (in $r$) and the report of this act in $r'$.

**d5** $\texttt{iA}(r,x,o) := \exists \bar{o}(\bar{o} = \mathbf{iasr}(r,x,o))$

**d6** $\texttt{iD}(r,x,o) := \exists \bar{o}(\bar{o} = \mathbf{idny}(r,x,o))$

**a14** $\texttt{IMP}(o_1,o_2) \rightarrow \neg\exists rt(\texttt{iA}(r,t,o_1) \wedge \texttt{iD}(r,t,o_2))$

**a15** $\texttt{INC}(o_1,o_2) \rightarrow \neg\exists rt(\texttt{iA}(r,t,o_1) \wedge \texttt{iA}(r,t,o_2))$

**a16** $\neg\exists rr'o(\texttt{A}(r,o) \wedge \texttt{iD}(r',r,o)) \wedge \neg\exists rr'o(\texttt{D}(r,o) \wedge \texttt{iA}(r',r,o))$

**a17** $\neg\exists rr'to(\texttt{iA}(r,t,o) \wedge \texttt{iD}(r',r,\mathbf{asr}(t,o))) \wedge \neg\exists rr'to(\texttt{iD}(r,t,o) \wedge \texttt{iA}(r',r,\mathbf{asr}(t,o)))$

Let us go back to the example where both $\texttt{A}(r,\mathbf{asr}(t,o))$ and $\texttt{A}(r,\mathbf{dny}(t,o))$ hold. First, $\texttt{A}(r,\mathbf{asr}(t,o))$ does not imply $\texttt{iA}(r,t,o)$. This is motivated by the fact that $\texttt{A}(r,\mathbf{asr}(t,o))$ does not presuppose the interpretation of $t$ by $a(r)$; the report might in this case simply document the observation $\mathbf{asr}(t,o)$ without

[27] The proposed characterization embraces a sort of pretense mediated version of the *Reality Assumption* (Friend, 2017), which has the known issue that everything $a(r)$ believes can be also the subject of a commitment. Our formulation mitigates the problem, still conjunctions of formulas in $\beta_\mu(a,t)$ could be included in $o$. We do not consider this problem in the following.

direct access to the text $t$. Vice versa, suppose to have $\texttt{iA}(r,t,o)$, which seems to require that $a(r)$ is committed to $\mathbf{asr}(t,o)$, namely, that $\texttt{iA}(r,t,o)$ implies $\texttt{A}(r,\mathbf{asr}(t,o))$. However, the implication does not hold. Indeed, according to the previous semantic characterization of (interpretative) illocutionary observations, $\texttt{iA}(r,t,o)$ means that $\beta_\mu(a(r),t)$ implies $o$, not that $\beta(a(r))$ implies $\mathbf{asr}(t,o)$. To capture this interpretation, one might introduce a sort of *updating function* of $\beta(a(r))$ that, when $\texttt{iA}(r,t,o)$ holds, introduces the observation corresponding to $\mathbf{asr}(t,o)$ in $\beta(a(r))$. We will refrain from conducting this analysis here and limit ourselves to the introduction of (a18). Accordingly, when $r$ reports an interpretative committment to the fact that the text $t$ asserts $o$, the report $r$ cannot denies $\mathbf{asr}(t,o)$ (dually for $\texttt{iD}$).

**a18** $\neg\exists rto(\texttt{iA}(r,t,o) \wedge \texttt{D}(r,\mathbf{asr}(t,o))) \wedge \neg\exists rto(\texttt{iD}(r,t,o) \wedge \texttt{D}(r,\mathbf{dny}(t,o)))$

Note that, (a14) – which, as said, relies on the assumption that interpreters solve possible inconsistencies of texts – does not allow to have both $\texttt{iA}(r,t,o)$ and $\texttt{iD}(r,t,o)$. This means that when $\texttt{A}(r,\mathbf{asr}(t,o))$ and $\texttt{A}(r,\mathbf{dny}(t,o))$ hold, then at least one of $\mathbf{asr}(t,o)$ and $\mathbf{dny}(t,o)$ is not ascribable to an interpretation act of $t$ made explicit in $r$, i.e., it must be "imported" from another report (possibly produced by the same author of $r$). Such "import" can be iterated, e.g., a first report imports from a second report that in its turn imports from a third report, etc. Intuitively, one could think that this importation chain must stop when $\mathbf{asr}(t,o)$ (or $\mathbf{dny}(t,o)$) is found on an interpretative act. We do not consider this aspect here.

We can now make explicit in what sense all observations are *entertained*, see (a19) where $\texttt{ARG}^*(x,o)$ is defined in (d7) and it stands for "$x$ is one of the arguments of the observation $o$." Each observation is either directly included in a report $r$ making public part of the beliefs of $a(r)$ and their interpretations of texts, or it is an argument of another observation, i.e., it is included in a complex observation, e.g., *ASR*- and *DNY*-observations or *SUP*- and *DEF*-observations (see Sect. 5). Note that (d8) here contemplates only the illocutionary observation-kinds we discussed; it needs to be extended when additional illocutionary observation-kinds are introduced in the observational vocabulary.

**d7** $\texttt{ARG}^*(x,o) := \bigvee_{i=1}^{\alpha}(\texttt{ARG}_i(x,o))$

**d8** $\texttt{IN}(o,r) := \texttt{REP}(r) \wedge \texttt{ARG}_1(r,o) \wedge (\mathit{ASR}(o) \vee \mathit{DNY}(o) \vee \mathit{iASR}(o) \vee \mathit{iDNY}(o))$

**a19** $\texttt{OBS}(o) \rightarrow (\exists r(\texttt{IN}(o,r)) \vee \exists \bar{o}(\texttt{ARG}^*(o,\bar{o})))$

### *4.4. Disagreement, Controversiality, and Disputability*

We will now introduce some syntactic definitions that formally capture some properties of (and relations among) reports, observers, and observations. The introduced notions are employed to analyze the collected set of observations (see the guiding example).

Let us start with the principle of *consistency* assumed in the assertion logic by Rescher (1968): an assertor cannot assert a proposition and its negation. As said, our theory does not include negations of observations, therefore this principle must be restated by relying on the notion of (in)coherence. Although the coherence of individual reports is ensured by several constraints (see Sect. 4.2 and Sect. 4.3), different reports produced by the same agent may contain incoherent claims, mainly due to changes in agents' beliefs. Accordingly, assertors are "locally" coherent but they may be incoherent across reports.

*Disagreement* between reports is definable following what done to warrant the coherence of reports in Sect. 4.2 and Sect. 4.3: two reports disagree when they assert incoherent observations or when one denies an assertion implied by the other one, see (d9). Similarly, one can mimic (a14)-(a15) to take into account *iASR*/*iDNY*-observations: (d10) considers a specific text while (d11) is general. One can take the disagreement on the interpretation of a text $t$ as a sign of *controversiality* of $t$, see (d12). Recall

that our theory assumes interpreters being able to solve the possible inconsistencies of texts, therefore a single report cannot make evident the controversiality of a text. First, as said, inconsistencies of texts can be taken into account in different ways, e.g., by introducing a mechanism to manage inconsistencies (e.g., paraconsistent logic). Second, one author can have incoherent views on a text across reports. Third, it is still possible to have genuine disagreements between interpreters, due to their personal beliefs or resolution of inconsistencies.

**d9** $\texttt{rDIS}(r_1, r_2) := \exists o_1 o_2(\texttt{A}(r_1, o_1) \wedge ((\texttt{A}(r_2, o_2) \wedge \texttt{INC}(o_1, o_2)) \vee (\texttt{D}(r_2, o_2) \wedge \texttt{IMP}(o_1, o_2))))$ (*disagreement between reports*)

**d10** $\texttt{iDIS}(r_1, r_2, t) := \exists o_1 o_2(\texttt{iA}(r_1, t, o_1) \wedge ((\texttt{iA}(r_2, t, o_2) \wedge \texttt{INC}(o_1, o_2)) \vee (\texttt{iD}(r_2, t, o_2) \wedge \texttt{IMP}(o_1, o_2))))$ (*interpretative disagreement between reports with respect to text t*)

**d11** $\texttt{iDIS}(r_1, r_2) := \exists t(\texttt{iDIS}(r_1, r_2, t))$ (*interpretative disagreement between reports*)

**d12** $\texttt{CNT}(t) := \exists r_1 r_2(\texttt{iDIS}(r_1, r_2, t))$ (*controversiality of text*)

The *disputability* of an observation can be reduced to disagreement about it, see (d13) and (d14) for a non-mediated and a mediated (by texts) form of disputability, respectively. A similar analysis can be done by focusing on observers rather than observations. An observer is incoherent when they are the author of reports which are in disagreement. Similarly, for the disagreement between observers. Incoherence and disagreement between observers can be formally defined by introducing in our theory a syntactic counterpart of $a(r)$.

**d13** $\texttt{DSP}(o) := \exists r_1 r_2 \bar{o}((\texttt{A}(r_1, o) \wedge \texttt{A}(r_2, \bar{o}) \wedge \texttt{INC}(o, \bar{o})) \vee (\texttt{A}(r_1, o) \wedge \texttt{D}(r_2, \bar{o}) \wedge \texttt{IMP}(o, \bar{o})) \vee (\texttt{D}(r_1, o) \wedge \texttt{A}(r_2, \bar{o}) \wedge \texttt{IMP}(\bar{o}, o)))$ (*disputability of observation*)

**d14** $\texttt{iDSP}(o) := \exists r_1 r_2 t_1 t_2 \bar{o}((\texttt{iA}(r_1, t_1, o) \wedge \texttt{iA}(r_2, t_2, \bar{o}) \wedge \texttt{INC}(o, \bar{o})) \vee (\texttt{iA}(r_1, t_1, o) \wedge \texttt{iD}(r_2, t_2, \bar{o}) \wedge \texttt{IMP}(o, \bar{o})) \vee (\texttt{iD}(r_1, t_1, o) \wedge \texttt{iA}(r_2, t_2, \bar{o}) \wedge \texttt{IMP}(\bar{o}, o)))$ (*interpretative disputab. obs.*)

### *4.5. Accessibility between Reports*

Van Ditmarsch et al. (2007) introduce the notion of *public announcement*. In some sense our reports can be considered as (complex) public announcements. However, in the proposal of Van Ditmarsch et al. (2007), a public announcement is always considered as truthful and impacting all agents (usually becoming *common knowledge*). Differently, in our theory, reports make explicit the point of view of their authors, i.e., a report is only contextually truthful (justified) in the sense that the author is sincere when they assert or deny an observation, i.e., the author's belief-base allows to infer what is claimed. Furthermore, in our scenarios, it is not plausible to assume that all agents have access to the entirety of existing reports. This means that (*i*) what is claimed in a report is not necessarily integrated in the belief-base of all agents present in the framework; (*ii*) when this integration exists, it keeps track of the fact that a report expresses the point of view of the author, it is not a truth.

We represent *accessibility* between reports with $\texttt{ACC}(r_1, r_2)$, standing for "the report $r_1$ *accesses* the report $r_2$." The intended meaning of accessibility is that when $\texttt{ACC}(r_1, r_2)$ holds, every observation with form $\mathbf{asr}(r_2, o)$, $\mathbf{iasr}(r_2, t, o)$, $\mathbf{dny}(r_2, o)$, or $\mathbf{idny}(r_2, t, o)$ is "imported as it is" into $\beta(a(r_1))$, i.e., $\beta(a(r_1))$ keeps track of the observations made public in the report $r_2$ maintaining the information about their source (i.e., $r_2$ itself). From a syntactic perspective, given that the reports only partially reveal the knowledge base of their authors, we cannot presuppose that all the observations imported from $r_2$ in $\beta(a(r_1))$ are made public in $r_1$. Syntactically, such importing mechanism can thus be only partially characterized by

assuring that what is stated in $r_2$ is not denied in $r_1$, see (a20)-(a23). Furthermore, in an empirical setting where reports are made public at a given time, $\texttt{ACC}(r_1, r_2)$ implies that $r_2$ must be published before $r_1$ making accessibility asymmetric.[28] Finally, (a24) assures that to interpret a report one needs to access it.

**a20** $\texttt{ACC}(r_1, r_2) \wedge \texttt{A}(r_2, o) \rightarrow \neg\texttt{D}(r_1, \textbf{asr}(r_2, o))$

**a21** $\texttt{ACC}(r_1, r_2) \wedge \texttt{D}(r_2, o) \rightarrow \neg\texttt{D}(r_1, \textbf{dny}(r_2, o))$

**a22** $\texttt{ACC}(r_1, r_2) \wedge \texttt{iA}(r_2, t, o) \rightarrow \neg\texttt{D}(r_1, \textbf{iasr}(r_2, t, o))$

**a23** $\texttt{ACC}(r_1, r_2) \wedge \texttt{iD}(r_2, t, o) \rightarrow \neg\texttt{D}(r_1, \textbf{idny}(r_2, t, o))$

**a24** $(\texttt{iA}(r_1, r_2, o) \vee \texttt{iD}(r_1, r_2, o)) \wedge \texttt{REP}(r_2) \rightarrow \texttt{ACC}(r_1, r_2)$

It is relevant to underline the difference between accessing vs. interpreting a report: when $a$ accesses $r$, $a$ just *registers* the point of view of $a(r)$ into $\beta(a)$ keeping track of its source; when $a$ interprets $r$, $a$ *embraces* the point of view of $a(r)$, and such an embracement might entail to renounce to part of $\beta(a)$.

**Example cnt. (reports)**. We start by considering some reports concerning, in a more or less direct way, the tale X.10 (tx10) and its characters Gualtieri (gua) and Griselda (gri). Consider the report $\texttt{r}_1$ making public some of the direct interpretations of tx10 by $a(\texttt{r}_1)$ listed in Table 1. According to the intended meaning of interpretative assertions, $a(\texttt{r}_1)$, by interpreting tx10, infers various basic observations like those relative to the moral conditions of Gualteri (f19) and Griselda (f20). Notice that, in our theory, basic observations like **unjust**(gua) cannot exist in isolation but they must be always included within illocutionary observations, see (a19).

| $\texttt{r}_1$ | |
|---|---|
| **iasr**($\texttt{r}_1$, tx10, **unjust**(gua)) | **f19** |
| **iasr**($\texttt{r}_1$, tx10, **just**(gri)) | **f20** |
| **iasr**($\texttt{r}_1$, tx10, **feudalLord**(gua)) | **f21** |
| **iasr**($\texttt{r}_1$, tx10, **peasant**(gri)) | **f22** |
| **iasr**($\texttt{r}_1$, tx10, **feudal**(gua, tx10)) | **f23** |

| $\texttt{r}_2$ | |
|---|---|
| **idny**($\texttt{r}_2$, tx10, **unjust**(gua)) | **f24** |
| **asr**($\texttt{r}_2$, **asr**(tx10, **high**(gri))) | **f25** |
| **asr**($\texttt{r}_2$, **asr**(tx10, **unjust**(gua))) | **f26** |
| **asr**($\texttt{r}_2$, **iasr**($\texttt{r}_1$, tx10, **peasant**(gri))) | **f27** |
| **iasr**($\texttt{r}_2$, tx10, **married**(gri, gua)) | **f28** |

Table 1
Reports $\texttt{r}_1$ and $\texttt{r}_2$.

Consider now the report $\texttt{r}_2$ and assume $\texttt{ACC}(\texttt{r}_2, \texttt{r}_1)$. First, according to (a20)-(a23), which regulate ACC, $\texttt{r}_2$ does not deny any assertion in $\texttt{r}_1$. In particular, (f24) is not a denial of (f19) but it makes explicit that $\texttt{r}_2$ and $\texttt{r}_1$ disagree on the interpretation of tx10 about the moral condition of Gualtieri, see (d10) (where $\texttt{IMP}(\textbf{unjust}(\texttt{gua}), \textbf{unjust}(\texttt{gua}))$ trivially holds). Consequently, such reports (as well as they authors) plainly disagree in the sense defined in (d11). This also means that **unjust**(gua) is interpretatively disputable, see (d14). Formulas (f22) and (f25) can be seen as another source of disagreement between $\texttt{r}_1$ and $\texttt{r}_2$ (and their authors) given that, by (f15), $\texttt{INC}(\textbf{peasant}(\texttt{gri}), \textbf{high}(\texttt{gri}))$. We refrain from providing a definition of this form of disagreement, although this could be easily added.

Second, (f26) is not incoherent with (f24); it simply makes explicit that $a(\texttt{r}_2)$ considers **unjust**(gua) with respect to tx10 and then that $a(\texttt{r}_2)$ is aware that there is at least a position on tx10 different from the one that $a(\texttt{r}_2)$ expresses through (f24). Assertions like (f26) depend on what others have said, and

[28] One could assume an introspective reasoning of reports' authors where ACC is reflexive. We do not consider this case here.

$a(\mathtt{r}_2)$ may have simply imported this belief in their report. The assertion in (f27) can be seen as an import of (f22) due to $\mathtt{ACC}(\mathtt{r}_2, \mathtt{r}_1)$.

Consider now the texts *Boccaccio Medievale* (bmd) by Branca, *Boccaccio e la Codificazione della Novella* (bcn) by Picone, *Lai de Fresne* (lai), and hagiographic literature (hag) together with the reports $\mathtt{r}_3$ and $\mathtt{r}_4$ in Table 2 concerning their interpretations.[29] Furthermore, assume that $\mathtt{ACC}(\mathtt{r}_3, \mathtt{r}_1)$. These reports contain subtle distinctions. In the following analysis we use the notions introduced in Sect. 4.4 but we also suggest possible extensions aimed at capturing complex links between reports and observations.

| $\mathtt{r}_3$ | |
|---|---|
| **iasr**($\mathtt{r}_3$, bmd, **similar**(gri, mary)) | **f29** |
| **iasr**($\mathtt{r}_3$, tx10+hag, **similar**(gri, mary)) | **f30** |
| **iasr**($\mathtt{r}_3$, hag, **peasant**(mary)) | **f31** |
| **iasr**($\mathtt{r}_3$, bmd, **asr**(tx10+hag, **similar**(gri, mary))) | **f32** |

| $\mathtt{r}_4$ | |
|---|---|
| **idny**($\mathtt{r}_4$, bcn, **similar**(gri, mary)) | **f33** |
| **iasr**($\mathtt{r}_4$, bcn, **similar**(gri, fre)) | **f34** |
| **iasr**($\mathtt{r}_4$, bcn, **def**(**similar**(gri, fre), **similar**(gri, mary))) | **f35** |

| $\mathtt{r}_5$ | |
|---|---|
| **iasr**($\mathtt{r}_5$, $\mathtt{r}_1$, **asr**(tx10, **peasant**(gri))) | **f36** |
| **iasr**($\mathtt{r}_5$, $\mathtt{r}_3$, **asr**(hag, **peasant**(mary))) | **f37** |
| **asr**($\mathtt{r}_5$, **asr**(tx10, **peasant**(gri))) | **f38** |
| **dny**($\mathtt{r}_5$, **similar**(gri, mary)) | **f39** |

Table 2
The reports $\mathtt{r}_3$, $\mathtt{r}_4$, and $\mathtt{r}_5$.

First, observe that $\mathtt{r}_3$ lists two direct sources for **similar**(gri, mary), see (f29) and (f30). That is, according to $a(\mathtt{r}_3)$, the direct interpretation of either bmd or tx10+hag[30] is enough to infer the similarity between Griselda and the Virgin Mary.

Second, considering (f29) and (f33), by (d14), **similar**(gri, mary) results iDSP-disputable, although $\mathtt{r}_3$ and $\mathtt{r}_4$ do not disagree since their observations concern different texts.

Third, (f31) together with $\mathtt{ACC}(\mathtt{r}_3, \mathtt{r}_1)$ could be seen as a weak sign of alignment between $\mathtt{r}_3$ and $\mathtt{r}_1$ since, in the context of tx10+hag, one could suppose that $a(\mathtt{r}_3)$, by accepting **similar**(gri, mary) and **peasant**(mary) (and given the fact that hag does not talk about Griselda) also accepts **peasant**(gri), an observation on which $\mathtt{r}_1$ is committed via (f22). To formally infer this alignment, a link between being similar and being peasant is required, but we can at least say that the situation is compatible with such alignment. Note that (f32) differs from (f29); it affirms that, according to $a(\mathtt{r}_3)$, bdm allows to

[29] Formula (f35) in $\mathtt{r}_4$ will be discussed at the end of Sect. 5 after the introduction of support and defeat observations.

[30] The notation tx10+hag indicates the text that is the mereological sum of tx10 and hag, i.e., "+" is just a way to refer to tx10 and hag together.

infer not only **similar**(gri, mary) but also **asr**(tx10+hag, **similar**(gri, mary)), i.e., bdm may have "imported" **similar**(gri, mary) from tx10+hag.

Consider now the report $\mathtt{r}_5$ in Table 2. Formula (f36), together with (f38), suggests that $a(\mathtt{r}_5)$ embraced, or was already in agreement with, the point of view of $\mathtt{r}_1$ about Griselda being peasant in tx10. Vice versa, (f37)-(f39) can indicate $(i)$ that $a(\mathtt{r}_5)$ did not embrace the point of view of $\mathtt{r}_3$ about the fact that in hag the Virgin Mary is peasant; or $(ii)$ that the fact that both Griselda and the Virgin Mary are peasant is not enough for $a(\mathtt{r}_5)$ to infer their similarity. □

## 5. Observations of Support and Defeat

We introduce in this section two other kinds of complex observations to make sense of the *arguments* (we could also say *evidences*) that one may put forward to *support* or *defeat* other observations. For instance, the observation ascribing a genre to a literary text may be supported by other observations concerning the linguistic features or the content of the text (Sanfilippo et al., 2023). Supporting or defeating observations are not always conclusive, because they may not be sufficient to infer or negate other observations; in a sense, they merely provide "hints" in favor or against an observation. They can be viewed as weaker forms of correlation and incoherence, respectively, with a subjective dimension. Indeed, while correlations and incoherencies are accepted by all the observers adopting an observational vocabulary, support or defeat depend on the specific observer's point of view.[31]

For these purposes, we adapt the *bipolar argumentation framework* (Cayrol and Lagasquie-Schiex, 2013), a general theory that extends the theory of Dung (1995) by considering a set of *arguments* and two kinds of interactions between arguments: a negative interaction expressed by the *defeat* (*attack*) relation, and a positive interaction expressed by the *support* relation.

Following this approach, we introduce the complex observation-kinds *SUP* (for support) and *DEF* (for defeat) which are included in all observational vocabularies. The support-observation $\mathbf{sup}(o_1, o_2)$ is read as: "the observation $o_1$ supports the observation $o_2$." The defeat-observation $\mathbf{def}(o_1, o_2)$ is read as "the observation $o_1$ defeats the observation $o_2$." *SUP*- and *DEF*-observations have two arguments both of which are observations (see (a25) and (a26)). Hence, they correspond to relations between pieces of information capturing the intuition that the supporting (defeating) information is intended to increase (decrease) the plausibility of the supported (defeated) information.[32] Axioms (a27)-(a30) regulate the incoherence and correlation relations between *SUP*- and *DEF*-observations considering refinements and correlations. For the sake of clarity, according to (a27), when observation $o_1$ implies (in the sense of either refinement or correlation) observation $o_2$, and there is an observation $o_3$ such that it is supported by both $o_1$ and $o_2$, then there is a relation of correlation between the support observations. The remaining axioms (a28)–(a30) have a similar reading for defeat observations and incoherence relations between observations.

**a25** $SUP(o) \rightarrow \exists xy(\mathtt{ARG}(x,y,o)) \wedge \forall xy(\mathtt{ARG}(x,y,o) \rightarrow \mathtt{OBS}(x) \wedge \mathtt{OBS}(y))$

**a26** $DEF(o) \rightarrow \exists xy(\mathtt{ARG}(x,y,o)) \wedge \forall xy(\mathtt{ARG}(x,y,o) \rightarrow \mathtt{OBS}(x) \wedge \mathtt{OBS}(y))$

**a27** $\mathtt{IMP}(o_1,o_2) \wedge \exists o\bar{o}o_3(o = \mathbf{sup}(o_2,o_3) \wedge \bar{o} = \mathbf{sup}(o_1,o_3)) \rightarrow \mathtt{COR}(o,\bar{o})$

**a28** $\mathtt{IMP}(o_1,o_2) \wedge \exists o\bar{o}o_3(o = \mathbf{def}(o_2,o_3) \wedge \bar{o} = \mathbf{def}(o_1,o_3)) \rightarrow \mathtt{COR}(o,\bar{o})$

[31] However, note that defeat has a direction whereas incoherence is symmetric.

[32] Since $\mathbf{sup}(o_1,o_2)$ and $\mathbf{def}(o_1,o_2)$ are observations, one may have support of support ($\mathbf{sup}(o,\mathbf{sup}(o_1,o_2))$), support of defeat ($\mathbf{sup}(o,\mathbf{def}(o_1,o_2))$), etc., namely, sorts of *higher-order interactions*.

**a29** $\texttt{IMP}(o_1,o_2) \wedge \exists o\bar{o}o_3(o = \mathbf{sup}(o_2,o_3) \wedge \bar{o} = \mathbf{def}(o_1,o_3)) \rightarrow \texttt{INC}(o,\bar{o})$

**a30** $\texttt{IMP}(o_1,o_2) \wedge \exists o\bar{o}o_3(o = \mathbf{def}(o_2,o_3) \wedge \bar{o} = \mathbf{sup}(o_1,o_3)) \rightarrow \texttt{INC}(o,\bar{o})$

### *5.1. Conflictuality, Acceptability, and Admissibility*

Once the evidences exhibited by different scholars to support or defeat an observation are made explicit, it is possible to analyze the debates among multiple agents represented by networks of *SUP*- and *DEF*-observations. Given the intuition that support and defeat are the subjective and weak counterparts of correlation and incoherence, we will consider some principles of *assertion logics* (Rescher, 1968) and *argumentation frameworks* (Dung, 1995; Cayrol and Lagasquie-Schiex, 2013; Brewka, 2001) in light of the definitions introduced in Sect. 4.4.

Definitions (d15)-(d18) are the adaptations of (d9)-(d12).[33] Intuitively, two reports are in conflict when they assert observations where one defeats (according to a third report) the other or the first report asserts an observation that (according to a third report) supports an observation denied by the second report, see (d15). We say that the report $r$ is *conflictual* when $\texttt{rCNF}(r,r)$. Similarly for the interpretative counterpart (see (d16)) which allows us to define the conflictuality of texts (see (d18)). Note that, when for instance $\texttt{A}(r_1,o_1)$ and $\texttt{A}(r_2,o_2)$ hold, the simple existence of the observation of $\mathbf{def}(o_1,o_2)$ is not enough to conclude that $r_1$ and $r_2$ are conflictual: $\mathbf{def}(o_1,o_2)$ could only appear in another report $r'$ such that $\texttt{D}(r',\mathbf{def}(o_1,o_2))$ or $\texttt{iA}(r,t,\mathbf{dny}(t',\mathbf{def}(o_1,o_2)))$.

A report $r$ is *acceptable* if for all the observations $o_1$ defeating an assertion $o$ of $r$ there is an assertion $o_2$ of $r$ defeating $o_1$, see (d19). The intuition behind (d19) is that report $r$ is acceptable if it contains counter-arguments for all criticisms that can be raised against what is asserted in it. A report $r$ is *admissible* if it is conflict-free and acceptable, see (d20).

Disputability and interpretative disputability, see (d13) and (d14), can be adapted by taking into account defeat and support observations. A weak form of disputability is defined in (d21), analogously f or interpretative disputability.

**d15** $\texttt{rCNF}(r_1,r_2) := \exists o_1o_2(\texttt{A}(r_1,o_1) \wedge ((\texttt{A}(r_2,o_2) \wedge \exists r(\texttt{A}(r,\mathbf{def}(o_2,o_1)))) \vee (\texttt{D}(r_2,o_2) \wedge \exists r(\texttt{A}(r,\mathbf{sup}(o_1,o_2))))))$ (*conflictuality between reports*)

**d16** $\texttt{iCNF}(r_1,r_2,t) := \exists o_1o_2(\texttt{iA}(r_1,t,o_1) \wedge ((\texttt{iA}(r_2,t,o_2) \wedge \exists r(\texttt{A}(r,\mathbf{def}(o_2,o_1)))) \vee (\texttt{iD}(r_2,t,o_2) \wedge \exists r(\texttt{A}(r,\mathbf{sup}(o_1,o_2))))))$

**d17** $\texttt{iCNF}(r_1,r_2) := \exists t(\texttt{iCNF}(r_1,r_2,t))$ (*interpretative conflictuality between reports*)

**d18** $\texttt{CNF}(t) := \exists r_1r_2(\texttt{iCNF}(r_1,r_2,t))$ (*conflictuality of text*)

**d19** $\texttt{ACP}(r) := \forall r_1o_1o(\texttt{A}(r,o) \wedge \texttt{A}(r_1,\mathbf{def}(o_1,o)) \rightarrow \exists o_2(\texttt{A}(r,\mathbf{def}(o_2,o_1))))$ (*acceptability*)

**d20** $\texttt{ADM}(r) := \neg\texttt{rCNF}(r,r) \wedge \texttt{ACP}(r)$ (*admissibility*)

**d21** $\texttt{wDSP}(o) := \exists r_1r_2\bar{r}\bar{o}((\texttt{A}(r_1,o) \wedge \texttt{A}(r_2,\bar{o}) \wedge \texttt{A}(r,\mathbf{def}(o,\bar{o})) \vee (\texttt{A}(r_1,o) \wedge \texttt{D}(r_2,\bar{o}) \wedge \texttt{A}(r,\mathbf{sup}(o,\bar{o}))) \vee (\texttt{D}(r_1,o) \wedge \texttt{A}(r_2,\bar{o}) \wedge \texttt{A}(r,\mathbf{sup}(\bar{o},o))))$ (*disputability of observation*)

**Example cnt. (support and defeat observations)**. Consider the report $\texttt{r}_6$ in Tab. 3. It can be seen as the result of the importing of information from $\texttt{r}_3$ and $\texttt{r}_4$, i.e., $\texttt{ACC}(\texttt{r}_6,\texttt{r}_3)$, $\texttt{ACC}(\texttt{r}_6,\texttt{r}_4)$ and $a(\texttt{r}_6)$ includes the content of interpretative assertions (f29), (f34), and (f35) in their knowledge base. According to (d15) we have $\texttt{rCNF}(\texttt{r}_6,\texttt{r}_6)$, i.e., $\texttt{r}_6$ is conflictual and then non admissible (see (d20)). We also

[33] Note that rCNF is not symmetric although (d15) can easily be complicated to achieve symmetry. Analougously for (d16).

trivially have that the observations **similar**(gri, mary) and **similar**(gri, free) are disputable (see (d21)). Conversely, there is no interpretative conflictuality (see (d17)) because no information in $\mathtt{r}_6$ assures that both **similar**(gri, mary) and **similar**(gri, free) originate from interpretations of the same text (in our scenario, they originate from interpretations of different texts, see (f29) and (f34)).

Note that $\mathtt{r}_6$ is coherent. However, "internally" conflicting reports (such as $\mathtt{r}_6$) where an observer asserts $o_1$ and $o_2$ together with **def**$(o_1, o_2)$, or asserts $o_1$ and **sup**$(o_1, o_2)$ but denies $o_2$, can be seen as weak forms of incoherence leaving room to reject them. In this case, the example can be modified by adding a report $\mathtt{r}_7$ and moving from $\mathtt{r}_6$ to $\mathtt{r}_7$ $(i)$ (f42) or $(ii)$ (f41). In the first case we still have $\mathtt{rCNF}(\mathtt{r}_6, \mathtt{r}_6)$ but the conflictuality of $\mathtt{r}_6$ is due to the *DEF*-observation asserted in a different report, i.e., $\mathtt{r}_7$. Furthermore, neither $\mathtt{rCNF}(\mathtt{r}_6, \mathtt{r}_7)$ nor $\mathtt{rCNF}(\mathtt{r}_7, \mathtt{r}_6)$ holds. In the second case, $\mathtt{rCNF}(\mathtt{r}_6, \mathtt{r}_7)$ holds but $\mathtt{rCNF}(\mathtt{r}_6, \mathtt{r}_6)$ does not. In both cases, **similar**(gri, mary) and **similar**(gri, free) are still disputable observations. □

| $\mathtt{r}_6$ | |
|---|---|
| **asr**$(\mathtt{r}_6,$ **similar**$(\mathtt{gri}, \mathtt{mary}))$ | **f40** |
| **asr**$(\mathtt{r}_6,$ **similar**$(\mathtt{gri}, \mathtt{fre}))$ | **f41** |
| **asr**$(\mathtt{r}_6,$ **def**$($**similar**$(\mathtt{gri}, \mathtt{fre}),$ **similar**$(\mathtt{gri}, \mathtt{mary})))$ | **f42** |

Table 3
The report $\mathtt{r}_6$.

## 6. State of the Art

The theory of observations introduced in the previous sections can be compared with approaches for text annotation in natural language processing tasks. In these cases, annotators use languages of different sorts, from controlled vocabularies to axiomatized ontologies (Pustejovsky et al., 2017). Annotations can have different goals; e.g., they capture the linguistic features of a text or have an explicit interpretative dimension to model how an annotator interprets the content of a text (Gius and Jacke, 2017; Horstmann et al., 2020; Jacke, 2025).

There is a straightforward similarity between annotation languages and the observational vocabularies of our theory. An observational vocabulary can be understood as a formal language that is shared among annotators to express their perspectives on texts. In this view, what we call report is a coherent collection of annotations made by means of an observational vocabulary. As said, our theory can be extended by representing explicitly the time when an observation is made, and by formulating coherence criteria for reports on a dynamic scale. At the current state, a report records a coherent sets of observations (annotations) made public at a given time. Despite the similarity between annotation and observational vocabularies, notice that the latter bear a broader scope than the former, since they can be also used to express the results of apparatuses of different sorts including physical measurement tools (see below). In addition, as a relevant key aspect of our theory, it is important to stress that by having observations into the domain of quantification, we can use them to model observations of observations like assertion, denial, support or defeat, among others. As shown in the previous sections, this is particularly important in the humanities. Also, our approach considers an explicit multi-agent perspective where annotators are not requested to annotate a text in the same way, i.e., they do not need to reach a common ground

truth in their observations. Quite the opposite, they can use the same observational vocabulary to model alternative and possibly conflicting perspectives. This approach is similar to the perspectivist turn in machine learning (Cabitza et al., 2023).

For the sake of completeness, the presented theory can be extended to cover sensors that are designed to output information encoded in terms of an observational vocabulary – for example, the position of the index on a weighing scale is associated with a specific weight observation concerning the object under measurement – or perception of cognitive agents. Although sensing and perceiving usually provide information about objects in the world, they do not presuppose reasoning or rational justification. At least three different approaches can be considered for modeling the results of sensing and perceiving in our framework. $(i)$ Represent the output of sensors via standard assertions. In this case, the sensing results constitute the knowledge bases of sensors that trivially found their inferences on them. $(ii)$ Introduce the new observation kind *Sensed* with a single argument of kind `OBS` and assume that, in their reports, sensors assert *Sensed*-observations. In this way, we explicitly represent the fact that the sensors are direct producer of information via observations with form $\mathbf{asr}(r, \mathbf{sensed}(o))$. $(iii)$ Represent sensing illocutionary acts by introducing the new observation kind *Sensing* that involves reports produced by sensors and observations, such as $\mathbf{sense}(r, o)$.

Commenting on existing ontologies in the state of the art, there exist few works providing an explicit formal representation of what someone (or something) could claim with respect to a domain entity. Daquino and Tomasi (2015) present the Historical Context Ontology (HiCO) in OWL to represent scholarly statements in the humanities like interpretations of literary texts.[34] HiCO is developed "for representing the *context* of a claim. [...] For instance, being created by somebody, or being created at a certain time, are events related to an artefact that are claimed by an agent at a certain time, motivated with usage of primary sources, and recorded in a secondary source [...]" (from the online documentation, see footnote 34; emphasis is ours). Without entering into technical aspects relative to the differences with our approach, HiCO focuses on representing statements' meta-data; it has therefore a narrower scope in comparison to ours, since it does not address *what* a statement claims.

To model the contents of claims, some of the authors of HiCO developed an approach based on *named graphs* (Pasqual et al., 2024). They distinguish between different sorts of graphs in the Resource Description Framework (RDF) to separate hypothetical statements, whose truth-value is neither true, nor false, from statements whose truth-values were settled by scholars. The proposal seems to be mainly developed for query-answering rather than for reasoning, e.g., to detect incompatibilities between claims. On this respect, the notion of "collapse graph" introduced by Pasqual et al. (2024) to deal with the representation of the so-called "settled disputes", i.e., claims that have been disputed by someone, but for which a consensus is eventually reached, marks a major difference with our approach. While collapse graphs, and therefore the existence of settled claims in a given scenario, are the result of a modeler's choice, our theory can reason about the existence of such claims. In particular, the theory can deduce that an observation is object of common agreement (i.e., it is a settled claim) because all the available sources assert it. In addition, the representation of observations in our approach could be enriched with explicit reference to background knowledge; e.g., knowing that the date of birth of a person is unique would support the categorization of two different date-observations for the same person as conflictual.

Another approach is the CRM Argumentation Model (CRMinf) (Doerr et al., 2023), an extension of the CIDOC-CRM ontology (Bruseker et al., 2017) designed to support the documentation of argumentation and the resulting data. The class *I1 Argumentation* plays a important role in the scope of the CRMinf.

[34] Documentation: https://marilenadaquino.github.io/hico/.

An argumentation is an activity resulting in a belief which has a proposition set as content (*what* the belief states) and a belief value (e.g., true, false, more or less certain, etc.). In the paper laying down the foundations of CRMinf, Doerr et al. (2011) talk of beliefs as mental entities. From a modeling standpoint, this view raises problems of how to access them without "looking into agents' minds." To do this, CRMinf relies on what it calls proposition sets. Following the documentation: "This class [*I4 Proposition Set*] comprises the sets of formal, binary propositions that a *I2 Belief* is held about. It could be implemented as a named graph, a spreadsheet, or any other structured dataset" (Doerr et al., 2023, p.15). If we understand it correctly, proposition sets are meant to provide formal representations for agents' beliefs. In our case, we intentionally avoid to embrace a mentalistic attitude and avoid reference to agent's beliefs altogether, relying instead on the modeling of observations as pieces of publicly available information expressed through observational vocabularies. This is an important departure point since our goal is to represent the sort of claims that agents express in accessible sources.

Finally, Carriero et al. (2021) adopt a pattern for what they call "interpretation situations" in order to model the distinction between factual and hypothetical data. For instance, these situations can be used to represent the attribution of properties to domain entities like that of authorship for texts whose authors are unknown. Sartini et al. (2023) extended the approach for interpretations in art history.

From a formal perspective, there are several differences between the approaches just mentioned and ours. It is enough to notice that ontologies in the CRM library are conceptual models in semi-formal notations. HiCO and the approach of Pasqual et al. (2024) are RDF-based, hence their expressivity is much lower in comparison to the theory that we presented. Future research will involve developing a computational ontology in OWL based on the initial contribution by Sanfilippo et al. (2024b) and the theory presented in this paper in order make the ontology operational in applications. Note that, by adopting a foundational approach, we also depart from Carriero et al. (2021), who treat interpretation situations as merely OWL patterns for n-ary relations.

## 7. Conclusions

We presented a formal theory to model observations and deepened the representation of observations expressing interpretations of literary texts. The grounding idea is that for a single text, interpreters can provide different and possibly incompatible interpretations. The theory allows to formally represent (parts of) the interpretations and to compare them to find their commonalities and points of departure. In this view, the interpretation of a text by an interpreter is a sentence – expressed through a shared observational vocabulary – that according to the interpreter can be derived from the text, possibly by considering also the interpreter's beliefs applied to the text. Following Jacke (2014), the notion of derivation is understood in a broad sense, which could cover classic deduction as well as non-classic reasoning mechanisms.

It is worth stressing that, although the theory provides a way to make texts' interpretations formally representable, a position that can sound suspicious to some, it *does* take into account interpreters' subjective dimension. This subjectivity is reflected in multiple aspects of the theory; e.g., it considers the beliefs of single interpreters, their individual understandings of texts, including the way in which interpreters solve the possible ambiguities or inconsistency of texts.

Despite our efforts, the phenomenon of interpretation is complex and multifaceted, and further work is required. For instance, developing observational vocabularies is challenging, especially when analyzing scholarly texts adopting different theoretical backgrounds and using multiple, not always plain clear, conceptual systems (Pichler and Reiter, 2022; Jacke, 2025). Another challenge is to better understand

the way scholars argue in favor or against certain theses (Tomazzoli, 2025). Our observations of support and defeat can partially capture these aspects, but further work is needed to dig into argumentation mechanisms. Last but not least, the theory can be extended to take into account time and change over time. First, to model when a thing is observed to have a given (temporally variable) property, one could add a temporal argument to the corresponding observation-kinds. Second, to model when an illocutionary (interpretative) act is produced, i.e., to explicitly tell when an agent commits to an observation, one could rely on the production time of reports. Third, a temporal argument can be introduced in the accessibility relation between reports to make explicit when an agent imported a piece of information coming from another report.

**Acknowledgment**. We would like to thank the reviewers of Applied Ontology for their feedback on previous versions of this paper. Any remaining errors are our responsibility. 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

Austin, J.L. (1962). *How to do things with words*. Massachusetts: Harvard University Press. edited by J.O. Urmson and Marina Sbisà.

Bach, K. & Harnish, R.M. (1979). Communication and speech acts. *Cambridge, Mass.: Harvard*.

Barabucci, G., Tomasi, F. & Vitali, F. (2021). Supporting complexity and conjectures in cultural heritage descriptions. In *Proceedings of the International Conference Collect and Connect: Archives and Collections in a Digital Age (COLCO 2020), CEUR workshop proceedings* (Vol. 2810).

Benton, M.A. & Van Elswyk, P. (2020). Hedged assertion. In S.C. Goldberg (Ed.), *The Oxford handbook of assertion* (pp. 245–263). Oxford University Press Oxford.

Branca, V. (1956). *Boccaccio medievale*. G. C. Sansoni.

Brewka, G. (2001). Dynamic argument systems: A formal model of argumentation processes based on situation calculus. *Journal of logic and computation*, *11*(2), 257–282.

Bruseker, G., Carboni, N. & Guillem, A. (2017). Cultural heritage data management: the role of formal ontology and CIDOC CRM. *Heritage and archaeology in the digital age: acquisition, curation, and dissemination of spatial cultural heritage data*, 93–131.

Cabitza, F., Campagner, A. & Basile, V. (2023). Toward a perspectivist turn in ground truthing for predictive computing. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 37, pp. 6860–6868).

Carrara, M., Chiffi, D. & De Florio, C. (2024). Towards a polarized semantics for assertion and denial. *Logic Journal of the IGPL*, jzae090.

Carriero, V.A., Gangemi, A., Mancinelli, M.L., Nuzzolese, A.G., Presutti, V. & Veninata, C. (2021). Pattern-based design applied to cultural heritage knowledge graphs. *Semantic Web*, *12*(2), 313–357.

Cayrol, C. & Lagasquie-Schiex, M.-C. (2013). Bipolarity in argumentation graphs: Towards a better understanding. *International Journal of Approximate Reasoning*, *54*(7), 876–899.

Daquino, M. & Tomasi, F. (2015). Historical Context Ontology (HiCO):A conceptual model for describing context information of cultural heritage objects. In *Metadata and Semantics Research: 9th Research Conference, 2015, Proceedings* (pp. 424–436). Springer.

Daquino, M., Pasqual, V. & Tomasi, F. (2020). Knowledge Representation of digital Hermeneutics of archival and literary Sources. *JLIS. it*, *11*(3), 59–76.

Derrida, J. & Attridge, D. (1991). "THIS STRANGE INSTITUTION CALLED LITERATURE". AN INTERVIEW WITH JACQUES DERRIDA. In J. Derrida and D. Attridge (Eds.), *Acts of Literature* (pp. 33–75). Routledge.

Doerr, M., Kritsotaki, A. & Boutsika, K. (2011). Factual argumentation—a core model for assertions making. *Journal on Computing and Cultural Heritage (JOCCH)*, *3*(3), 1–34.

Doerr, M., Ore, C.-E. & Fafalios, P. (2023). Definition of the CRMinf. An Extension of CIDOC-CRM to support argumentation. Technical report, CIDOC CRM-SIG.

Dubois, D., Hájek, P. & Prade, H. (2000). Knowledge-driven versus data-driven logics. *Journal of logic, Language and information*, *9*, 65–89.

Dung, P.M. (1995). On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial intelligence*, *77*(2), 321–357.

Feagin, S.L. (1982). Incompatible interpretations of art. *Philosophy and Literature*, *6*(1), 133–146.

Friend, S. (2017). The real foundation of fictional worlds. *Australasian Journal of Philosophy*, *95*(1), 29–42.

Galván, L. (2017). Counterfactual claims about fictional characters: philosophical and literary perspectives. *Journal of Literary Semantics*, *46*(2), 87–107.

Gius, E. & Jacke, J. (2017). The hermeneutic profit of annotation: On preventing and fostering disagreement in literary analysis. *International Journal of Humanities and Arts Computing*, *11*(2), 233–254.

Hatzel, H.O., Stiemer, H., Biemann, C. & Gius, E. (2023). Machine learning in computational literary studies. *it-Information Technology*, *65*(4-5), 200–217.

Hempfer, K.W. (2024). *Fundamentals of Literary Theory*. Springer.

Horstmann, J., Nantke, J. & Schlupkothen, F. (2020). Undogmatic Literary Annotation with CATMA. *Annotations in Scholarly Editions and Research*, 157–176.

Jacke, J. (2014). Is There a Context-Free Way of Understanding Texts? The Case of Structuralist Narratology. *Journal of Literary Theory*, *8*(1), 118–139.

Jacke, J. (2025). Operationalization and Interpretation Dependence in Computational Literary Studies. *Journal of Computational Literary Studies*, *4*(1).

Janowicz, K., Haller, A., Cox, S.J., Le Phuoc, D. & Lefrançois, M. (2019). SOSA: A lightweight ontology for sensors, observations, samples, and actuators. *Journal of Web Semantics*, *56*, 1–10.

Kamp, H. & Reyle, U. (2013). *From discourse to logic: Introduction to model-theoretic semantics of natural language, formal logic and discourse representation theory*. Springer.

Krausz, M. (2010). *Is There a Single Right Interpretation?* Penn State Press.

Kripke, S.A. (1982). *Wittgenstein on rules and private language: an elementary exposition*. Cambridge, Mass.: Harvard University Press.

Kroon, F. & Voltolini, A. (2023). Fictional Entities. In E.N. Zalta and U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Fall 2023 ed.). Metaphysics Research Lab, Stanford University.

Lascarides, A. & Asher, N. (2007). Segmented discourse representation theory: Dynamic semantics with discourse structure. In *Computing meaning* (pp. 87–124). Springer.

Margolin, U. (2012). Mathematics and narrative: A narratological perspective. In *Circles disturbed: the interplay of mathematics and narrative* (pp. 481–507). Princeton University Press.

Marsili, N. (2024). The definition of assertion: Commitment and truth. *Mind & Language*, *39*(4), 540–560.

Masolo, C., Botti Benevides, A. & Porello, D. (2018). The interplay between models and observations. *Applied Ontology*, *13*(1), 41–71.

Masolo, C., Sanfilippo, E.M., Ferrario, R. & Pierazzo, E. (2021). Texts, Compositions, and Works: A Socio-Cultural Perspective on Information Entities. In *JOWO proceedings*. CEUR vol.2969.

Pasqual, V., Manzano, G., Uzeir, E., Tomasi, F., Vitali, F., et al. (2024). Evaluation of Expressing without asserting approaches in RDF. The case of Conjectures. In *IRCDL* (pp. 60–73).

Pianzola, F., Cheng, L., Pannach, F., Yang, X. & Scotti, L. (2025). The GOLEM Ontology for Narrative and Fiction. *Humanities*, *14*(10).

Pichler, A. & Reiter, N. (2022). From Concepts to Texts and Back: Operationalization as a Core Activity of Digital Humanities. *Journal of Cultural Analytics*, *7*(4).

Picone, M. (2008). *Boccaccio e la codificazione della novella: letture del "Decameron"*. Longo.

Priest, G. (2002). Paraconsistent logic. In *Handbook of philosophical logic* (pp. 287–393). Springer.

Pustejovsky, J., Bunt, H. & Zaenen, A. (2017). Designing annotation schemes: From theory to model. *Handbook of Linguistic Annotation*, 21–72.

Rescher, N. (1968). *Topics in philosophical logic*. Springer Science & Business Media.

Sanfilippo, E.M. (2021). Ontologies for information entities: State of the art and open challenges. *Applied ontology*, *16*(2), 111–135.

Sanfilippo, E.M. & Ferrario, R. (2024). D3.1 - Observations Modeling: State of the Art. Project Report, Project MITE – Make it explicit: Documenting interpretations of literary fictions with conceptual formal. Funded by European Union – Next Generation EU. https://www.loa.istc.cnr.it/mite/wp-content/uploads/2024/09/MITE_D3_1.pdf.

Sanfilippo, E.M., Sotgiu, A., Tomazzoli, G., Masolo, C., Porello, D. & Ferrario, R. (2023). Ontological Modeling of Scholarly Statements: A Case Study in Literary Criticism. In *Formal Ontology in Information Systems (FOIS 2023)*. Frontiers in Artificial Intelligence and Applications (Vol. 377, pp. 349–363). IOS Press.

Sanfilippo, E.M., Masolo, C., Ferrario, R. & Bottazzi, E. (2024a). Interpreting Texts and Their Characters. In *Formal Ontology in Information Systems (FOIS 2024)*. Frontiers in Artificial Intelligence and Applications (Vol. 394, pp. 119–133). IOS Press.

Sanfilippo, E.M., Masolo, C., Mosca, A., Tomazzoli, G., et al. (2024b). Operationalizing Scholarly Observations in OWL. In *Semantic Web and Ontology Design for Cultural Heritage 2024. Proceedings of the fourth edition of the International Workshop on Semantic Web and Ontology Design for Cultural Heritage* (Vol. 3809, pp. 1–12). CEUR.

Sartini, B., Baroncini, S., van Erp, M., Tomasi, F. & Gangemi, A. (2023). Icon: An ontology for comprehensive artistic interpretations. *ACM Journal on Computing and Cultural Heritage*, *16*(3), 1–38.

Schöch, C., Hinzmann, M., Röttgermann, J., Dietz, K. & Klee, A. (2022). Smart Modelling for Literary History. *International Journal of Humanities and Arts Computing*, *16*(1), 78–93.

Schwitzgebel, E. (2024). Belief. In E.N. Zalta and U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Spring 2024 ed.). Metaphysics Research Lab, Stanford University.

Searle, J. (1985). *Foundations of illocutionary logic*. Cambridge University Press.

Tomazzoli, G. (2025). La filoginia è nell'occhio di chi guarda? Verso un modello formale per l'argomentazione nella critica letteraria. In *Raccontare le donne (eds.) G. Paradisi, and G. Santini*. Spolia (forthcoming).

Trafford, J. (2014). An inferentialist approach to paraconsistency. *abstracta*, *8*(1).

Van Ditmarsch, H., van Der Hoek, W. & Kooi, B. (2007). *Dynamic epistemic logic* (Vol. 337). Springer Science & Business Media.

Williamson, T. (2002). *Knowledge and its Limits*. Oxford University Press, USA.

Willis, I. (2017). *Reception*. Routledge.

Wittgenstein, L. (2009). *Philosophical investigations*. London: Blackwell.