Formal Ontology in Information Systems
R. Ferrario and W. Kuhn (Eds.)

doi:10.3233/978-1-61499-660-6-197

Observations and Their Explanations

Claudio MASOLO
Laboratory for Applied Ontology, ISTC-CNR, Trento, Italy

Abstract. We introduce a first-oder theory where observations are reified into the domain of quantification. Observations have an epistemological nature, they describe how the world appears, not as the world is. Our primitive notions allow to represent how some observations are explained in terms of more simple ones or how they are aggregated into macro-indexes. We analyze in detail the cases of measurement and testing where observations are collected through calibrated devices and eventually aggregated into scores. Our framework is based on a decoupling between the observations and the propositions that belong to the temporally qualified A-box. It allows contradictory observations, but it requires these disagreements to be resolved via a merging process that identifies, among the contradictory observations, the most plausible one that can then be safely transferred into the A-box.

Keywords. Ontology of Observations, Measurement, Testing, Plausibility, Truth

1. Introduction

Conceptual modeling and knowledge representation mainly focus on characterizing how, according to some experts, a given domain is structured, i.e., to identify a set of concepts and relations together with the constraints that hold for this domain, in short, a model, a T-box. In this context, *data* are usually reduced to factual instantiations of the model, an A-box. Data sharing can then be achieved by standardizing, integrating, or aligning the involved models. The subjective or epistemological dimension is confined to the conceptualization of the domain since different models of the same domain are possible.

Recently an enormous volume of data collected by heterogenous sensors or resulting from complex analyses is made available on the web. The homogeneity of the data taken into account and the understanding of their provenance critically impact the quality, reliability, validity, and trustworthiness of the analyses performed on these data. This is especially relevant for the e-science community and, in particular, for large-scale and distributed collaborative science where data-exchange is sometimes directed to ensure the reproducibility of scientific analyses and experiments. This leads to the need of explicitly representing the nature of data, the way they have been acquired, produced, modified, etc. The (sharing of the) model of the domain is not enough, one needs (to share) a model of the data. Measurements, observations, and analyses have a subjective nature that transcends the conceptual apparatus necessary to express the snapshots of the domain, a step towards an operationalist or constructivist stance about data. Calibration and measurement procedures, instruments, and standards have been introduced to smooth this difficulty into an inter-subjective, mediated, and controlled access to the 'external world'. Economics, medicine, biology, psychology, sociology, but also physics and cognitive science, are deeply founded on data analysis and testing. A double subjectivity is present here: (*i*) the acquisition of raw data and (*ii*) their (often complex) transformation or aggregation into macro indexes or scores. Furthermore, statistical analysis invests the scores with a comparative nature: scores result from contrasting the raw data collected by testing a subject with (the distribution of) the raw data collected by testing the subjects that belong to a sample selected as representative of the assessed attribute.

The Semantic Web, Applied Ontology, Conceptual Modeling, and DB communities, started to pay attention to the nature and the provenance of data only quite recently with the intent to support the sharing and integration of data, to enable interoperability for sensors and sensing systems, and to produce detailed descriptions of scientific investigations. The approaches focused on provenance tend to introduce information about the life-cycle of data by means of annotations(-graphs). In this context, the Open Provenance Model[1] and the W3C PROV Data Model[2] result from standardization efforts aimed to establish a reference provenance model. However, as recognized by part of the DB community, the separation between the data- and the domain-model prevents a uniform approach where provenance-data are intrinsic to the schema rather than an external annotation. The approaches devoted to a conceptual analysis tend to extend foundational ontologies with notions able to characterize the semantics of data and provenance. Ontologies of observations and measurements mainly developed in the context of Geographical Information Systems [1,2] explicitly refer to observations and observation processes. A similar methodology guided the W3C Semantic Sensor Network Incubator group[3] in developing an OWL-2 ontology for describing sensors in terms of measurement processes, observations and deployments [3]. The focus here is on the nature of raw data. The DataTop ontology (based on [4]) and the Ontology for Biomedical Investigations (see [5]) address the need for the description of biological and clinical investigations, i.e., they are also concerned with the way raw data are elaborated.

Following this last kind of approaches, we introduce a first-order theory where observations are reified into the domain of quantification and provided with a precise identity criterion. In particular, in Section 2, we modify the framework presented in [6] by shifting from an ontological perspective that considers *states* as *truth-makers* of propositions to an epistemological one where *observations* support propositions: observations describe how the world *appears* (at a time) while states describe how it *is* (at a time).

As usual in the practice of ontological analysis, [1,2,3] model the provenance of observations by extending the foundational ontologies with primitive relations able to link the observations to the sensors that collected them, to the used procedures, etc. Here we follow a different approach: we (partially) capture the provenance of observations by taking into account how they are explained or justified in terms of more simple observations. This move allows us to also represent some weak forms of elaboration of data. The primitive of *explanation* is formally analyzed in Section 3 while Section 4 focuses on two specific kinds of explanation: *aggregation* and *evaluation*. In the first case, a new observation about x is obtained by aggregating a series of observations all concerning x with the goal of making explicit meaningful and cognitively effective information. In the second case, a new observation about x is explained in terms of the configuration of a

[1] http://openprovenance.org/
[2] http://www.w3.org/TR/prov-overview/
[3] http://www.w3.org/2005/Incubator/ssn/ssnx/ssn

different object, a *mediator*, connected to x. In Section 5, (weak) measurement is seen as a particular evaluative process where the mediators are calibrated measurement devices.

Furthermore, we do not assume that the explanations (and the explained observations) are always exact; the explanation relation represents just the fact that an observation has been justified in a given way. Section 5 shows that, in a scientific scenario, some incorrect explanations can be detected on the basis of calibration processes (theoretical laws). These mistakes correspond to evident misuses of devices (erroneous applications of laws). However, devices can be used in bad environmental conditions or they can malfunction even after calibration. In these cases, one can obtain contradictory observations—e.g., different synchronic observations about the weight of a person—and the wrong ones cannot be ruled out by using the previous technique. Despite that, by being individuals, contradictory observations do not necessarily generate a logical inconsistency. In addition, the 'world' of the (temporally qualified) propositions in the A-box—the factual instantiation of the temporally qualified part of the model—does not need to be necessarily *aligned* with the 'world' of observations. Section 6 sketches a strategy to resolve disagreements among measurements on the basis of some information about the devices. The adopted solution consists in individuating the *most plausible* (according to the additional information considered) measurement among the contradictory ones. The (data) consistency can then be preserved by importing in the A-box only the plausible measurements, i.e., by hiding the non-plausible ones from the A-box. This mechanism can be also used to hide from the A-box meaningless or cognitively irrelevant data used to compute *macro-indexes*, e.g., in the case of *testing*, the raw data clustered into *scores*.

2. Observations

We consider three disjoint basic categories: *time* (TME), *object* (OBJ), and *simple observation*, (SOB).[4] *Time* is linear and discrete but we leave open if TME-instances, called *times*, are punctual or extended atomic entities. *Objects*—also called *endurants* or *continuants*—are individuals that are wholly present at every time they exist, e.g., tables, persons, companies, bits of stuff. A *simple observation* corresponds to the *classification* under, the *attribution* of, a (unary) concept to (one) several objects during one or several times.[5] Both objects and simple observations are in time, they exist at least at one time (a1)-(a2), where $\varepsilon_t x$ stands for "*x exists* at time *t*".[6] The existence of an observation at a time means that the classification has been done considering how the objects are *at that time*, i.e., an observation that exists at t partially describes a t-snapshot of the world. The time t neither concerns when the classification has been done (even though the two times can coincide), nor whether or when a material support for the observation exists. Observations may also *persist* through time, objects may be classified under a concept for several times, i.e., relatively to a conceptual dimension, the objects under analysis may not change for a whole period.

The main components of the SOB-instances are (*i*) the objects and (*ii*) the concepts under which they are classified. The atemporal primitives $\multimap_{\mathtt{i}}$ hold between objects and

[4] We introduce *complex* observations at the end of the section.

[5] The framework can be easily extended to the classification of *events*. However, it is also possible to build events from observations as done in [6].

[6] We write $\mathrm{P}_t x$ instead of $\mathrm{P}xt$ to highlight the time-argument.

simple observations (a3) and identify *the* ith object involved in the observation, the ith *participant* (a4). We indicate with α the maximal arity of the considered concepts that corresponds to the number of the $\multimap_{\mathtt{i}}$ primitives necessary to distinguish the position of the objects in the concept. (d1)[7] defines n-ary participation while *general participation* abstracts from the position (d2).

Simple observations are organized according to a *finite* set $\bar{\mathcal{P}}$ of unary predicates (a5). Intuitively, each observation-*kind* in $\bar{\mathcal{P}}$ collects all the classifications under one concept. For instance, the observation of Tim being enrolled in the University of Trento now may be represented by $(\mathtt{tim},\mathtt{unitn}) \multimap s \wedge \bar{\mathtt{ENROLL}}s \wedge \varepsilon_{\mathtt{now}}s$; the detection of a change in an object requires at least two observations, e.g., $\mathtt{tim} \multimap s \wedge \mathtt{tim} \multimap r \wedge \bar{\mathtt{80KG}}s \wedge \bar{\mathtt{82KG}}r \wedge \varepsilon_t s \wedge \varepsilon_{t'} r \wedge t \neq t'$; an object can be the subject of several synchronous observations, e.g., $\mathtt{tim} \multimap s \wedge \mathtt{tim} \multimap r \wedge \bar{\mathtt{80KG}}s \wedge \bar{\mathtt{180CM}}r \wedge \varepsilon_t s \wedge \varepsilon_t r$. The $\multimap_{\mathtt{i}}$ primitives and the $\bar{\mathcal{P}}$-predicates are not temporally qualified, i.e., both the participants and the concepts are constant (essential) components of observations (see (a6) and (d3)), i.e., during their existence, observations cannot vary their participants or migrate from a kind to a different one.

d1 $x^n \multimap s \triangleq \bigwedge_{1 \leq \mathtt{i} \leq n}(x_{\mathtt{i}} \multimap_{\mathtt{i}} s) \wedge \bigwedge_{n < \mathtt{i} \leq \alpha} \neg\exists x(x \multimap_{\mathtt{i}} s)$ (*n-ary participation*)
d2 $x \multimap\!\!\diamond\, s \triangleq \bigvee_{1 \leq \mathtt{i} \leq \alpha} x \multimap_{\mathtt{i}} s$ (*general participation*)
d3 $x \circledS_{\varepsilon} y \triangleq \forall t(\varepsilon_t x \rightarrow \varepsilon_t y)$ (*temporal inclusion*)
a1 $\varepsilon_t x \rightarrow \mathrm{TME}t \wedge (\mathrm{OBJ}x \vee \mathrm{SOB}x)$
a2 $(\mathrm{OBJ}x \vee \mathrm{SOB}x) \rightarrow \exists t(\varepsilon_t x)$
a3 $x \multimap_{\mathtt{i}} s \rightarrow \mathrm{OBJ}x \wedge \mathrm{SOB}s$
a4 $x \multimap_{\mathtt{i}} s \wedge y \multimap_{\mathtt{i}} s \rightarrow x = y$
a5 $\mathrm{SOB}x \rightarrow \bigvee_{\bar{\mathtt{P}} \in \bar{\mathcal{P}}}(\bar{\mathtt{P}}x)$
a6 $x \multimap_{\mathtt{i}} s \rightarrow s \circledS_{\varepsilon} x$
a7 $\bar{\mathtt{P}}s \rightarrow \exists x^n(x^n \multimap s)$

Up to now, the formal properties of our observations are quite close to the ones considered for states in [6]. Things start to become different when one considers the taxonomical structure of the $\bar{\mathcal{P}}$-predicates that is specified via subsumption relations: $\mathsf{SUB}(\bar{\mathtt{P}},\bar{\mathtt{Q}})$, with $\bar{\mathtt{P}} \neq \bar{\mathtt{Q}}$, stands for "$\bar{\mathtt{P}}$ is *directly* (without intermediate subsumption-steps) subsumed by $\bar{\mathtt{Q}}$". We assume $(\bar{\mathcal{P}},\mathsf{SUB})$ to be a directed, connected, and *acyclic* graph where the nodes are predicates in $\bar{\mathcal{P}}$ and the arcs SUB-relations. Any $\mathsf{SUB}(\bar{\mathtt{P}},\bar{\mathtt{Q}})$ statement corresponds to a $\forall s(\bar{\mathtt{P}}s \rightarrow \bar{\mathtt{Q}}s)$ axiom, therefore the acyclic condition rules out necessarily co-extensive predicates from $\bar{\mathcal{P}}$. One can then define the following subsets of $\bar{\mathcal{P}}$ (where SUB^+ is the transitive closure of SUB):[8]

- $\bar{\mathcal{P}}_L = \{\bar{\mathtt{P}} \in \bar{\mathcal{P}} \mid$ there are no $\bar{\mathtt{Q}} \in \bar{\mathcal{P}}$ such that $\mathsf{SUB}(\bar{\mathtt{Q}},\bar{\mathtt{P}})\}$ (the leaves of $\bar{\mathcal{P}}$)
- $\mathsf{DSC}(\bar{\mathtt{P}}) = \{\bar{\mathtt{Q}} \in \bar{\mathcal{P}} \mid \mathsf{SUB}^+(\bar{\mathtt{Q}},\bar{\mathtt{P}})\}$ (the descendants of $\bar{\mathtt{P}}$)

Figure 1 depicts an example of taxonomy where SUB is represented by a vertical line with the bottom predicate subsumed by the top one.[9] In this example, $\bar{\mathcal{P}}_L = \{\bar{\mathtt{ROUND}}, \bar{\mathtt{SQUARE}}, \bar{\mathtt{GLUED}}, \bar{\mathtt{CRIMSON}}, \bar{\mathtt{OLIVE}}, \bar{\mathtt{EMERALD}}\}$, and $\mathsf{DSC}(\bar{\mathtt{COLORED}}) = \{\bar{\mathtt{RED}}, \bar{\mathtt{CRIMSON}}, \bar{\mathtt{YELLOW}}, \bar{\mathtt{OLIVE}}, \bar{\mathtt{GREEN}}, \bar{\mathtt{EMERALD}}\}$.

---

[7] x^n is a shortcut for $x_1,\ldots,x_n$.
[8] We started from SUB (rather than SUB^+) to minimize the constraints to be introduced by the user.
[9] Note that $\bar{\mathtt{PHYSICAL}}s \rightarrow \mathtt{SOB}s$ holds but this implication is captured by including $\bar{\mathtt{PHYSICAL}}$ in $\bar{\mathcal{P}}$ and not via an explicit SUB relation. For this reason, we used a different representation in the Figure 1.

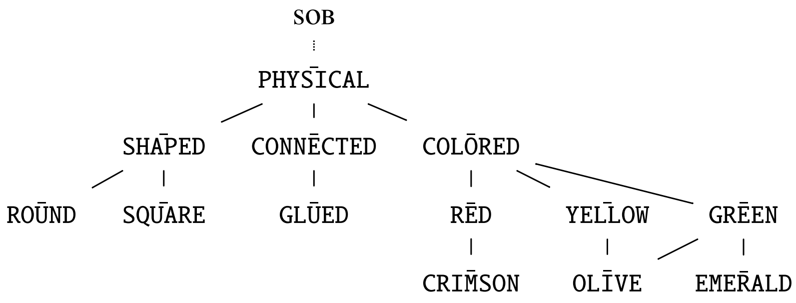

Figure 1. Example of SUB-structuring of the predicates in $\bar{\mathcal{P}}$.

(a7)—to be added for each predicate in $\bar{\mathcal{P}}_L$—guarantees that all the instances of a leaf have the same 'arity', the same number of participants.[10] However, this constraint does not hold in general. For instance, the $\overline{\mathtt{PHYSICAL}}$-instances usually have different arities, e.g., $\overline{\mathtt{RED}}$ is unary but $\overline{\mathtt{GLUED}}$ is binary.

Intuitively, given the participants and the time, at most one observation of a specific kind exists, see (f1) where temporal overlap is defined in (d4). However, (f1) seems too strong for the following reason. Consider the taxonomy in Figure 1 and assume $a \multimap s \wedge a \multimap s' \wedge \overline{\mathtt{ROUND}}s \wedge \overline{\mathtt{OLIVE}}s' \wedge s \otimes_{\varepsilon} s'$. The observations s and s' have the same participants, different leaf-kinds, but a common father, namely $\overline{\mathtt{PHYSICAL}}$. Because of that, (f1) would imply $s = s'$ even though s and s' regard two different aspects of a.

d4 $x \otimes_{\varepsilon} y \triangleq \exists t(\varepsilon_t x \wedge \varepsilon_t y)$ (*temporal overlap*)
f1 $\bigvee_{\bar{\mathtt{P}} \in \bar{\mathcal{P}}} (\bar{\mathtt{P}}s \wedge \bar{\mathtt{P}}s') \wedge \bigwedge_{1 \leq \mathtt{i} \leq \alpha} \forall x (x \multimap_{\mathtt{i}} s \leftrightarrow x \multimap_{\mathtt{i}} s') \wedge s \otimes_{\varepsilon} s' \rightarrow s = s'$

One then needs to better qualify which $\bar{\mathcal{P}}$-predicates provide the identity criterion for observations. One possibility is to substitute $\bar{\mathcal{P}}$ with $\bar{\mathcal{P}}_L$ in (f1). However, in this case, the modified (f1) would not apply to the observations that are not instances of leaf-kinds—e.g., in Figure 1, $\overline{\mathtt{RED}}$-observations that are not $\overline{\mathtt{CRIMSON}}$-observations—i.e., it would not be a true identity criterion for all the SOB-instances. The refinement of (a5) with $\textsc{sob}s \rightarrow \bigvee_{\bar{\mathtt{P}} \in \bar{\mathcal{P}}_L} (\bar{\mathtt{P}}s)$ would assure the modified (f1) to be a true identity criterion but it would imply that in Figure 1, for instance, $\overline{\mathtt{RED}}$ is equivalent to $\overline{\mathtt{CRIMSON}}$ and $\overline{\mathtt{GREEN}}$ to the disjunction of $\overline{\mathtt{OLIVE}}$ and $\overline{\mathtt{EMERALD}}$, i.e., 'being green' becomes only an abstraction from 'being olive' and 'being emerald'. All observations would then be maximally specified, i.e., it would not be possible, for instance, to observe a green object without observing its exact shade. In an applicative or epistemological perspective, this assumption is manifestly too strong. For this reason, we prefer to assume a weaker constraint. To introduce it, first we identify the set of the *minimal* kinds of a state s:

– $\mathsf{min}(s) = \{\bar{\mathtt{P}} \in \bar{\mathcal{P}} \mid \bar{\mathtt{P}}s$ and there are no $\bar{\mathtt{Q}} \in \mathsf{DSC}(\bar{\mathtt{P}})$ such that $\bar{\mathtt{Q}}s\}$ (*minimal kind*)

what cardinality can be greater than 1. E.g., assume to delete $\overline{\mathtt{OLIVE}}$ from the taxonomy in Figure 1, and consider an observation s (an olive observation) that is an instance of both $\overline{\mathtt{YELLOW}}$ and $\overline{\mathtt{GREEN}}$ but not of $\overline{\mathtt{EMERALD}}$. In this case $\mathsf{min}(s) = \{\overline{\mathtt{YELLOW}}, \overline{\mathtt{GREEN}}\}$.[11]

[10] Stronger constraints may be added to characterize the kind of objects admitted to play a given role in the relation. These constraints rely on the nature of the specific predicates $\bar{\mathcal{P}}$.

[11] Note that if $\mathsf{min}(s)$ contains different leafs then they have the same arity.

Now we can introduce the identity criterion (a8). The idea is that the *resolution* has an impact on the identity of observations. For instance, (a8) does not apply to a minimal $\mathtt{R\bar{E}D}$-observation s and a minimal $\mathtt{CRI\bar{M}SON}$-observation s' because $\mathsf{min}(s) \cap \mathsf{min}(s') = \emptyset$. The difference resides in the fact that s is less precise than s', i.e., these two observations refer to different levels of resolution, they capture different epistemological situations.

a8 $\bigvee_{\bar{\mathtt{P}} \in (\mathsf{min}(s) \cap \mathsf{min}(s'))} (\bar{\mathtt{P}} s \wedge \bar{\mathtt{P}} s') \wedge (\bigwedge_{1 \leq \mathtt{i} \leq \alpha} \forall x (x \multimap_{\mathtt{i}} s \leftrightarrow x \multimap_{\mathtt{i}} s')) \wedge s \otimes_{\varepsilon} s' \rightarrow s = s'$
(*identity criterion for observations*)

By the fact that the observations are covered by the predicates in $\bar{\mathcal{P}}$, axiom (a5), and that the SUB-graph is acyclic, it follows that for all observation s, $\mathsf{min}(s) \neq \emptyset$. Thus, (a8) applies to all the observations. To refer to the observation s that exists at time t, has participants $x_1, \ldots, x_n$, and such that $\bar{\mathtt{P}} \in \mathsf{min}(s)$, we will use the shortcut $\mathbf{p}_t(x^n)$. An observation with several minimal kinds can then be described in different ways. E.g., suppose $\mathsf{min}(s) = \{\mathtt{YEL\bar{L}OW}, \mathtt{GR\bar{E}EN}\}$, one may have $s = \mathbf{yellow}_t(x) = \mathbf{green}_t(x)$. Analogously, in the case of persistent observations, one may have $\mathbf{yellow}_t(x) = \mathbf{yellow}_{t'}(x)$ and, putting the two cases together, $\mathbf{yellow}_t(x) = \mathbf{green}_{t'}(x)$. We denote with $\mathbf{P}$ the set of the *description-functions* $\mathbf{p}$ that correspond to all the predicates in $\bar{\mathcal{P}}$.

From the above constraints one can conclude that observations are not propositions. Observations are in time and their identity depends on time. In addition, despite their epistemological nature, observations are not *private* (mental entities); the meaning an observation conveys can be shared by, and communicated across, different subjects.[12]

Complex observations. The category of complex observations (OBS), simply called observations, is the closure of simple observations under mereological sum. We consider a parthood relation on OBS—$x \sqsubseteq y$ stands for "x is part of y"—that satisfies the axioms for a classical atomic extensional mereology closed under the mereological sum (the sum of $x_1, \ldots, x_n$ is noted $x_1 + \ldots + x_n$), see [7,6] for the technical details. (a9) enforces simple observations and mereological atoms—see (d5)—to coincide, thus observations are uniquely decomposable into simple observations. Complex observations correspond to conjunctions of classifications under simple concepts not to classifications under complex concepts. For instance, one could distinguish $\mathbf{p}_t(a) + \mathbf{q}_t(a)$—a conjunction of classifications—from $[\mathbf{p} \wedge \mathbf{q}]_t(a)$—a classification under a conjunction of concepts.

With a slight abuse of notation, (d7) extends the existence ε to complex observations—where the atomic part is defined in (d6). An observation x is *completely existent* at t, $\hat{\varepsilon}_t x$, if all its (atomic) parts exist at t (d8). At a time t, the participants in an observation are the participants in its atomic parts that exist at t (d9).

d5 $\Lambda x \triangleq \mathrm{OBS} x \wedge \neg \exists y (y \sqsubseteq x \wedge x \neq y)$ (*atom*)

d6 $x \,{}_{\Lambda}{\sqsubseteq}\, y \triangleq \Lambda x \wedge x \sqsubseteq y$ (*atomic part*)

d7 $\varepsilon_t x \triangleq \exists s (s \,{}_{\Lambda}{\sqsubseteq}\, x \wedge \varepsilon_t s)$ (*existence of observations*)

d8 $\hat{\varepsilon}_t x \triangleq \forall s (s \,{}_{\Lambda}{\sqsubseteq}\, x \rightarrow \varepsilon_t s)$ (*complete existence*)

d9 $x \multimap_t y \triangleq \exists s (s \,{}_{\Lambda}{\sqsubseteq}\, y \wedge x \multimap s \wedge \varepsilon_t s)$ (*general participation*)

a9 $\Lambda x \leftrightarrow \mathrm{SOB} x$

[12] [2] distinguishes private *qualia* from *measurements* on the basis of their communicability through semantic reference spaces and signs. In our framework, this distinction reduces to the way observations are explained, see Section 4.

3. Direct Explanation

Explanation is a cognitive process that induces a *simple* observation to emerge from a (possibly) complex one: $x \lessdot_t s$ stands for "at time t, the simple observation s is *directly explained* by the complex observation x" (a10). We consider here only *direct* explanations, i.e., explanations with no intermediate steps (a11), the basic blocks for building explanation-chains. Explanation can be seen as a (temporally qualified) *specific* dependence. In particular, here we focus on a *synchronic* dependence, i.e., both the *explanans* and the *explanandum* must *completely* exist when the explanation holds (a12). The interesting case of explanations that contemplate temporal patterns of diachronic observations, forms of historical dependence, is ruled out at this stage.

a10 $x \lessdot_t y \rightarrow \text{OBS}x \wedge \text{SOB}y \wedge \text{TME}t$

a11 $x \lessdot_t y \rightarrow \neg\exists z(x \lessdot_t z \wedge z \lessdot_t y)$

a12 $x \lessdot_t y \rightarrow \hat{\varepsilon}_t x \wedge \varepsilon_t y$

Observations represent both direct classifications or sensations, and classifications resulting from high-level cognitive processes like reasoning, measuring, reporting, etc. These processes rely on simple observations to build explicit, concise, meaningful, and cognitively effective classifications. For instance, in *data analysis*, one starts from *raw data* to build macro-indexes that can be further elaborated to generate more complex indexes. In metrology, explanations can be used to represent the way data are collected, their origin or provenance, the involved devices or observers (see Section 5). Logically, $x \lessdot_t s$ can be seen as a form of inference, a dependence of the information contained in s on the one in x. However, our explanations are not necessarily truth-preserving, they just allow to take track of the justifications advocated for the explanandum. This does imply neither the truth of the explanans nor the validity of the explanation, e.g., one can encounter faulty reasonings. The check of the *validity* of an explanation requires additional information (see Section 5 for some simple cases).[13]

f2 $x \lessdot_t y \wedge \hat{\varepsilon}_{t'} y \rightarrow x \lessdot_{t'} y$

f3 $x \lessdot_t y \wedge \hat{\varepsilon}_{t'} x \rightarrow x \lessdot_{t'} y$

a13 $x \lessdot_t y \wedge z \multimap y \rightarrow z \multimap x$

The temporal qualification of explanation is necessary because observations can persist through time, therefore, at different times, they could have different explanations, i.e., (f2) does not hold in our framework. For instance, the fact that an object weights 1kg during a period of time can be explained by several measurements collected by different instruments at different times. Second, even at a single time, an observation can have alternative explanations, it can be supported by multiple justifications. For instance, it is possible to have both $\mathbf{r}_t(x,d)+\mathbf{p}_t(d) \lessdot_t \mathbf{9.109\times10^{-31}kg}_t(x)$ and $\mathbf{1.602\times10^{-19}coulomb}_t(x) \lessdot_t \mathbf{9.109\times10^{-31}kg}_t(x)$, i.e., the mass of an electron x has been determined in two alternatives ways: by observing a specific configuration of the scale d when connected to x (the first explanation) and by considering a physical law (the second explanation). Third, there are observations, called *primitive* observations, that lack an explicit explanation. Primitive observations supply a starting point to explanation-chains. Intuitively, they represent phenomenological conscious sensations or, in measurement

[13] We do not consider the source of an explanation, a possible interesting extension of our framework.

processes, the simple readings of the outputs of the technical devices by the operators.[14] Fourth, one is tempted to assume that if $x \prec_t y$ holds then, at any time x completely exists, x continues to explain y (f3). We do not commit to this view because explanations have to be intended as *explicit* statements resulting from underlying cognitive processes. The explanation $x \prec_t y$ represents an explicit commitment to the justification, at t, of y by means of x. Both x and y could exist at t' without any explanatory commitment between then. This is a central aspect of our framework. Fifth, intuitively, to justify an observation y about the object o, the explanans x should concern o, i.e., the participants in y should participate also in x. Even though it is maybe possible to conceive explanantia that do not explicitly refer to the objects involved in the explananda—e.g., in the case of conventional signals—these implicit justifications have a very weak explanatory power. For this reason we prefer to avoid them by embracing (a13).

4. Evaluation and Aggregation

We focus here on two special kinds of explanation (of observations with one participant): *aggregation* and *evaluation* that have the form in (f4) and (f5), respectively.

f4 $(\mathbf{p1}_t(x) + \ldots + \mathbf{pn}_t(x)) \prec_t \mathbf{q}_t(x)$ (*aggregation*)

f5 $(\mathbf{r}_t(x,m) + \mathbf{p}_t(m)) \prec_t \mathbf{q}_t(x)$ (*evaluation*)

In (f4), the explanandum $\mathbf{q}_t(x)$ is justified in terms of several observations all about x. It reveals, in a concise way, a 'conjunction' of, a relevant pattern of, classifications of x. Aggregations may be used to represent theoretical laws or abstractions—e.g., the density of x is explained in terms of its weight and volume.

In (f5), $\mathbf{q}_t(x)$ is explained in terms of (*i*) a relational observation concerning both x and m and (*ii*) a classification of m. The observation $\mathbf{p}_t(m)$ is a sort of *proxy* for $\mathbf{q}_t(x)$, i.e., m is a *mediator* able to transduce, by interacting with x, a property of x into a property of m. By connecting the object x to the mediator m in a qualified way, some observations about x can be indirectly obtained by observing m. Evaluations may be used to represent *measurements*, e.g., the weight of x is explained in terms of the position of the pointer of a scale m where x has been put. For evaluations, the source relation σ can be defined as in (d10): $\sigma_t^m s$ stands for "at time t, m is a *source* or *origin* of the observation s". The source m coincides with the mediator, *what* or *who* interacted with the participant of s during the measurement process. Note that some observations may have several sources that can be further characterized by information about its kind, its reliability, etc., see [3].

d10 $\sigma_t^m s \triangleq \bigvee_{\mathbf{r},\mathbf{p} \in \mathbf{P}} (\exists x s_1 s_2 (x \multimap s \wedge s_1 = \mathbf{r}_t(x,m) \wedge s_2 = \mathbf{p}_t(m) \wedge s_1 + s_2 \prec_t s))$ (*source*)

(f6) shows as n evaluations can be aggregated into a unique classification. The composition of evaluations with aggregations can then capture quite complex classificatory phenomena. In addition, notice that the observations that are part of the explanans in (f4),

[14]Note that some observations, e.g., $\mathbf{red}_t(x)$, could be the result of both a direct (phenomenological) perception and a measurement process. To represent this double nature one could introduce the new kind of observations $\mathtt{EX\bar{I}ST}$. The identification and tracking of objects is one of the main problem in (cognitive) science. $\mathtt{EX\bar{I}ST}$-observations could then be intended as non-conceptual identifications of objects, i.e., observations that cannot be conceptually explained, see [8]. In this way, the double nature of $\mathbf{red}_t(x)$ is captured by $\mathbf{exist}_t(x) \prec_t \mathbf{red}_t(x)$ and $\mathbf{r}_t(x,d) + \mathbf{p}_t(d) \prec_t \mathbf{red}_t(x)$, where d is a device. We leave open whether the propositions $\varepsilon_t x$ correspond to the observations $\mathbf{exist}_t(x)$.

(f5), and (f6) are not necessarily primitive. For instance, $\mathbf{p}_t(m)$, $\mathbf{pi}_t(x)$, or $\mathbf{ri}_t(x_i,m)$ could be, in their turn, the result of an evaluation, aggregation, or both.

f6 $(\mathbf{r1}_t(x,m_1)+\mathbf{p1}_t(m_1)) \prec_t \mathbf{q1}_t(x) \wedge \ldots \wedge (\mathbf{rn}_t(x,m_n)+\mathbf{pn}_t(m_n)) \prec_t \mathbf{qn}_t(x) \wedge (\mathbf{q1}_t(x)+\ldots+\mathbf{qn}_t(x)) \prec_t \mathbf{s}_t(x)$

We conclude this section with two remarks. First, as noted before, explanations can be incorrect, e.g., $(\mathbf{3g}_t(x)+\mathbf{1m}^{\mathbf{3}}_t(x)) \prec_t \mathbf{1g/m}^{\mathbf{3}}_t(x)$. To filter out incorrect explanations additional information is required, e.g., in the example, the definition of density (see Section 5). Second, (f4) and (f5) could be interpreted in an ontological way. For instance, (f4) can be specialized as: $(\mathbf{p}_t(x)+\mathbf{p}_t(x)) \prec_t \mathbf{q}_t(x)$ that, from the axioms governing the mereological sum, is equivalent to $\mathbf{p}_t(x) \prec_t \mathbf{q}_t(x)$. At least two interpretations are possible:[15] (*i*) we are in presence of an ontological correlation, e.g., all the objects with mass 9.109×10^{-31}kg have an electric charge of 1.602×10^{-19} coulomb, or (*ii*) we are in presence of a generalization, e.g., all the scarlet objects are red. Furthermore, (f5) can be specialized as: $(\mathbf{inheres}_t(x,m)+\mathbf{p}_t(m)) \prec_t \mathbf{p}_t(x)$. Here, the mediator m is an *individual quality* or *trope* [9] that inheres in x. The classification of x depends on the classification of its trope, e.g., x appears red because of its red-trope.

5. Weak Measurement

We analyze some constraints to rule out explanations flawed by *material* errors. In the case of aggregations, scientific laws can be captured by introducing constraints with the form (f7) (by (a12) s exists at t and by (a13) x is the only participant in s). In the case of *functional* laws we have a unique $\bar{\mathtt{Q}}\mathtt{i}$, e.g., $(\mathbf{3g}_t(x)+\mathbf{1m}^{\mathbf{3}}_t(x)) \prec_t s \rightarrow \mathtt{3G\bar{R}/M^3}s$. Admittedly, these constraints are quite unpractical—contrast with mathematical equations—but it is still possible to rely on them especially in the qualitative cases.

f7 $(\mathbf{p1}_t(x)+\cdots+\mathbf{pn}_t(x)) \prec_t s \rightarrow \bar{\mathtt{Q}}\mathtt{1}s \vee \ldots \vee \bar{\mathtt{Q}}\mathtt{M}s$

More interesting is the case of *qualitative* evaluations, a perspective explicitly addressed by the theory of *weak measurement* introduced by Finkelstein [10] and further elaborated by Mari [11]. In this weak perspective, measurement does not necessarily involve *quantities* (interval or ratio properties) but also *qualities* (nominal or ordinal properties). Qualitative classification plays a fundamental role in disciplines like psychology, medicine, or sociology where non-physical properties can be attributed to the subjects via the administration of *tests*. Measurement becomes "uncorrelated with quantification: the measurability of a property is a feature derived from experiment, not algebraic constraints" [11, p.2894]. The basic formula of the quality calculus is $p = \{p\}\,\mathbf{in}\,[p]$, where p is a property of an object (e.g., the color or the shape of an object), $[p]$ is a *classification system*, a system of properties all related to the same quality (e.g., the color-properties), and $\{p\}$ is an element of $[p]$, it individuates the 'position' of the object under measurement in the system $[p]$ (e.g., scarlet for colors), see [12] for more details.[16] Weak mea-

[15] From an ontological perspective, the explanations with form $\mathbf{exist}_t(x) \prec_t \mathbf{q}_t(x)$ introduced in the previous footnote could represent the fact that $\mathbf{q}_t(x)$ corresponds to an essential property of x, i.e., the mere existence of x explains its having a given property.

[16] Usually, the classification systems are structured. The *domains* of *conceptual spaces* [13] and the *quality spaces* of DOLCE-CORE [14] can then be seen as classification systems.

surement (as well as standard measurement) commits to *individual properties* or *tropes* of objects. The value $\{p\}$ is attributed to an individual property of the object under measurement, e.g., it is the *color of the object* that is scarlet. The object is (indirectly) scarlet just because its color is scarlet (see the end of Section 4). However, as shown in [15], by explicitly taking into account the measurement devices, one can avoid this commitment and consider measurement as a (partial) mapping from objects to properties in the system $[p]$. Given the fact that the evaluation form (f5) explicitly refers to the mediators, here we embrace this less committed approach. In particular, the *interaction* function (see [15,16]) that maps the objects into the states of the device is captured by the explanans in (f5) that, in addition, explicitly represents also the way the object under measurement is linked to the device.

According to Mari [11], the distinction between *measurement* and *evaluation* is based on the *objectivity* and *inter-subjectivity* of the results. The measurement results must regard (as much as possible) a property the object under measurement has independently from (*i*) its other properties, (*ii*) the measuring device, and (*iii*) the environment;[17] and they need to be *sharable* by different subjects at different times and in different places. To achieve the required objectivity and inter-subjectivity, measurement relies on *calibrated* devices. The devices *transduce* the interaction with a object into an internal state (of the system constituted by the device and the object under measurement) that is empirically accessible via the pointer. Calibration allows to provide a meaning to the positions of the pointers. Once a physical or theoretical *reference system* isomorphic to $[p]$ is established, the procedure of calibration (see [15,16] for details) establishes a one-to-one correspondence between the positions of the pointers and the properties in $[p]$, i.e., the positions of the pointers stand for properties. The device becomes a sort of physical embodiment of the classification system.

Let us analyze the conditions an evaluation $(\mathbf{r}_t(x,m)+\mathbf{p}_t(m)) \lessdot_t \mathbf{q}_t(x)$ must satisfy to be classified as a measurement. First, we need to assure that m is a measurement device, i.e., an object with the design characteristics previously discussed. Here we do not explicitly consider these characteristics, we simply introduce n kinds $\mathtt{Di}$ of stable and calibrated devices all subsumed by OBJ. Furthermore, each kind of devices needs to be characterized in terms of (*i*) the possible positions of the pointers, and (*ii*) the possible ways an object can be connected to the device. To each kind of devices $\mathtt{D}$, we associate a $\bar{\mathcal{D}} \subseteq \bar{\mathcal{P}}$ that characterizes the configurations (of the pointers) of the devices and a $\overline{\mathcal{RD}} \subseteq \bar{\mathcal{P}}$ that specifies how the input objects must be connected to the devices.

Second, we need to represent the classification system $[p]$. To account for the possibility to classify objects at different *resolutions*, we allow $[p]$ to contain multi-resolution values, i.e., the properties in $[p]$ can be *taxonomically* structured. In our framework, a multi-resolution system can be represented by a taxonomically structured $\bar{\mathcal{S}} \subseteq \bar{\mathcal{P}}$. For instance, in Figure 1, one can consider $\bar{\mathcal{S}} = \{\bar{\mathtt{COLORED}}, \bar{\mathtt{RED}}, \bar{\mathtt{CRIMSON}}, \bar{\mathtt{YELLOW}}, \bar{\mathtt{OLIVE}}, \bar{\mathtt{GREEN}}, \bar{\mathtt{EMERALD}}\}$ that contains predicates at different levels of resolution, e.g., $\bar{\mathtt{COLORED}}$, $\bar{\mathtt{RED}}$, and $\bar{\mathtt{CRIMSON}}$. Flat (non-taxonomically structured) systems can also be considered, e.g., $\bar{\mathcal{S}} = \{\bar{\mathtt{CRIMSON}}, \bar{\mathtt{OLIVE}}, \bar{\mathtt{EMERALD}}\}$.

[17] When used and initialized in a correct way, the design of the devices guarantees their *selectivity*—i.e., the internal state of the system composed by the device and the object under measurement is independent from environment conditions as much as possible—and their lack of *invasivity*—the devices interact with the objects without changing them, at least with respect to the quality under measurement. In addition, the pointers are designed in order to avoid possible reading errors.

Third, for the representation of the calibration, we need to individuate the properties in $\bar{\mathcal{S}}$ the observation-kinds in $\bar{\mathcal{D}}$ stand for. We need then an embedding of $\bar{\mathcal{D}}$ into $\bar{\mathcal{S}}$. This embedding represents the calibration of D-devices with respect to the system $\bar{\mathcal{S}}$, i.e., we explicitly represent neither the calibration process nor the reference system that allows for the calibration.[18] Given a D-device d, such that $\bar{\mathtt{P}} \in \bar{\mathcal{D}}$ and $\bar{\mathtt{R}} \in \overline{\mathcal{RD}}$, and a classification system $\bar{\mathcal{S}}$, these calibration constraints have the form in (f8) where $\bar{\mathtt{Q}} \in \bar{\mathcal{S}}$.

f8 $(\mathbf{r}_t(x,d)+\mathbf{p}_t(d)) \prec_t s \land x \multimap s \rightarrow \bar{\mathtt{Q}}s$ (with $\bar{\mathtt{P}} \in \bar{\mathcal{D}}$, $\bar{\mathtt{R}} \in \overline{\mathcal{RD}}$, $\bar{\mathtt{Q}} \in \bar{\mathcal{S}}$)

Once the calibration constraints are available, it becomes trivial to filter out the evaluations that are not measurements. If the observations $\mathbf{r}_t(x,m)$ and $\mathbf{p}_t(m)$ are not primitive, the calibration and filtering process can be applied also to them.[19] Vice versa, the check of the correctness of 'pure' evaluations—e.g., when the mediator is a person, group, or institution—could involve social or historical behaviors of the mediator that are very difficult to be analyzed and represented.

6. The Interplay Between Observations and True Propositions

We have seen that the calibration constraints help in discovering material mistakes. Consider now two devices d_1 and d_2 of kind D1 and D2, both calibrated with the classification system $\bar{\mathcal{S}}$. Furthermore, assume that the object x is correctly connected to d_1 by $\bar{\mathtt{R1}} \in \overline{\mathcal{RD}}1$ and to d_2 by $\bar{\mathtt{R2}} \in \overline{\mathcal{RD}}2$ and that the calibration constraints are satisfied. In the situation represented in (f9), all these conditions do not guarantee the identity of $\mathbf{q1}_t(x)$ and $\mathbf{q2}_t(x)$, it is still possible to have two different measurements of x relative to $\bar{\mathcal{S}}$.

f9 $(\mathbf{r1}_t(x,d_1)+\mathbf{p1}_t(d_1)) \prec_t \mathbf{q1}_t(x) \land (\mathbf{r2}_t(x,d_2)+\mathbf{p2}_t(d_2)) \prec_t \mathbf{q2}_t(x)$

This difference can be due to the resolution of the devices, e.g., $\mathbf{q1}_t(x) = \mathbf{scarlet}_t(x)$ and $\mathbf{q2}_t(x) = \mathbf{red}_t(x)$, or to the kind of receptors the devices are equipped with. For instance, in Figure 1, $\bar{\mathtt{OLIVE}}$ is subsumed by both $\bar{\mathtt{YELLOW}}$ and $\bar{\mathtt{GREEN}}$, therefore $\mathbf{q1}_t(x) = \mathbf{yellow}_t(x)$ and $\mathbf{q2}_t(x) = \mathbf{green}_t(x)$ can be justified by the lack of information about the exact shade of x: d_1 classifies an olive shade as yellow, while d_2 as green. Disagreements like $\mathbf{q1}_t(x) = \mathbf{olive}_t(x)$ and $\mathbf{q2}_t(x) = \mathbf{crimson}_t(x)$ are less easy to be justified because, intuitively, *being olive* and *being crimson* are incompatible properties, no calibrated devices should, in principle, produce these results. However, in the scientific and ordinary practice, sometimes devices are used in a wrong way, in extreme environmental conditions, or they are just malfunctioning. Thus, an epistemological approach cannot exclude the previous kind of contradictory observations.

Our framework does not contain disjointness constraints that concern the leaves of the $\bar{\mathcal{P}}$-taxonomy, therefore the existence of both $\mathbf{olive}_t(x)$ and $\mathbf{crimson}_t(x)$ does not generate a logical inconsistency. This is a prerequisite to manage the disagreement among observations inside the theory. On the other hand, one would also represent the fact that

[18] A device could be used to measure different qualities. E.g., a ruler can be used to measure both the depth and the height of an object. The way the device is used is then important to individuate the properties it is measuring. We will not consider this aspect here, i.e., we assume that every device is connected to a unique $\bar{\mathcal{S}}$.

[19] The measurements taken at time t are often preceded by precise procedures to correctly set the used devices. To express these procedures on can use sums of diachronic observations (see [6]). However, the dependence of the evaluation on these sums of diachronic observations would require the introduction of diachronic explanations. We do not consider this extension here.

being olive and *being crimson* are, in an ontological perspective, incompatible properties. To represent ontological knowledge, we first introduce, in the vocabulary, a set $\mathcal{P}$ of temporally qualified predicates, i.e., predicates with a TME-argument. Then, we add ontological constraints in the T-box that characterize the ontological nature of the $\mathcal{P}$-predicates. For instance, in the previous example, we introduce the predicates OLIVE and CRIMSON and add the constraint $\mathtt{OLIVE}_t x \rightarrow \neg\mathtt{CRIMSON}_t x$.

The $\mathcal{P}$-predicates are in a one-to-one correspondence with the $\bar{\mathcal{P}}$-predicates,[20] and, intuitively, the temporally contingent $\mathcal{P}$-propositions are the counterparts of the observations, e.g., the proposition $\mathtt{80KG}_t(\mathtt{luca})$ corresponds to the observation $\mathbf{80kg}_t(\mathtt{luca})$. [6] assumes the one-to-one meta-correspondence (f10) between $\mathcal{P}$-propositions and $\bar{\mathcal{P}}$-observations. In particular, all the observations are transferred into the *temporally qualified* A-box, the observations are then the *truth-makers* of the $\mathcal{P}$-propositions.

f10 $\mathrm{P}_t x^n \leftrightarrow \exists s(\bar{\mathrm{P}} s \wedge \varepsilon_t s \wedge x^n \multimap s)$

However, in our epistemological perspective, (f10) is not acceptable because the import of contradictory observations would generate logical inconsistencies. For instance, the conjunction of the two $\mathcal{P}$-propositions that correspond to $\mathbf{olive}_t(x)$ and $\mathbf{crimson}_t(x)$ is inconsistent with the ontological constraint $\mathtt{OLIVE}_t x \rightarrow \neg\mathtt{CRIMSON}_t x$. Our idea it then to filter and clean the chaotic factual knowledge represented by means of the observations to make it consistent with the ontological knowledge. Still we want all the true propositions in the A-box to be grounded on observations but we need to avoid inconsistencies.

One possibility consists in substituting (f10) with $\mathrm{P}^d_t x \leftrightarrow \exists s(\bar{\mathrm{P}} s \wedge \sigma^d_t s \wedge x \multimap s)$—i.e., contextualizing to their sources all the propositions that correspond to observations—and then resolve the disagreement at the level of the A-box by introducing specific axioms that aggregate source-dependent propositions into source-independent ones. This would mean that (*i*) a subset of the A-box has an epistemological nature—it reflects the point of view of the devices on the world—and (*ii*) that the source-independent propositions do not correspond to any observation. In the following, we prefer to explore a solution that approaches the resolution of the disagreement among measurements at the level of observations. By relying on explanations and on some information about the devices, we identify the *most plausible* measurement among a set of (possibly) contradictory ones. It is then possible to restrict the application of the (meta-)correspondence (f10) only to the most plausible measurements. By breaking the right-to-left arrow in (f10)—i.e., by decoupling observations and true propositions—we can control which observations are imported into the the A-box. We shift towards a *verificationist* approach to truth: propositions must be verifiable, they are true only if they are verified, and *truth* "is constrained by our abilities to verify, and is thus constrained by our epistemic situation." [17]. Below we only sketch this solution focusing on disagreements among measurements.

We have seen that a measurement of an object x is explained in terms of the output of a calibrated device connected to x. Here we exploit the idea that these explanations—that correspond to cognitive processes—can be, in their turn, observed, i.e., we have observations about the way observations are explained in terms of other observations. This requires the explanation relation to be moved from the ontological to the epistemological realm, i.e., the primitive relation $\lessdot$ must be replaced by a new kind of observations, an extreme move for which we can provide only a partial technical analysis.

[20] Predicates with a bar apply to observations while predicates without a bar apply to objects (at a time).

First, we extend $\bar{\mathcal{P}}$ with the new kind of observations $\bar{\mathtt{EXP}}$ (not subsumed by $\bar{\mathtt{PHYSICAL}}$ in Figure 1) that have two participants: $\mathbf{exp}_t(s, s')$ is the observation about the fact that, at t, the observation s directly explains the observation s'. Because $\bar{\mathtt{EXP}}$-observations are about observations, (a3) must be modified to allow observations to participate (in the sense of $\multimap$) in observations. Second, to resolve a disagreement among measurements we have to individuate the most plausible one. To do that, we first add to $\bar{\mathcal{P}}$ the kind $\bar{\mathtt{PLAUS}}$ that collects all the observations about the plausibility of measurements. Then, we explain $\bar{\mathtt{PLAUS}}$-observations as in (f11), where we assume that (*i*) all the devices are calibrated and relative to the same measurement system $\bar{\mathcal{S}}$ and (*ii*) the additional information about the devices is given by $\mathbf{mr}_t(d_i, d_1, \ldots, d_{i-1}, d_{i+1}, \ldots, d_n)$, i.e., the device d_i is more reliable than the others. In (f11), the plausibility of s_i is explained by observing that s_i is the measurement collected by the most reliable device. Clearly, alternative aggregation techniques coming from *judgment aggregation* [18], *belief merging* [19], or *merging of (populated) ontologies* [20] may be considered here. For instance, the plausibility of measurements could be individuated by the agreement of the majority of (or, of the most reliable) devices or even taking into account statistical analyses.

f11 $\mathbf{exp}_t(\mathbf{exp}_t(\mathbf{r1}_t(x, d_1) + \mathbf{p1}_t(d_1), s_1) + \ldots + \mathbf{exp}_t(\mathbf{rn}_t(x, d_n) + \mathbf{pn}_t(d_n), s_n) + \mathbf{mr}_t(d_i, d_1, \ldots, d_{i-1}, d_{i+1}, \ldots, d_n), \mathbf{plaus}_t(s_i))$

(f10) can then be easily modified to transfer into the A-box only plausible measurements, i.e., the measurements that participate in a $\bar{\mathtt{PLAUS}}$-observation. To guarantee the consistency of the A-box, we assure that, given a measurement system $\bar{\mathcal{S}}$, there exists a unique $\bar{\mathtt{PLAUS}}$-observation about a given object (a14).[21]

a14 $(s = \mathbf{plaus}_t(s_i) \wedge s' = \mathbf{plaus}_t(s_j) \wedge x \multimap s_i \wedge x \multimap s_j \wedge \bigvee_{\bar{\mathtt{P}}, \bar{\mathtt{Q}} \in \bar{\mathcal{S}}}(\bar{\mathtt{P}}s_i \wedge \bar{\mathtt{Q}}s_j)) \rightarrow s = s'$

The approach above sketched introduces an *hiding* mechanism, there are observations that are not accessible from the temporally qualified A-box. Some $\mathcal{P}$-propositions could then be intended as *abstractions* from the universe of observations, as *macro-indexes* that cluster observations. Usually, the macro-indexes have a cognitive function, e.g., they resolve some disagreements or they hide some details to allow an effective management of the information. The case of *testing* is particularly interesting because the intermediate observations have no conceptual relevance. Testing usually involves at least two main steps: (*i*) the *administration* of the test to a subject to obtain the *raw data*; and (*ii*) the *aggregation* of the raw data into the *score* that represents how much of an attribute is present in a subject. However, the raw data taken in isolation are often meaningless or they only provide partial insight to the measured attribute. They usually are not even values of an attribute of the subject, e.g., a response time for a given task. In many cases, they need to be combined and mapped to the *score* of the subject for the assessed attribute. One can then assume that only this score needs to be translated into the A-box. $\mathcal{P}$-propositions are still grounded on observations, but these observations (*i*) can be the result of a complex explanation-path; and (*ii*) they are just a subset of the whole universe of observations.

[21] Without this constraint one would need to solve possible disagreements between plausible measurements. In addition, note that (a11) and (a14) neither guarantee that all the measurements relative to $\bar{\mathcal{S}}$ has been considered, nor prevent possible inconsistencies with, for instance, physical laws or aggregations.

7. Conclusion

We followed an epistemological approach to model observations and data. We shown that the explanation relation makes possible the introduction of evaluative and aggregative processes that play a fundamental role in measurement and testing and that allow to explicitly account for the provenance and elaboration of (at least some) observations. One original aspect of the proposed framework regards the decoupling between the universe of the observations and the temporally qualified A-box. By controlling the observations that are imported in the A-box, a chaotic 'soup' of possible incompatible observations and a clean, ontologically founded, and consistent A-box can coexist. However, this coexistence requires to abandon the realm of pure truth to enter the one of plausibility.

Acknowledgements. We thank Alessander Botti Benevides, Laure Vieu, and the reviewers for their useful comments.

References

[1] W. Kuhn. A functional ontology of observation and measurement. In K. Janowicz, M. Raubal, and S. Levashkin, editors, *GeoSpatial Semantics - Third International Conference (GeoS 2009)*, Lecture Notes in Computer Science 5892, pages 26–43. Springer Verlag, 2009.

[2] F. Probst. Observations, measurements and semantic reference spaces. *Applied Ontology*, 3(1–2):63–90, 2008.

[3] K. Janowicz and M. Compton. The stimulus-sensor-observation ontology design pattern and its integration into the semantic sensor network ontology. In K. Taylor, A. Ayyagari, and D. D. Roure, editors, *Proceedings of the 3rd International workshop on Semantic Sensor Networks (SSN10)*. CEUR, 2010.

[4] B. Batrancourt, M. Dojat, B. Gibaud, and G. Kassel. A core ontology of instruments used for neurological, behavioral and cognitive assessments. In A. Galton and R. Mizoguchi, editors, *Proceedings of FOIS 2010*, pages 185–198. IOS Press, 2010.

[5] R. Brinkman, M. Courtot, D. Derom, J. Fostel, Y. He, P. Lord, J. Malone, H. Parkinson, B. Peters, P. Rocca-Serra, et al. Modeling biomedical experimental processes with obi. *Journal of biomedical semantics*, 1(1):1–11, 2010.

[6] A. B. Benevides and C. Masolo. States, events, and truth-makers. In P. Garbacz and O. Kutz, editors, *Proceedings of FOIS 2014*, pages 93–102. IOS Press, 2014.

[7] R. Casati and A. Varzi. *Parts and Places*. MIT Press, Cambridge, MA, 1999.

[8] Z. W. Pylyshyn. *Things and Places. How the Mind Connects with the World*. MIT Press, 2007.

[9] C. Daly. Tropes. In D. Mellor and A. Oliver, editors, *Properties*, pages 140–159. OUP, 1997.

[10] L. Finkelstein. Widely, strongly and weakly defined measurement. *Measurement*, 34:39–48, 2003.

[11] L. Mari. A quest for the definition of measurement. *Measurement*, 46:2889–2895, 2013.

[12] L. Mari and A. Giordani. Quantity and quantity value. *Metrologia*, 49:756–764, 2012.

[13] P. Gärdenfors. *Conceptual Spaces: The Geometry of Thought*. Bradford Books, 2000.

[14] S. Borgo and C. Masolo. Foundational choices in DOLCE. In S. Staab and R. Studer, editors, *Handbook on Ontologies*, pages 361–381. Springer Verlag, 2nd edition, 2009.

[15] C. Masolo. Founding properties on measurement. In A. Galton and R. Mizoguchi, editors, *Proceedings of FOIS 2010*, pages 89–102. IOS Press, 2010.

[16] A. Frigerio, A. Giordani, and L. Mari. Outline of a general model of measurement. *Synthese*, 175(2):123–149, 2009.

[17] M. Glanzberg. Truth. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Fall 2014 edition, 2014.

[18] C. List. The theory of judgment aggregation: An introductory review. *Synthese*, 187(1):179–207, 2012.

[19] S. Konieczny and R. Pérez. Merging information under constraints: a logical framework. *Journal of Logic and Computation*, 12(5):773–808, 2002.

[20] D. Porello and U. Endriss. Ontology merging as social choice: judgment aggregation under the open world assumption. *Journal of Logic and Computation*, 24(6):1229–1249, 2014.