Modeling Concept Drift for Historical Research in the Digital Humanities

Claudio MASOLO [a,1], Emilio M. SANFILIPPO [b,c], Marion LAMÉ [b,c] and Perrine PITTET [c]

[a] *ISTC-CNR Laboratory for Applied Ontology, Trento, Italy*

[b] *Le Studium Institute for Advanced Studies, Orléans and Tours, France*

[c] *CESR UMR 7323 – University of Tours, France*

Abstract. In Digital Humanities one often needs to deal with the modeling of *concept drift*, i.e., the mechanism underlying changes in concepts' intensions. This particularly applies to historical research, e.g., when experts study changes in language or entire conceptual frameworks. The purpose of the paper is to address some challenges concerning the ontological modeling of both concepts and concept drift. We sketch an initial formal approach which allows to consider the temporal dimension of concept characterization and, consequently, temporal relations between multiple concepts. This eventually leads us to discuss some preliminaries ideas on the formal treatment of experts' intentional attitudes towards the transmission and manipulation of concepts inherited from the past.

Keywords. Concept, concept drift, history, history of ideas

1. Introduction

Domain experts working in the Digital Humanities, especially in research contexts related to history and history of ideas, often need to study the relations holding between concepts traced at different periods in time [9]. For instance, a domain expert may be interested in understanding how the concept of *science* adopted in ancient Greece relates to the concept of *science* used by contemporary Western scientists.[2] Simplifying the research methodology for the sake of easiness, by investigating and studying multiple sources, historians first try to define (or characterize) the concepts at stake and then understand their links by taking into account multiple parameters (e.g., related to the cultural bias of contemporary historians when approaching past societies) [7]. In this context, an important research question concerns whether concepts can change in time while preserving their identity or whether their characterizations are rather fixed.

In studies related to history of ideas, computer science, and philosophy, the phenomenon of concept change is also called *concept drift* [1,3]. From a theoretical per-

[1]Corresponding Author: Laboratory for Applied Ontology, via alla cascata 56/C, 38123, Trento, Italy; E-mail: masolo@loa.istc.cnr.it.

[2]Depending on the application domain, concepts analyzed by historians can be very broad including politics, science, culture, agriculture, artefacts, etc.

spective, it is a matter of debate what concept drift amounts to, for instance, what are the conditions that a concept needs to satisfy to keep its identity in time. From a computer science perspective, a formal model of concept drift for historical modeling purposes is still lacking. This model could be useful to help historians in their analysis of concepts, but also to make accessible and comparable the data originating from historical research.

We present in the following a preliminary analysis of concept drift ultimately aimed at the design of ontology-based systems for historical data. By reviewing previous works on the history of ideas, we sketch a formal model of concept drift which stresses the *intentional nature* of concepts and is able to take into account their originating contexts.

The remaining of the paper is structured as follows. Section 2 presents some general aspects of concept theories. Given the plethora of work done in different studies, our purpose is just to introduce some characteristics commonly ascribed to concepts. Section 3 clarifies the distinction between *term drift* and *concept drift* which is sometimes blurred in the literature. Section 4 reports some discussions in the state of the art on the history of ideas about concept drift. We consider this literature because it explicitly addresses issues concerning the understanding and modeling of concept drift that are relevant from a computer science perspective (e.g., how multiple *versions* of the same concept relate to each other). Section 5 introduces some basic formal tools that we use to analyze the state of the art and to frame our proposal. Section 6 adds some preliminary ideas for a new approach to model concepts for historical research that relies on the way the authors intentionally refer to pre-existing concepts. Section 7 concludes the paper.

2. Concepts

The nature of concepts is highly debated across cognitive sciences, philosophy, and linguistics [10]. Independently from this heterogeneity, scholars agree that a core functionality of concepts is to *classify* (*categorize*) other entities. For example, the concept *person* classifies human beings. It is also common to find the distinction between concepts' *extension* and *intension*, where the former refers to the (set of) individuals classified by a concept while the latter to its definition or characterization (that, in principle, allows to individuate concepts' instances). A concept's intension is not reducible to its extension, first, because the extension can change without affecting the intension (e.g., if Paul dies, the definition of *person* does not change); second, because concepts with different intensions (e.g., *being an animal with heart* and *being an animal with kidney*) can have the same extensions.

Intensionally, concepts are usually seen as *complex* entities characterized in terms of other concepts (see Sect. 5 for more details). The way concepts are characterized and how such characterization is exploited in the classification (categorization) process cause some disagreements among concepts theories. In the *classical view*, a concept is reduced to a conjunction of necessary and sufficient conditions; accordingly, an individual is either categorized or not under a concept, hence there is no uncertainty and all the instances of a concept are equally treated. By contrast, categorization under everyday concepts often presents typicality effects and a certain degree of indeterminacy. This pushed cognitive scientists to develop alternative models to match empirical data (see [10] for a review). For instance, in the *prototype theory* the categorization of an individual under a concept depends on the degree of similarity between the individual and the prototype of the concept. The degree of similarity is calculated by considering (i) the *salience* weights contextually assigned to the attributes of a given concept (e.g., for apples, *shape* may be more salient than *color*); and (ii) the *typicality* weights contextually assigned to attribute-values (e.g., *red* apples are more typical than *brown* apples).

The approaches in the field of the history of ideas only seldom endorse an explicit theory of concepts. There is indeed the underlying assumption that concept drift can be described without a strong commitment on the nature of concepts. The approaches presented in [1] and [8], which we will further analyze in the paper, assume that concepts are intentionally introduced at given times for specific aims and are characterized by simpler concepts called *features*. In the following sections, we start by analyzing the temporal and structural aspects of concepts and then see how these aspects can be used to compare alternative positions on concept drift. Before that, an informal clarification on concept and term drift is due.

3. Term and Concept Drift

For the sake of clarity, we distinguish between *term drift* and *concept drift*.

Term drift can be understood in different ways. A first case is when a term w_1 (w for word) used at time t_1 with meaning m is replaced at t_2 with a term $w_2 \neq w_1$ which however keeps m as meaning. This phenomena can be due to multiple factors, for instance, linguistic changes that occur within the language used at different times by different agents. A typical example is about morphological changes in the spelling of words, e.g., the way in which toponyms change in time [5]. A second case of term drift occurs when the same term w is used at different times but with different meanings, i.e., w has meaning m_1 at t_1 and meaning $m_2 \neq m_1$ at t_2. An example coming from the analysis of cooking texts in the French tradition concerns the term 'blanc manger'. During the Middle Age, this term had the meaning of a salty dish prepared with either fish or white meat while nowadays it means a white pudding (a dessert) [9]. The first case is a syntactic change of the term that is however used with the same meaning; the second case is a shift in meaning—a *semantic change* in Geeraerts' terms [6, ch.1][3]—likely due to the way in which the same term happens to be differently used across history.

Theories of *concept drift* focus on analyzing the mechanisms underlying the way in which concepts (possibly) change over time. A core issue is whether a concept keeps its identity when its *characterization*, i.e., its *intention*, changes. This issue recalls investigations in formal ontology about the identity of (physical) objects when some of their properties or components change through time [4]. A plethora of positions exist and some philosophers argue that identity and persistence conditions for objects cannot be set in a pure *a priori* way, but reference to domain knowledge is necessary [12]. For instance, just as it is the task of mathematicians to establish the conditions for two sets being identical or different, it is up to physicists to establish when two material bodies traced at different times are numerically the same individual. The case of concepts drift, considering the abstract and complex nature of concepts, is still more challenging. In Betti and van den Berg's words [1]: "The question of what concepts are and how they retain their identity over time is fundamentally a philosophical one [...] and there is no

[3] Geeraerts [6, ch.1] discusses and classifies several approaches to semantic change.

conclusive reason to take the belief that concepts retain their identity over time to be true."

From these considerations, it should be clear that term drift and concept drift cannot be identified: the first phenomenon concerns changes in linguistic entities and/or their meanings, while the second one is about concepts and their intensions, which do not necessarily bear a linguistic nature.

4. Mapping the Debate on Concept Drift

The debate on concept drift can be mapped into two main segments. In the first approach, concepts are intensionally static and invariable. For example, it is not possible for a concept C to be characterized by the features a and b at time t, and by the features a and c at time t'. Rather, we have two different concepts, namely, C_1 existing at t and characterized by a and b, and $C_2 \neq C_1$ existing at t' and characterized by a and c. One can however study the degree of similarity between (the characterizations of) C_1 and C_2. In the example, C_1 and C_2 share the feature a while being different concepts.

In the second approach, concepts can change (at some degree) their intensions while retaining identities. Mink [11], for instance, commenting on Lovejoy's view on the non-changing nature of concepts, argues that "[i]t is not necessary to mount a massive attack on Lovejoy's claim that unit-ideas remain identical in all appearances; one needs only put it under the lens of the historical sense for the anti-historical nature of this methodological belief to reveal itself." In Mink's view, therefore, dismissing the dynamic nature of concepts amounts to dismissing the very nature of historical analysis. In the previous example one can consider a single *dynamic* concept C that changes its characterization from t to t': C retains the feature a from t to t' but the feature b is substituted with c.

The dynamic approach is forced to specify the *persistence* conditions of concepts to limit the way in which they can change. To deal with this issue and allow for restricted concepts change, Kuukkanen [8] argues that when studying and comparing multiple concepts for historical analysis purposes, it is necessary to distinguish between concepts' *core* and *marginal* features. Core features are those that a concept must maintain during its whole life, whereas marginal features can be lost or acquired. For instance, following Kuukkanen, *being material* and *being indecomposable* are core feature of the concept of *element*, whereas *being found in all bodies* or *being the ultimate constituents of bodies* are marginal features (endorsed only by some authors).

On the basis of the core vs. marginal features dichotomy, Kuukkanen distinguishes three possible temporal behaviors of concepts: (i) *concept stability*: a concept keeps through time both its core and marginal features; (ii) *concept change*: a concept keeps through time its core features but it undergoes changes in its marginal features; (iii) *concept replacement*: the core of a concept definition changes in time. For instance, suppose a concept C has only one core feature, say, a. When C has the features a and b at all times at which it exists it is stable. When C is defined by a and b at time t, and by a and c at t' we have a genuine concept change: C changes but only marginally while keeping its core feature a. Vice versa, C defined by a and b at time t is replaced by C' at t' defined by c and d because it lost its core feature a.

From a general perspective, first, it is unclear how to establish that a concept replaces another. E.g., in the case of two concepts C and C' characterized by a and b, c

and d, respectively, it seems to be an intentional choice of the historian to consider C' as replacing C, since there is nothing in the features suggesting the replacement, i.e., the model does not make explicit why C' (and not, for example, C'' characterized by b and c) replaces C. One possibility we consider in Sect.6 is to complement the feature-based characterization of concepts with an intentional stance on concepts creation and transmission mechanisms.

Second, the role of core and marginal features and the way they are identified is not clear. Kuukkanen himself recognises that "there is no naturally carved line between the central/essential versus marginal features" [8, p. 370-71]. We think that this distinction can be left out without the loss of any expressive power. For instance, studying the concept of *science* in Kant and Brentano, a historian may first identify two stable concepts, S_K (Kant's view) and S_B (Brentano's view). Supposing that S_K and S_B share some features, Kuukkanen would probably introduce a new *dynamic* concept S where the core features are the ones shared by both S_K and S_B, whereas the marginal features are the ones characterizing only S_K and only S_B. Alternatively, it is possible to introduce a more general *stable* concept S' characterized only by the features shared by both S_K and S_B. In this case the links between S', S_K, and S_B are not lost, S_K and S_B are different *specializations* of S'. One can still compare concepts on the basis of their characterizations committing neither to the core vs. marginal dichotomy, nor to the dynamicity of concepts.

Betti and van den Berg [1] seem to adopt this second approach. In addition, they propose to characterize concepts within broader conceptual structures called *models*. The idea, inherited from cognitive science, is that models provide context-dependent networks of concepts whose inter-relations contribute to the definition of the concepts themselves. In this framework, specialization between concepts does not reduce to set-theoretical relations between their features, but it is instead based on more general links between the models that characterize the concepts. It is then possible that two specializations of a concept do not share any feature. Unfortunately, similarly to Kuukkanen's work, this approach remains at a high abstraction level.

In the next section, we use some formal tools to better analyze and compare the approaches in [8] and [1]. Despite the high-level similarities, there are relevant differences between these approaches which can be better grasped by formal means. More generally, our analysis aims at shedding some light on concepts' characterization and concept drift.

5. Formal Tools for Concept and Concept Drift Modeling

We start by analyzing the relation between a concept C and its *intensional* characterization ϕ. For this purpose, we introduce a *partial* function γ that associates to a concept C, at any time (or context) t at which it exists, its characterization ϕ. Concepts are assumed to be in time (or to exist in contexts) and, usually, to have a bounded life.

Before discussing the possibility for a concept to undergo a change in its characterization (Sect. 5.3), hence the possibility to have $\gamma(C,t) = \phi$ and $\gamma(C,t') = \phi'$ with $\phi \neq \phi'$, we clarify the nature of both the characterization ϕ (Sect. 5.1) and the function γ (Sect. 5.2). Finally, following [1], Sect. 5.4 introduces *models*, i.e., *networks* of concepts, to define (and compare) concepts in relational and structural terms.

5.1. Characterizations

In Sect. 2, we saw that concepts are commonly defined, or simply characterized, in terms of simpler concepts called *features*. In a non relational perspective, characterizations can be simply modeled as (finite) sets of features, i.e., $\phi = \{F_1, \dots, F_n\}$, that can be easily linked by means of set-theoretic relations.

To represent the distinction introduced by Kuukkanen [8] between *core* and *margin* features, each characterization ϕ can be partitioned into a set ϕ^c of core features and a set ϕ^m of margin features, i.e., $\phi = \phi^c \cup \phi^m$ and $\phi^c \cap \phi^m = \emptyset$. Characterizations can then be linked by considering the set-theoretical relations between the core-subsets and the margin-subsets separately.

Betti and van den Berg [1] refer also to the determinate-determinable distinction [13]. More precisely, they discuss the possibility to taxonomically organize features by means of an (*intensional*) *subsumption* relation, here noted $\sqsubseteq_{\mathsf{ft}}$: $F \sqsubseteq_{\mathsf{ft}} F'$ stands for "the feature F is subsumed by, specializes, the feature F'." Subsumption allows to express new links between the characterizations. For instance, the set-theoretical inclusion $\phi \subseteq \phi'$ between characterizations, stating that ϕ' is more specific than ϕ because it contains more features, can be generalized by considering $\sqsubseteq_{\mathsf{ft}}$: $\phi' \sqsubseteq \phi$ (note the different order of ϕ and ϕ') if and only if for all features $F \in \phi$ there exists a feature $F' \in \phi'$ such that $F' \sqsubseteq_{\mathsf{ft}} F$, i.e., all the features in ϕ are specialized by a feature in ϕ'. Given the fact that $F \sqsubseteq_{\mathsf{ft}} F$, when $\phi \subseteq \phi'$, for each $F \in \phi$ it is enough to consider $F' = F$ to see that $\phi \subseteq \phi'$ is just a special case of $\phi' \sqsubseteq \phi$.

As mentioned in Sect. 2, cognitive approaches to concepts consider additional information about features (usually called attribute-values): (*i*) features are partitioned by attributes or kinds; (*ii*) features are structured by means of geometric or topological relations; (*iii*) features and attributes are weighted, respectively, in terms of their *typicality* and *salience* in the context of a given concept; (*iv*) *subsumption* (defined between features of the same kind) is usually distinguished from *correlation* (defined between features in different kinds).[4]

Depending on the information one disposes of, it is possible to establish relations $R(\phi, \phi')$ between characterizations with different levels of precision and granularity. As we will see in the next sections, by relying on richly structured features, subtle analyses about the (dis-)similarity of charactizations and concepts are possible.

5.2. The Function γ

The function γ individuates how a concept is characterized at a given time or context, i.e., considering the previous discussion, what are the features associated to it.

Betti and van den Berg [1] claim that features are *components* of concepts, i.e., there exists a *parthood* relation defined between features and concepts. The function γ can be hence intended to collect all the components that C has at t, i.e., $F \in \gamma(C,t)$ if and only if $Part(F,C,t)$. In this view, concepts are complex entities composed by features

[4] Cognitive approaches are commonly interested in determining the similarity between objects, or the degree of classification of an object under a concept characterized by, e.g., a given prototype, while they do not usually consider similarity between prototypes.

(i.e., they are mereological sums of features, see [12] for details on mereologies).[5] Still it remains unclear whether features have to be intended as properties of the *instances* of the concepts C they characterize, i.e., whether $F \in \gamma(C,t)$ and $C(x,t)$ imply $F(x,t)$, where $C(x,t)$ ($F(x,t)$) stands for "at t, the entity x is an instance of the concept C (the feature F)". Kukkanen [8] talks of features as components, too, but he seems to accept among features also *properties of concepts* so that it is possible to have $F \in \gamma(C,t)$ and $F(C,t)$, i.e., at t, the concept C is an instance of the feature F, C has the property F.

Both [1] and [8] agree on ascribing to features and concepts the same nature; features are just simpler concepts. However, there is no further discussion about what kinds of properties are included among the features. This makes also unclear whether the features characterizing a concept C completely determine C or it is rather possible to have different concepts with exactly the same features.[6]

Given these remarks, we commit only to a weak notion of characterization where features only partially characterize what concepts are, i.e., concepts do not reduce to sets of features (Sect. 6 further discusses this view). Furthermore, we do not assume that features are enough to determine the *extension* of concepts. Even though [1,8] do not dedicate much attention to the classificatory role of concepts (not indeed fundamental for the remaining of our discussion), this aspect is crucial for cognitive theories, which usually rely on elaborated classification rules (grounded on features together with their taxonomical organization, salience weights, etc.). Note however that the intension of a concept is usually intended to influence, if not determine, its extension, i.e., the extension depends on the intension. By avoiding to specify the way the features of a concept are used to classify individuals we loose this link.

5.3. Analyzing Relations Between Concepts: the Case of Concept Drift

Let us consider again the concept of *science* discussed by Betti and van den Berg [1], which we take as a useful example throughout this section to discuss about concept drift. Presumably, as scholars interested in the history of ideas, they started from a detailed analysis of Kant's and Brentano's positions about science. The first analytical step consists therefore in characterizing the two positions as precisely as possible; let us refer to them by S_K and S_B, respectively.[7] Before studying how S_K and S_B are inter-linked, e.g., if they are the *same* concept or not, the historian must consider how S_K and S_B are characterized by their authors and, in particular, what are their features. Suppose that S_K exists at least at context *kant* with $\gamma(S_K, kant) = \phi_K$ and S_B exists at least at context *brentano* with $\gamma(S_B, brentano) = \phi_B$. The case where $\phi_K \neq \phi_B$ makes evident that scholars concluded that the concepts of science of Kant and Brentano have different peculiarities even though they can be potentially assimilated into a general concept of science.

---

[5] We intend features as mereologically *atomic*, i.e., simple concepts without components such that $\gamma(F,t) = \{F\}$ holds during the whole life of any feature F. If features are complex concepts, in their turn, they have simpler features as components, i.e., it is possible to have $Part(F,C,t)$ and $Part(F',F,t)$ with $F' \neq F$. Usually *Part* is considered as transitive, therefore it follows that $Part(F',C,t)$, i.e., F' is also a feature of C. At this point, unless infinitely decomposable concepts exist, it is enough to reiterate the decomposition process until reaching undecomposable concepts and include only them among the features that are used in the characterizations ϕ.

[6] By assuming an *extensional* mereology [12] to model parthood relations between concepts and their features (as components), the possibility of having different concepts with the same features would be excluded.

[7] To simplify our analysis, we assume that Kant and Brentano did not change their positions about science. In presence of changes, the analysis needs to consider all these positions.

The second analytical step aims at understanding the *degree of similarity* between the characterizations ϕ_K and ϕ_B. An interesting case is when ϕ_K and ϕ_B share some common features ($\phi_K \cap \phi_B \neq \emptyset$) but no characterization is included in the other ($\phi_K \nsubseteq \phi_B$ and $\phi_B \nsubseteq \phi_K$). When the features in $\phi_K \cap \phi_B$ are relevant, Kuukkanen would probably commit to the existence of a general dynamic concept S of science with core features included in $\phi_K \cap \phi_B$ and marginal features including $(\phi_K \setminus \phi_B) \cup (\phi_B \setminus \phi_K)$. S, such that $\gamma(S, kant) = \phi_K$ and $\gamma(S, brentano) = \phi_B$, would persist through contexts *kant* and *bolzano* maintaining its core features but changing the marginal ones. Vice versa, Betti and van den Berg would probably introduce a third stable concept S characterized only by $\phi = \phi_K \cap \phi_B$, i.e., such that both ϕ_K and ϕ_B are specializations of ϕ ($\phi \subseteq \phi_K$ and $\phi \subseteq \phi_B$ trivially hold).

In the first analysis it is not clear whether there exists only the dynamic (and general) concept S that changes from Kant to Bolzano or, in addition to S, one needs to consider also the stable concepts of science of Kant and Bolzano. The second analysis clearly commits to three different (stable) concepts: S_K, S_B, and S. The link between S_K and S_B is here established through the intermediation of S that generalizes both S_K and S_B. By relying on the general specialization relation $\sqsubseteq$ (see Sect. 5.1), this indirect link can be introduced even when $\phi_K \cap \phi_B = \emptyset$ (where the approach based on core-features is not applicable).[8]

The individuation of the persistence conditions of S or its level of generality is usually a matter of historical investigation, it is not intrinsically determinated by the characterizations of S_K and S_B. There is a gap between the concepts S_K and S_B on the one hand, and the concept S on the other hand, that may be introduced by historians for analytic purposes. This gap raises the problem of establishing the temporal or contextual extension of concepts. For instance, in the case of the first analysis, we suggested that S exists both at contexts *kant* and *bolzano*. In this view, S is not new, the scholar historically retraces it according to her examination of the definitions provided by Kant and Bolzano. The second analysis is more flexible because the generalization links between S and S_K/S_B are acontextual: $S_K \sqsubseteq S$ and $S_B \sqsubseteq S$ may hold even though S_K exists only at the context *kant*, S_B exists only at the context *brentano*, and S exists only at the context *betti_vanderberg* (i.e., S has been 'created' by Betti and van der Berg). This situation is not precluded by the first analysis, but in this case S would not change from Kant to Bolzano, actually it would not exist at these contexts. Vice versa, in the second analysis one could also pretend that the stable concept S generalizing both S_K and S_B exists at the contexts *kant* and *brentano*.

In general, by starting from stable concepts, historians may adopt some rules to determine continuities among them—to determine the degrees of similarity between their characterizations—and consequently introduce static or dynamic concepts that generalize or group together the concepts that result similar. For instance, starting from ϕ_K and ϕ_B one may individuate a set of shared features or, by using the taxonomy of features, to find an indirect link through a common generalization. Historians could then experiment different rules and study their impact in terms of individuating possible continuities among the original concepts.

However, the analysis of the previous example suggests that the commitment towards dynamic concepts does not provide any analytical advantage to historians. Actually it seems to be less flexible for representing the temporal or contextual extensions

[8] Structural similarities between characterizations allow to individuate more abstract links, see Sect. 5.4.

of concepts. From this perspective, our claim is that the introduction of new dynamic or static concepts can be avoided without the loss of any analytical power. The continuity between the original (stable) concepts can be captured by means of *clusters* of characterizations individuated by means of the selected similarity relation R (usually an *equivalence* relation). Clusters result explicitly determined by scholars through the adoption of precise rules. However, scholars do not have to necessarily commit to the existence of a dynamic or static concept collecting or generalizing all the characterizations in a given cluster. In this perspective, historians can consider different relations R among concepts, explore the effects of these choices in terms of the obtained clusters, and *possibly* introduce concepts that correspond or generalize the characterizations in the clusters. The discussed mechanism facilitates and makes explicit the analytical process under the decisions of historians without forcing a strong commitment on the nature and the persistence of concepts. The existence of static or dynamic concepts and their reduction to given kinds of clusters concern the metaphysical realm rather than the historical one.

5.4. Structural Approaches

The previous sections analyze the characterization of concepts in terms of (possibly varying) features. However, as noted in [1], concepts are often part of a network of interlinked concepts, which contributes to relationally characterize concepts themselves. We explore in this section how to formally represent networks to augment concepts characterization.

Let us assume to represent a network of concepts by a relational structure $\langle \mathscr{C}, R_1, \ldots, R_n \rangle$ where $\mathscr{C}$ is a set of concepts and R_is are relations defined on $\mathscr{C}$. Similarly to concepts, networks are in time and may therefore undergo some changes. Our analysis is however restricted to *stable* networks where R_is are not temporally qualified and the time(s) at which the network exists is indicated by a subscript, e.g., $\langle \mathscr{C}, R_1, \ldots, R_n \rangle_t$ exists at t and it characterizes all the concepts in $\mathscr{C}$ at t. In this sense, one could rethink the γ relation to associate, at t, to each concept in $\mathscr{C}$, the relational structure $\langle \mathscr{C}, R_1, \ldots, R_n \rangle_t$, i.e., all the concepts in $\mathscr{C}$ are simultaneously characterized by the same network.

This approach is more powerful than the one purely based on features. For instance, $\mathfrak{S}_1 = \langle \mathscr{C}_1, R_1 \rangle_{t_1}$ with $\mathscr{C}_1 = \{Red, Blue, Ferrari\}$ and the characterization relation $R_1 = \{\langle Red, Ferrari \rangle\}$ captures the situation where $\gamma(Ferrari, t_1) = \{Red\}$. Similarly, $\mathfrak{S}_2 = \langle \mathscr{C}_2, R_2 \rangle_{t_2}$ with $\mathscr{C}_2 = \mathscr{C}_1$ and $R_2 = \{\langle Blue, Ferrari \rangle\}$ capture the situation where $\gamma(Ferrari, t_2) = \{Blue\}$. However, $\mathfrak{S}_1$ ($\mathfrak{S}_2$) characterizes also *Red* (*Blue*) to be a feature of *Ferrari* at t_1 (t_2). The tracking of concepts from $\mathfrak{S}_1$ to $\mathfrak{S}_2$ is here obtained by the identity mapping $\mu : \mathscr{C}_1 \to \mathscr{C}_2$ and by the correspondence between R_1 and R_2. However additional relations can be introduced among the R_is and different kinds of mappings can be established (see Sect. 6).

Note that in this view the R_is apply to concepts rather than to the individuals that concepts classify. Differently, the notion of *model* considered in [1] refers to *schemata* (i.e., a sort of conceptual graphs) where the relations apply to individuals. For instance, a model where *Finger* and *Hand* are linked by the relation *Part* constrains the instances of *Finger* to be part of the instances of *Hand*[9], a situation different from the previous one where the feature *Red* is part of the concept *Ferrari* (that, in conceptual modeling languages, is usually represented by means of *attributes* of concepts). Betti and van den

[9]This is one of the possible semantics for this model. Different conceptual modeling languages have different semantics and some of them, e.g., UML, do not have a clear semantics.

Berg argue that by using models it is possible to study the "history of concepts that do not share any features in Kukkanen's sense, but which occupy a similar place in the network of concepts specified in a certain model" [1, p.829]. In order to understand what 'similar place' means, consider a second model where *Palm* and *Hand* are linked by the relation *Part*. Intuitively, palms and fingers have something in common, i.e., they are both part of hands, they are both instances of the *relational* concept of *Being part of an hand*. A third abstract model can be introduced with a general concept *Being part of an hand* linked by the parthood relation to *Hand*. In this model, *Being part of an hand* represents a 'place' that abstracts from the different kinds of parts of hands. We are here in presence of a mechanism similar to the one analyzed in the case of taxonomically organized features, however *Finger* and *Palm* are here generalized *in the context* of the relation their instances have with the instances of *Hand*. It should be clear that this comparison does not allow *per se* to understand historical links between concepts. However, the characterization of concepts by means of models allows for more abstract clustering mechanisms; hence, by relying on the mappings between concepts and relations in different networks, this approach captures abstract structural similarities among concepts.

6. The Intentional Dimension of Concepts

Up to now, the similarity relations grounding the construction of clusters rely on the features of concepts and/or on the relations the concepts have with the other concepts in the networks. The similarity between concepts reduces therefore to the similarity between their features and/or the places they occupy in the conceptual networks.

Part of the analysis of historians is however also driven by explicit or implicit reference done by concepts' creators to pre-existing concepts [2]. Scholars are indeed often interested in historical evidence about the background or the influential sources of an author and, especially, in the *intentional* dimension at the basis of the development of a concept. Concepts' authors may not introduce concepts from scratch and may rely on existing concepts in the light of their aims. In this perspective, a certain author could intend to specialize, generalize, modify, radically transform or discard a pre-existent concept to achieve her goals.[10] These links—studied and traced by historians—can be seen as integral information about the *intentional process* of developing concepts, and finally about concepts themselves.

The intentional links can be modeled by introducing a set Δ of *diachronic* relations between concepts. Each relation $\delta \in \Delta$ with form $\delta(C_i, t, C_j, t')$ intuitively models the fact that the concept C_i, as it is at time or context t, is intentionally δ-linked to the pre-existing concept C_j, as it is at time or context t'. For instance, Brentano could explicitly refer to the notion of science of Kant with the intention of introducing a totally different notion. Or, the historian could retrace that Brentano known the work of Kant and that he started from this work to develop his own concept of science. First note that each $\delta \in \Delta$ is a relation between concepts, not between their instances. Thus, as observed in Sect. 5.4, these relations are not part of models as intended in [1]. Second, by linking concepts that exist at different times or contexts these relations cannot be part of any $\langle \mathscr{C}, R_1, \ldots, R_n \rangle_t$ structure where the R_is are always synchronic (even when they apply to concepts). Thus,

[10] Think about the classical debates in philosophy, e.g., on the concept of *universal* where a philosopher reinterprets what done by previous philosophers in order to propose a new philosophical theory.

the relations in Δ capture *intentional dependencies* that complement the characterization of concepts based on features and models with a *historical* perspective.

Δ-relations may be exploited to refine the similarity relations R and to individuate (dis-)continuities in the history of concepts in terms of the intentions of creators. There may be also cases in which authors could explicitly claim conceptual innovations or transformations which are not directly reflected in the created concepts, i.e., their intentions can be decoupled from the way in which concepts are characterized in terms of features and models. The act of claiming some intentions, expressed via the Δ-relations, about a concept C_i (at t) with respect to a pre-existing concept C_j (at t') does not automatically translate into a relation between the features or the models in $\gamma(C_i,t)$ and $\gamma(C_j,t')$. They rather make explicit the intentions of the authors that may be only marginally translated into the produced characterization $\gamma(C_i,t)$. The intentions of authors may then play a central role for understanding the created concepts and for studying their behavior through time, a crucial aspect in a historical perspective.

In general, one could think to represent the intentional dimension of concepts by means of features. Rather than $\delta(C_i,t,C_j,t')$ one could introduce in $\gamma(C_i,t)$ the feature F_δ such that $F_\delta(C,t)$ if and only if $\delta(C,t,C_j,t')$. This kind of features is taken into account neither by Kuukkanen [8] nor by Betti and van den Berg [1] and, as discussed in Sect. 5.2, they would be properties of concepts (rather than properties of their instances). Note that, using a similar strategy, the ‘places’ in the networks of concepts could also be reduced to features. A deeper analysis on the ontological nature of features and on the possibility to attribute them an exclusive role in the characterization of concepts is here required. An alternative, and maybe more interesting, possibility consists in introducing the Δ-relations as *modal* mappings between the concepts in the models that characterize the concepts C_i and C_j. The general idea is that the concepts in a model could be mapped to the concepts in another model in different ways, according to different modalities and intentionalities that can then be exploited by in the similarity relations R. Again, further work on a deeper analysis of the nature of these intentional relations and on the way they can be used is required.

7. Conclusions

We discussed in the paper some theoretical issues concerning the modeling of concepts and concept drift for historical research in the context of the Digital Humanities. As said in Sect. 1, this analysis is a first step for a robust modeling approach leading to an ontology-based information system for the organization of historical data.

To sum up the results of our study, the phenomenon of concept drift, understood as the possibility for a concept to change while keeping its identity, remains an open topic for theoretical research. The challenge is to define concepts’ persistence conditions, which is not a simple task given the abstract (and perhaps even vague) nature of concepts. Our idea is that, in a historical perspective, this debate is secondary. We proposed to first study the concepts introduced by the authors by considering the way in which they are characterized in terms of features and/or models. Second, to introduce similarity relations R among characterizations (and consequently among concepts) that can be used to indivuate *clusters* of characterizations (and concepts). The study of these clusters and the possibility of experiment alternative similarity relations generating different clusters offer to historians a powerful analytical tool that is uncommitted with respect to the fact that these clusters correspond or not to (new or already existing) concepts. In this approach, one is not committed towards a specific stance in concept drift while being able at the same time to investigate possible continuities or discontinuities among various concepts.

Finally, differently from the state of the art, we proposed to take into account the creation process from which concepts originate by considering, in particular, the authors' intentional attitude towards pre-existing concepts from which the new concepts derive. This allows to make sense of the fact that a concept is created by 'manipulating' some pre-existing concepts. In this view, the building process of a concept, the way it originates from pre-existing concepts, becomes part of the characterization of the concept itself.

Further work is necessary at both the theoretical and implementation level. From the former perspective, as said, a deeper understanding of the nature of features and of concepts' characterization is due. In particular, it may be interesting to understand the kind of relations holding between concepts when some derive from the others. From the latter perspective, concepts' characterization (in terms of the function γ) can be useful to semi-automatically organize or cluster concepts. This may lead to an algorithmic procedure to aid decision making for concepts organization in historical research.

References

[1] A. Betti and H. van den Berg. Modelling the history of ideas. *British Journal for the History of Philosophy*, 22(4):812–835, 2014.

[2] J. P. Burkholder. The uses of existing music: Musical borrowing as a field. *Notes*, 50(3):851–870, 1994.

[3] A. Fokkens, S. Ter Braake, I. Maks, D. Ceolin, et al. On the semantics of concept drift: Towards formal definitions of semantic change. In E. S. Darányi, L. Hollink, A. M. Peñuela, and E. Kontopoulos, editors, *Proceedings of the 1st Workshop on Detection, Representation and Management of Concept Drift in Linked Open Data (Drift-a-LOD)*, volume 1799, pages 10–17. CEUR, 2016.

[4] A. Gallois. Identity over time. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2016 edition, 2016.

[5] P. Garbacz, A. Ławrynowicz, and B. Szady. Identity criteria for localities. In S. Borgo, P. Hitzler, and O. Kutz, editors, *Proceedings of the 10th International Conference on Formal Ontology in Information Systems: (FOIS 2018)*, volume 306, pages 47–54. IOS Press, 2018.

[6] D. Geeraerts. *Theories of lexical semantics*. Oxford University Press, 2010.

[7] R. Koselleck. *Futures past: on the semantics of historical time*. Columbia University Press, 2004.

[8] J.-M. Kuukkanen. Making sense of conceptual change. *History and theory*, 47(3):351–372, 2008.

[9] M. Lamé. Heterotoki: non-structured and heterogeneous terminology alignment for digital humanities data producers. In A. Poggi, editor, *Proceedings of the First International Workshop on Open Data and Ontologies for Cultural Heritage*, volume 2375. CEUR, 2019.

[10] E. Margolis and S. Laurence. *Concepts: core readings*. Mit Press, 1999.

[11] L. O. Mink. Change and causality in the history of ideas. *Eighteenth-Century Studies*, 2(1):7–25, 1968.

[12] P. Simons. *Parts: A study in ontology*. Clarendon Press, 1987.

[13] J. Wilson. Determinables and determinates. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2017 edition, 2017.