

(2014) *Proceedings of the Aristotelian Society* 64(2)  
Post-talk author pre-print  
Please refer to published version for page numbers

## **Exploitable Isomorphism and Structural Representation**

**Nicholas Shea**

### **Abstract**

An interesting feature of some sets of representations is that their structure mirrors the structure of the items they represent. Founding an account of representational content on isomorphism, homomorphism or structural resemblance has proven elusive, however, largely because these relations are too liberal when the candidate structure over representational vehicles is unconstrained. Furthermore, in many cases where there is a clear isomorphism, it is not relied on in the way the representations are used. That points to a potential resolution: that an isomorphism must be used, hence usable, if it is to be an ingredient in a theory of content. This paper argues that the class of *exploitable* isomorphisms can indeed play a content-constituting role.

### **Contents**

- I Introduction
- II Case study
- III Unexploited Isomorphism
- IV Application to Cartographic Maps
- V Exploitable Isomorphisms
- VI Exploiting Isomorphisms Contrasted With Exploiting Correlations
- VII Beyond Strict Isomorphisms
- VIII Conclusion

## I Introduction

A cartographic map has a spatial structure that mirrors the spatial structure of the locations it represents. That fact seems important to the way that the map functions as a representation. When we trace out alternative routes between two points on a map and select the shortest, we rely on the fact that the shortest distance on paper corresponds to the shortest distance on the ground. Generalising away from the special case where spatial relations correspond to spatial relations (first order resemblance), any relation over a set of representations could in principle correspond to a relation over the set of things represented. When the relation over the representations not just corresponds to but also represents a relation over the entities represented, we have a case of *structural representation* (Ramsey 2007, pp. 77-92; Shagrir 2012; Swoyer 1991).

### *Structural representation*

A collection of representations in which a relation between representational vehicles represents a relation between the entities they represent.

Cartographic maps are structural representations because spatial relations on a map represent spatial relations on the ground. However maps are artefacts. Their intentionality may depend on us, their users. Are there structural representations whose representational properties do not depend on an intelligent interpreter? This paper argues that there are. It also addresses a further question. There has been considerable controversy over whether an isomorphism or second order resemblance between a set of putative representations and a domain of objects could give rise to or be the basis of representational content (Cummins 1989; Gallistel 1990; Godfrey-Smith 1996, pp. 184-87; Millikan 1984, p. 107; Millikan 2013; O'Brien and Opie 2004; Ramsey 2007). This paper offers a positive answer. I argue that, in some cases, the existence of a correspondence between a relation over a set of putative representations and a relation over the items represented can be part of what makes it the case that the representations have the content they do.

An obstacle to relying on isomorphism and related notions in a theory of content is the liberality of the isomorphisms that exist between two sets of equal size (Millikan 2000; Neander forthcoming, ch. 10; Shea 2013; Suárez 2010). A structure-preserving isomorphism is just a 1-1 map in which a structure on the elements in the domain corresponds to a structure on the elements onto which they map. Consider a collection of entities to be represented  $y_i$  and a relation  $R'$  over them. Any set of putative representations of the same cardinality will map onto the  $y_i$  1-1, and will do so in many ways. For each mapping there is a relation over the putative representations that corresponds to  $R'$ . The existence of an isomorphism is thus a very undemanding constraint. Although some embrace this liberality (Cummins 1989), a more satisfying answer would cut down the class of candidate admissible isomorphisms that a theory of content can appeal to.

When the candidate relations on both sides of the isomorphism are specified, for example that spatial relations between representations should correspond to spatial

representations between represented locations, then the existence of an isomorphism is an extremely demanding constraint. Even the most accurate map does not satisfy this requirement, since there are always slight inaccuracies (for example due to projecting a curved world onto a flat sheet). So spatial relations on maps are simply not isomorphic in the mathematical sense to spatial relations on the ground. Representation can at most require an isomorphism that holds approximately. However, the move to approximate isomorphism compounds the liberality problem. The bare existence of an approximate correspondence between some structure over putative representations and another structure over a domain of entities they might represent is an extremely undemanding constraint.

This paper identifies another, less-appreciated problem. In many cases where there is an isomorphism between representations and representeds, the isomorphism has no significance for the organism or system making use of the representations. It is not relied upon by the system. It enters into no computations, nor does it form part of an account of how the system manages to represent correctly when it does. According to one influential idea, many kinds of representational content depend in part on there being an *exploitable relation* between representations and the things they represent, a relation that is made use of when the system relies on those representations in generating behaviour (Godfrey-Smith 2006). (Others, too, make appeals to the way representations are interpreted or used in order to home in on the operative isomorphism: Cummins 1996; Gallistel 1990, p. 28; Suárez 2003, 2010). In many cases, even when there is a reasonably natural isomorphism between a set of representations and the things they represent, that relation is not being exploited at all in the way the system makes use of the representations.

Only a very restricted class of isomorphisms are candidates for being exploited in this way. A restriction to such *exploitable* isomorphisms cuts down dramatically on the liberality of the isomorphism concept, making it a plausible ingredient in a theory of content for some kinds of representational system. This paper uses some stylised examples to illustrate the contrast between exploited and unexploited isomorphisms, and argues that the class of exploitable isomorphisms can play a substantive role in a theory of content.

Section 2 uses rat navigation as a case study to illustrate the contrast between exploited and unexploited isomorphisms and goes on to argue that the salient isomorphism in the well-known honeybee nectar dance is not in fact exploited. Section 3 uses the case of cartographic maps to show that the requirement that an isomorphism be exploited is not unreasonably strong. Section 4 delineates the class of exploitable isomorphisms more precisely and argues that they are a plausible ingredient in a theory of content for some kinds of system. Section 5 asks whether making use of an exploitable isomorphism is just making use of a source of correlational information, and section 6 briefly examines some ways in which the isomorphism requirement needs to be relaxed if it is to be relied on in theories of content for real natural systems.

I first set aside some related issues that would distract from the main focus. Our topic is not whether there can be analogue representations (Goodman 1976;

Lewis 1971). Structural representation of the kind considered here could be digital in the sense of not admitting of continuous values; nor do I discuss Fodor's "picture principle" – that parts represent parts (Fodor 2007). The structural representations considered here are not constrained to be 'informationally rich' (Braddon-Mitchell and Jackson 2007, pp. 179-80), in the sense that they can give some information only by giving a lot of information (in the way that cartographic maps do). Nor will I delve into the mental imagery debate (Block 1983). I also set aside linguistic representations and the question of whether the subject-predicate structure of language mirrors something in the world.

## II Case study

To illustrate the contrast between exploited and unexploited isomorphisms, I use an idealised example based on the way spatial navigation works in rats. Place cells in the rat hippocampus fire when and only when the rat is in a particular location within a spatial domain (O'Keefe and Burgess 2005; O'Keefe and Nadel 1978). Cell A might fire when the rat is in one corner of an enclosed arena, cell B when it is two thirds of the way down another side. To simplify considerably, suppose these cells were driven only by perceived features, independent of the rat's orientation or direction of motion: by the sights and smells observed from a given location. The rat could learn by instrumental conditioning what to do at each location – to press a lever when in one corner and to pull a chain when two thirds of the way down another side – by associating motor programs with place cell firing. When the rat reached a location, firing of the corresponding place cell would then trigger appropriate behaviour. That mechanism depends on place cell firing correlating reliably with location, but not on any relations between different place cells, nor on spatial relations between locations.

Since the rat follows continuous routes through its arena, place cells that correlate with nearby locations will tend to fire one after the other, and so could become associated through classical conditioning. There would then be a *co-activation* structure over the place cells. The firing of one place cell would potentiate the firing of place cells which correlate with nearby locations. Although the spatial organisation of place cells within the hippocampus has no spatial significance (they do not form a 'topographic map' as some cells in visual cortex do), there is evidence that place cells have a co-activation structure that corresponds to the proximity of their preferred locations (Dragoi and Tonegawa 2013). However, in our simplified example that isomorphism is not being used to guide behaviour. The rat is relying on the correlation between each individual place cell and a particular location, but the co-activation structure over the place cells has no significance for it. It is a case of unexploited isomorphism.

A necessary condition for an isomorphism to be exploited is that relations between putative representations should make a difference to downstream computations or behaviour. An *exploited isomorphism* is one where such a relation is made use of in virtue of its correspondence to a relation between the entities represented. An adequate characterisation of 'use in virtue of correspondence' would

require a full theory of content, but the necessary condition is enough for present purposes.

To see how this necessary condition could be met, let's enrich the case in line with further empirical results. Place cell activity also occurs 'offline', that is based on the endogenous dynamics of hippocampal neurons when decoupled from perceptual input, for example when the rat is asleep (Dragoi and Tonegawa 2011, 2013). There is evidence that this happens when a rat is faced with a choice of where to go: before the rat sets off, place cells fire offline in sequences that correspond to continuous paths through spatial locations. Now suppose, much more speculatively, that offline place cell firing were to precess through two or more sequences before the rat sets off, each ending at the same final place cell associated with reward. Suppose further that there is a system which somehow keeps track of how long it took for each of the offline chains of firing to reach the final reward-associated place cell.

If the rat has a disposition to pick a route that corresponds to the quickest chain of place cell firing, that disposition would cause the rat to follow the shorter route to the rewarded location. This ability to choose the shorter route does depend on the fact that a relation over the place cells (co-activation) corresponds to a relation over the locations with which they correlate (spatial proximity). The relation between place cell firing acts as input to a computational process of choosing the shorter chain, whose utility lies in the fact that co-activation relations over the place cells correspond to spatial relations over locations. So in this case the correspondence between co-activation structure and spatial structure *is* being exploited. It is also a plausible case of structural representation.

### **III Unexploited Isomorphism**

In many cases where there is a salient isomorphism, the correspondence is not in fact exploited. One of Millikan's central examples is the honeybee nectar dance, in which there is a very obvious isomorphism between dances and locations (Millikan 1984). The angular separation between two dances corresponds to the angular separation between two nectar sources; and differences in the number of waggles performed correspond to spatial distances between nectar sources. In fact, neither of these isomorphisms is plausibly a case of structural representation.

For simplicity, consider only one of these: the correspondence between distance of nectar and number of waggles. More waggles corresponds to closer nectar locations, so let us suppose that incoming bees produce dances, based on the distance they have travelled returning home from a nectar source, according to the simple formula  $\text{waggles} = 300\text{m}/\text{distance}$  (rounded to the nearest whole number). 'Consumer' bees observing these dances evolved a disposition to fly off a corresponding distance when foraging. That pattern of behaviour is useful because each individual dance type correlates with a particular distance of nectar (e.g. 5 waggles correlates with nectar at 60m).

Had there just been one dance of three waggles for nectar approximately 100m away, a disposition to fly off 100m on observing such a dance would still have been

evolutionarily advantageous. Perhaps incoming bees could have produced dances according to a different system, with more waggles corresponding to more distant nectar: waggles = distance/50m, say. Then if consumer bees had evolved a disposition to fly off a distance corresponding to the number of observed waggles times 50m, that would have been evolutionarily beneficial.

Indeed, even a more 'gruesome' mapping from dances to distances would still be useful, provided consumer bees evolved a corresponding disposition to behaviour, for example flying off 300m for 1 waggle, 60m for 2 waggles, 150m for 3 waggles, 75m for 4 waggles, and so on. Such a system is unlikely to evolve in natural systems, but if there were such a system, with incoming bees' dispositions to produce dances matching consumer bees' dispositions to fly off for different distances, then a teleosemantic theory of content would deliver corresponding correctness conditions for each of the dances. In this case too there is a relation on the dances that corresponds to spatial relations between locations (Shea 2013), although it is not a natural relation between the representational vehicles (dances). In this case, although there is a correspondence between a relation on the dances and a relation on the locations, no use is being made of it.

But surely the real case is different, because there a natural relation on the dances (more waggles) corresponds to a natural relation on locations of nectar (closer distance)? This is indeed an important difference, but it is still not a case of structural representation. In the real case the mapping from representations to represented locations is a simple function from natural properties to natural properties: location = 300m/ number of waggles. As a result, it is relatively straightforward for a single mechanism to achieve the needed mapping, both on the input side and the output side. I have argued elsewhere that whether a representational system is arbitrary is a matter of degree, based on the ease with which the needed mapping from content to representation to appropriate behaviour can be achieved by the type of system in question (Shea 2011, p. 186). By that criterion, the actual bee dance system is relatively un-arbitrary, whereas the gruesome mapping suggested above is more arbitrary. That is one significant difference.

Another difference is that, because a single mechanism has evolved to deal with a range of different dances, the mechanism has a relational evolutionary function: to fly off a distance given by the formula  $300\text{m} / \text{number of waggles}$ . The evolutionary function of a particular type of dance derives from that relational evolutionary function, e.g. the function of the consumer in response to a 2 waggle dance is to fly off 150m. Now suppose that for some reason a 4 waggle dance has never been performed in the history of the bees. If an occasion arises to perform it, then the bees have a mechanism that can cope: producer bees dancing 4 waggles when nectar is at 75m and consumer bees flying off 75m when they observe 4 waggles. That new dance type still has an evolutionary function, deriving from the overall relational function of the mechanism. When there is nectar at 75m for the first time and the bees forage for it successfully, that is not by chance. It is because they have a mechanism that makes use of a simple function from natural properties of the

representations (waggles) to natural properties of the locations represented (distances). That is an important feature of the case.

However, this is not yet a case where a relation between representations is being used to represent a relation between the items represented. At no stage are two different dances compared or the relation between them computed. The fact that the 4 waggle dance is 2 waggles more than the 2 waggle dance is not being computed nor being used to guide behaviour in any other way. That is most obvious in our hypothetical example with the gruesome mapping. In that case the correlation between individual dance types and distances was being made use of, but the gruesome relation between dances obviously was not. The real system is just the same in this respect. The non-arbitrariness of the mapping makes it easier to evolve and develop a mechanism that responds appropriately to a range of dances, but that mechanism does not depend on comparing different dances and using the relation as a proxy for spatial separation between two sources of nectar. The isomorphism between dances and locations is not exploited in the synchronic control of behaviour.

Recall that, to be a structural representation, a relation between two or more representations must represent a relation between the entities represented. With the bee dance, no relation between dances is acting as input to subsequent computations or is conditioning the consumer bees' behaviour. Nor does an explanation of the success of the bees' behaviour depend on their being able to represent the relations between two different sources of nectar (at least in the aspects of bee navigation described above). So it is unlikely that relations between locations are being represented. If so, this is not a case of structural representation.

#### **IV Application to Cartographic Maps**

Cartographic maps are another case where an isomorphism is being exploited. Furthermore, they illustrate the possibility that the relations between putative representations are in a sense prior, in that they play a role in fixing the content of the so-related representations. Maps are of course a special case, in that they rely on first order resemblance. Nevertheless, they provide a good illustration of the more general point.

A map need not have a co-ordinate grid, scale bar or north arrow. Consider an array of three points on a page, labelled London, Oxford and York. Two points are not enough, but with three points the map can represent relative spatial relations between the cities, e.g. that York is three times as far from London as Oxford is, and in such-and-such relative direction. The relations between points are essential to the individual points being involved in any truth-evaluable content at all, but still the points represent entities and the relations between them on the map represent spatial relations between cities on the ground.

Consider a map where we replace the singular terms with generic symbols, say for churches, pubs and schools. The map makes claims about relations between property instances, e.g. that there is a school twice as far from a church as a pub. Which pub, church and school are represented by the points depends in part on the

relations between the points on the map. It may also depend on facts about the map as a whole, for example what it is being used for or how it was produced. However, the relations are playing a role, too, such that were the relations between the church, pub and school symbols on the map different, they could represent different entities in the world. The same applies for other generic symbols, like coloured lines for roads and rivers and coloured shading for forests. Unlike a point labelled 'London', the relations between these generic symbols on the map are playing a role in determining which entities are picked out. The idea of structural representation does not require that the content of the related representations is prior to the content carried by the relations between them.

There are important issues about exactly how maps represent and precisely which kinds of correctness conditions should be associated with different kinds of maps (Blumson 2012; Camp 2007; Rescorla 2009a, b). Without committing to any detailed claims about the right way of characterising content, nor about the compositional semantics that underpins content, I just rely on the claim that relations between points represent relations between represented entities, however these other issues turn out. I also remain neutral on the representational significance of absence: whether the absence of a symbol at a point on a map represents that no entity of the kind represented elsewhere on the map (e.g. city, church, forest) is present at that location. Cartographic maps are paradigmatic structural representations. The obvious isomorphism between space and space is readily exploitable, for example to work out which route to follow. When the exploitable correspondence is relied upon, it is an exploited isomorphism.

## **V Exploitable Isomorphisms**

Having defined a structural representation as one in which relations between representations represent relations between the entities they represent, I argued that there are real cases of structural representation. Furthermore, there can be structural representations that do not depend on an intelligent interpreter for their content. However, a relation between representations is unlikely to represent unless the system is sensitive to it and uses it in a way that makes a computational and/or behavioural difference – unless it is exploited. Applying that test, some cases where a correspondence between representations and representeds is salient turn out not to be ones where the relation between representations is used, and so are unlikely to be structural representations. In this section I argue that these considerations can be turned into a plausible constraint on the admissible class of isomorphisms that are candidates to figure in a theory of content, thus cutting down considerably on the problematic liberality of the isomorphism concept.

There are two sides of the isomorphism to consider, representations and representeds. On one side, relations between representations need to be usable by the system. We intelligent human interpreters can in principle learn any relation over representations, for example the relation of successor in the counting sequence of number words, which corresponds to no natural relation between the words (phonetic

entities) or numerals (inscriptions). Without an intelligent interpreter and within simpler systems, a relation between representational vehicles has to be more natural if it is to have behavioural or computational significance. It has to be a relation that the organism can make use of in downstream processing. Whether that is so will be relative to the processing resources of the system. In the brain, relations between the pattern and precise timing of the depolarization of neurons can be made use of, similarity in the colours of the neural cell bodies cannot.

Gallistel (1990) argues that which relations between representations are relevant is fixed relative to an interpreter and its 'interpretative code'. However, for our purposes it is no good to allow the 'code' used by the interpreter entirely to determine the relations on representations that are admissible, because then we are back to the possibility that any relation over representations, no matter how arbitrary, could in principle be the subject of an isomorphism – leading us back to the liberality problem. If the isomorphism is to count as a relation that downstream processes can exploit, then it should, in a substantial sense, pre-exist the use made of it, and not depend entirely on an interpreter that has just the right 'code' so as to react to relations between representations in just the right way. If a structure over representations is arbitrary and gruesome, then a mathematical isomorphism between such a relation and some structure in the world would not be a useful exploitable resource.

On the other hand, some natural relations between representations can readily act as input to downstream computations. Relations like the difference in firing rate between two neurons is readily the subject of downstream computations. Arranging neuronal receptivity such that differences in firing rate correspond to some natural relation of interest or, in the example above, so that co-activation structure corresponds to spatial proximity, is a very considerable computational achievement. In the same way, producing a sheet of paper on which spatial relations between marks on the paper correspond to spatial relations between the towns, mountains, rivers etc. of a country is a major task. That is why the survey of India was such a notable achievement. The pebbles in Horseguards Parade were already isomorphic to the geography of India, but that isomorphism is completely useless because no one has the appropriate 'interpretative code'. The cartographers' achievement is precisely to set up an isomorphism where the relation over representations is usable, projectable and relatively uniform.

It is not straightforward to specify which relations between putative representational vehicles are detectable, computationally tractable, projectable and hence usable, since this raises general issues about the naturalness of properties and the projectability of predicates. However, the distinction appealed to here is not a special resource needed just by theories of content. Relying on it, only projectable relations between putative representations are candidates for an exploitable isomorphism. One rough and ready test of whether the relation is pre-existing, or whether it is entirely established by the interpreter, is to ask whether it would extend to new cases without the interpreter having to learn afresh what to do in respect of them.

Furthermore, many projectable relations are unexploitable because downstream processes cannot be sensitive to them, as in the example of the colour of neurons. Neural processes are obviously not suited to making colour comparisons between upstream cell bodies. Exploitability in this sense is a matter of degree, depending on how difficult it is for computational processes to adapt so as to be sensitive to a feature of the representations. Another important matter of degree is the question of how widespread the relation is. If a projectable relation that obtains between a few representations corresponds to a relevant relation between entities represented, that could in principle be made use of; but a correspondence that extends uniformly across a whole range of possible representations will be much more useful (as with the co-activation relation discussed above, which extends across a large array of place cells).

Turning now to the other side of the isomorphism, there are also constraints on the entities and relations in the world to which putative representations are mapped under the isomorphism. Most importantly, they must be objects, properties and relations that are of some significance to the organism. Furthermore, in the kinds of lower level cases we are considering here, arbitrary or gruesome relations between representeds are unlikely to be significant to the organism. So the class of exploitable isomorphisms can be delineated as follows:

*Exploitable Isomorphism*

- (i) A 1-1 map between a set of putative representational vehicles in an organism or other system, and a set of entities to be represented,
- (ii) in which a projectable relation on the vehicles to which operations of the system could be sensitive
- (iii) corresponds to<sup>1</sup>
- (iv) a relation, of significance to the system, on the entities to be represented.

There is not scope here to offer a full theory of content based on exploitable isomorphisms. Clearly the bare existence of an exploitable isomorphism is not enough on its own to give rise to correctness conditions and/or satisfaction conditions. In some kinds of system it might plausibly be combined with natural teleology to give

---

<sup>1</sup> Relation R on vehicles  $v_i$  corresponds to relation R' on entities  $x_i$  under the 1-1 map I in which  $I(v_i)=x_i$   
iff  
 $\forall i,j \{R(v_i,v_j) \leftrightarrow R'(x_i,x_j)\}$   
*(Mutatis mutandis for other polyadicities.)*

rise to correctness conditions.<sup>2</sup> However, the idea that representation depends in part on use being made of an exploitable relation between representations and the world extends beyond teleosemantics. My claim is that exploitable isomorphisms are one such relation. As well as the way they are used, the way a system of representations is produced may play a role in determining which exploitable relation gets to count as the content. The claim here is that exploitable isomorphisms are suited to being partly constitutive of content in a way that the more general class of isomorphisms is not.

Are structural representations, as we have defined the term, the only place where exploitable isomorphisms play a content-constituting role? Recall that the structure that does the representing has to consist of a set of relations over representational vehicles. So some of the proper parts of the structure that does the representing must themselves be representation (although their content need not be prior to the relations which relate them, as we saw in relation to maps above). Representations in general can of course have structure without their parts being representations. The letter 'A' has a structure but many of its proper parts are not representations at all. Computer icons are another possible case, like the icon that represents the instruction 'print'; although here the structure probably has representational significance for the computer user, since it is structure that makes the icon look like a printer. Nevertheless, perhaps there are representational vehicles that have structure, where the whole has representational content and none of the proper parts are representations. They would not then count as structural representations. Could it be that isomorphism or structural correspondence plays a role in fixing the content of representations of this kind? Nothing I have said rules out such cases, although the considerations I pointed to about exploitability make them unlikely. If relations between parts of a representation are to act as input to downstream computations, or are to be the basis for conditioning behaviour, then it seems likely that the so-related parts would thereby be constituted as having some kind of representational content. Whether that is so in general is an issue I leave aside here.

## **VI Exploiting Isomorphisms Contrasted With Exploiting Correlations**

Another kind of exploitable relation is correlation. The simplest kinds of exploitable correlations concern variable matters of fact. There is range of tokens  $R_i$  and variable set of condition in the world  $C_i$ , such that when one of the  $R_i$  is tokened it changes (say raises) the probability that one of more of the conditions  $C_i$  obtains. In our recurring example, the availability of nectar differs over time and a particular dance raises the probability that there is nectar available now at a particular location.

The structure of a cognitive map can also correlate in this way if it changes over time in step with changes to the structure of the environment. When place cell A tends to activate place cell B, that raises the probability that their preferred locations

---

<sup>2</sup> In (Shea 2007) I argued that correlational information can be combined with evolutionary functions to form the basis of correctness conditions for some simple representational systems.

are nearby in physical space. If the spatial relations change over time then so gradually will the co-activation structure. Even when the physical environment is static for the whole life of the organism, there is a sense in which the relations between representations correlate with spatial relations between locations – had the environment been different, the relations between representations would have been different (provided some mechanism of plasticity is involved in establishing the exploitable isomorphism). Against a background of possible physical spaces and possible representational structures, the actual representational structure is probability-raising.

So are exploitable isomorphisms nothing more than exploitable correlations? One question is whether in all exploitable isomorphisms the relation between representations carries correlational information about the entities represented. I want to leave that issue to one side, since it is clear that there are many cases where exploitable isomorphisms do carry correlational information of a relevant kind. My claim is that exploitable isomorphisms form a special class that is worth recognising, even when they also carry correlational information.

One reason is that the relation between representations can then be used not just to represent an undifferentiated state of affairs (with which it correlates), but to represent a relation between the entities represented by its relata. Another is that the relation between representations can have a common significance across a very wide range of representations. Adding an extra symbol to a map of  $n$  existing symbols immediately adds at least  $n$  claims about relations to places already represented on the map (Braddon-Mitchell and Jackson 2007, pp. 179-80). The same is true when adding a new place cell into a co-activation structure. The common significance of the relation across a whole array of representations facilitates computation, for example calculating a shortest route. None of that follows from having a series of sources of correlational information about related matters of fact. A related point is that a relation between representations can be used to calculate offline, as described in the shortest route example, and so can be used in a very basic form of conditional reasoning. So to the extent that exploitable isomorphisms are instances of correlational information, they are cases of a very special kind, and so can do significant additional work. Their existence could then plausibly be an ingredient in content-determination, either instead of or in addition to the role of correlational information. The following examples illustrate the differences.

Desert ants of the genus *Cataglyphis* reliably exhibit the remarkable ability to perform path integration (Wehner, Michel and Antonsen 1996). After foraging for a period following a tortuous, criss-crossing route, when the ant finds food it is able to follow a ‘beeline’ to its home burrow: it sets off in the right direction and follows a straight line for the right distance before beginning a search for local cues to the burrow entrance. This ability depends on ‘dead reckoning’ – integrating information about speed and direction of movement – and does not depend on using landmarks, since an ant translocated to a new location makes a beeline to the location where its burrow would be had it too been translocated.

Stylising the example for simplicity, suppose the ant can use the polarisation of sunlight to detect its current orientation in relation to the sun, with one set of cells firing maximally when facing the sun and dropping off to a background tonic level of firing when at  $90^{\circ}$  and reducing further to a minimum when at  $180^{\circ}$  to it; and another set firing maximally when perpendicular, with the sun on the right hand side, passing through tonic firing when facing directly towards or away from the sun, and dropping off to a minimum when perpendicular to the sun on the left. Suppose further that activity from these two sets of cells is fed into a set of four accumulators, gated by haptic input from the feet. One accumulator takes the activity of the ‘facing the sun’ cell above its tonic rate, times the current level of haptic activity, and adds that up. The other three do the same for the three other primary directions. Then each accumulator will correlate with the ant’s total net displacement along one each of four perpendicular axes.

So far we only have correlations, but there is also a relation over the accumulators that is very relevant to the ant’s needs. To set off in the right direction for home the ant just needs to orient itself so that the current input is the inverse of the relative levels of the accumulators.<sup>3</sup> To make a straight line for home it just needs to travel in that direction until its accumulators run down to zero, and then stop and make a local search for its burrow. That (hypothetical) mechanism makes use of a relation between internal vehicles: a ratio of differences. The relation is made use of because it correlates with a fact of significance to the ant, the direction of its home burrow. However, the relation is being used just for the correlational information it carries. It is not being used because of its significance in relation to two entities being represented. This is not a case where a relation between representations is used to represent a relation between entities represented. So it is not a case of structural representation.

There are many cases where two sources of information are combined to produce a more reliable correlation or a stronger correlation with some further condition (Ernst and Banks 2002). That is a ubiquitous feature of hierarchical information processing in the brain. The new correlate is based on a relation between two existing information-carriers, which in many cases are themselves representations. So a relation between representational vehicles is being used. In some cases, as with the desert ant example, it will be used simply for the correlational information it carries. In others it may represent a relation between the entities represented. For example, in visually tracking a small array of objects the visual system registers the egocentric spatial position of each of them (Pylyshyn 1989). Relations between these registers (“FINSTs”) could be used to keep track of the spatial separation between the represented objects, so as to form our expectations about when they will collide (Spelke and Van de Walle 1993). That would be to use a

---

<sup>3</sup> More precisely, the ratio of the firing rate of parallel cells to perpendicular cells needs to equal the ratio of { $180^{\circ}$  accumulators minus  $0^{\circ}$  accumulators} to { $270^{\circ}$  accumulators minus  $90^{\circ}$  accumulators}.

relation between representations to represent a relation between the entities represented – the array of FINSTs would then be a structural representation.

Another empirical example is analysed by Shagrir (2012). Fixation and horizontal saccades of the eyes are plausibly underpinned by a recurrent neural network with a single line attractor whose points correspond to eye positions. Shagrir argues that motor neurons program a saccade between the current position and a target position by taking as input the distance in activation space between the points corresponding to the current position and the target position (2012, p. 14). That would be to make use of the correspondence between distance in activation state space and spatial position of the eyes. So it would be another case of structural representation. (By contrast, some of the other computations mentioned by Shagrir do not depend on exploiting structural correspondence.)

In all of the cases above where an isomorphism is actually being exploited, the relation between representational vehicles also carries correlational information about the relation which it is used to represent. However, being used to represent a relation between represented entities is something over and above simply being used for the correlational information that the relation carries.

## **VII Beyond Strict Isomorphisms**

Our focus on isomorphisms is a simplification – useful because it highlights the central issue about how structure can stand for structure – which would need to be relaxed in several ways to be turned into a plausible ingredient in a theory of content for many real systems (cp. Shagrir 2012, p. 3). There is only scope here to gesture in the direction of these modifications. First, the 1-1 mapping requirement is doubtless too strong. In real cases there may be representational redundancy, such that two or more different representations in fact represent the same entity, yet still many of the relations over the array of representations remain useful. Homomorphism may therefore be the more appropriate tool. Another idealization is that the isomorphism should hold precisely, as between a natural, computable relation over representations and a natural, significant property over representeds. For example, if we required spatial relations between points on a map to be genuinely isomorphic to spatial relations between locations on the ground, there would be no room for a partially accurate map. Even the slightest divergence from exact correspondence would destroy the isomorphism. So we need some account of an isomorphism or homomorphism holding only approximately. As we saw above, the class of isomorphisms or homomorphisms that hold only approximately is extremely wide, so the fact that such relations obtain is not, taken alone, a plausible basis for the metaphysics of content.

Given a set of represented entities and a represented relation, it is reasonably straightforward to calculate the accuracy of the whole array. For each represented relation, calculate how closely it matches the actual relation between the entities represented. The sum (possibly weighted) of these values over the whole array measures the accuracy of the structural representation.

As is the case with correlations (Godfrey-Smith 1989), the array of representations may not represent that which would make it maximally accurate. Since other resources will be appealed to by any plausible theory of content, in addition to exploitable approximate homomorphisms, content need not be fixed so as to maximise accuracy. For example, a map of the churches in a city might be more accurate if the points were taken to represent pubs (since every church has a pub next door but many churches have been converted to other uses since the map was made). Nevertheless, facts about how the map is used might constitute it as having church-related content not pub-related content.

The representational significance of the relation between representational vehicles is also up for grabs, and affects the accuracy of the structural representation. For each potential assignment of content to relations between representations, we can calculate how accurate the structural representation would be on that assignment. In fact, both the content of the points and the content of the relations can be allowed to vary independently, telling us for a given structure how accurate it would be as a representation of a whole suite of entities and their potential relations.<sup>4</sup> Each of these is, to a greater or lesser extent, an exploitable feature of a given fixed structured set of representations.

A structured set of putative representations may have a large number of approximate exploitable homomorphisms to natural structures in the environment of significance to an organism. We saw in section 4 that structures vary in exploitability depending on how available for downstream processing the relations between representations are on the one hand, and how natural and significant the relations between representeds are on the other. To that we should add variation in how accurate they would be as a structural representation given the way the world is. *Ceteris paribus*, structures will be more useful as representations of sets of conditions about which they would be more accurate.

## VIII Conclusion

The requirement that a structure over representations be used if it is to have representational significance is an important constraint. It follows that, at least in simple systems of the kind considered here, only isomorphisms that are exploitable are plausible ingredients in a theory of content. This cuts down substantially on the otherwise problematic liberality with which isomorphisms subsist between putative representations, and entities and relations in the world. It is a considerable achievement to arrange things so that a projectable relation to which the system can be sensitive over a set of putative representational vehicles corresponds to a

---

<sup>4</sup> Furthermore, downstream processing could be sensitive to the fact that the representational significance of a relation varies somewhat for different relations in the array, in the way that longitude gets distorted on a Mercator projection of the globe. So approximate homomorphisms in which the relation varies over the representational space are also to some extent exploitable.

significant natural relation on a set of entities relevant to the system. There being such a correspondence could, then, be part of what makes it the case that some kinds of representation have the content they do.

Nicholas Shea  
Department of Philosophy  
King's College London

### Acknowledgements

The author would like to thank the following for helpful comments: Rosa Cao, Peter Godfrey-Smith, Phyllis Illari, Kate Jeffrey, Karen Neander, Gerard O'Brien, David Papineau, Ron Planer, Oron Shagrir, Mauricio Suárez, and the audience at the meeting of the Aristotelian Society.

### References

- Block, Ned 1983: 'Mental pictures and cognitive science'. *The Philosophical Review*, 92, pp. 499-541.
- Blumson, Ben 2012: 'Mental Maps'. *Philosophy and Phenomenological Research*, 85, pp. 413-34.
- Braddon-Mitchell, David and Jackson, Frank 2007: *The philosophy of mind and cognition, 2nd ed.* Oxford: Blackwell.
- Camp, Elisabeth 2007: 'Thinking with maps'. *Philosophical Perspectives*, 21, pp. 145-82.
- Cummins, Rob 1989: *Meaning and Mental Representation*. Cambridge, MA: MIT Press.
- Cummins, Rob 1996: *Representations, Targets and Attitudes*. Cambridge, MA: MIT Press.
- Dragoi, George and Tonegawa, Susumu 2011: 'Preplay of future place cell sequences by hippocampal cellular assemblies'. *Nature*, 469, pp. 397-401.
- Dragoi, George and Tonegawa, Susumu 2013: 'Distinct preplay of multiple novel spatial experiences in the rat'. *Proceedings of the National Academy of Sciences*, 110, pp. 9100-05.
- Ernst, M. O. and Banks, M. S. 2002: 'Humans integrate visual and haptic information in a statistically optimal fashion'. *Nature*, 415, pp. 429-33.
- Fodor, Jerry 2007: 'The revenge of the given'. In *Contemporary debates in philosophy of mind*. McLaughlin, Brian and Cohen, Jonathan (eds) Oxford: Blackwell pp. 105-16.
- Gallistel, C R 1990: *The Organization of Learning*. London / Cambridge MA: MIT Press.
- Godfrey-Smith, P. 1989: 'Misinformation'. *Canadian Journal of Philosophy*, 19(4), pp. 533-550.
- Godfrey-Smith, Peter 1996: *Complexity and the Function of Mind in Nature*. Cambridge: Cambridge University Press.

- Godfrey-Smith, Peter 2006: 'Mental representation, naturalism and teleosemantics'. In *New Essays on Teleosemantics*. Papineau, David and Macdonald, Graham (eds) Oxford: OUP pp. 42-68.
- Goodman, Nelson 1976: *Languages of art: An approach to a theory of symbols*. Hackett publishing.
- Lewis, David 1971: 'Analog and digital'. *Nous*, 5, pp. 321-27.
- Millikan, Ruth Garrett 1984: *Language, Thought and Other Biological Categories*. Cambridge, MA: MIT Press.
- Millikan, R. G. 2000: 'Representations, Targets and Attitudes'. *Philosophical and Phenomenological Research*, 60(1), pp. 103-111.
- Millikan, R. G. 2013: 'Reply to Nick Shea'. In *Millikan And Her Critics*. Ryder, D., Kingsbury, J. and Williford, K. (eds) Oxford / Malden MA: Wiley-Blackwell
- Neander, K. forthcoming: *Mental Representation*. London / Cambridge MA: MIT Press.
- O'Brien, G. and Opie, J. 2004: 'Notes towards a structuralist theory of mental representation.' In H. Clapin, P. Staines and P. Slezak (Eds.), *Representation in Mind: New Approaches to Mental Representation*. Amsterdam: Elsevier.
- O'Keefe, John and Burgess, Neil 2005: 'Dual phase and rate coding in hippocampal place cells: theoretical significance and relationship to entorhinal grid cells'. *Hippocampus*, 15, pp. 853-66.
- O'Keefe, John and Nadel, L 1978: *The hippocampus as a cognitive map*. Oxford: Clarendon Press.
- Pylyshyn, Zenon 1989: 'The role of location indexes in spatial perception: A sketch of the FINST spatial-index model'. *Cognition*, 32, pp. 65-97.
- Ramsey, William M 2007: *Representation Reconsidered*. Cambridge: Cambridge University Press.
- Rescorla, Michael 2009a: 'Cognitive maps and the language of thought'. *The British Journal for the Philosophy of Science*, 60, pp. 377-407.
- Rescorla, Michael 2009b: 'Predication and cartographic representation'. *Synthese*, 169, pp. 175-200.
- Shagrir, O. 2012: 'Structural Representations and the Brain'. *British Journal for the Philosophy of Science*, 63, pp. 519-45.
- Shea, Nicholas 2007: 'Consumers Need Information: supplementing teleosemantics with an input condition'. *Philosophy and Phenomenological Research*, 75, pp. 404-35.
- Shea, Nicholas 2011: 'What's Transmitted? Inherited Information'. *Biology and Philosophy*, 26, pp. 183-89.
- Shea, Nicholas 2013: 'Millikan's isomorphism requirement'. In *Millikan And Her Critics*. Ryder, D., Kingsbury, J. and Williford, K. (eds) Oxford / Malden MA: Wiley-Blackwell pp. 63-80.
- Spelke, Elizabeth S. and Van de Walle, Gretchen 1993: 'Perceiving and reasoning about objects: Insights from infants'. *Spatial representation: Problems in philosophy and psychology*, pp. 132-61.
- Suárez, M. 2003: 'Scientific representation: against similarity and isomorphism'. *International Studies in the Philosophy of Science*, 17(3), pp. 225-244.
- Suárez, M. 2010: 'Scientific Representation'. *Philosophy Compass*, 5(1), pp. 91-101.
- Swoyer, Chris 1991: 'Structural representation and surrogative reasoning'. *Synthese*, 87, pp. 449-508.

Wehner, Ruediger, Michel, Barbara and Antonsen, Per 1996: 'Visual navigation in insects: coupling of egocentric and geocentric information'. *Journal of Experimental Biology*, 199, pp. 129-40.