A question of priority

van Rooij, R.; Schulz, K.

DOI
10.1007/978-94-017-8813-7_13
Publication date
2014
Document Version
Author accepted manuscript
Published in
Formal approaches to semantics and pragmatics

A Question of Priority*

Robert van Rooij and Katrin Schulz

Inst. for Logic, Language and Computation
Universiteit van Amsterdam
Oude Turfmarkt 141-147
Amsterdam, The Netherlands
0031-20-5254541

Abstract

Properties as set of individuals, or of features? Worlds, or propositions? Time-points, or events? Preference, or choice? Natural kinds, or similarity? In modern analytic philosophy it is standard to take (i) individuals as basic, and properties as defined in terms of them; (ii) worlds as basic, and propositions as defined in terms of them; (iii) time-points as basic, and intervals as constructions out of them; (iv) preference as basic, and optimal choice as defined in terms of them; and (v) natural kinds as basic, and similarities as defined in terms of them. In this paper we show that in all cases the other direction is possible as well. Most of the constructions used are well-known. But by putting them collectively on the table we hope to show that the constructions have something in common, and that it is not always clear which perspective is ontologically less committing.

1 Properties: sets of individuals or of features?

Logic started with Aristotelian syllogistic reasoning. Almost all modern text-books use Venn diagrams, and thus an *extensional semantics*, to decide whether a syllogistic inference is valid. It is basically assumed that the terms denote non-empty sets of individuals. Somewhat more generally, the idea is to start with a partially ordered set $\langle \mathcal{A}, \leq \rangle$ (where '$\leq$' is a reflexive, transitive, and

*We would like to thank the anonymous reviewers for this volume, as well as the volume editors.

anti-symmetric relation on $\mathcal{A}$) such that all subsets B of $\mathcal{A}$ have a greatest lower bound. (b is the greatest lower bound of B iff (i) $\forall x \in B : b \leq x$, and (ii) $\forall y \in \mathcal{A}$: if $\forall x \in B : y \leq x$, then $b \leq y$.). The greatest lower bound of $\mathcal{A}$ itself we denote by $\perp$. We assume that every term T of the language denotes an element of $A - \{\perp\}$, i.e., $[[T]] \in \mathcal{A} - \{\perp\}$. A universal sentence like 'All S are P', represented as SaP, is counted as true iff $[[S]] \leq [[P]]$. A particular sentence like 'An S is P', represented by SiP, is counted as true iff the greatest lower bound of $[[S]]$ and $[[P]]$ is not equal to $\perp$, $glb\{[[S]], [[P]]\} \neq \perp$.

According to Leibniz, an extensional semantics is not the most natural way to interpret the meaning of syllogistic terms and syllogistic inference. Rather – or so he proposed – we should start out with an *intensional* semantics. According to it, SaP is true if the *intension* of P is contained in the *intension* of S. Following Rescher (1954), we will think of the intension of S as the set of attributes associated with S, and will call a semantics *intensional* iff it doesn't crucially refer to individuals. One way to think of this – in fact, this is arguably exactly how Leibniz thought of it – is to start out with a (semi-)lattice, instead of with a partial order. Indeed, let us start with a semi-lattice $\langle \mathcal{A}, \circ \rangle$, where '$\circ$' is a binary operation on A that is idempotent ($\forall x \in A : x \circ x = x$), commutative $\forall x, y \in A : x \circ y = y \circ x$), and associative ($\forall x, y, z \in A : (x \circ y) \circ z = x \circ (y \circ z)$). In terms of this structure, we can say that SaP is true iff $[[S]] \circ [[P]] = [[S]]$. To also interpret particular sentences, we need to assume that the semi-lattice also has an *extreme element* e, meaning that $\forall x \in \mathcal{A} : x \circ e = e$. Now we can SiP is true iff $[[S]] \circ [[P]] \neq e$.

What should be the intuitive interpretation of '$\circ$'? Meet, or join? In fact, it doesn't matter much. It is clear is that if we interpret $\langle \mathcal{A}, \circ \rangle$ as a meet semi-lattice, it corresponds exactly with the partially ordered set $\langle \mathcal{A}, \leq \rangle$ closed under greatest lower bounds. The extreme element e of the semi-lattice would intuitively correspond with $\perp$, a primitive notion of *inconsistency*. But we might as well interpret $\langle \mathcal{A}, \circ \rangle$ as a join semi-lattice, which would correspond with the same partially ordered set closed, this time, under smallest upper bounds. In fact, if we want to interpret Leibniz' semantics intensionally, the latter interpretation is the way to go. But what should in this case the extreme element e be associated with? By making use of an explicit intensional interpretation, we will explain that also now e should correspond with a primitive notion of inconsistency.

To give a set theoretic intensional semantics for syllogistics without negative terms, we have to start at least with a primitive set of attributes A and an interpretation function '$[[\cdot]]$' that assigns sets of attributes to terms. It is quite clear how to provide a semantics for sentences of the form SaP. This universal sentence is true iff $[[P]] \subseteq [[S]]$. But the problem is how to provide a semantics for particular sentences: SiP. The first idea that came to Leibniz's mind given in set-theoretic terms would be to say that SiP is true iff $[[S]] \cap [[P]] \neq \emptyset$. But this idea is clearly non-sensical: some bike is red, but there is nothing in the intension of 'red' that is also in the intension of 'bike', or so it seems. Or even more obviously, the sentence 'No gold is silver' is obviously true. According to the above suggestion this is true iff there is no attribute, or property, that gold and silver share. But there is obviously one: metal. What has to be assumed, rather, is the following idea: the intensions of 'red' and 'bike' are *not incompatible*. What this means is that also for our intensional interpretation, we must assume that a *primitive* relation of *(in)compatibility*. And the relation is one *between primitive features*.[1]

The fact that we have to assume such a notion of (in)compatibility already suggests why Leibniz had a hard time to come up with a satisfying characteristics for even simple syllogistic logic. Just like Wittgenstein when he was writing his Tractatus, also Leibniz thought of his simple terms, or attributes, as being *logical independent* of each other, i.e., their being mutually compatible with all other simples (cf. Ishiguro, 1972, p. 54): only if the simples are logically independent of each other is it possible to construct a language where inference and equivalence can be checked 'from the surface'. To check validity we don't have to know what the interpretation of the different terms is. But if all terms are interpreted by sets of these simple attributes, a sentence like 'No gold is silver' can never be true. In our following interpretation, we assume an incompatibility relation $\perp$.

Let $M = \langle A, [[\cdot]], \perp \rangle$ be a model, with A a set of attributes, $[[\cdot]]$ an interpretation function which assigns to each primitive term T a subset of A, $[[T]] \subseteq A$, and $\perp$ a symmetric and irreflexive relation between elements of A. If for two elements $x, y \in A$ it holds that $x \perp y$, we say that the attributes x and y are incompatible. We will denote by Δ the set of subsets of A which contain such mutually incompatible elements: $\Delta = \{S \subseteq A : \exists x, y \in S : x \perp y\}$. We assume that for each primitive term T, $[[T]] \notin \Delta$, and that the set of supersets of $[[T]]$ does not equal the set of all maximally consistent sets of A. Now we say that $[[SaP]] = 1$ iff $[[S]] \cup [[P]] = [[S]]$ iff $[[S]] \supseteq [[P]]$, and $[[SiP]] = 1$ iff $[[S]] \cup [[P]] \notin \Delta$. Thus, SiP is true iff S and P do not contain mutually incompatible attributes. SoP and SeP are interpreted as the negations of SaP and SiP, respectively. If we say that $\phi_1, ..., \phi_n \models \psi$ iff for all models in which the premisses are true, the conclusion is true as well, this semantics validate all and only all arguments in classical syllogistic style if and only if they are

[1]If we think of the extensional counterpart, this means that 'some bike is red' is true not because there actually exists a red bike, but rather that it is possible that such a bike exists. And indeed, what Leibniz considers to be the extension of a term (a set of individuals scattered around all worlds) is very much what in possible worlds semantics is its intension (cf. Leibniz (1966b) and Ishiguro (1972, p. 49)).

traditionally counted as valid.

2 Worlds and individuals: primitives, or maximal sets?

In modal logic it is standard to think of worlds as primitive entities. David Lewis (1973, 1986) even believed that they are universes that really exist. But if we want to prove completeness results one does this by thinking of worlds as maximally consistent sets of sentences or propositions. In that case one *defines* worlds in terms of other primitives: propositions and a notion of (in)consistency. And indeed, a number of authors have proposed to think of possible worlds in exactly this way, as total state descriptions (e.g. Carnap, 1947). Lewis (1973) used a primitive similarity relation between possible worlds to account for counterfactual conditionals. Proponents of premisse semantics (Veltman, 1976; Kratzer, 1981) have argued that it is more illuminative to *construct* such a comparative similarity relation in terms of sets of propositions. Turner (1981) went even so far that these propositions themselves should be taken as primitive, instead of the possible worlds. We have seen above that Leibniz favored an interpretation of syllogistic logic in which the terms denote sets of features, instead of sets of individuals. But how then to think of individuals? Although it is not very clear whether he takes individuals to be primitive as well, he at least also thinks of individuals as maximally consistent sets of properties (Leibniz, 1686). Indeed, in agreement with the ontology of syllogistic reasoning, Leibniz doesn't seem to make a clear ontological distinction between properties and individuals, it is just that individuals are the maximally consistent ones.

For Leibniz, all properties were on a par. Indeed, properties were taken to be closed under logical operations like complementation, disjunction, and conjunction.[2] The early Russell was strongly influenced by Leibniz, and during his 'realistic' period when Russell believed in the existence of real universals, he also wanted to define (in Russell, 1912) individuals in terms of properties. However, in constrast to Leibniz, he didn't want to use *all* properties, only the natural ones: the universals. In distinction with standard properties, universals are *not* taken to be closed under complementation and disjunction. Thus, also Russell defined individuals somehow as maximal sets of properties, constrained by a notion similar – though not identical – to that of consistency. The notion now was that of the primitive relation of '*compresence*'. Even though constructivists have to take a notion like 'consistency' to be primitive, we have

[2]This is what he believed, but he was not able to work out a full semantics for syllogisms with complex terms.

an intuitive understanding of what it is supposed to be. For 'compresence' this is somewhat more difficult. Still, the notion of compresence is very much like consistency: it is a reflexive and symmetric relation, though now restricted to properties. Intuitively, the compresence relation should just like consistency not be transitive: we can imagine two distinct individuals, one having only universals P and Q, and a second only having Q and R. If compresence were transitive, it would mean not only that we also had an individual with universals P and R, but also that if we think of individuals as *maximal* compresent sets of properties, there would in fact be only one individual having all properties P, Q, *and* R. There could be no two different individuals sharing a single property. But thinking of the 'compresence' relation as being non-transitive still gives rise to problems. Intuitively, we can imagine three distinct individuals, one having only universals P and Q, one only having properties Q and R, and a third only having properties P and R. But if we think of individuals as *maximal* sets of compresent properties, this is impossible: in such cases there also *has to be* a fourth individual having all three properties. For the construction of individuals out of natural properties this problem didn't receive a lot of attention. The very similar problem discovered in the trial to construct universals, or natural properties, out of particulars together with a primitive similarity relation, however, received a lot of attention. We will discuss this problem in a later section. Before that, however, we will consider first a more successful constructive move made by Russell to think of instants as maximal sets of intervals.

3 Instants, or Events?

Events It is standard to take instant time points, and the 'before'-relation between them, to be basic, and to define intervals in terms of them as convex sets of instants. Russell (1914, 1936), however, proposed to go the other way around: temporal instants should be constructed from what he calls events. His motivation is that he wants to show that our conception of abstract instants is derived from (reconstructed from) the events and the temporal relationships we perceive. Similar things have been done by Whitehead and Wiener, and more recently by Kamp, van Benthem, and Thomason, and after that in AI by people like Allen and Hayes. They all start with an *event structure*. An event structure, $\langle E, \prec \rangle$ is just a set of events that is temporally ordered: $e \prec e'$ means that e is temporally completely before e'. According to some, events should just be strictly partially ordered (irreflexive and transitive). Most authors, however, assume something stronger. They assume that events give rise to what is known as an *interval order*:

Definition 1 *An <u>interval order</u> is a structure $\langle X, R\rangle$, with R a binary relation on X that is irreflexive, and satisfies the interval order condition (IO):*
(IR) $\forall x : \neg R(x,x)$.
(IO) $\forall x, y, v, w : (R(x,y) \wedge R(v,w)) \rightarrow (R(x,w) \vee R(v,y))$.

Notice that any interval order is also a strict partial order, because from (IR) and (IO) one can immediately derive that the ordering is also transitive.

The difference between starting with a strict partial order or with an interval order is that according to the former it might be unclear how some events are temporally related to one another, while this is (almost) impossible according to the latter approach. To see this, let us define a relation '$\sim$' as follows: $e \sim e'$, iff$_{def}$ $e \not\prec e' \wedge e' \not\prec e$. Notice that from this definition it follows that (i) $\sim$ is reflexive and symmetric, but need not be transitive, (ii) $\prec$ and $\sim$ are disjoint, and (iii) $\prec \cup \sim$ is complete. If we take $\langle E, \prec\rangle$ to be an interval order, '$\sim$' intuitively represents 'temporal overlap'. But if $\langle E, \prec\rangle$ is just a strict partial order, it might also be that $e \sim e'$ because the events are temporally *incomparable*. To illustrate this, it is possible (Fishburn) to represent interval orders in terms of, yes, *intervals* of the real line: f is a function from events to intervals of $\mathbf{R}$. $x \prec y$ then means that all points of $f(x)$ are before all points of $f(y)$, and $x \sim y$ means that $f(x)$ and $f(y)$ have a non-empty intersection. Such a representation is *not* possible for strict partial orders: even if $x \prec y$ and $v \prec w$ it might still be possible that both x *and* y are incomparable with both v *and* w.

In terms of event structures we can define a relation of *temporal inclusion* between events '$\sqsubseteq$'. If $\langle E, \prec\rangle$ is an interval order, we can define $e \sqsubseteq e'$ iff$_{df}$ iff $\forall e'' : e'' \sim e \rightarrow e'' \sim e'$. It is easy to see that now '$\sqsubseteq$' is reflexive and transitive, and that $\sqsubseteq$ means 'temporally included' as intended if '$\sim$' means 'temporal overlap'. But the latter is only the case if $\langle E, \prec\rangle$ is an interval order.

We can also define a notion of temporal inclusion between events if $\langle E, \prec\rangle$ is a strict partial order. Define $e \sqsubseteq e'$ iff$_{def}$ $\forall e''[e' \prec e'' \rightarrow e \prec e''] \wedge \forall e''[e'' \prec e' \rightarrow e'' \prec e]$. It is easy to prove that '$\sqsubseteq$' is reflexive and transitive, and thus a *pre-order*. But it need not be a satisfy antisymmetry, and thus '$\sqsubseteq$' does not (necessarily) give rise to a partial order.

Events structures can be *atomic* or give rise to *endless descent*. Event structure $\langle E, \prec\rangle$ is atomic iff $\forall e \in E : \exists e' \sqsubseteq e : \forall e'' \sqsubseteq e' : e'' = e'$. An event structure gives rise to endless descent iff $\forall e \in E : \exists e' \sqsubseteq e : e' \neq e$.

Intervals In terms of the pre-order, '$\sqsubseteq$', we can define a new relation, '$\approx$', as follows: $e \approx e'$ if$_{def}$ $e \sqsubseteq e' \wedge e' \sqsubseteq e$. It is obvious that this relation is an equivalence relation. Now we can define *intervals* as equivalence classes of events, and we can derive a new structure, $\langle [E]_{\approx}, <, \sqsubseteq^{*}\rangle$, with $I < J$ iff$_{def}$

$\exists e \in I, e' \in J : e \prec e'$ and $I \sqsubseteq^* J$ iff$_{def}$ $\exists e \in I, e' \in J : e \sqsubseteq e'$. One can show that if $\langle E, \prec\rangle$ is a strict partial order/interval order, then (i) $\langle [E]_{\approx}, \sqsubseteq^*\rangle$ is a partial order, and (ii) $\langle [E]_{\approx}, <\rangle$ is a strict partial order/interval order. An interval structure is atomic, or gives rise to endless chains iff the corresponding event structure is.

Instants Now Russell defines instants as maximal sets of pairwise overlapping intervals.

Definition 2 *Let $\Sigma_I = \langle [E]_{\approx}, <\rangle$ be an interval structure. An instant t is a subset of $[E]_{\approx}$ such that: (i) $\forall I, J \in t : I \sim J$ and (ii) $\forall I \notin t : \exists J \in t : I \not\sim J$ (in other words, an instant is a maximal subset of $[E]_{\approx}$ where $\forall I, J \in t : I \sim J$)*

We denote the set of instants of Σ_I as $T(\Sigma_I)$.

Definition 3 *Let t and t' be any two instants of $T(\Sigma_I)$. Then: $t <^* t'$ iff$_{def}$ $\exists I \in t : \exists J \in t' : I < J$. We call $\tau(\Sigma_I) = \langle I(\Sigma_I), <^*\rangle$ the instant structure derived from Σ_I.*

Most important in Russell's construction is the following theorem:

Theorem 1 *$\tau(\Sigma_I)$ is a linear order, if Σ_I is an interval order. (It is a strict partial order if Σ_I is).*

A linear order demands that all elements of a set are comparable. In this sense, linear orders are very informative. Intuitively, however, it seems that the ordering between instants is stronger than a linear order. For one thing, the order seems to be *dense*. An ordering $\langle X, R\rangle$ is dense iff $\forall x, y : R(x, y) \rightarrow \exists z : R(x, z) \wedge R(z, y)$. The set of natural numbers is not dense, because there is no natural number between, say, 1 and 2, but the set of rational numbers is. Russell showed that the density of the ordering relation between instants can be derived from some constraints on the ordering between events (intervals). To account for density, Russell (1936) proposed the following constraint on interval orders:

Let $\Sigma_I = \langle [E]_{\approx}, <\rangle$ be an interval order. The following condition suffices to make the ordering on instants $\tau(\Sigma_I) = \langle I(\Sigma_I), <^*\rangle$ being dense:[3]

$$\text{For all } I, J \in [E]_{\approx} : I < J \rightarrow \exists K, L \in [E]_{\approx} : I < K \sim L < J.$$

[3]Interestingly, Allen & Hayes define the notion of 'meet' ':', as follows:
$I : J$ iff$_{def}$ $I < J \wedge \neg\exists K, L(I < K \wedge K \sim L \wedge L < J)$.

Although this condition is sufficient, it is *not necessary* for $\tau(\Sigma_I)$ being dense.[4] If an ordering relation is not dense, it is *discrete*. There exists an alternative way to say that $\tau(\Sigma_I) = \langle I(\Sigma_I), <^* \rangle$ is dense/discrete. It is easy to show that in case $\langle E, < \rangle$ is atomic, the ordering relation between instants is discrete, and in case the interval structure gives rise to endless descent, the ordering relation between instants is dense.

From instants to intervals and back Suppose we start with a primitive instant structure $\langle T, <^{**} \rangle$, where T is a set of instants, and $<^{**}$ a primitive relation between instants. It is very natural to assume that $\langle T, <^{**} \rangle$ is a strict partial order or even a linear order. Now we can define intervals as non-empty convex subsets of T, where I is *convex* iff $x \in I \wedge z \in I \wedge x <^{**} y <^{**} z \rightarrow y \in I$. More in particular, assume that i and j are points, then we can define an interval as follows: $\{t \in T : i <^{**} t \wedge t \leq^{**} j\}$.[5] Let us now take $I(T)$ to be the set of intervals so constructed. Define the interval structure based on $\sigma_T = \langle T, < \rangle$, $\iota(\sigma_T)$, as the structure $\langle I(T), <, \sqsubseteq \rangle$ where $I < J$ iff $\forall i \in I, j \in J : i <^{**} j$ and $I \sqsubseteq J$ iff $I \subseteq J$. Now one can show that in the structure $\langle I(T), <, \sqsubseteq \rangle$, '$<$' is an interval order if '$<^{**}$' is a linear order, $\sqsubseteq$ a partial order that satisfies (CONJ), and where '$<$' and '$\sqsubset$' satisfy (MON) and (CONV). Moreover, $\langle I(T), < \rangle$ is atomic if $\langle T, <^{**} \rangle$ is discrete, and it has endless descent if $\langle T, <^{**} \rangle$ is dense.

Above we have seen that from interval orders we can derive a linearly ordered instant structure. Thus, this can also be done for $\iota(\Sigma_T)$. Call the result $T(\iota(\Sigma_T))$. From $T(\iota(\Sigma_T))$ we can derive an interval order again, and it is possible to show that $\iota(\Sigma_T)$ is isomorphic to this new interval order. From this new interval order we can derive an instant structure again, and it is possible to show that this new instant structure is isomorphic to $T(\iota(\Sigma_T))$. This process can be continued indefinitely.

4 Orderings, or choice?

Maximizing choice functions In the theory of choice it is standard to take a comparative preference order to be basic. Analoguously, in the possible world theory of counterfactuals of David Lewis (1973), a comparative similarity relation between worlds is taken to be primitive. However, a much discussed

[4] For a necessary condition, see Lück (2006). In this paper it is also proven under which circumstances one can generate a *continuous* order of instants.

[5] Of course, it is not necessarily to define intervals as having open beginnings and closed ends. The other way is possible as well. Just to assume that it any convex set is an interval doesn't give rise to endless descent even if $\langle T, <^{**} \rangle$ is dense.

topic in the theory of choice is how a preference order among options can be derived on the assumption that the notion of *choice* is primitive. In the semantic analysis of counterfactuals (and of belief revision, for instance) a similar question has been addressed: can we define an ordering from natrual constraints on choice functions? And, of course, this is how Stalnaker (1968) started.

Assuming a choice function that selects an element from each finite set of options, one can easily show how we can generate a linear order by putting constraints on how this function should behave on different sets of options. Let us define a *choice structure* to be a triple $\langle X, O, C\rangle$, where X is a non-empty set, the set O consists of all finite subsets of A, and the choice function C assigns to each finite set of options $o \in O$ an element of o, $C(o)$, satisfying the following condition:

(LIN) $\forall o, o' \in O$: If $(C(o) \in o'$ and $C(o') \in o)$, then $C(o) = C(o')$.

If we say that $x > y$, iff$_{def}$ $C(\{x, y\}) = x$, one can easily show that the ordering as defined above gives rise to a *linear order*

Arrow (1959) already showed how we can generate a strict weak ordering by putting other constraints.

Definition 4 *A* <u>*strict weak order*</u> *is a structure* $\langle X, P\rangle$*, with* P *a binary relation on* X *that is irrreflexive (IR), transitive (TR), and almost connected (AC):*
(IR) $\forall x : \neg P(x, x)$.
(TR) $\forall x, y, z : (P(x, y) \wedge P(y, z)) \rightarrow P(x, z)$.
(AC) $\forall x, y, z : P(x, y) \rightarrow (P(x, z) \vee P(z, y))$.

In this case, the choice function C assigns to each finite set of options $o \in O$ a *subset* of o, $C(o)$. Arrow (1959) stated the following principle of choice (C), and the constraints (A1) and (A2) to assure that the choice function behaves in a 'consistent' way:

(C) $\forall o \in O : C(o) \neq \emptyset$.
(A1) If $o \subseteq o'$, then $o \cap C(o') \subseteq C(o)$.
(A2) If $o \subseteq o'$ and $o \cap C(o') \neq \emptyset$, then $C(o) \subseteq C(o')$.

If we say that $x > y$, iff$_{def}$ $x \in C(\{x, y\}) \wedge y \notin C(\{x, y\})$, one can easily show that the ordering as defined above gives rise to a *strict weak order*.

Condition (A1) is better known as Sen's *Property* α. Condition (A2) is also known as Sen's Property β^+. Taken together with "(EMPTY) If $o \subseteq o'$ and $C(o') = \emptyset$, then $C(o) = \emptyset$" (also assumed by Lewis, and which follows from

(C)), it implies both (II) and (III) discussed below. Arrow formulated the choice function as the combination of (A1) and (A2), and called it the axiom of independence of irrelevant alternatives:

(A) If $o \subseteq o'$ and $o \cap C(o') \neq \emptyset$, then $C(o') \cap o = C(o)$.

While condition (A1) expresses some kind of 'contraction consistency' in proceding from larger menus to smaller ones, the following condition proceeds from smaller menus to larger ones:

(II) $C(o) \cap C(o') \subseteq C(o \cup o')$.[6]

The following axiom is known as *Aizerman's axiom*:

(III) If $o \subseteq o'$ and $C(o') \subseteq o$, then $C(o) \subseteq C(o')$

Taken together with (A1), the superset axiom implies (III). Condition (III) is independent of condition (II), even in the presence of condition (A1).

Suppose that O consists of all finite subsets of I. We can state facts like the following: Ordering $<$ is *acyclic*, if C satisfies EMPTY and (A1); the ordering is *transitive*, if C also satisfies (III). If C satisfies EMPTY and (A2), then $<$ is almost connected. Another sufficient condition for $<$ to be almost connected is for C to be closed under arbitrary union and satisfying (EMPTY) and (A2). Thus, for natural properties the preference relation has, there correspond 'natural' constraints on choice functions. It is not clear what should be taken as primitive.

Satisficing choice functions We would like to derive the meaning of 'better than' in terms of the meaning of 'best' – as is assumed if agents are taken to be utility maximizers –, but rather to derive the meaning of 'better than' in terms of the context-dependent meaning of 'good'.[7] What is crucial for the interpretation of the results of our paper is that although 'good' seems to obey axiom (A2), axiom (A1) seems much too strong: (A1) demands that if both x and y are considered to be good in the context of $\{x, y, z\}$, both should considered to be good in the context $\{x, y\}$ as well. But that is exactly what we don't want for a context dependent notion of 'good': in the latter context, we want it to be possible that only x, or only y, is considered to be good.

[6] This axiom is a finitary version of Sen's Property γ.

[7] Interestingly enough, this is exactly analogue to what Klein (1980) intended to do in linguistics: the meaning of 'taller than' (or 'better than') should be defined in terms of the meaning of 'tall' (or 'good'), not that of 'tallest' (or 'best').

We should conclude that if we want to characterize the behavior of 'good', we should give up on (A1). Unfortunately, by just constraints (C) and (A2) we cannot guarantee that the comparative relation 'better than' behaves as desired. In particular, we cannot guarantee that it behaves almost connected.

To assure that the comparative behaves as desired, we add to (C) and (A2) the Upward Difference-constraint (UD), proposed by Van Benthem (1982). To state this constraint, we define the notion of a difference pair: $\langle x, y\rangle \in D(o)$ iff$_{def}$ $x \in C(o)$ and $y \in (o - C(o))$. Now we can define the constraint:

(UD) $o \subseteq o'$ and $D(o') = \emptyset$, then $D(o) = \emptyset$.

In fact, van Benthem (1982) states the following constraints: No Reversal (NR), Upward Difference (UD), and Downward Difference (DD) (where o^2 abbreviates $o \times o$, and $D^{-1}(o) =_{def} \{\langle y, x\rangle : \langle x, y\rangle \in D(o)\}$):

(NR) $\forall o, o' \in O : D(o) \cap D^{-1}(o') = \emptyset$.
(UD) $o \subseteq o'$ and $D(o') = \emptyset$, then $D(o) = \emptyset$.
(DD) $o \subseteq o'$ and $D(o) = \emptyset$, then $D(o') \cap o^2 = \emptyset$.

One can show that if constraints (NR), (UD) and (UD) are satisfied, the preference relation '>' as defined before still has the same properties as before: it is still predicted to be a strict weak order. In van Rooij (2012) it is shown that other constraints on the satisfying choice function gives rise to other ordering structures, making it less than obvious, again, to assume that either the one (preference order), or the other (choice), should be taken as primitive.

5 Universals, or Similarity ?

There are not many laws named after philosophers. But Leibniz' law is an exception: it states that 'two' objects are *identical* if and only if they are indescernible, i.e. when they share *all* their properties. The notion of 'similarity' is closely related with that of 'identity', and Leibniz expressed ideas about this notion as well: he claimed that x is similar to y if and only if x and y share *at least one* property. Goodman (1972) famously argued that the notion of similarity thus defined is useless. Assuming that properties are sets, and that all sets are on equal footing, it indeed follows immediately that any two objects have a property (set) in common. On a similar assumption one can also easily prove that even the comparative notion of similarity 'y is more similar to x than z' is useless: there are exactly as many sets of which x and y are elements than there are of which x and z are elements. Thus, we should not work with any notion of 'overall similarity'. At best, we should have a *relative* notion of similarity: 'x is similar to y *in respect* r'. Unfortunately, according to Goodman, once we introduce such 'respects' the notion of 'similarity' plays no role anymore: the respects do all the work. Suppose, for instance, that we say that we take r to be 'red'. What then would be the use of similarity? 'x is similar to y *in respect* r' would now be true just because both x and y are red.

There is a lot to say about Goodman's arguments, most obviously his equation of properties with sets. But we are not so much interested in the issue how to *define* a notion of similarity. Our major concern is what we can do with it, once we have such a notion. We will discuss whether we can explain natural properties, or universals, in terms of them.

The problem of universals What makes it that we can 'group' several objects or individuals together under a general term, and that it is more natural to divide the world up in one way than in another? This is basically the very old problem of universals that is with us ever since Plato and Aristotle wrestled with it. For Plato and Aristotle the answer to the problem was (relatively) simple: we divide the things around us up in the way we do because this is the way *reality* is cut to its joints. According to *realists* like them, we classify the world up in cats, dogs, humans and trees, because this is the way the world is divided, and x and y are both cats, because they both have a property in common: cathood. This picture is indeed very natural. The animal kingdom, for instance, is divided in species and what these species are has to be *discovered* by the methods of classification. Individuals do not belong to the same species because they are similar, but they are similar because they belong to the same species, just as Leibniz had it. Thus, a particular classification could be shown to be wrong. Natural as it may seem, the realist' position also gives rise to difficulties: how should we think of these species exactly? Is it something that exists in addition to the cats themselves? And can it even exist without being instantiated? This last question seems unnatural for species, but is more natural for the properties denoted by other general terms, including adjectives. Plato answered all these questions positively, but never made it very clear what it means to 'have' a property. Aristotle, on the other hand, was a more down-to-earth philosopher than Plato, and didn't like uninstantiated universals. But could, as a result, not make very clear what he assumed species and other universals to be.

Suspicious of abstract objects for which they could not determine identity conditions, nominalists rejected the existence of universals. According to them, nature produces individuals and nothing more. Species have no actual existence in nature, they are not the 'objects' we denote by general terms.

General terms have just been invented in order that we be able to refer to a great numbers of individuals collectively. General terms can be used because we classify the world. This classification is conventional and can be neither right nor wrong. It is not a theory, but merely a way of summarizing information in an inteligible form. One assesses its value by consideration of its *usefulness*.

Still, it appears that the conventional classification of the objects around us in groups has to be based on something. The animals of the world can be classified in infinitely many ways. How come that all cultures divide up this class in similar ways in terms of cats, dogs, etc.? One suggestion would be that this says a good deal about the ways we humans are able to classify. This is no doubt true, but if we gathered this ability by natural selection (Quine, 1969), it must say a good deal about the world as well: why else could this way to categorize be so useful? A nominalist, however, doesn't want to say that this means that our categories correspond with real species out there in the world, as the realist has it. He will say at most that what we find in the world is a notion of 'similarity'. His project is to explain how our categorization and our use of general terms works by distingishing 'natural' groupings from 'unnatural' ones, by defining 'natural' groups (or sets) in terms of a primitive notion of similarity.

Resemblance Nominalism According to nominalism, properties are just sets. But not all sets are *real* properties, the properties that *cut nature at its joints*. These real properties are defined in terms of similarity. This type of nominalism has a long history, but the first one who seriously tried to work it out was Carnap in his *Logische Aufbau der Welt*. In this famous book he proposed that real properties are *maximal resemblance sets*.

Let us start with a similarity structure $\langle X, R \rangle$, where R is some kind of similarity relation between the objects in X. Let us call each set like P for whom it holds that $\forall x \in P : \forall y \in P : R(y, x)$ a *resemblance set*, an element of $\mathcal{R}$. The maximal resemblance set are those elements of $\mathcal{R}$ such that there is no other element $Q \in \mathcal{R}$ such that $P \subset Q$. Thus, a maximal resemblance set is a *maximal* set of individuals each of which resemble each other. Can we think of such a maximal resemblance set as a natural property?

Suppose that similarity is an equivalence relation '$\approx$'. The similarity structure is then one of the form $\langle X, \approx \rangle$. As is well-known we can determine properties as equivalence classes: $Q = \{[x]_\approx : x \in X\}$. We can also go from partition to equivalence relation: $x \approx y$ iff $\exists q \in Q : x, y \in q$. But it is interesting to observe that the set of maximal similarity sets as defined above are exactly the set of equivalence classes that partition X, if the similarity relation is an equivalence relation. Is it natural to call these sets *natural properties*? Not really. For (i) there are intuitively natural properties P and P' that can overlap each other, $P \cap P' \neq \emptyset$, and (ii) there are intuitively natural properties P and P' where the one is a proper subset of the other, $P \subset P'$. Neither of those possibilities is allowed, if we assumed that the similarity relation involved is an equivalence relation, i.e., a relation that is *reflexivie*, *symmetric*, and *transitive*.

Carnap was aware of this, and for this reason he assumed that the similarity relation '$\sim$' is just reflexive and symmetric. In that case the similarity structure $\langle X, \sim \rangle$ is sometimes called a *tolerant space*. In terms of it one can determine similarity sets and maximal similarity sets as before.[8] But now the set of maximal similarity sets need not form a partition. (Note: Russell's construction of instants as maximal sets of overlapping events is special case. Though now this is a partition). Instead, what results is just a *cover*, where Q is a **cover** of X iff (a) Q is a set of subsets of X; (b) $\emptyset \notin Q$ and (c) $\forall x \in X : \exists q \in Q : x \in q$ (i.e. $\bigcup Q = X$). Now we say that $\langle X, Q \rangle$ is a **property structure** over X iff (i) $X \neq \emptyset$ and (ii) Q is a cover of X. Then we can determine a similarity relation as follows: $x \sim y$ iff$_{df}$ $\exists q \in Q : x, y \in q$.

Consider the set of objects $X = \{1, 2, 3, 4\}$ with the set of properties Q consisting of two properties: $\{1, 2, 3\}$ and $\{3, 4\}$.

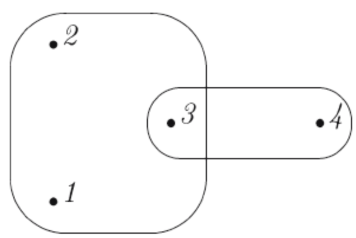

Figure 1: Set of properties

This set of properties gives rise to the following similarity structure (closed under reflexivity and symmetry): $\langle X, \{\langle 1, 2\rangle, \langle 2, 3\rangle, \langle 1, 3\rangle, \langle 3, 4\rangle\}\rangle$.

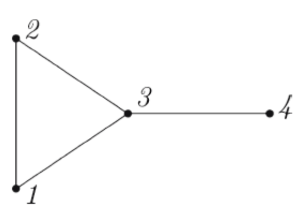

Figure 2: Corresponing similarity relation

[8]The existence of maximal similarity sets is, in general, guaranteed by Zorn's Lemma.

In this case, Q is the set of maximal similarity sets w.r.t. the similarity structure. Unfortunately, it is not in general the case that we can recover the original set of properties as the set of maximal similarity sets. We have the *no-uniqueness* problem: not any set of properties can be adequatly represented by a similarity set. There are two problems: *the problem of imperfect community* and the *companionship difficulty*.

Imperfect community: Suppose that we start with cover $Q = \{\{1,2,4\}, \{2,3,5\}, \{4,5,6\}\}$. The similarity relation is then $\langle\{1,2,3,4,5,6\}, \{\langle 1,2\rangle, \langle 1,4\rangle, \langle 2,3\rangle, \langle 2,5\rangle, \langle 3,5\rangle, \langle 2,4\rangle, \langle 4,5\rangle, \langle 4,6\rangle, \langle 5,6\rangle\}\}$. (as before, closed under reflexivity and symmetry). Let $r(x)$ be the function that maps x to the maximal similarity sets in which x is part. Now $\{r(x) : x \in X\} = \{\{1,2,4\}, \{2,4,5\}, \{2,3,5\}, \{4,5,6\}\} \neq Q$. Goodman states that this Q exihibits the difficulty of imperfect community': the similarity relation defined via Q gives rise to some maximal similarity sets that are not properties.

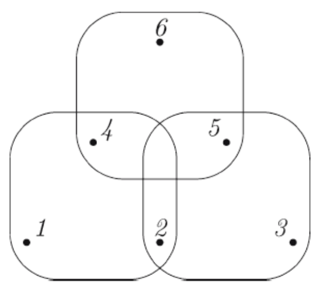

Figure 3: Imperfect community 1: properties

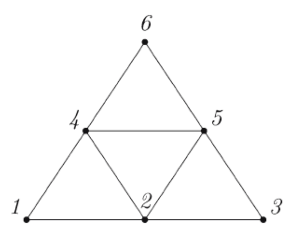

Figure 4: Imperfect community 1: similarity

In the above example we had more maximal similarity sets than original properties. But at other times, the converse problem will appear: we end up with less maximal similarity sets than we had original properties. As a minimal example, consider cover $Q = \{\{1,2\}, \{2,3\}, \{1,3\}\}$. The similarity structure then is $\langle\{1,2,3\}, \{\langle 1,2\rangle, \langle 2,3\rangle, \langle 1,3\rangle\}\rangle$. But the set of maximal similarity sets derived from this is just $\{\{1, 2, 3\}\} \neq Q$.

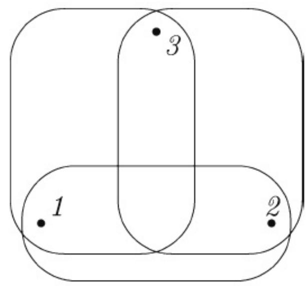

Figure 5: Imperfect community 2: properties

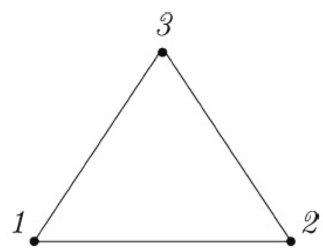

Figure 6: Imperfect community 2: similarity

Companionship difficutly. Let us we start with cover $Q = \{\{1, 2\}, \{1, 2, 3\}\}$. The similarity structure then is $\langle\{1, 2, 3\}, \{\langle 1, 2\rangle, \langle 2, 3\rangle, \langle 1, 3\rangle\}\rangle$. But the set of maximal similarity sets derived from this is just $\{\{1, 2, 3\}\} \neq Q$.

Goodman (1960) famously argued that because of these problems it is hopeless to start with a binary similarity relation as basic: in this way we cannot *recover* the basic properties. But in a sense this criticism is *not fair*. For a real nominalist, there are no basic properties, and similarity is all there is. So, in a sense the above problems cannot even be stated. What can be stated at most is that not all covers of X can be sets of basic properties. But in fact, as pointed out by Leitgeb (2007), this is true for purely cardinality reasons:

The non-uniqueness problem is unsolvable, due to *cardinality problems*: the set of all covers of X is much greater than the set of all possible similarity relations on X. Thus, a unique binary similarity relation *cannot* be found for each cover. The problem is solvable, however, if we limit ourselves to covers Q that satisfy certain constraints. Berge (1989) proved that if Q satisfies the

following constraints, constructing properties as maximal similarity sets is ok and we don't have a non-uniqueness problem.
$\langle X, \sim \rangle$ is faithful w.r.t. $\langle X, Q \rangle$ iff

(i) $\forall X, Y, Z \in Q, \exists P \in Q : (X \cap Y) \cup (X \cap Z) \cup (Y \cap Z) \subseteq P$, (imperfect community) and

(ii) $\neg \exists X, Y \in Q : X \subset Y$ (companionship)

These constraints show the **main problem** of Carnap's analysis: mainly due to (ii) such a Q cannot be thought of as set of natural properties, because it can't account for laws.

Goodman and Quine concluded that Carnap's binary similarity relation was *too weak*, we need at least a 3-place *comparative* similarity relation, or a 4-place similarity relation. Moreover, Goodman argued that we need similarity *in respect*.

Rodriguez-Pereyra Rodriguez-Pereyra claims that we can solve Goodman's problem by making use of a more general similarity relation: similarity also relating *pairs*.

Suppose 1,2, and 3 all resemble each other (have property P) and so do 4 and 5 (have property Q). Then the pairs $\langle 1,2 \rangle$ and $\langle 2,3 \rangle$ resemble each other, but $\langle 1,2 \rangle$ and $\langle 4,5 \rangle$ do not. (In this whole section it is assumed that this is equivalent to say that the sets $\{1,2\}$ and $\{2,3\}$ resemble each other, but $\{1,2\}$ and $\{4,5\}$ do not.) But then also $\langle \langle 1,2 \rangle, \langle 2,3 \rangle \rangle$ resembles $\langle \langle 1,2 \rangle, \langle 1,3 \rangle \rangle$ but not $\langle \langle 1,2 \rangle, \langle 4,5 \rangle \rangle$. (Meaning that the sets $\{\{1,2\},\{2,3\}\}$ and $\{\{1,2\},\{1,3\}\}$ resemble each other, but the sets $\{\{1,2\},\{2,3\}\}$ and $\{\{1,2\},\{4,5\}\}$ do not.) Thus, $\langle x, y \rangle$ resembles$_1$ $\langle u, v \rangle$ iff x and y resemble both u and v. Similarly $\langle X, Y \rangle$ resembles$_1$ $\langle U, V \rangle$ iff X and Y resemble both U and V.

Definition: A set of objects P is a *perfect community* iff (i) all its members resemble each other, (ii) all pairs of members of P resemble each other, (iii) all pairs of pairs of members of X resemble each other etc.

Take the *imperfect community*: $Q = \{\{1,2\},\{2,3\},\{1,3\}\}$. If we want to represent it with just a similarity relation between individuals it gives rise to similarity structure $\langle \{1,2,3\}, \{\langle 1,2 \rangle, \langle 2,3 \rangle, \langle 1,3 \rangle\} \rangle$. The only maximal similarity set here is $\{1,2,3\}$, i.e., we cannot recover Q.
The *perfect community* $Q' = \{X'_0 = \{1,2,3\}\}$ gives rise to the same similarity structure and can be reconstructed.

But now let us also assume that we have a similarity relation between pairs. Such a relation can contain more information then one that only relates individuals. In particular, we can construct a similarity structure $\langle X, \sim' \rangle$ that faithfully represents the above mentioned perfect community Q' as follows: $\langle \{1,2,3\}, \{\langle 1,2\rangle, \langle 2,3\rangle, \langle 1,3\rangle, \langle\langle 1,2\rangle, \langle 2,3\rangle\rangle, \langle\langle 1,2\rangle, \langle 1,3\rangle\rangle, \langle\langle 1,3\rangle, \langle 2,3\rangle\rangle\}\rangle$. The similarity structure $\langle X, \sim \rangle$ that faithfully represents Q, on the other hand, is just $\langle \{1,2,3\}, \{\langle 1,2\rangle, \langle 2,3\rangle, \langle 1,3\rangle\}\rangle$ and does not contain similarities between pairs. Notice that $P = \{1,2,3\}$ is a perfect community with respect to $\langle X, \sim' \rangle$ but not with respect to $\langle X, \sim \rangle$, because there are at least two pairs of members of P that do not resemble each other: $\langle 1,2\rangle \not\sim \langle 1,3\rangle$. Thus, we can recover Q by looking not at the maximal similarity sets, but rather at the perfect communities.

There is a simpler way to arrive at the same result. The reason is that Pereira Rodriguez seems to assume that if $\langle x,y\rangle \sim \langle x,z\rangle$ then it also holds that $\langle x,y\rangle \sim \langle x,z\rangle$ and $\langle y,z\rangle \sim \langle x,z\rangle$.[9] Now suppose that we start with a PR-similarity structure $\langle X, \sim_{pr}\rangle$. Define for each $x \in X$ the set $s^i(x)$ for each positive natural number i as follows: $s^1(x) = \{\{x,y\} : y \in X \ \&\ x \sim_{pr} y\}$; $s^{n+1}(x) = \{X \cup Y : X, Y \in s^n(x) \ \&\ X \sim_{pr} Y\}$. $s(x)$ is now defined as the fixed point of this sequence. The set of properties Q is now defined as $Q = \{s(x) : x \in X\}$.

To go the other way, we have to define the similarity relation in terms of a cover Q such that after recovering the similarity relation again, we can recapture the same cover Q again. Define for each $x \in X$, $f(x)$ as $\{q \in Q : x \in q\}$. For pairs and higher we define it as follows: $f(\langle x,y\rangle) = f(x) \cap f(y)$. Notice that in the perfect community (i) $\forall x \in X : f(x) = \{\{1,2,3\}\}$, but also (ii) $f(\langle 1,2\rangle) = f(\langle 1,3\rangle) = f(\langle 2,3\rangle) = \{\{1,2,3\}\}$. In the imperfect community, however, $f(1) = \{\{1,2\},\{1,3\}\} \neq f(2) = \{\{1,2\},\{2,3\}\} \neq f(3) = \{\{1,3\},\{2,3\}\}$. But this means that $f(\langle 1,2\rangle) = \{\{1,2\}\} \neq \{\{2,3\}\} = f(\langle 2,3\rangle)$. In general, we can determine the similarity relation from a cover Q by defining the similarity as follows: $x \sim^Q y$ iff$_{df}$ $f^Q(x) \cap f^Q(y) \neq \emptyset$, where x can either be an element of X, or a pair, or a pair of pairs, etc. Then we would like it to be the case that for any cover Q of X, $\sim^Q$ as defined above faithfully represents the cover: we look at all maximal perfect communities that we get from $\sim^Q$ and see whether this is the same as Q. This is not yet the case, because we haven't solved yet the companionship difficulty.

Companionship difficulty. Let us start with cover $Q = \{\{1,2\},\{1,2,3\}\}$. The similarity structure $\langle X, \sim^Q\rangle$ that we derive now is $\langle \{1,2,3\}, \{\langle 1,2\rangle, \langle 2,3\rangle, \langle 1,3\rangle, \langle\langle 1,2\rangle, \langle 2,3\rangle\rangle, \langle\langle 1,2\rangle, \langle 1,3\rangle\rangle, \langle\langle 1,3\rangle, \langle 2,3\rangle\rangle\}\rangle$. But the set of perfect communities derived from this is just $\{\{1,2,3\}\} \neq Q$. Thus, we have not solved

[9] This is so, because in order for $f(\langle 1,2\rangle) \cap f(\langle 2,3\rangle) \neq \emptyset$ it must be that $\exists X : X \in f(1) \cap f(2) \cap f(3)$ such that $\{1,2,3\} \subseteq X$, see below.

the companionship difficulty yet. For the companionship difficulty, Rodriguez needs *degrees of resemblance*. He says that $sim(x, y) = n$ iff x and y share n properties. Thus, let us make a distinction between $\sim_1$ and $\sim_2$. In general, we say that $x \sim_n^Q y$ iff$_{df}$ $|f(x) \cap f(y)| \geq n$. Notice that $f(1) \cap f(2) = \{\{1, 2\}, \{1, 2, 3\}\}$, whereas $f(1) \cap f(3) = \{\{1, 2, 3\}\}$. Thus, whereas $1 \sim_2^Q 2$, it is not the case that $1 \sim_2^Q 3$, although $1 \sim_1^Q 3$. Now we represent cover Q by the following similarity structure $\langle X, \sim_1^Q, \sim_2^Q\rangle$: $\langle\{1, 2, 3\}, \{\langle 1, 2\rangle, \langle 2, 3\rangle, \langle 1, 3\rangle, \langle\langle 1, 2\rangle, \langle 2, 3\rangle\rangle, \langle\langle 1, 2\rangle, \langle 1, 3\rangle\rangle, \langle\langle 1, 3\rangle, \langle 2, 3\rangle\rangle\}, \{\langle 1, 2\rangle\}\rangle$. We determine the perfect communities w.r.t. each $\sim_n$. Notice that $\{1, 2\}$ and $\{1, 2, 3\}$ are perfect communities of level 1, whereas only $\{1, 2\}$ is a perfect community of level 2. Let us say that P is a basic property if there is an n such that (i) P is a perfect community of level n and (ii) there is no perfect community P' of level n such that $P \subset P'$. Thus, we now see that $\langle X, \sim_1^Q, \sim_2^Q\rangle$ gives rise to the following basic properties: $\{\{1, 2\}, \{1, 2, 3\}\}$. This is the same as Q, as desired.

Alternatively, define for each $x \in X$ the set $s_n^i(x)$ for each positive natural number i and level m as follows: $s_m^1(x) = \{\{x, y\} : y \in X \ \& \ x \sim_m y\}$; $s_m^{n+1}(x) = \{X \cup Y : X, Y \in s^n(x) \ \& \ X \sim_m Y\}$. $s_m(x)$ is now defined as the fixed point of this sequence. The set of properties Q of level m is now defined as $Q = \{s_m(x) : x \in X\}$.

Of course, once we allow for degrees we basically take resemblance to be a 3- or even 4-place relation: 'b is more similar to b than c' and 'a is more similar to b as c is similar to d' become meaningful. Rodriguez-Pereyra's solution of the companionship difficulty is not the most interesting feature of his proposal. What is interesting about his proposal is how he solves the problem of imperfect communities. However, even if Rodriguez-Pereryra can solve both of these problems, that is still not enough to guarantee a 1-1 relation between sets of properties and similarity structures.

Mere intersection difficulty. Is it the case that every cover Q can be faithfully represented by a similarity structure? This depends on whether covers are closed under intersection yes or no. If yes (which I think is natural), we are ready. If no, we have a problem. Consider the following cover of $X = \{1, 2, 3, 4, 5, 6\}$: $Q = \{\{1, 2, 3\}, \{1, 2, 3, 4, 6\}, \{1, 2, 3, 4, 5\}\}$. Notice that from this Q we derive a similarity structure $\langle X, \sim_1^Q, \sim_2^Q, \sim_3^Q\rangle$ from which we derive the following set of properties: $\{\{1, 2, 3\}, \{1, 2, 3, 4, 6\}, \{1, 2, 3, 4, 5\}, \{1, 2, 3, 4\}\}$: $\{1, 2, 3\}$ is a maximal perfect community of level 3, $\{1, 2, 3, 4, 6\}$ and $\{1, 2, 3, 4, 5\}$ are maximal communities of level 1, and $\{1, 2, 3, 4\}$ is a maximal community of level 2. But this is not the same as Q, because $\{1, 2, 3, 4\}$ does not belong to Q. Notice that this set is the intersection of $\{1, 2, 3, 4, 6\}$ and $\{1, 2, 3, 4, 5\}$.

Similarity relation between sets Can't we simplify what Rodriguez Pereyra wanted to do? Why not simply say – as already suggested above – that we allow for similarity relations between *sets* of individuals? Thus, not only can there be similarity relations between different sets of exactly two elements, but there can be similarity relations between sets of all non-zero cardinality. In this way, we don't need to go any 'higher', i.e., assume similarity relations between pairs of pairs of .... pairs of individuals. So, here is the idea.[10] Suppose we start with a cover Q. Now we define the similarity relation between non-empty sets as follows: $p \sim p'$ iff$_{def}$ $\exists q \in Q : p, p' \subseteq q$.[11] Notice that this similarity relation is certainly reflexive and symmetric. Is it also transitive? No, for consider $Q = \{\{1,2,3\},\{3,4\}\}$. In this case we have, for instance, that $\{1\} \sim \{3\}$ and that $\{3\} \sim \{4\}$, but it is not the case that $\{1\} \sim \{4\}$. Thus, we can go from a cover to a similarity structure, although this similarity is now between sets, rather than between individuals. Of course, this new similarity relation would be isomorphic to a similarity relation between individuals, if we limited ourselves to singleton sets. But we did not, and thus our new similarity relation contains possibly more information.

Is this extra information enough to solve the imperfect community problem? Let us look at our simplest example again: $Q = \{\{1,2\},\{1,3\},\{1,3\}\}$. The similarity relation that this gives rise to contains (if we forget about the reflexive relations) only $\{1\} \sim \{2\}$, $\{1\} \sim \{3\}$, and $\{2\} \sim \{3\}$. This is different with the cover $Q' = \{\{1,2,3\}\}$, which gives rise to a similarity relation also connecting $\{1,2\} \sim \{2,3\}$, $\{1,2\} \sim \{3\}$, and $\{1,3\} \sim \{2,3\}$, for instance. But whether we have solved the imperfect community problem depends on how we are now going to define (sparse, or natural) properties.

The proposal is very straightforward. First, we are going to define what it means to be a similarity* set. A similarity* set X is now not just a set that obeys constraint (i) all its elements (or better, singleton sets) are similar to each other, but also constraint (ii) for all non-empty subsets p, p' of $\bigcup X$ it holds that $p \sim p'$.[12] After this strengthening of the notion of a similarity set, we go on as before. First, we collect all maximal similarity* sets. Let us call MAX the set of all maximally similarity* sets. The properties induced by the similarity structure are then defined as $\{\bigcup X : X \in MAX\}$.

Let us first see whether the similarity relation induced by cover Q indeed

[10] Only after writing this paper we discovered Paseau (2012), where something very similar was worked out very precisely. Paseau argues that resemblance similarity can be saved, but that the cost of assuming similarity relations between *sets* of individuals is probably a too high price to pay for a nominalist.

[11] The empty set will be similar to no other set.

[12] Because if $p \in X$, it follows that $p \subseteq \bigcup X$, we can do with only condition (ii). Notice that (ii) entails that all individuals in $\bigcup X$ resemble each other.

avoids inducing imperfect community $\{1,2,3\}$ as a property. We have seen that the similarity relation Q gives rise to contains only $\{1\} \sim \{2\}$, $\{1\} \sim \{3\}$, and $\{2\} \sim \{3\}$. Notice that although $\{1,2,3\}$ is a similairty set, it is not a similarity* set, because it doesn't hold, for instance, that $\{1,2\} \sim \{3\}$. With $\{1,2,3\}$ out of the way, we can see that now the maximal similarity* sets are $\{1,2\}, \{1,3\}$, and $\{2,3\}$, just as desired. It is different for the similarity relation induced by $Q' = \{\{1,2,3\}\}$. Because all non-empty subsets p and p' of $\{1,2,3\}$ are by construction similar to each other, we end up with only one maximal similarity* set, X, such that $\bigcup X = \{1,2,3\}$.

What we would really like to know, however, is whether the set of properties $Q' = \{\bigcup X : X \in MAX\}$ induced from the similarity relation that is itself induced by cover Q is such that $Q' = Q$. To prove this, we should prove that $Q \subseteq Q'$ and that $Q' \subseteq Q$.

Let us start with an arbitrary element $q \in Q$. it holds by construction for non-empty subsets p and p' of q that $p \sim p'$. But this means that the set X of all non-empty subsets of q is a similarity* set. Will X also be a *maximal* similarity* set? If we forget about covers like $Q = \{\{1,2\}, \{1,2,3\}\}$ that give rise to Goodman's companionship difficulty, it is clear that it is. But obviously $\bigcup X = q$, and thus $q \in Q'$.

To go the other way around, let us assume that $q' \in Q'$. This means that there is a maximal similarity* set X such that $q' = \bigcup X$. Form the fact that X is a similarity* set it follows that (i) $\forall p, p' \in X : p \sim p'$, and (ii) $\forall p, p' \subseteq \bigcup X : (p \neq \emptyset \wedge p' \neq \emptyset) \rightarrow p \sim p'$. Because if $p \in X$, it follows that $p \subseteq \bigcup X$, we can do with only condition (ii). But the similarity relation defined in terms of Q was exactly defined like (ii), with an element Q instead of $\bigcup X$. This means that $\bigcup X = q'$ is an element of Q.

But a resemblance nominalist doesn't really start with a cover Q, but rather with a similarity relation. Can he really start out with a set of individuals S and then a arbitrary similarity relation between non-empty subsets of S? Suppose, for instance, that $S = \{1,2,3\}$, and that the similarity relation only connected $\{1,2\} \sim \{3\}$ (ignoring reflexivity and symmetry as usual). It is clear that such a similarity relation does only give rise to singleton sets as similarity* sets. But what would it mean then that $\{1,2\} \sim \{3\}$? It would have no meaning. We propose to get rid of such similarity relations by the following constraint: $\forall p, p'$: if $p \sim p'$, then $\forall r_{\neq\emptyset}, r'_{\neq\emptyset} \subseteq p \cup p' : r \sim r'$. Notice that this constraint is indeed obeyed in the above examples.

We believe that to solve the companionship difficulty staring from covers like $Q = \{\{1,2\}, \{1,2,3\}\}$, we can use a trick very similar to what Rodriguez-Pereryra made use of: just assume that we we have more than 1 similarity relation, or better, perhaps, that we start with a three- or four-place comparative similarity relation: in this case, all non-empty subsets of $\{1,2\}$ are at least as similar to each other as all non-empty subsets of $\{1,2,3\}$ are, and $\{1\}$ and $\{2\}$ are more similar to each other than $\{1\}$ and $\{3\}$ are, for instance. One way to go is to start with a four-place comparative similarity relation $q <_p r$ (q is more similar to p than r is), and define $<_p$ in terms of original cover Q as follows: if $\exists q, q' \in Q : q \subset q'$, then $\forall p, p'_{\neq\emptyset} \subseteq q : \forall r_{\neq\emptyset} \subseteq p : \forall s_{\neq\emptyset} \subseteq q' - q : p' <_p (r \cup s)$.

6 Conclusion

In this paper we have shown, or reminded, the reader that it is in many cases less than obvious which notions should be taken as primitive and what should be constructed out of what. In many cases, both directions are possible. Some constructions give rise to technical difficulties, but these can in many cases be solved by assuming a somewhat richer ontology – as for instance in the previous section – or stronger constraints on the initial orderings – as in the case of events. Which direction the construction should go (if we want construction at all) depends on how 'natural' the primitives and constraints on the constructions one starts out with to get what one wants are taken to be. To think, for instance, of propositions as primitives, and to define worlds and similarities between them in terms of it, allows one to make more distinctions, but is somewhat harder to work with. The same holds for the view according to which properties are sets of features. But once one has decided on taking features to be basic, it is also natural to think of individuals in terms of (sets of sets of) features as well. It was not the purpose of this paper to argue what the primitives should be, but just to point out certain options.

References

[1] Allen, J.F. and P.J. Hayes (1985), 'A common-sense theory of time', in A. Joshi (ed.), *Proceedings of the the Ninth Joint Conference on Artificial Intelligence*, Los Altos.

[2] Aristotle. (1941), *Basic Works*, (ed.) R. McKeon, Random House.

[3] Armstrong, D.M. (1989), *Universals: An Opinionated Introduction*, Westview Press, Oxford.

[4] Arrow, K. (1959), 'Rational choice functions and orderings', *Economica*, **26**: 121-127.

[5] Benthem, J. van (1982), 'Later than late: on the logical origin of the temporal order', *Pacific Philosophical Quarterly* **63**: 193-203.

[6] Benthem, J. van (1991), *The Logic of Time*, Synthese Library.

[7] Berg, C. *Hypergraphs*, North Holland, Amsterdam.

[8] Carnap, R. (1923), 'Die Quasizerlegung. Ein Verfahren zur Ordnung nichthomogener Mengen mit den Mitteln der Beziehungslehre', unpublished manuscript RC-081-04-01, University of Pittsburgh.

[9] Carnap, R. (1928), *Der logische Aufbau der Welt*, Berlin, Weltkreis.

[10] Carnap, R. (1947), *Meaning and Necessity*, University of Chicago.

[11] Gärdenfors, P. (2000), *Conceptual Spaces*, MIT, Cambridge.

[12] Fishburn, P. C. (1970), 'Intransitive indifference with unequal indifference intervals', *Journal of Mathematical Psychology*, **7**: 144-149.

[13] Glashoff, K. (to appear), 'An intensional Leibniz semantics for Aristotelian logic', *The Review of Symbolic Logic*.

[14] Goodman, N. (1953), *The Structure of Appearance*, 2nd edition, Indianapolis: The Bobbs-Merrill Company, Inc.

[15] Goodman, N. (1972), 'Seven strictures on similarity', in Goodman (1972), *Problems and Projects*, Indianapolis: The Bobbs-Merrill Company, Inc, 437-46.

[16] Hazen, A.P. and L. Humberstone (2004), 'Similarity relations and the preservation of solidity', *Journal of Logic, Language and Information*, **13**: 25-46.

[17] Ishiguro, H. (1972), *Leibniz' Philosophy of Logic and Language*, Duckworth, London.

[18] Kamp, H. (1979), 'Events, instants and temporal reference', in R. Bäuerle et al (eds.), *Semantics form Different Points of View*, Springer, Berlin, pp. 131-175.

[19] Klein, E. (1980), 'The semantics of positive and comparative adjectives', *Linguistics and Philosophy*, **4**: 1-45.

[20] Kratzer, A. (1981), 'Partition and revision: the semantics of counterfactuals', *Journals of Philosophical Logic*, **10**: 242-258.

[21] Leibniz, G. (1686), *Discours de Métaphysique*, (first published in English by Open Court, 1902).

[22] Leibniz, G. (1966a), 'Rules from which a decision can be made, by means of numbers, about the validity of inferences and about the forms and moods of categorical syllogisms', in Parkinson (ed.), *Leibniz. Logical Papers*, pp. 25-32.

[23] Leitgeb, H. (2007), 'A new analysis of Quasianalysis', *Journal of Philosophical Logic*, **36**: 181-226.

[24] Lewis, D. (1973), *Counterfactuals*, Oxford, Blackwell.

[25] Lewis, D. (1983), 'New work on the theory of universals', *Australian Journal of Philosophy*, **64**.

[26] Lewis, D. (1986), *On the Plurality of Worlds*, Oxford, Blackwell.

[27] Lück, U. (2006), 'Continuous Time Goes by Russell', *Notre Dame Journal of Formal Logic*, **47**: 397-434.

[28] Mormann, T. (2009), 'New work for Carnap's quasi-analysis', *Journal of Philosophical Logic*, **38**: 249-282.

[29] Paseau, A. (2012), 'Resemblance theories of properties', *Philosophical Studies*, **157**: 361-382.

[30] Plato. *Republic*, (1941), trans. F. Cornford, Oxford University Press.

[31] Quine, W.V. (1969), 'Natural kinds', in W.V. Quine (ed.), *Ontological Relativity and Other Essays*, Columbia University Press, New York, pp. 114-138.

[32] Rodriguez-Pereyra, G. (1999), 'Resemblance nominalism and the imperfect community', *Philosophy and Phenomenological Research*, **59**: 965-982.

[33] Rodriguez-Pereyra, G. (2002), *Resemblance Nominalism: A solution to the problem of universals*, Clarendon Press, Oxford.

[34] Rooij, R. van (2012), 'Semi-orders and satisficing behavior', *Synthese*, 2011, Vol. 179, 1-12.

[35] Russell, B. (1912), *The Problems of Philosophy*, Home University Library.

[36] Russell, B. (1914), *Our Knowledge of the External World*, London and New York, Routledge.

[37] Russell, B. (1936), 'On order in time', in *Proceedings of the Cambridge Philosophical Society*, vol. 3, pp. 216-228.

[38] Sen, A. K. (1971), 'Choice functions and revealed preference', *The Review of Economic Studies*, **38**: 307-317.

[39] Stalnaker, R. (1968), 'A theory of conditionals', in N. Rescher (ed.) *Studies in Logical Theory*, American Philosophical Quarterly Monograph Series, 2, Oxford, Blackwell, pp. 98-112.

[40] Thomason, S.K. (1984), 'On constructing instants from events', *Journal of Philosophical Logic*, **13**: 85-96.

[41] Turner, R. (1981), 'Counterfactuals without possible worlds', *Journals of Philosophical Logic*, **10**: 453-493.

[42] Veltman, F. (1976), 'Prejudices, presuppositions, and the theory of counterfactuals', in J. Groenendijk & M. Stokhof (eds.), *Amsterdam Papers in Formal Grammar. Proceedings of the 1st Amsterdam Colloquium*, Amsterdam, pp. 248-281.

[43] Wiener, N. (1914), 'A contribution to the theory of relative position', *Proc. Camb. Phil. Soc.*, **17**: 441-449.