Extensional and Intensional Collectives and the *De re* / *De dicto* Distinction

Antony Galton [a] Zena Wood [b]

[a] *College of Engineering, Mathematics, and Physical Sciences, University of Exeter, UK,*
apgalton@ex.ac.uk
[b] *Department of Computing and Information Systems, University of Greenwich, UK,*
Z.Wood@greenwich.ac.uk

Abstract. Expressions designating collectives, such as "the committee" or "the ships in the port", may be interpreted *de re* or *de dicto*, depending on context, according as they pick out collectives defined by their members or collectives defined by some criterion for membership. We call these E(xtensional)-collectives and I(ntensional)-collectives respectively, and in this paper we explore in depth the relationship between these two categories. In particular, we identify important respects in which they differ, regarding the nature of the dependence of the collective on its members, the nature of the parthood relation in which the members stand to the whole, and, from an application perspective, the different methods used for identifying the two forms of collective from large spatio-temporal data-sets.

Keywords: collectives, collections, de re, de dicto

1. Introduction

In previous work [35] we considered a wide range of collective phenomena and proposed a number of principles for their classification. One of these principles concerned membership criteria, and we proposed a top-level division into collectives whose membership is constant and collectives whose membership is variable.[1] It seems to us now that there is a more fundamental distinction that was overlooked, or at least very much marginalised, in [35]. For amongst collectives whose membership is constant, we can distinguish those whose membership merely happens to be constant (because although there is nothing to prevent replacement of members, in fact this does not occur) from those for which constancy of membership is an essential feature, so that the result of any replacement would no longer be the same collective. The former kind of collective, but not the latter, can be regarded as a limiting case of the range of collectives whose membership actually does vary: they are collectives which have, as it were, a "variation parameter", which happens to assume the value zero, whereas for collectives whose membership is necessarily constant such a parameter is absent.

A remark in [35] is pertinent here. We noted that there is a 'potential ambiguity' in descriptions for collectives, giving as an example:

1. *In five years' time all the committee will be female.*
2. *In five years' time all the committee will be dead.*

The difference here is explained as arising from the possibility of understanding 'all the committee' either *de dicto* or *de re*. These two readings are distinguished by how they track the referent of a phrase like 'all the committee' across different contexts. In the *de dicto* reading, in any given context, 'all the committee' is understood to refer to the people who are members of the committee *in that context*; in (1) above, the context is given by 'in five years' time'. In the *de re* reading, whatever the context, 'all the committee' refers to the people who are members of the committee in the context in which the phrase is uttered,

---

[1]In [35] we attributed this distinction to Simons [29], who introduced a distinction between 'classes' and 'groups' (pp. 144–147), but we now believe that our reading of Simons was mistaken.

without reference to whether they are members of the committee in the context that is being referred to — and in the case of (2), being dead, they will certainly not be members of the committee in five years' time.

To bring together the observations of the previous two paragraphs, then, we see that the *de re* reading of a collective description picks out a collective of the kind whose membership is essentially constant, whereas the *de dicto* reading picks out a collective of the kind whose membership is variable — which does not mean that it *must* vary, only that it could do so without necessarily compromising the identity of the collective. In the philosophical literature the discussion of *de re* and *de dicto* interpretations of referring expressions has concentrated on cases where this ambiguity arises in the context of belief, desire, and intention. An example is Quine's "Ralph believes that someone is a spy", where the ambiguity turns on whether there is a particular person whom Ralph believes to be a spy — the *de re* reading — or whether Ralph merely believes in the existence of some spy or other — *de dicto* [24]. Another example is "Mary wants to marry a doctor". More generally, the possibility of this kind of dual interpretation arises in any *intensional* context, that is any context for which the substitution of co-referential terms can result in a difference in truth value; and temporal terms such as 'in five year's time' also supply such contexts.[2]

The ability of apparently straightforward expressions to admit of both *de re* and *de dicto* readings can lead to some subtle issues of interpretation, depending on imponderables such as "frame of mind", as nicely illustrated by this quotation from Sharvy:

> I might say that I have had this beard for five years, but this hair on my face for less than a week. But in a fit of materialism I might say that I have had this beard for less than a week; and in the opposite frame of mind, I might say that I have had this hair on my face for five years. [28]

Phrased in terms of collectives, we could say that whereas normally 'this beard' refers to a variable-membership collective of facial hairs, while 'this hair on my face' refers to a fixed-membership collective, Sharvy in his fit of materialism takes 'this beard' to refer to the fixed-membership collective, whereas in the opposite frame of mind he takes 'this hair on my face' to refer to the variable-membership collective.

The possibility of interpreting a collective description such as 'all the committee' in two ways links this whole area of discussion to another issue that has been of importance both in philosophy and in applied ontology, that of *multiplicativism*. The classic example concerns the statue and the clay: what is the relationship between a statue and the clay from which it is moulded? They cannot be identical, because they have different histories: the clay was there before the statue was there. But if they are different then it seems puzzling that there can be two distinct entities existing in exactly the same place at the same time, coinciding in all their physical properties. A popular answer is to say that the statue and the clay are not identical, but that the statue is *constituted* by the clay. The difficulty is then to come up with a workable definition of this constitution relation. There has been extended philosophical discussion of this problem (see [31] for an excellent summary), and it is one respect in which two of the best-known modern formal ontologies, BFO [3] and DOLCE [21], diverge, the former repudiating, and the latter embracing, the notion of constitution.

Our two readings of 'the committee' lead to a similar situation. For there seem to be two distinct collectives that are referred to by this description. On the one hand, we have a particular collection of individuals who currently constitute the committee (and note how natural the everyday use of the word 'constitute' seems in this context). This is the collective picked out by the *de re* reading in (2), where it is said of that particular collection of individuals that they will all be dead in five years' time. On the other hand, we have the committee itself, as an institution, by which we understand an entity which may be constituted by different individuals at different times. This is the collective picked out by the *de dicto* reading in (1), which implies that all the men who are currently on the commitee will be replaced by women over the next five years.[3] If we have two distinct collectives here (a group and a class, in Simons'

---

[2]An example of *de re* / *de dicto* ambiguity arising in a temporal context, that does not involve collectives, is "In five year's time the poorest man in town will be a millionaire", which can either be read as saying that the person who is currently the poorest man in town will become a millionaire in the next five years or, more dramatically, that over the next five years the town as a whole will become so prosperous that everyone, even the poorest person, will be a millionaire.

[3]A *de re* reading of (1) is also possible: all the men on the committee will change sex in the next five years.

terminology), how are they related? As (1) and (2) show, the two collectives have different histories, but at the present moment they *coincide*. But this coincidence relation does not seem to be entirely symmetrical: the fixed membership collection of individuals given by the *de re* reading seems more fundamental than the variable membership collective given by the *de dicto* reading — it is natural to say that the latter is constituted by the former, in much the same way as the statue is constituted by the clay, and Theseus' ship is constituted by a consignment of wood.

Against this, an opponent of multiplicativism might argue that even at the time of supposed coincidence the membership of the two committees is not identical. Whereas the members of the *de re* committee are just what we supposed them to be, namely people, it might be argued that the members of the *de dicto* committee are not people *tout court* but people *qua* holders of certain roles in terms of which the committee is constituted. Minimally, this might involve just one role, "committee member", but in typical committees this generic role may be differentiated into sub-roles such as chair, secretary, treasurer, and so on. In [22], an individual *qua* holder of a role is called a role-holder; this could be specified by means of an ordered pair $\langle individual, role \rangle$. If the role-holder is regarded as distinct from its first component (the "bare" individual considered without reference to the role) then the *de re* and *de dicto* committees, so far from ever having the *same* membership, must of necessity be always disjoint. This way one can avoid embracing multiplicativism if that doctrine is held to be repugnant,[4] although whether or not it really succeeds in doing so is a moot point, since distinguishing between people and people *qua* holders of certain roles could itself be seen as a kind of multiplicativism.

Our purpose in this paper is to explore some important ramifications of the distinction we have highlighted here between two kinds of collective. We shall adopt the following terminology: Collectives whose membership is necessarily fixed, in the sense that the collective itself is defined in such a way that having precisely those members is a necessary condition for it to exist at any given time, will be called *extensional collectives*; these are picked out by *de re* readings of collective descriptions. In contrast to these, collectives whose membership is potentially variable, which are picked out by *de dicto* readings of collective descriptions, will be called *intensional collectives*. For convenience, we abbreviate these terms to *E-collectives* and *I-collectives* respectively. E-collectives and I-collectives seems to differ in many respects; arguably, they are so different in nature that it is misleading to apply the same term 'collective' to both of them. One possibility would be to reserve 'collective' for I-collectives, and use 'collection' for E-collectives, as in [11]; but 'collective' and 'collection' sound so similar that this might invite the risk of confusion, whereas 'I-collective' and 'E-collective', while equally similar-sounding, at least overtly proclaim their status as terms of art, alerting the reader to attend closely to the distinction intended. For now, therefore, we will stick with this terminology.

Although we have introduced the distinction between E-collectives and I-collectives with reference to properties of the members of the collective (e.g., in our committee example, whether they are male or female, or alive or dead), having established the distinction we could go on to consider properties that are ascribed to a collective as a whole rather than its members. When a committee makes a decision, for example, this is not simply a case of each member of the committee making a decision; rather, what the members of the committee do is to collectively agree (possibly through a majority vote) that such a decision should be considered to have been enacted by the committee, and this collective agreement is not something that can be ascribed, as such, to the individual members. While it is undoubtedly true that all the most interesting examples of collective attribution relate to I-collectives, collective properties can be ascribed to E-collectives too: as when, for example, a particular cohort from some degree course hold a reunion five years after graduating — it is essential that it is those particular members that are involved, so it is an E-collective here, but "holding a reunion" is not, of course, something that can be ascribed to the members individually.

The organisation of the rest of this paper is as follows. In §2, we review some relevant literature, indicating particularly where people have pointed towards distinctions similar to the one we are highlighting here. In §3 we explore the implications of this distinction for a number of areas that have been seen as

---

[4]Something similar was suggested to us by Barry Smith (personal communication, February 2016.)

important for ontology, and in particular we show how it aligns with a number of other distinctions that have played a prominent part in ontological theorising. In §4, we take a more applied focus and illustrate the importance of the distinction between extensional and intensional collectives for understanding different methods for analysing the collective behaviour in spatio-temporal datasets. Finally, in §5, we provide a summary of our conclusions.

Before concluding this introduction, we should clarify that in this paper we understand a collective to be a continuant entity existing as a whole at each time of its existence (albeit possibly with different members at different times). We thereby exclude from consideration examples such as "The wives of Henry VIII": in so far as Henry may be thought to have had a collective of wives, this was only ever a one-member collective (in accordance with the custom of monogamy still prevalent in England), though the membership of that collective did change on five separate occasions. But this is to stretch a point: we do not generally find it useful to refer to collectives which cannot have more than one member at a time. We also exclude from consideration artificial collectives defined by explicitly specifying different extensions at different times, for example a collective defined as comprising Tom, Dick and Harry on even dates and Anne, Jane and Mary on odd dates; such a collective, were it to be countenanced, might be regarded as an extensional collective with variable membership, but again, we do not normally specify collectives in this way in practice.[5]

2. Related Work

Because membership of collections or collectives may be considered to be a kind of part–whole relation, the extensive literature on the ontology of parthood is of relevance to the concerns of this paper. One of the most widely-cited papers on parthood, by Winston *et al.* [32], proposes a classification comprising six different types of "meronymic" part–whole relation, of which one, Member / Collection, is particularly pertinent to our present concerns. Collections are sharply differentiated from classes: "membership in a class is determined on the basis of similarity to other members, while membership in a collection is determined on the basis of spatial proximity or by social connection", collections of the latter kind being referred to as "groups". But within collections, no attempt is made to distinguish those which are defined extensionally by their members from those whose membership can vary.

Several of the authors discussed below cite a work by Odell [23], whose classification also comprises six types of part–whole relation. The last two of these are *member–bunch*, where the members of the "bunch" (generally denoted with a collective noun) can change over time, and *member–partnership*, for which changing a member does destroy the whole. These clearly foreshadow our distinction between I-collectives and E-collectives respectively. Keet and Artale [17] incorporate this distinction into their own taxonomy of part–whole relations, not as items on an equal footing but with the member–partnership relation, notated *member-of′*, regarded as a specialisation of the member–bunch relation *member-of*, the defining characteristic of the former being that "the whole is existentially dependent on the part and vice versa". They cite a marriage partnership as an example of this — although we should point out that it is John-*qua*-Mary's-husband that is dependent on the partnership, not John-*qua*-human-being (in the terminology of [22], the former, but not the latter, is a *role-holder* in the context of the marriage partnership).

Bittner *et al.* [6] provide an axiomatic framework for what they call the "foundational relations" of ontology, with particular reference to biomedical applications. The framework recognises three basic entity types, individuals, universals, and collections, and postulate a number of primitive relations connecting them. Of particular importance are the relations in which an individual can stand to entities of each of the three types: an individual can be *part of* another individual, an *instance of* a universal, and a *member of* a collection. Of these, the first two, but not the third, are regarded as time-dependent; that is, an

---

[5]And even if we did, it could be argued that, far from being extensional, this specifies an *intensional* collective, whose membership criterion is given by the property "Equal to Tom, Dick or Harry if the date is even and equal to Anne, Jane or Mary if the date is odd".

individual can gain and lose parts, a universal can gain and lose instances, but a collection cannot gain or lose members. Thus, for these authors, the term "collection" refers to our E-collectives only; this is in line with the use of the term "collection", as opposed to "collective", in [11]. At any one time the totality of instances of a universal at that time is a collection, called the *extension* of that universal. At different times the universal may have different extensions, but at times when it has no instances, it has no extension — there is no place for an "empty collection" in this framework. Unlike in some other schemes that have been proposed in the literature, singleton collections are allowed; the extension of a universal with only one instance would be an example of this. The axioms imply that "two universals are identical if and only if they have identical extensions at every time"; this means, in effect, although Bittner *et al.* do not themselves offer this characterisation, that each universal is uniquely defined by a partial function from times to collections. Not every such function defines a universal, however, and it is here, if anywhere, that we might find a home for the notion of an I-collective in their framework, for it is true of I-collectives, as for universals, that each such collective defines a partial function from times to collections. One might not wish to go so far as to say that every such function defines a collective; what we are mainly interested in is functions picked out by ordinary language descriptions such as "The member states of the EU". Since Bittner *et al.* do not define what they mean by a universal, it is impossible to say, on the basis of [6], whether this particular description picks out a universal, or whether the authors would recognise any class of entity occupying this "middle ground" between collections and universals.

An important paper by Rector, Rogers, and Bittner [25] discusses collectivity, granularity and parthood in the context of biomedical applications, although their observations are wide-ranging and in many cases of more general applicability. The key distinction made in this paper is between what the authors call "granular" and "determinate" parthood, where the former is an irreflexive, asymmetric, and non-transitive relation between physical objects and "collectives", whereas the latter is a reflexive, antisymmetric, transitive relation between physical objects. A collective is thus characterised as an entity to which physical objects stand in the relation of granular parthood, and these objects are called the "grains" of the collective (deliberately eschewing the more standard "member" in order to prevent confusion with the "member" relation of set theory). They go on to note that the grains of a collective all "play the same role" in the collective, and that the identity of a collective is not determined by its grains. As an example they contrast the skin cells of a hand, which form a collective in their sense, and the fingers of a hand, which do not — the fingers being determinate parts of the hand.

It is clear that in [25] the term "collective" is largely restricted to our I-collectives; again, this is in line with the use of "collective", as opposed to "collection", in [11], E-collectives being, for them, merely scattered physical objects of which the "members" (in our sense) are determinate parts. They note that in biomedical contexts, at least, it is convenient to treat collectives as forming "amounts of matter". The parthood relation appropriate to amounts of matter is "ingredient of", which relates the components of a mixture to the mixture as a whole. This enables them, for example, to speak of an amount of blood as a mixture whose ingredients include amounts of matter that are collectives of red and white blood cells.

It would be a mistake, however, to align these authors' distinction between determinate and granular parthood too closely with our distinction between membership of E-collectives and I-collectives respectively. In their §4.3.2, they allow that 'there is no problem with the notion of "determinate collective"', to which may be referred 'collective effects of determinate collections of entities — the collective grip of the fingers, acuity of the eyes, the total capacity of the plates in a dinner service, etc.', and at first sight this might seem to imply that their use of the term "collective" can encompass both E-collectives and I-collectives. At least in most cases, however, it would seem that the collectives in question should be regarded as intensional, since, for example, a dinner service can survive the loss of some of its parts, and parts that are lost or destroyed can be replaced, without compromising the identity of the service itself.

In the same section, they suggest that the same entity could be both 'a determinate and granular part of the same whole', and it may be that this should be interpreted as a gesture towards the possibility of reading, say, "the salt in the pot" either *de dicto* or *de re*, the former referring to something that would not be destroyed by the removal or addition of a grain (these being granular parts of the whole), the latter referring to something that would be (since the salt grains are now determinate parts). In this case,

however, it is arguable that in these two usages "the salt in the pot" does not refer to 'the same whole'. This question is, as the authors note, intimately bound up with the issue of DOLCE's multiplicativism vs BFO's unitarism, referred to above. On the multiplicativist view, we can say that at any one time the salt in the pot *qua* intensional collective is *constituted* by the salt in the pot *qua* extensional collective — a different whole since it has different persistence conditions. On the unitarist view, one possibility would be to countenance some notion of temporally-indexed identity, allowing us to say that at any moment the (intensional) salt in the pot is (temporally) identical to the (extensional) salt in the pot. Another possibility would be to invoke the notion of roles and role-holders, as suggested above, although this may seem somewhat strained in the present example since it would require us to postulate role-holders such as "grain of salt *qua* occupant of the pot". The authors of [25], since they include members from both "camps", refrain from coming down on one side or the other of this debate.

Parthood is also central to a series of articles [13; 14; 15; 16] in which Guizzardi has explored several kinds of parthood in the light of different theories of the part-whole relation. In [13], he enlists the *de re*/*de dicto* distinction, as well as the distinction between (specific) ontological dependence and generic dependence, to resolve some subtle issues concerning the modal properties of part-whole relations. An *essential part* of an object is necessarily a part of the object so long as the object exists (as, for example, your brain is a necessary part of you), whereas a *mandatory part* of an object is a type such that so long as the object exists it must have *some* object of that type as a part (so, given the possibility of heart transplants, a heart is a mandatory part of you, but no heart is an essential part of you). An object is specifically dependent on its essential parts and generically dependent on its mandatory parts. These part-whole relations concern wholes considered as instances of *kinds*, where a kind is defined as "a type defining all the essential properties for the individuals it classifies". Thus *person* is a kind, and it is *qua* person that you have a brain as an essential part and a heart as a mandatory part. In the statement "Every person has a brain", it makes no difference whether we interpret "person" *de re* or *de dicto*, since applicability of the designation (*dictum*) "person" is exactly coextensive with the lifetime of the object (*res*) which it designates. But Guizzardi also considers *phased sortals*, which "possibly apply to an individual only during a certain phase of its existence", for example *boxer*, and it is in connection with these that the *de re* and *de dicto* distinction becomes important. He notes that the statement "Every boxer necessarily has a hand" is true in the *de dicto* reading of "boxer" but not in the *de re* reading: it is only *qua* boxer that an individual necessarily has a hand, since loss of the hand would necessarily result in the individual's no longer being a boxer. The *de re* reading, on the other hand, applies to the individual *qua* person, where *person* is the kind of which the phased sortal *boxer* is a subtype; and since a person can survive the loss of a hand, our boxer does not, *qua* person, necessarily have a hand. In Guizzardi's terminology, a boxer has a hand as an *immutable* part.

Guizzardi distinguishes "four distinct ontological types" of the part-whole relation (derived by conflating some of the six relations listed by [32]), namely *subquantity–quantity*, *member–collective*, *subcollective–collective*, and *component–functional complex*. Collectives provide the focus for [16], in which it is made clear that, for Guizzardi, the term "collective" should be restricted to collections of entities having a "uniform structure". By this he means that "all member parts play the same role type with respect to the whole", as for example all the trees in a forest play a role of type "forest member". A collection whose members play different roles — for example a string quartet, with roles first violin, second violin, viola, and cello — would, for Guizzardi, not be a collective but a functional complex. In addition he insists that a collective must have at least two members; this is in order to avoid the "ontological exuberance" such as we find in set theory, where the sets $\{a\}$, $\{\{a\}\}$, $\{\{\{a\}\}\}$, ..., are all distinct while in a certain sense having the same "content".[6] In all these respects, Guizzardi offers a more restricted notion of collectives than are allowed by, for example, [35], but within this more restricted setting he does find room to distinguish between intensional and extensional collectives, for which he uses precisely that terminology. He explicitly links the property of extensionality with the notion of essential parthood: an

---

[6]It is not clear, though, how Guizzardi's stipulation prevents this: the sets $\{\{a, b, c\}, \{d, e, f\}\}$ and $\{\{a, b\}, \{c, d\}, \{e, f\}\}$ are also distinct but arguably have the same content.

extensional connective is one for which the *member–collective* relation is an essential parthood relation. That is, the members of the collective are essential parts (indeed, it would be appropriate here to speak of "essential members") of it.

Interestingly, although Guizzardi makes use of the *de re / de dicto* distinction in [13], and discusses collectives in [16], he does not appear to bring the two topics together in the way that we have done here. This may be perhaps explained by a significant disanalogy between the cases he considers and our examples. In the case of Guizzardi's boxer, the relevant distinction is between the applicability of the term "boxer" and the existence of the entity to which, at some time, it applies: the relevant constraint is that the temporal interval associated with the former must be wholly contained within that associated with the latter. With our committee, on the other hand, the distinction is between the application of the term "committee" and the existence of the collection of individuals to which, at some time, it applies; and these are only subject to the much weaker constraint that the associated temporal intervals overlap.

Formal aspects of essential, mandatory and immutable parthood are explored further by Artale *et al.* [4]. They make use of *status classes* [30] to help in elucidating these. The statuses considered are "scheduled", "active", "suspended", and "disabled", which they apply not just to classes but to relations. This enables them to characterise immutable parts and wholes in terms of the possible temporal relations between the intervals over which the whole and its relation to the part have these different statuses. Although they do not explicitly consider collectives, similar considerations might be applicable. Consider an I-collective whose members are mandatory but not essential, such as a chess-set:[7] if the black queen goes missing, it must be replaced by the same or another one for the chess set to continue to exist. This is comparable to the dinner service mentioned by Rector *et al.* During the period between the loss of the queen and its subsequent replacement, however, it is not appropriate to say that the chess-set has been destroyed ("disabled"); rather, it is "suspended". In the terminology of [22] we would say that the role of the black queen in that chess set still exists but is currently unfilled.

Bottazzi *et al.* [7] are concerned with "intentional collectives", where "intentional" (with a "t") is not to be confused with "intensional" (with an "s") as used above; nor is "intentional" to be confused with "intention" in the sense of what one intends to do, but is rather related to *intentionality* in the sense of "aboutness", as described by Searle [27]. In this paper, as distinct from [6] and [11], the term "collection" is used to denote an I-collective; for example, "a collection of books in a library remains the same entity even if some books are lost and others are acquired over time". But collections, for Bottazzi *et al.*, are not just functions from times to sets of members, because they "depend . . . on the *role(s)* played by their members". These roles are characterised by means of *descriptions* which provide *unity criteria* for the collection, for example "book in this library", and it would seem that such descriptions correspond to the *dicta* whose *de dicto* readings are used to specify I-collectives in our sense. In common with other authors, they are at pains to emphasise that collections are not to be identified with mathematical sets, the main points of difference being that, unlike sets, collections (a) can have variable membership, (b) cannot be arbitrarily combined with others to form new collections, (c) must possess identity criteria for membership, (d) cannot have less than two members, and (e) are located in space and time. As Bottazzi *et al.* use the term, *collectives* are "something more than collections . . . but something less than social groups". In fact they define a collective as a special kind of collection, namely "a collection of agents unified by a plan" — for example "people waiting at a bus stop", an example they attribute to Sartre. There appears to be an element of circularity in their definitions of the key terms here, however, since a plan is defined as a description representing an action schema that, amongst other things, is conceived by a cognitive agent, and a cognitive agent is defined as something that can "conceive a plan". We shall not pursue this further, however, as their distinction between collections and collectives is tangential to the main themes of our paper.

---

[7]Note that for Guizzardi this would not be a collective but a functional complex, since its members have different roles with respect to the whole and therefore should rather be described as components.

3. E-collectives and I-collectives

As discussed in §1, we use the term 'E-collective' to denote a collective which necessarily always has the same members. This means that the collective can only exist so long as those members exist. A consequence of this is that the collective is *specifically dependent* on its members, meaning that the former cannot exist without the latter existing. Using the predicate EC to mean "is an E-collective", this can be expressed by the formula

$$EC(x) \wedge Member(y, x) \rightarrow \Box(Exists(x) \rightarrow Exists(y)) \tag{1}$$

More specifically, we can say that if y is ever a member of x then it is necessarily a member of x at all times at which that collective exists. To express this, we need to relativise existence and membership to times, adding an additional temporal argument-place to each of the predicates $Exists$ and $Member$. Note that, membership being a specialised form of parthood, this is fully in keeping with the widespread practice of relativising parthood to time, as for example discussed in some detail in the context of biomedical ontologies by Schulz *et al.* [26] and in a more general context by [5]. Our formula now becomes:

$$EC(x) \wedge Member(y, x, t) \rightarrow \Box \forall t'(Exists(x, t') \rightarrow Member(y, x, t')) \tag{1'}$$

A question which will assume some importance for us is whether it is also true conversely that an E-collective exists at all times at which all its members exist, i.e., whether we have

$$EC(x) \rightarrow [\forall y(\exists t' Member(y, x, t') \rightarrow Exists(y, t)) \rightarrow Exists(x, t)] \tag{2}$$

Given (2), if we regard the present members of the committee as forming an E-collective, that collective will continue to exist, with the same membership, so long as those members all remain alive, independently of whether they remain members of the I-collective which is the committee itself.[8]

None of this, however, forces us to regard the present members of the committee as forming an E-collective at all; we could just decide to recognise the existence of the I-collective (i.e., the committee, understood *de dicto*) and leave it at that. What are the conditions for an E-collective to exist? We can consider three different conditions, in increasing order of stringency.

The least stringent, or most liberal, condition states that for *any* set of individuals, so long as there is at least one time at which all those individuals exist simultaneously, there is an E-collective which they form, and which exists at any time at which they all exist. Thus characterised, an E-collective may be regarded as a *bare plurality*, defined solely by its members. Insofar as it is appropriate to apply the term "collective" to such a thing at all, it represents a limiting case of the range of applicability of that term. In depending entirely on its members, a bare plurality resembles a mathematical set, but as has been widely acknowledged, it cannot be equated with the set of its members. The existence of an E-collective, unlike that of a mathematical set, is time-dependent: an E-collective exists at a time if and only if all its members exist at that time:[9]

$$EC(x) \rightarrow \exists S(Set(S) \wedge \forall t(Exists(x, t) \leftrightarrow \forall y(y \in S \rightarrow Exists(y, t)))) \tag{3}$$

Mathematical sets, by contrast, are thought to exist eternally, if at all. Still, when understood in this most liberal sense, E-collectives do share some of the ontological *blandness* of sets: insofar as such a bare

---

[8]Bittner *et al.* [6] take a somewhat different view: They would say that the (E-)collective is *fully present*, *partly present*, or *non-present* at t according as all, some, or none of its members exist at t.

[9]Note that in this formula, the use of the connective '↔' implies that we are taking the most liberal view regarding the existence of E-collectives. On any more stringent view, the most we can assert is the formula obtained by replacing this biconditional by a one-way (left-to-right) conditional.

plurality has collective properties, these are solely dependent on the properties of its individual members. The unity, such as it is, of an E-collective consists solely in the fact of its having just those members.

The question of whether or not every bare plurality can be regarded as forming an E-collective is similar to the question of whether every bare plurality has a mereological sum. The collective and the sum, assuming both exist, must coincide spatially, but they are not the same thing. The sum of two bicycles, for example, is also the sum of two frames, four wheels, four pedals, and so on; but the collective whose members are the two bicycles is disjoint from (i.e., shares no members with) the collective whose members are the components of those bicycles. Moreover, because at least some components of a bicycle are replaceable without prejudice to the bicycle's identity, these two collectives need not have the same sum throughout their history.

The second, more stringent, condition we might postulate for the existence of an E-collective is that a set of individuals forms an E-collective so long as at some time they constitute some I-collective. On this understanding, an E-collective is what is picked out by the *de re* reading of a collective designator that primarily picks out, through its *de dicto* reading, some I-collective. The six current members of the committee, considered *de re*, form an E-collective because they currently constitute an I-collective (the committee understood *de dicto*). At a later time when these individuals no longer constitute the committee, the E-collective they form still exists, and can indeed be referred to as, say, the committee of 2015, a designation that persists, so that, for example, we might say "By 2020 the committee of 2015 were all dead". But whereas *any* six individuals picked at random form an E-collective under the first criterion, they will not do so under the second unless they also, at some time, form an I-collective.

The third, most stringent condition for E-collectivity requires that there is some principle of unity that unites them at *all* times. The possible principles of unity form a heterogeneous bunch, ranging from properties wholly intrinsic to the collective (e.g., that the individuals in question consider themselves to form that collective) to, at the other extreme, properties wholly extrinsic to the collective (e.g., that the individuals are considered by some external agent to form a collective) — with a variety of intermediate cases such as, for example, collectives exhibiting spatial proximity arising as a result of some external causal influence.[10] Under this most stringent condition, an E-collective must, throughout its existence, coincide with a particular I-collective; if the membership of the latter changes, the former goes out of existence. Thus an E-collective, considered in this way, is a *phase* in the existence of an I-collective, and the life of the I-collective consists of a succession of such phases, with one phase giving way to the next each time the membership of the collective changes. We could understand "the committee of 2015" in this way, assuming that there was no change of membership of the committee during 2015: the committee of 2015 first comes into existence when the members are elected onto the committee, and it goes out of existence when the following year's committee is elected. (By contrast, under either of the more liberal conceptions of E-collective, it persists beyond that point.)

It is also possible, of course, that an I-collective never does change its membership. In this case its life consists of just one phase, and the only distinction between the I-collective and the E-collective is in their differing modal properties (e.g., the I-collective *could have* outlived the E-collective). An example of a collective of this kind is, perhaps, the triumvirate of Octavian, Antony, and Lepidus.

If we accept that the members of a collective are *parts* of the collective, as suggested by [32] and endorsed by [35], then the members of an E-collective (under any of the above conceptions) must be *essential parts* in the sense of [4; 13] (see §2). If the members of an E-collective are essential parts of it, there must be a close relationship between the collective and the mereological sum of its members, since any object must be an essential part of the mereological sum of any plurality to which it belongs. The series of successively more restrictive conceptions of E-collectivity we considered above would then correspond to a similar series of attitudes that one might take to the existence of mereological sums — with the most liberal stance on collectives matching the doctrine of unrestricted mereological fusion that is accepted by many (but by no means all — cf. [9]) ontologists.

---

[10]Compare the discussion of coherence in [35], where collectives bound by a wholly extrinsic principle of unity are called "pure fiat", as opposed to more "bona fide" collectives whose unity criteria are more intrinsic to the collectives themselves.

To summarise, then, an E-collective necessarily always has the same members, which are essential parts of the E-collective, and the E-collective is specifically dependent on them. Moreover, an E-collective must coincide spatially with the sum of its members so long as both these entities — the collective and the sum — exist.

An I-collective, by contrast, has none of these properties. It does not necessarily always have the same members; if it so happens that such a collective does always have the same members, this circumstance has arisen as it were by chance: had things turned out differently, then it might have had different members for some or all of its history. We cannot say this of an E-collective. The members of an I-collective are not, or not all, essential parts, meaning that they can leave the collective without destroying the collective's identity.[11] Again, the relationship between an I-collective and the mereological sum of its members is much looser. At any one time, the collective spatially coincides with the mereological sum of its members at that time, but as soon as one of the members leaves the collective, or a new member joins, the collective and the original mereological sum part company. Finally, the dependence of an I-collective on its members is not specific; but it could be generic. More exactly, to specify an I-collective, it would be usual to specify what type of members it has — e.g., people, sheep, books. Then to say that the collective is generically dependent on that type, means that the collective itself cannot exist unless entities of that type exist [11]. Whether or not we accept this depends on whether or not we wish to allow there to be empty collectives. If we wish to allow a collective to exist even when it has no members (e.g. "the queue at the post-office counter") then we cannot say that the collective is dependent on members, since it can exist even when it has none. This is a point we will take up in more detail below.

Just as there are different conceptions of E-collective depending on how liberal we want to be in recognising arbitrary pluralities as forming such collectives, we can also distinguish different degrees of liberality in our conception of I-collectives. The most liberal conception would be that, for any predicate ϕ ascribing a property to an individual at a time, there is an I-collective whose members, at time t, are precisely the individuals x such that $\phi(x, t)$. The predicate ϕ may be regarded as supplying the unifying principle for the collective, but it is evident that some predicates provide greater unity than others. The more unifying predicates will typically be relational in nature, involving not just properties of the members as individuals but also how they are related to other individuals. For example, in a marriage partnership, the defining property is that each member of the partnership has undertaken certain commitments towards the other; in a football team, each member has entered a commitment to further the goals of the team as a whole over and above his or her own personal goals. In what follows, while we allow that predicates of this kind may be involved in the definition of a collective, we will not focus on these specifically.

It should be noted that for many I-collectives the relevant predicate is self-referential, e.g., along the lines of 'x has whatever it takes to be member of c at t'. How else, for example, could one define, say, the London Symphony Orchestra as a collective: its members at a time are precisely those people who are signed-up members of the orchestra at the time. Although they may not exactly have signed a piece of paper saying "At time t, x is a member of the London Symphony Orchestra", they will presumably have undergone some form of initiation into the orchestra that is tantamount to this, so the predicate ϕ in this case simply states that an individual has joined the orchestra, and has not left it.[12]

We shall use the terms *unrestricted E-collectivity* and *unrestricted I-collectivity* to refer to the most liberal interpretations above. Thus, under unrestricted E-collectivity, for any set of individuals, at any time at which all the members of the set exist, there is an E-collective which has exactly those members; and under unrestricted I-collectivity, for any property that can be predicated of individuals at times, there is an I-collective whose members at any given time are all the individuals which satisfy that property at that time. (It is a separate question whether an I-collective exists when no individuals satisfy its defining property.)

---

[11]There are cases where a collective that is properly regarded as intensional nonetheless has some essential parts: thus "Professor X's research group" presumably has Professor X as an essential part even though none of the other members are essential.

[12]In a similar way, we can only really define the state of wedlock in terms of having previously undergone an appropriate marriage ceremony, which means that we cannot without circularity define "getting married" as "entering into wedlock".

As already indicated, these unrestricted notions of E- and I-collectivity are limiting cases, and as such may prove unsuitable for most practical applications where notions of collectivity are required. From a formal point of view, however, as so often happens, it is precisely these limiting cases that are most tractable, and the formalisations of them that we provide below are offered as a kind of "base line" from which, perhaps, one might specify the departures that are needed for extending the formalisation to other cases of collectivity that one might encounter in various applications.

We suggested in §1 that at any time of its existence an I-collective is constituted by the E-collective formed from its members at that time. Those individuals are temporary members of the I-collective but essential (and therefore permanent) members of the E-collective. Let us try to formalise the relationships amongst these three types, i.e., I-collectives, E-collectives, and individuals which can be members of either of these types. The account here is similar in spirit to that of [10], but goes further than that account in explicitly recognising I-collectives and E-collectives as distinct types.

We shall adopt the following notations:

– The E-collective whose members are precisely those of the set S will be denoted $ec(S)$. Unrestricted extensional collectivity may then be expressed by the formula

$$Member(a, ec(S), t) \leftrightarrow a \in S \land \forall x(x \in S \rightarrow Exists(x, t)) \tag{4}$$

If S is given as, say, $\{a_1, \ldots, a_n\}$, we shall write $ec(a_1, \ldots, a_n)$ rather than $ec(\{a_1, \ldots, a_n\})$.[13]

– The I-collective whose members at any time t are precisely those individuals satisfying the condition $\phi(x, t)$ will be denoted $ic(\phi)$. Unrestricted intensional collectivity is expressed by:

$$Member(a, ic(\phi), t) \leftrightarrow \phi(a, t) \tag{5}$$

– We shall write $X \approx_t Y$ to mean that collectives X and Y coincide (i.e., have the same membership) at time t:[14].

$$X \approx_t Y =_{\text{def}} \forall x(Member(x, X, t) \leftrightarrow Member(x, Y, t)) \tag{6}$$

Under the more restricted interpretations of extensional and intensional collectivity, there is no guarantee, given a set S or a predicate ϕ, that any such collectives as $ec(S)$ or $ic(\phi)$ exist. In this case the biconditionals in (4) and (5) would have to be replaced by conditionals.

Under a certain very natural assumption, we can now derive that for an I-collective, at any time that it exists there is an E-collective which it coincides with. The assumption we need for this is that the predicate ϕ by which the I-collective is defined has the property that

$$\forall x \forall t(\phi(x, t) \rightarrow Exists(x, t)) \tag{7}$$

i.e., at any time the property can only hold of an individual at a time if that individual exists at that time. We then have

$$Member(a, ic(\phi), t) \leftrightarrow \phi(a, t) \quad \text{(by (5))}$$
$$\leftrightarrow \phi(a, t) \land \forall x(\phi(x, t) \rightarrow Exists(x, t)) \quad \text{(by (7))}$$
$$\leftrightarrow a \in \{x \mid \phi(x, t)\} \land \forall x(x \in \{x \mid \phi(x, t)\} \rightarrow Exists(x, t))$$
$$\leftrightarrow Member(a, ec(\{x \mid \phi(x, t)\}), t) \quad \text{(by (4))}$$

---

[13]This means that for a set $\{S\}$ the notation $ec(S)$ is strictly speaking ambiguous since it could denote either the collective whose members are the members of S or the collective whose only member is the set S. This is unlikely to cause problems in practice.

[14]If one adopts the view described above, whereby the members of I-collectives are role-holders, then the definition of coincidence will need to be modified. We believe it should be possible to do this in such a way that the result (8) below still follows once suitable adjustments of notation have been made.

and therefore, by (6),

$$ic(\phi) \approx_t ec(\{x \mid \phi(x,t)\}) \tag{8}$$

Because any I-collective must coincide with some E-collective at each time that it exists, one might propose that I-collectivity can be reduced to E-collectivity in the following sense: an I-collective is completely determined by a *partial function* from times to E-collectives. Thus for each time t in some range we specify an E-collective $c(t)$, and identify the I-collective determined by c as $ic(\lambda x.\lambda t.Member(x, c(t)))$. Of course, in practice the normal way of specifying the function c would be by means of some predicate $\phi(x, t)$, bringing us back to the normal way of defining I-collectives. Roughly speaking: the intension determines the extension, and not the other way round.

This suggests that we might give *in*tensional definitions of *ex*tensional collectives, using the formula

$$ec(S) = ic(\lambda x.\lambda t.[x \in S \wedge \forall x(x \in S \rightarrow Exists(x, t))]),$$

which follows from the definitions above. Again, this only works if we allow unrestricted I-collectivity, for otherwise we might reject the property $\lambda x.\lambda t.[x \in S \wedge \forall x(x \in S \rightarrow Exists(x, t))]$ as a legitimate defining property for an I-collective.

It thus seems that under unrestricted E- and I-collectivity the distinction between E-collectives and I-collectives collapses, or rather, E-collectives become a limiting case of I-collective, as suggested by Simons [29]. But if, as suggested above, we have reason to doubt that E-collectivity and/or I-collectivity should be unrestricted, then this becomes less tenable, and we should rather regard the two forms of collectivity as quite distinct.

4. Identifying Spatial Collectives in Data

From a computational perspective, an important problem concerning collectives is to find algorithmic ways of identifying collective behaviour when presented with a data-set recording the positions of individuals at different times. Problems of this type have been studied as part of the relatively new research field of *movement pattern analysis* [2; 8; 12; 18; 20; 19; 33; 34]. In general, to identify any kind of collectivity from masses of individual data points, one searches for *commonalities* shared by subsets of the data. Thus for example Laube and Imfeld [18] define various kinds of *incidents*, that is, "patterns occurring among multiple objects", for example *concurrence*, defined as an "incident of n MPOs [i.e., moving point objects] showing the same motion attributes value at time t", for which they give as an example a group of deer simultaneously moving in the same direction. Dodge *et al.* [8] similarly define a *moving cluster* as "a set of objects that stay close to each other while taking the same path for a specific duration". Of particular pertinence to the present paper, they distinguish between *fixed* and *varying* moving clusters, depending on "whether the participating entities stay the same or change during the observed period". While this distinction might suggest our distinction between E- and I-collectives, this appearance is a little misleading since the category of fixed moving clusters will include collectives in which the constancy of membership is merely contingent as well as those in which it is necessitated by the way in which the collective is defined.

In this section we demonstrate how the ideas discussed in the foregoing sections impact on movement pattern analysis, building on work previously reported in [33], and in the concluding section we frame this explicitly in terms of E-collectives and I-collectives. We shall assume that the data consists of records of the form $\langle i, \vec{p}, t \rangle$, which states that the individual with identifier i was in position $\vec{p}$ at time t. The position $\vec{p}$ would normally be given as a tuple of coordinates, depending on the spatial reference frame used.

It is not to be supposed that *all* collectives reveal themselves through the position-histories of their members. We use the term *spatial collective* to describe those that do. It is not possible to give a rigorous, hard-and-fast criterion to demarcate spatial from non-spatial collectives, but roughly speaking we can characterise a spatial collective as a collective whose condition of unity is spatial, i.e., involves the position

or motion of its members.[15] What this means is that the members of a spatial collective can be expected to exhibit some commonality with respect to their spatial behaviour. Different kinds of commonality will define different kinds of spatial collective. From now on we shall assume that all the collectives we are dealing with are spatial collectives, so that their members are spatio-temporally located physical objects.

The forms of commonality considered in [33] are *common position* and *common motion*. Each of these can take two different forms, depending on whether we focus on the individuals or their position or motion values. Thus we have a two-by-two cross-classification as shown in Table 1.

Table 1
Four criteria for detecting spatial collectivity

	Common position	Common motion
Individual-based	PI	MI
Value-based	PV	MV

The meanings of these four categories can be spelt out as follows:

PI "Common position, individual-based": A sufficiently large set of individuals is regarded as exhibiting spatial collectivity of type **PI** if their positions are sufficiently similar at sufficiently many times.
MI "Common motion, individual-based": A sufficiently large set of individuals is regarded as exhibiting spatial collectivity of type **MI** if their velocities are sufficiently similar at sufficiently many times.
PV "Common position, value-based": If for a given position, there are sufficiently many times at which sufficiently many individuals are sufficiently close to that position, then those individuals are regarded as exhibiting spatial collectivity of type **PV**.
MV "Common motion, value-based": If for a given velocity, there are sufficiently many times at which sufficiently many individuals have velocities sufficiently close to that velocity, then those individuals are regarded as exhibiting spatial collectivity of type **MV**.

These are, of course, rather imprecise statements, in that what is meant by each of the occurrences of "sufficiently" is left open to interpretation. This vagueness is deliberate: the idea is that one can adjust the various parameters here to suit the data and application in question. As emphasised by Andrienko and Andrienko [1; 2], it is unrealistic to expect humanly-meaningful analysis of spatio-temporal datasets to be accomplished in a fully automated way: the presence of a human analyst is required, e.g., to determine suitable settings for such parameters as (in our case) the minimum number of individuals that can constitute a collective, or the minimum number of time-steps over which commonality must be manifested in order to furnish evidence of collectivity.

To be more precise, we can designate *thresholds* as follows:

– *Cardinality threshold* θ_k, the minimum number of individuals needed to count as "sufficiently many".
– *Position proximity threshold* θ_p, such that a group of individuals will count as having "sufficiently similar position" so long as no two of them are separated by a distance greater than θ_p.
– *Motion proximity threshold* θ_v, such that a group of individuals will count as having "sufficiently similar motion" so long as no two of them have velocities differing by more than θ_v.
– *Frequency threshold* θ_f, the minimum number of times a condition must recur in order to count as occurring "sufficiently often".

---

[15]Dodge *et al.* [8] draw a distinction between *groups* and *cohorts*, where the members of a group "share a behaviorally relevant functional relationship (e.g., a flock of sheep, a wolf pack)" while those of a cohort "merely have a factor in common that may be statistically relevant, such as similar age ... or sex". A spatial collective in our sense might be either a group or a cohort, depending on whether the spatial attributes that unify it reflect a "functional relationship" or are "merely statistical". Distinguishing these types on the basis of the data alone can be problematic.

Criteria for the different forms of spatial collectivity may now be stated as follows. Here we write $\vec{p}(i, t)$ for the unique $\vec{p}$ such that the data-set contains the triple $\langle i, \vec{p}, t \rangle$, and $\vec{v}(i, t)$ for the velocity of i at time t, as estimated from the positions of i at t and the nearest times to t for which positions of i are recorded in the data-set.[16]

PI There is a set I of individuals and a set T of times such that

    1. $|I| \geq \theta_k$
    2. $|T| \geq \theta_f$
    3. $\forall t \in T \, \forall x, y \in I \left(|\vec{p}(x, t) - \vec{p}(y, t)| \leq \theta_p \right)$

MI There is a set I of individuals and a set T of times such that

    1. $|I| \geq \theta_k$
    2. $|T| \geq \theta_f$
    3. $\forall t \in T \, \forall x, y \in I \left(|\vec{v}(x, t) - \vec{v}(y, t)| \leq \theta_v \right)$

PV There is a position $\vec{p}$ and a set T of times such that

    1. $|T| \geq \theta_f$
    2. $\forall t \in T (|\{x \; : \; |\vec{p}(x, t) - \vec{p}| \leq \theta_p\}| \geq \theta_k)$

MV There is a velocity $\vec{v}$ and a set T of times such that

    1. $|T| \geq \theta_f$
    2. $\forall t \in T (|\{x \; : \; |\vec{v}(x, t) - \vec{v}| \leq \theta_v\}| \geq \theta_k)$

For each of these types we can define two variants [33], which are called the *continuous* and *intermittent* cases. In the former case, the times in set T are consecutive: that is, any time represented in the data set that falls between two members of T must itself be in T. In the latter case, this does not hold. The idea is to distinguish collectives which exhibit continuous spatial unity from those whose members come together only occasionally. This distinction may be more or less useful, depending on the nature of the data set. If the data set is *synchronous*, that is, it records the position of each individual at each time, then continuity or lack of it can be informative; if on the other hand it is *asynchronous*, so that each individual is recorded at a different set of times, then continuity or lack of it is harder to interpret.

To illustrate these different types of collectivity, we use a simple artificial example in the spirit of Figure 4 in [19]. In Figure 1 are shown, in the upper diagram, the position-histories of six individuals labelled A–F moving in a one-dimensional space comprising 14 distinguishable locations over a period of 12 timesteps, and in the lower diagram, the corresponding velocity-histories, derived from the position-histories using the formula given in footnote 16.[17]

With this data, suppose we set our thresholds to be

$$\theta_k = 2, \theta_p = 0, \theta_v = 0, \theta_f = 5.$$

Then, using the "continuous" variants of the criteria rather than the "intermittent", we have

– Under criterion **PI**, we can identify a collective with members {A,B} subsisting over timesteps 2–7, in the process moving from position 3 to position 6 and resting there.
– Under criterion **PV**, we can identify a collective occupying position 6 over timesteps 5–9, with membership varying from {A,B,C} to {A,B} and then {B, D}.

---

[16] If t^- and t^+ are respectively the latest time earlier than t and the earliest time later than t for which a position is recorded for i, then a reasonable estimate is

$$\vec{v}(i, t) = \frac{t^+ - t}{t^+ - t^-} \cdot \frac{\vec{p}(i, t) - \vec{p}(i, t^-)}{t - t^-} + \frac{t - t^-}{t^+ - t^-} \cdot \frac{\vec{p}(i, t^+) - \vec{p}(i, t)}{t^+ - t},$$

i.e., an appropriately weighted average of the average velocities over the intervals $[t^-, t]$ and $[t, t^+]$.

[17] Since a velocity at time t can only be computed if positions are available for times $t - 1$, t, and $t + 1$, the velocity histories comprise 10, rather than 12, timesteps.

TIMESTEP

POSITION	0	1	2	3	4	5	6	7	8	9	10	11
1	A											A
2		A										
3	B C	B	A B								A	
4		C		A B						A		
5			C		A B				A		B	BD
6	DEF	DE	DE	CE	C	ABC	AB	AB	BD	BD	D	
7		F		D	E		C	D				
8			F		D	E	D	C	C			
9				F		D	E			C		E
10					F			E			CE	
11						F			E	E		C
12						F						
13								F			F	F
14									F	F		

TIMESTEP

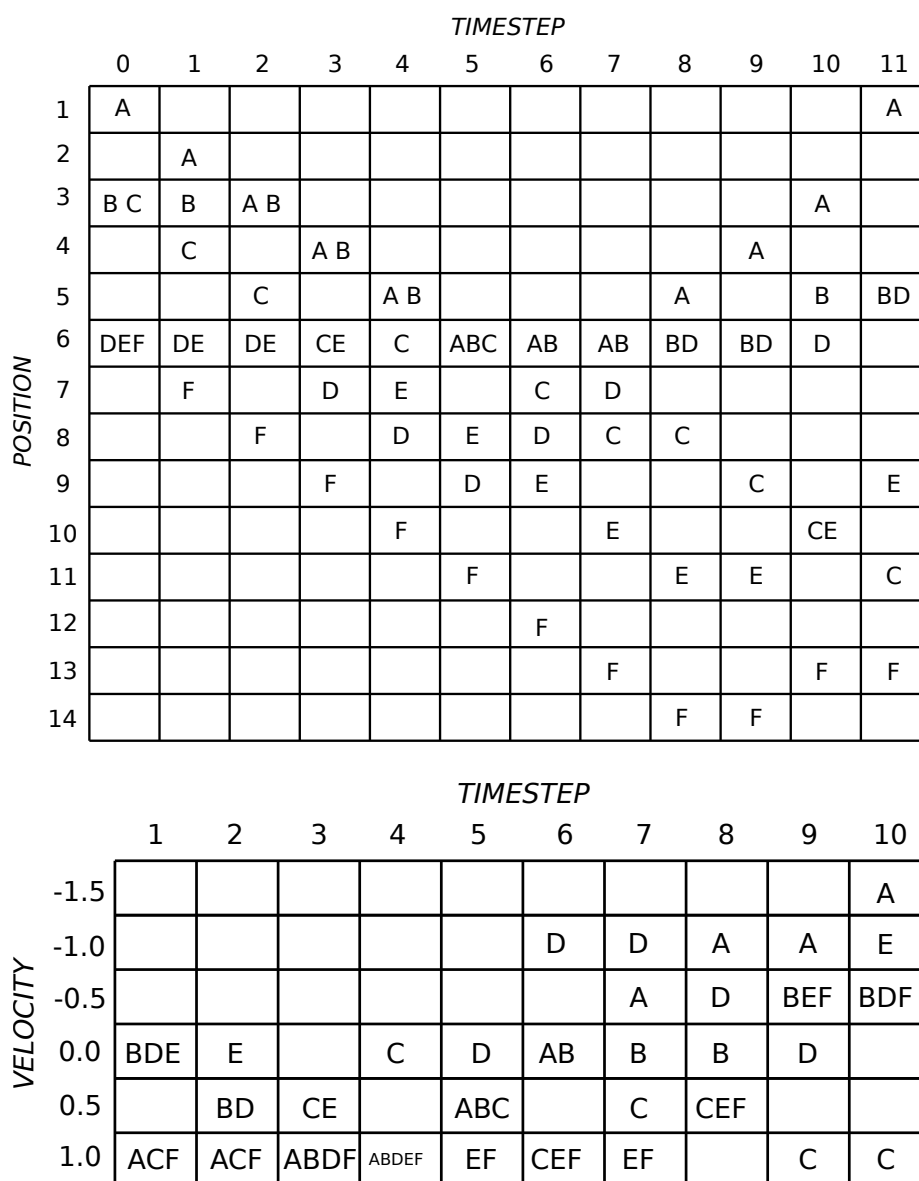| VELOCITY	1	2	3	4	5	6	7	8	9	10
-1.5										A
-1.0						D	D	A	A	E
-0.5							A	D	BEF	BDF
0.0	BDE	E		C	D	AB	B	B	D	
0.5		BD	CE		ABC		C	CEF		
1.0	ACF	ACF	ABDF	ABDEF	EF	CEF	EF		C	C

Fig. 1. Top: Position histories for six individuals; Bottom: Velocity histories computed from the position histories

– Under criterion **MI**, we can identify a collective with members {E,F} exhibiting a common motion pattern over timesteps 4–9, with velocity initially 1, and eventually −0.5; during this motion E and F are separated by a distance of 3, and therefore would not be picked up as exhibiting collective behaviour using criterion **PI**.

– Under criterion **MV**, we can identify a collective comprising individuals moving with velocity 1 over timesteps 1–7, its membership varying but with F as a constant member. Qualitatively we might describe this collective as a steady stream of individuals moving in the positive direction (downwards on the page) — but not always the same individuals.

Collectives encountered in real life can often be characterised as exhibiting one or other of these forms of collectivity, and in principle could, given suitable data, be identified using the corresponding criteria. Some examples are:

– A choir on tour consists of a particular group of individuals who come together at a sequence of distinct venues, thus manifesting type **PI**. If they all travel together as well, then it is the continuous

variant, but if they travel between the venues separately, it is intermittent.

– Certain constellations are not just chance alignments of stars in the night-sky but form genuine clusters sharing a common origin and retaining a certain coherence due to their mutual gravitation; five of the seven bright stars forming the Big Dipper are of this type.[18] Such groups are detected by their sharing a common proper motion against the background of relatively fixed more distant stars and galaxies. They therefore exhibit pattern **MI**.

– At the Louvre in Paris, throughout the day there is always a crowd in front of the *Mona Lisa*, close to the painting itself; but there is a continuous turn-around of individuals constituting that crowd. This is an example of the continuous variant of type **PV**. For an intermittent example, consider the crowd gathering round, for example, the Glockenspiel in Bahnhofstraße in Zürich, which only operates twice a day — so most of the time the crowd is absent.

– In a shower of rain, at any time there are a large number of drops falling from the clouds to the ground, all with very nearly the same velocity. But at different times during the shower different individual drops are involved. This provides an example of type **MV**.

These types do not exhaust the possibilities. In particular, we could define some more sophisticated variants of the "common motion" criteria, e.g., in which what we are looking for is not a particular velocity but rather a particular *pattern* of motion, as for example exhibited by cars going round a roundabout, or a common *description* of the motion, e.g., moving towards a particular position, which entails different velocities if different individuals are coming from different directions (*convergence* behaviour).

We have applied these ideas to analysis of real-world data. Here we give only a brief summary of the results that are relevant to this paper; for technical details of how the analysis was carried out, see [33; 34]. The data were collected from ships in the Solent, the stretch of water separating the Isle of Wight from the British mainland. Each of the ships was fitted with an Automatic Identification System (AIS) which enabled it to generate records including its position and bearing at different times.

The simplest and most obvious case concerns the criterion **PI**, which can be used to identify groups of ships which travel together, maintaining proximity over a period of time. What the criterion cannot tell us, of course, is whether this behaviour arises from coordinated action on the part of the ships — which thereby form a *convoy* — or whether it arises because the ships happen, independently, to be following the same course at the same time. At any rate, discovery of such a group of ships provides a prima facie case for further investigation to distinguish these two cases.

More interestingly, by applying the **PV** condition to the data, it was possible to identify various locations characterised by the frequent congregation of ships. These included the main ports, as well as pilot stations where ships stop to pick up a pilot who will guide them into port. Of course, these locations are not themselves collectives, but there is an obvious correspondence between, say, a port and the I-collective that may be designated "the ships in the port".

Application of the **MV** condition identifies the ships following the shipping lanes. As with the port, we have here a situation where spatial features (the shipping lanes themselves) are identified through the collective behaviour of individual entities located in or moving through them. Thus the behaviour of collectives can provide information about their environment [34].[19]

5. Conclusions

In this paper we have drawn connections between a number of seemingly disparate issues that have repeatedly surfaced in ontology: the distinction between *de re* and *de dicto* readings of referring expressions, the distinction between specific and generic dependence, the various kinds of parthood, and application-oriented issues concerning the identification of collective phenomena from raw data. In the previous sec-

---

[18] See www.astronomy.ohio-state.edu/˜pogge/Ast162/Movies/umapm.mov.

[19] In [33], an additional criterion, not considered here, was used: *common formation*, which identifies groups of individuals which maintain the same *relative* positions over an interval.

tion we showed how various criteria used for identifying connectives divide sharply into those which pick out E-collectives and those which pick out I-collectives. Now, in one final example we shall close the circle by explicitly linking these ideas to the *de re / de dicto* distinction we began with.

Consider, therefore, a school in which there are several classes and several classrooms. Each classroom belongs to one of the teachers, so that classes move from room to room for lessons with different teachers. Now consider the statement

> *In twenty minutes' time the children in room A will be in room B.*

This is quite plausible under a *de re* reading of the phrase "the children in room A", but under a *de dicto* reading it is necessarily false. The children in room A can be regarded as a collective in two different ways, depending on how we track them over time. On the one hand, we can latch onto those particular children, identifying them as a *class*, and follow them from room to room. The phrase "The children in room A" only serves to identify that class at certain times, namely when they are in room A. It is thus an identifier rather than a descriptor; what it identifies is an E-collective. On the other hand, we can stay in room A and observe the classes as they come and go. At different times the phrase "The children in room A" picks out different classes, and we are interested in all of them. Here the phrase serves as a descriptor, which identifies an I-collective.

Now consider which of the criteria discussed in §4 are appropriate for detecting these collectives. The E-collective comprising a particular group of children (which could be identified as the ones in room A at 12 o'clock on Monday or, say, the ones in class 4B this year) is detectable using the **PI** criterion. It is a group of individuals who are repeatedly in close proximity to one another. Once they have left school and gone their separate ways, they will no longer be detectable under this criterion; we might say that they still form a collective ("class of '13", say), but they no longer form a *spatial* collective — except perhaps occasionally if they come together in class reunions. On the other hand, the I-collective, comprising at different times the children who are in room A at those times, will be detectable using the **PV** criterion. They are the occupants of a spatial location in which individuals repeatedly come together in close proximity. We see here a clear alignment between, on the one hand, *de re* readings of collective designators and the individual-based detection criteria, and on the other, *de dicto* readings of the designators and the value-based detection criteria.

We can make a further interesting connection here, with the distinction between *Lagrangian* and *Eulerian* approaches in fluid dynamics. The Lagrangian approach is individual-based — in the fluid-dynamical context this means following a particular portion of fluid as it flows through the system under study. The Eulerian approach is location-based — here we fix on a location and observe the flow of fluid through that location. If we take the designator "the fluid at location *l*", then on the Lagrangian approach we are interpreting this *de re* whereas on the Eulerian approach we are interpreting it *de dicto*. If the "fluid" in question is aggregated from many individual units, then the Lagrangian approach studies the E-collectives formed by specific groups of such units, whereas the Eulerian approach studies the I-collectives formed by the units which successively occupy a given location. This dichotomy does not just apply to fluid dynamics, but can be applied, e.g., to studying the flow of vehicles in a traffic network. It would be interesting to see to what extent the terminology and methods associated with these two approaches could be adapted to the more general study of extensional and intensional collectives.

The various correlates of the distinction between E-collectives and I-collectives that we have discussed are summarised in Table 2.

References

[1] G. Andrienko, N. Andrienko, S. Rinzivillo, M. Nanni, D. Pedreschi, and F. Giannotti. Interactive visual clustering of large collections of trajetories. In *IEEE Visual Analytics Science and Technology (VAST 2009)*, pages 3–10. IEEE Computer Society Press, Atlantic City, NJ, USA, 2009.

[2] N. Andrienko and G. Andrienko. Designing visual analytics methods for massive collections of movement data. *Cartographica*, 42(2):117–138, 2007.

E-COLLECTIVE	I-COLLECTIVE
De re interpretation of collective descriptor	*De dicto* interpretation of collective descriptor
Specifically dependent on its members	Generically dependent on its member type
Members are essential parts	At least some members are non-essential (but possibly mandatory) parts
Spatial collectives detected using individual-based criteria (**PI**, **MI**)	Spatial collectives detected using value-based criteria (**PV**, **MV**)
Lagrangian dynamics	Eulerian dynamics

Table 2

Correlates of the distinction between extensional and intensional collectives

[3] Robert Arp, Barry Smith, and Andrew D. Spear. *Building Ontologies with Basic Formal Ontology*. The MIT Press, Cambridge MA and London UK, 2015.

[4] A. Artale, N. Guarino, and M. Keet. Formalising temporal constraints on part-whole relations. In G. Brewka and J. Lang, editors, *Proceedings of the Eleventh International Conference on Principles of Knowledge Representation and Reasoning*, pages 673–683. AAAI Press, 2008.

[5] T. Bittner and M. Donnelly. A temporal mereology for distinguishing between integral objects and portions of stuff. In R. C. Holte and A. Howe, editors, *Proceedings of the 22nd AAAI Conference on Artificial Intelligence*, pages 287–292. AAAI, 2007.

[6] T. Bittner, M. Donnelly, and B. Smith. Individuals, universals, collections: On the foundational relations of ontology. In A. Varzi and L. Vieu, editors, *Proceedings of the Third Conference on Formal Ontology in Information Systems (FOIS 2004)*, pages 37–48, Amsterdam, 2004. IOS Press.

[7] E. Bottazzi, C. Catenacci, A. Gangemi, and J. Lehmann. From collective intentionality to intentional collectives: an ontological perspective. *Cognitive Systems Research*, 7(2-3):192–208, 2006.

[8] S. Dodge, R. Weibel, and A.-K. Lautenschütz. Towards a taxonomy of movement patterns. *Information Visualization*, 7:240–252, 2008.

[9] C. Elder. Against universal mereological composition. *Dialectica*, 62(4):433–454, 2008.

[10] A. Galton. How is a collection related to its members? In B. Smith, R. Mizoguchi, and S. Nakagawa, editors, *Interdisciplinary Ontology, Volume 3*, pages 9–17, Keio University, Tokyo, 2010. Open Research Center for Logic and Ontology.

[11] A. Galton. On generically dependent entities. *Applied Ontology*, 9(2):129–153, 2014.

[12] J. Gudmundsson, M. van Kreveld, and B. Speckmann. Efficient detection of patterns in 2D trajectories of moving points. *Geoinformatica*, 11:195–215, 2007.

[13] G. Guizzardi. Modal aspects of object types and part-whole relations and the *de re/de dicto* distinction. In J. Krogstie, A. L. Opdahl, and G. Sindre, editors, *Advanced Information Systems Engineering, Proceedings of the 19th International Conference, CAISE'07, Trondheim, Norway, June 11-15, 2007*, pages 5–20. Springer, 2007.

[14] G. Guizzardi. The problem of transitivity of part-whole relations in conceptual modelling revisited. In E. Yu, J. Eder, and C. Rolland, editors, *Proc. 21st International Conference on Advanced Information Systems Modelling (CAISE'09), Amsterdam, The Netherlands*, http://ceur-ws-org, 2009. CEUR Workshop Proceedings.

[15] G. Guizzardi. On the representation of quantities and their parts in conceptual modelling. In A. Galton and R. Mizoguchi, editors, *Proc. Sixth International Conference on Formal Ontology in Information Systems (FOIS'10), Toronto, Canada*. IOS Press, 2010.

[16] G. Guizzardi. Ontological foundations for conceptual part-whole relations: The case of collectives and their parts. In H. Mouratidis and C. Rolland, editors, *Advanced Information Systems Engineering*, volume 6741 of *Lecture Notes in Computer Science*, pages 138–153. 23rd International Conference, CAISE 2011, London, UK, June 20-24, 2011, Springer, 2011.

[17] C. M. Keet and A. Artale. Representing and reasoning over a taxonomy of part-whole relations. *Applied Ontology*, 3:91–110, 2008.

[18] P. Laube and S. Imfeld. Analyzing relative motion within groups of trackable moving point objects. In M. J. Egenhofer and D. M. Mark, editors, *Geographical Information Science, Proceedings of the Second International Conference GIScience 2002*, pages 132–144, Berlin, Heidelberg, New York, 2002. Springer.

[19] P. Laube, S. Imfeld, and R. Weibel. Discovering relative motion patterns in groups of moving point objects. *International Journal of Geographical Information Science*, 19(6):639–668, 2005.

[20] P. Laube, M. van Kreveld, and S. Imfeld. Finding REMO—Detecting relative motion patterns in geospatial lifelines. In P. Fisher, editor, *Developments in Spatial Data Handling: Proceedings of the 11th International Symposium on Spatial Data Handling*, pages 201–214, Berlin, Heidelberg, New York, 2004. Springer.

[21] C. Masolo, S. Borgo, A. Gangemi, N. Guarino, and A. Oltramari. WonderWeb deliverable D18. Technical report, Laboratory for Applied Ontology, ISTC-CNR, Trento, Italy, 2003.

[22] R. Mizoguchi, A. Galton, Y. Kitamura, and K. Kozaki. Families of roles: A new theory of occurrent-dependent roles. *Applied Ontology*, 10(3-4):367–399, 2015.

[23] J. J. Odell. Six different kinds of composition. In *Advanced Object-Oriented Analysis and Design Using UML*. Cambridge University Press, Cambridge, 1998. Originally published in *Journal of Object-Oriented Programming*, January 1994.

[24] W. V. O. Quine. Quantifiers and propositional attitudes. *Journal of Philosophy*, 53(5):177–187, 1956.

[25] A. Rector, J. Roger, and T. Bittner. Granularity, scale and collectivity: When size does and doesn't matter. *Journal of Biomedical Informatics*, 39(3):333–349, 2006.

[26] S. Schulz, A. Kumar, and T. Bittner. Biomedical ontologies: What *part-of* is and isn't. *Journal of Biomedical Informatics*, 39:350–361, 2006.

[27] J. Searle. *The Construction of Social Reality*. The Free Press, New York, 1995.

[28] R. Sharvy. Things. *The Monist*, 53(3):488–504, 1969. Reprinted in Freeman, E., and Sellars, W. (eds.) *Basic Issues in the Philosophy of Time*, Open Court, Illinois, 1971, pages 164–180.

[29] P. Simons. *Parts: a Study in Ontology*. Clarendon Press, Oxford, 1987.

[30] S. Spaccapietra, C. Parent, and E. Zimanyi. Modeling time from a conceptual perspective. In G. Gardarin, J. C. French, N. Pissinou, K. Makki, and L. Bouganim, editors, *Proc. International Conference on Information and Knowledge Management CIKM'98*, pages 432–440. ACM, 1998.

[31] R. Wasserman. Material constitution. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Summer 2013 edition, 2013.

[32] M. E. Winston, R. Chaffin, and D. Herrmann. A taxonomy of part-whole relations. *Cognitive Science*, 11:417–444, 1987.

[33] Z. Wood. *Detecting and Identifying Collective Phenomena within Movement Data*. PhD thesis, University of Exeter, 2011.

[34] Z. Wood. What can spatial collectives tell us about their environment? In *Proceedings of the 2014 IEEE Symposium on Computational Intelligence and Data Mining*, pages 329–336, 2014.

[35] Z. Wood and A. Galton. A taxonomy of collective phenomena. *Applied Ontology*, 4:267–292, 2009.