Refining OntoClean. Identity Criteria and Grounding

Massimiliano CARRARA\textsuperscript{a} and Ciro DE FLORIO\textsuperscript{b}

\textsuperscript{a} FISPPA Department, University of Padua, Padua Italy
\textsuperscript{b} Faculty of Economics, Catholic University of Milan, Milan Italy

Abstract. In this paper we introduce some logical and philosophical refinements to OntoClean, first by developing some formal constraints on identity criteria, secondly by specifying a kind of identity criteria, two level identity criteria, whose role is to explain an identity among some entities referring to some other, more basic, entities. Using such refinement we add a formal constraint to the stock of OntoClean meta-constraints (OC+). We, then, observe that two level identity criteria have an intuitive reading in terms of dependence of a kind of entities on some other entities, possibly specified in terms of a grounding relation. Are identity criteria grounding principles? In the second part of the paper we discuss this option.

Keywords. OntoClean, identity criteria, constraints on logical adequacy of identity criteria, two-level identity criteria, grounding

1. Introduction

OntoClean (see \cite{9}, \cite{10} and \cite{11}) is surely one of the main important contributions to a formal foundation for ontologies in knowledge representation. It has been developed and applied in a variety of papers and researches.\textsuperscript{3} OntoClean analyzes ontologies using some formal, domain-independent properties, i.e. metaproperties such as identity, rigidity, unity and dependence.

Aim of this paper is to introduce some logical and philosophical refinements to OntoClean, refinements we hope could be useful for new developments and research of this important tool for ontologies. Specifically, we concentrate here on identity criteria and on a related topic, grounding.

The paper is organized as follows. First (2) we briefly introduce identity criteria-- an OntoClean formal, domain-independent property of identity. Then (3), we describe some formal constraints on identity criteria so that one can say what does it mean they are logically adequate. They are specified on the basis of the logical form of identity criteria and some properties induced by it. Moreover, we observe (4) that identity criteria have to play an explanatory role for identity. To this purpose we introduce a specific kind of identity criteria: two level identity criteria, whose role is just to explain an identity among

\textsuperscript{1}Copyright © 2020 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
\textsuperscript{2}Corresponding Author email address: massimiliano.carrara@unipd.it
\textsuperscript{3}Google scholar reports 2.340 items just searching the word “OntoClean”
some entities referring to some other, more basic, entities. We enrich OntoClean with a principle for this kind of criteria, (OC+). A first reading of two level identity criteria is in terms of a relation of dependence of a kind of entities on some other entities. Is it possible to read such dependence relation in terms of a grounding principle? In (5) we analyze this option. We give a largely negative answer to the above question.

2. Identity Criteria

The credit for introducing identity criteria is usually attributed to Frege. In his *Foundations of Arithmetic*, he introduces this idea in a context where he wonders how we can grasp or formulate the concept of number (see [8], sec. 62). This is the standard Fregean quotation:

If we are to use the symbol $a$ to signify an object, we must have a criterion for deciding in all cases whether $b$ is the same as $a$, even if it is not always in our power to apply this criterion. (see [8], sec. 62)

Two famous examples of identity criteria provided by Frege are as follows:

- **Directions**: If $a$ and $b$ are lines, then the direction of line $a$ is identical to the direction of line $b$ if and only if $a$ is parallel to $b$.
- **Hume’s principle**: For any concepts $F$ and $G$, the number of $F$-things is equal to the number of $G$-things if and only if a one-to-one correspondence exists between $F$-things and $G$-things.

In ([9], [10] and [11]) Guarino and Welty suggest that an identity criterion, associated to a property $K$, has to answer the following question:

(OQ) If $a$ and $b$ are $K$s, what are necessary and sufficient conditions for the object $a$ to be identical with $b$?

In Guarino and Welty’s opinion identity criteria are «the criteria we use to answer questions like, “is that my dog?”» ([9]: 154). That means identity criteria should be useful to justify identity judgements. A general, formal way to represent identity criteria (hereafter: IC) is as follows:

$$\forall x \forall y ((K(x) \wedge K(y)) \rightarrow (x = y \leftrightarrow R(x,y)))$$

$R$ represents the identity condition; $R$ holds between a pair of $K$s $x$ and $y$, iff $x$ and $y$ are identical. Given that $x = y$ is an equivalence relation, the right side of the biconditional $R$ must be an equivalence relation. Following the Fregean example of directions (IC) is also formulated in the following way (without a reference to $K$):

$$\forall x \forall y(x' = y' \leftrightarrow R(x,y)),$$

where “$x’$” and “$y’$” are terms representing entities of the kind $K$ suitably connected with $x$ and $y$. Frege’s criterion of identity for directions is an example of (IC*):

$$\forall x \forall y((K(x) \wedge K(y)) \rightarrow (x = y \leftrightarrow R(x,y)))$$
∀x∀y((o(x) = o(y) ↔ P(x,y)))

(O)

where x and y range over lines, o stands for ‘the direction of’ and P means ‘is parallel to’. As an example the direction of line a is identical to the direction of line b if and only if a is parallel to b. In (O), the identity sign is flanked by terms constructed with a functional letter, and the right-hand side of the biconditional introduces a relation among entities (lines) different from the entities for which the criterion is formulated (directions).

As we will see in the following, the logical structure of IC raises some interesting questions about their applicability in ontology. Specifically, we will take into exam the so-called cases of two-level Identity Criteria like (O), that is, cases in which the identity of certain entities is explained by the reference to a particular (equivalence) relation between other, more basic, entities. Before that, in what follows, we consider some formal constraints on (IC).

3. On some formal constraints for identity criteria

An implicit assumption of OntoClean treatment of identity criteria (see, for example, \[9\]) is that logical adequacy is a necessary constraint for them. What does it mean to say that identity criteria are logically adequate?

First, observe that those formal constraints concern the logical adequacy of the relation R in (IC), which is the identity condition of the criterion. In other words, given an identity statement a = b, R is a relation that holds between a and b. R is other than identity and analyzes what it is for the referents of a and b to be identical. How should it be characterized so to have logically adequacy? In what follows we propose some logical properties of R.

Non-vacuousness: the identity condition should not be vacuously satisfiable. Consider the following example (see \[12\], 32-33): let PO be the set of physical objects, S the set of relevant abstract objects, R(x,y) the identity condition for PO, and R'(x,y) the identity condition for S. Then let (IC) be defined as:

∀x∀y(x ∈ PO ∨ x ∈ S) ∧ (y ∈ PO ∨ y ∈ S) → (x = y ↔ (R(x,y) ∨ R'(x,y))))

(1)

The condition given above for the identity of x and y is not associated with a kind of entity in a metaphysically interesting sense, since the members of the alleged kind physical objects or set do not share any essential property.

In a strong metaphysical realist perspective the identity condition R should identify instances of the same kind of those objects that share all the essential properties associated with that kind. From such a perspective, the identity condition can be thought of as a property of properties. Lombard calls such a property determinable since it determines a class of properties, or determinates, having that property. An example of a determinable is ‘being a spatio-temporal property’, which can be considered a good candidate for an identity criterion for objects: if o and o’ are physical objects, then they are identical iff they are alike with respect to all the properties that are spatio-temporal properties. A criterion of identity for K-objects, to be acceptable, must provide a determinable such that
it makes non-vacuously sense to attribute determinates falling under the determinable to each K-object.

**Informativeness:** $R$ should contribute to specifying the nature of the kind $K$ of objects for which $R$ acts as an identity condition. The identity condition does not completely characterize the nature of instances of $K$: to decide identity questions concerning items of $K$ we need the concept of $K$, which is not provided by the identity criteria. Nevertheless, an identity criterion should specify some non-trivial essential properties of objects of kind $K$. This means that the form of the relation cannot be tautological, for instance, of the form:

$$S(x, y) \lor \neg S(x, y)$$

where $S(x, y)$ is an arbitrary binary predicate.

**Partial exclusivity:** an identity condition for a kind $K$ of objects cannot be so general that it can be applied to other kinds of objects. The example provided by Lombard is the following:

If $x$ and $y$ are both non-physical objects, then $x$ and $y$ are identical iff they have the same individual essence.

Now, properties falling under the wide property ‘having an individual essence’ do not apply only to non-physical objects, and can be part of the identity conditions for many kinds of objects. Living beings, for example, instantiate properties usually considered individual essences, being an individual $x$ generated by gametes $y$ and $z$ for example, but they are not non-physical objects.

**Minimality:** the identity condition for $K$-objects is required to specify the smallest number of determinables such that the determinates falling under them turn out to be necessary and sufficient to ensure identity between two objects of kind $K$. The determinables specified in the identity condition cannot be superfluous. Consider the following example (12, 38):

If $x$ and $y$ are both sets, then $x$ and $y$ are identical iff they have the same members and are liked by the same people.

The above criterion suggests that it is part of the very idea of sets that they are liked by people. But, clearly, it is not so. In order to rule out such cases the formal requirement of minimality for identity criteria is introduced.

**Non-circularity:** the identity condition for $K$-objects cannot make use of the concept of identity itself; otherwise it is circular. There has been a long debate about the circularity of IC. Consider the criterion of identity for events proposed by Davidson (3):

If $x$ and $y$ are events, then $x$ is identical to $y$ iff $x$ and $y$ have the same causes and effects.
One could argue that the above formulation is not formally circular, since the identity predicate does not occur in the right part of the bi-conditional. However, it has been observed that whether an event $e$ has the same causes and the same effects as an event $e'$ can depend on the solution to an identity question concerning entities of the same kind. On the right side of the bi-conditional, causes and effects are mentioned; since those are considered to be events, the identity criterion for events turns out to involve identity between events. In fact, to determine whether two events are the same, one is first required to determine the identity of the events taken as their causes or effects. One can thus conclude that identity is already presupposed. This criticism corresponds to the denial that it is possible to give an explicative criterion of identity for objects of a certain kind, such as events. In fact, the formulation of such a criterion would involve a quantification over all objects for which the criterion is specified, and quantification presupposes the determinacy of the identity of the objects quantified (see on this [13]).

Non-totality: given at least two objects belonging to some kind $K$, $R$ cannot be a property that every two $K$-objects share. Formally:

\[(C1) \quad R \subset K \times K\]

(C1) says that the relation $R$ is a proper subset of the set $K \times K$: that is, there is some pair of $K$-objects such that the objects of the pair are not in the extension of $R$.

K-Maximality: $R$ must be maximal with respect to $K$. In other words, $R$ is required to be the widest dyadic property that makes an identity condition true.

A dyadic property $G$ is wider than a property $G'$ iff for any $x$ and $y$, if $G'(x,y)$ then $G(x,y)$, but not vice versa.

In other words, the ordered pairs of $G'$ are a subset of the set of ordered pairs of $G$. In such a way we always obtain a condition for an ultimate kind or ultimate sortal (concept) $K$ (here, for the sake of simplicity, we use the term ‘kind’ and ‘sortal’ as synonymous.

The reason for introducing the formal constraint of K-maximality is this. Consider what Wiggins calls “a structure comprising only sortals” where sortals or kinds stand in relation to one another and have common members. Take two sortals $C(1)$ and $C(2)$. We have three cases.

**Case 1.** Neither $C(1)$ nor $C(2)$ is a restriction of any other sortal and each is an ultimate sortal. If they have common members, then, “because they will cover identities relating to these common members [...] $C(1)$ must be identical with $C(2)$ or extensionally equivalent to it” ([17], 67).

**Case 2.** Either $C(1)$ or $C(2)$ is an ultimate sortal and the other is not. In this case, if there are common members the non-ultimate one gives a restriction to the other.

**Case 3.** $C(1)$ and $C(2)$ have common members but no concept subsumes the other. In this case we have cross-classification, but some ultimate sortal will subsume both $C(1)$ and $C(2).$ “This picture of things”, Wiggins argues, “founded in the nature of sortal[s] . . . and the absoluteness of identity, concedes everything that deserves to be conceded to the over-stringent demand that sortal[s] . . . should form a hierarchy” ([17], 67).
**Uniqueness:** R is unique with respect to K. This means that if there are relations \( R_1, R_2, \ldots, R_n \) such that (i) each \( R_i \) satisfies IC and (ii) each \( R_i \) is independent of each \( R_j \) – that is, every \( R_i \) is neither narrower nor wider than each \( R_j \) – then at most one among the relations in \( R_1, R_2, \ldots, R_n \) provides a correct identity criterion for K-objects.

**Equivalence:** R must be an equivalence relation. On the left side of the bi-conditional in IC there is an identity relation that is an equivalence relation; consequently, the right side of the conditional is supposed to present an equivalence relation as well. R must then be reflexive, symmetric and transitive.

**Congruence:** \( a \) is the same K as \( b \) iff the way in which \( a \) is K-related to \( b \) via R is sufficient for whatever is true of \( a \) to be true of \( b \) and for whatever is true of \( b \) to be true of \( a \).

### 4. Identity Criteria and Functional Entities

As we have seen before, the use of identity criteria in OntoClean entails a series of logical requirements which must be satisfied.

It is important to notice, however, that since the introduction of IC an important IC feature is the explanatory role which the identity condition plays. In other terms, IC should provide an explanation of why items belonging to K are identical.

Notice that there are at least two general forms of IC. Frege’s preferred example is the so called Identity Criterion for directions. Roughly, the idea is to give the necessary and sufficient condition for two directions are the same:

\[
\forall x \forall y ((o(x) = o(y) \iff P(x, y)))
\]  

(O)

As it is easy to notice, (O) is an instance of the following general schema:

\[
\forall x \forall y (x' = y' \iff R(x, y)),
\]  

(IC*)

where “\( x' \)” and “\( y' \)” are terms representing entities of the kind K suitably connected with \( x \) and \( y \).

(IC*) is the logical form of a two-level identity criterion (see [18], 145-146). The crucial point is that in the case of two-level identity criteria the conditions of identity concern objects which are not of the same kind of objects for which the identity criterion is provided. On two-level identity criteria Williamson points out that: «The idea of a two-level criterion of identity has an obvious advantage. No formula could be more basic (in any relevant sense) than ‘\( x = y \)’, but some might be more basic than ‘\( ox = oy \)’, by removing the symbol ‘\( o \)’ and inserting something more basic than it». ([18], 147)

If identity criteria play an explanatory role, Williamson’s remark seems to be plausible: in order to explain an identity we have to refer to some other, more basic, level of entities.

It is worth to notice that, in this case, (O) has not only a control-function on the inflation of our ontology (according to Quine’s slogan: No Entity without Identity) but
it is to *introduce* new entities in the domain. So, two levels identity criteria seem to be explanatory; however, this entails that entities on which identity is defined are presented as *functions*. Let us ask a question similar to Guarino and Welty’s one:

Given that \( x \) and \( y \) are directions, is \( x \) the same direction of \( y \)?

The answer is that \( x \) and \( y \) are the same direction if and only if the *lines of which* they are directions are parallel. This confirms Williamson’s remark: the explanation of the identity is genuine since it exploits more basic entities, i.e., lines. In other terms, it is because the lines are parallel that the direction of \( a \) is identical to the direction of \( b \). But this means that some kinds of entities are intrinsically *functional*: they are always items of something. OntoClean could be, then, enriched by a principle as follows (OC+):

\[
x \in K_f \leftrightarrow \exists y (x = f(y))
\]

that is, \( x \) is an entity which belong to a certain *functional kind* iff there exists another entity of which \( x \) is the value of the function \( f \); \( x \) is an entity belonging to kind *directions* iff there exists a line which \( x \) is the direction of.

This equivalence is an *admissibility* principle that can be added to OntoClean general framework: a certain kind (for instance, *directions*) is admissible in dependence of the existence of other entities. In order to underline the dependence relation surrounding identity criteria, let us consider these two scenarios related to the well known identity criterion for directions:

\[
\forall x \forall y (L(x) \land L(y) \rightarrow (d(x) = d(y) \leftrightarrow P(x,y)))
\]

(3)

Within \( S_1 \) there is no line and therefore it is meaningless to refer to \( P(a,b) \), given that there is no entity which satisfies the general property of being a line. So: no lines, no directions. Within \( S_2 \), there are two items which satisfies the property of being a line and moreover, it holds that \( \neg P(a,b) \). But then, through identity criteria for directions we have two "new" entities, since \( d(a) \neq d(b) \). But how to specify and characterize the intuitive idea of *dependence* introduced in our model of two-level identity criterion as
explanatory principles? In the next paragraph we propose a good candidate for this job: the *grounding* relation.

5. Grounding and Identity Criteria

Grounding is, for many scholars in metaphysics, the new black. In the last two decades, there has been a noteworthy amount of works dedicated to grounding (although this concept is not entirely new, as often happens in philosophy. One can argue that grounding originates within Aristotle’s notion of *ἀιτια*).

Grounding is one of the most discussed notions in contemporary philosophy. Roughly, grounding is a type of non-causal, primitive relation or operation, such that the grounded entities are somehow explained, determined or constituted by the grounding entities.

The “grounding revolution” ([15]: 91) contributed to clarify concepts such as priority and fundamentality, usually conveyed by locutions as “in virtue of” and “because” (see [6], [2], and [14] – at the moment the most complete overview of the logic of grounding). As crucial as they are in the philosophical debate, these expressions were rarely analyzed in a systematic way.

More recently, grounding has been applied, and is expected to be applicable, in a vast range of philosophical disciplines, from metaphysics to philosophy of mind, of logic and of language, and metaethics. One can draw several instances of grounding from these disciplines: for example, the fact that a flower is scarlet is taken to ground the fact that it is red; mental facts are taken to be grounded in neuro-physiological facts; and the truth of “snow is white” would be grounded in the fact that snow is white, and not vice versa.

The closeness between the notion of grounding and that of dependence is straightforward and, therefore, it is natural, at least *prima facie*, to consider identity criteria as grounding principles.

In general, to say that at a certain time *t*, *x* and *y* are distinct particulars or the same particular items seems to imply that there is something in virtue of which *x* and *y* are distinct particulars/are the same particular, i.e. a fact that grounds the distinctness of the two particulars at play.

In our case, the condition of identity explains, grounds, the identity at issue. By using > to indicate grounding relation, we have that:

\[
\forall x \forall y (L(x) \land L(y) \rightarrow (P(x, y) > d(x) = d(y))) \quad (4)
\]

Now, we have all the ingredients for the refinement we propose to OntoClean; identity criteria, construed as grounding methodological principles, are able to account for kinds the instances of which depend on other entities. For that reason, the identity of these items is not, so to speak, primitive but it is explained by the suitable identity relations. Therefore, we can establish a form of *admissibility criterion* which concerns some categories of entities. In other words, we set a criterion that, if satisfied, allows a certain kind of entities to be included in an ontology. From an intuitive point of view, the

---

*On this point, literature is still thin; see, for instance, [7], [1], [3].*
criterion is inspired by Quine’s slogan quoted at the beginning of this article: no entity without entity.

This is just one half of the story since the grounding construal of identity criteria does not simply allow the admission of entities (for instance, numbers or directions or events or people) but it provides an illustration of the grounding relations between the “new” entities and the fundamental ones in virtue of which the former can be introduced.

So, let us see how a provisional form of this criterion could be laid down for the well known case of directions:

\[
\vdash (d, D) \leftrightarrow \exists o_1 \exists o_2 (R(o_1, o_2) > d(o_1) = d(o_2)) \tag{5}
\]

Here, \(\vdash\) is a meta-predicate that expresses the relation of instantiation between an individual \(x\) and a kind \(K\). So, in our example, the entity \(d\) is an instance of kind \(D\) (i.e. Directions) iff there exist at least two objects \((o_1\) and \(o_2)\) and their parallelism relation explains (grounds) the fact that the direction of \(o_1\) is identical to the direction of \(o_2\).

So far, so good. There are, of course, many questions on the table. In the following, we will take into account just one problem concerning the inner logical structure of grounding. The problem concerns the discrimination of specific facts connected by a grounding relation. The majority of the accounts on grounding state that grounding is an irreflexive relation: in other terms, if the fact \(A\) grounds the fact \(B\), then \(A\) must be distinct from \(B\). Otherwise we would have a self-grounding fact. But then, we have a sort of puzzle. Let us consider, with Rosen, a classical case of definition of square in terms of an equilateral rectangle:

After all, if our definition of square is correct, then surely the fact that \(ABCD\) is a square and the fact that \(ABCD\) is an equilateral rectangle are not different facts; they are one and the same. But then the grounding-Reduction Link must be mistaken, since every instance of it will amount to a violation of irreflexivity. ([16], 124)

But the same problem affects also the identity criteria, interpreted as grounding principles. Consider the case of directions and let us reflect on what it means that it is a fact that \(Pur(a, b)\). Very likely, this is a (geometrical) scenario in which at least two items are in a certain spatial relation. But in this specific scenario, the direction of \(a\) is the direction of \(b\); in other words, it seems that there is not a further grounded fact based in the identity of the directions.

However, if the two alleged facts (about respectively, parallelism and directions) are, actually, the same geometrical scenario, then they cannot be stuck together by a grounding relation. There are some possible answers to this problem; one is sketched by Rosen himself:

We can resist this [critique] by insisting that the operation of replacing a worldly item in a fact with its real definition never yields the same fact again. It yields a new fact that ‘unpacks’ or ‘analyzes’ the original. ([16], 124)

Of course this train of thought should be accompanied by a more careful analysis of the operation of “unpacking”. Another path is proposed in [1] where an account of conceptual grounding has been provided. In the following we want to sketch a view similar to the latter.
OntoClean is a methodology for validating the ontological adequacy of taxonomic relationships. It means that, within OntoClean various ways "to carve reality at its joint" can be discussed and compared. The use of identity criteria with a grounding interpretation is a path for showing the grounding relations between various modes of presentation of bunch of entities.

The key concept is here that of mode. The same scenario can be intended and represented in many different ways; the extended example of Guarino and Welty (9, 8ff) clearly shows how to confront an intuitive and largely imprecise taxonomy with a far more refined, cleaned taxonomy. Through identity criteria as grounding principles we are able to fix the admissible ways in which one can represent a certain scenario: of course we can say that, within $S_1$ the lines $a$ and $b$ are parallel. But we can introduce another kind of entities, that is, directions; in other words, the same scenario can be adequately described in another mode: in a direction-mode in addition to a parallelism-mode.

Let us notice that, since the identity condition at issue is understood in terms of grounding, the two modes (direction-mode and parallelism-mode) are not, so to speak, on a par; on the contrary, the latter is more fundamental than the former. It explains, and so it justifies also, why we can safely speak about directions.

6. Conclusions

In this paper we first have developed some formal constraints on identity criteria such that, following OntoClean advice, one can say when they are logically adequate. Moreover, we have introduced a specific kind of identity criteria, two level identity criteria, whose role is just to explain an identity among some entities referring to some other, more basic, entities. Using them we have added a formal constraint to the stock of OntoClean meta-constraints (OC+). Two level identity criteria have an intuitive reading in terms of dependence of a kind of entities on some other entities, a reading possibly specified in terms of grounding. Are identity criteria grounding principles? In the second part of the paper, we argued that there are problems with a grounding construal of identity criteria; we sketched a possible way out, considering grounding in a more conceptual fashion.

Some further developments of the paper regard (a) the connection between the grounding relation and the dependency relation and (b) an analysis of the role of modes in OntoClean, including a formal framework able to characterize dependence relations between modes of presentation of the same scenario.

References