RESEARCH ARTICLE

A Causal-Mentalist View of Propositions

Jeremiah Joven B. Joaquin* – James Franklin**

Received: 3 May 2020 / Revised: 20 September 2020 / Accepted: 24 January 2021

Abstract: In order to fulfil their essential roles as the bearers of truth and the relata of logical relations, propositions must be public and shareable. That requirement has favoured Platonist and other non-mental views of them, despite the well-known problems of Platonism in general. Views that propositions are mental entities have correspondingly fallen out of favour, as they have difficulty in explaining how propositions could have shareable, objective properties. We revive a mentalist view of propositions, inspired by Artificial Intelligence work on perceptual algorithms, which shows how perception causes persistent mental entities with shareable properties that allow them to fulfil the traditional roles of (one core kind of) propositions. The clustering algorithms implemented in perception produce outputs which are (implicit) atomic propositions in different minds. Coordination of them across minds proceeds by game-theoretic processes of communication. The account does not rely on any unexplained notions such as mental
content, representation, or correspondence (although those notions are applicable in philosophical analysis of the result).

**Keywords**: Propositions; causal-mentalist view; cluster analysis; game theory; perception algorithms; Platonism; symbol grounding.

1. Introduction

Propositions play several theoretical roles. They are thought of, for example, as the primary bearers of truth and falsity, the *relata* of logical relations such as entailment, the objects of certain intentional states such as belief, and contents of linguistic acts such as assertion (McGrath and Frank 2018; Briggs and Jago 2012). There is disagreement, however, as to the nature of the posited entity that performs these roles. Platonist, possible-worlds, deflationary, and naturalist accounts compete.

We argue for one form of naturalist theory of propositions. In our view, propositions are non-abstract, structured entities, and we identify them with certain types of persistent mental entities created and coordinated by cognitive algorithms common to different minds. Work in Artificial Intelligence has discovered the kind of algorithms needed to create such mental entities, while the causal story of how they are created and work in game theory has shown how such entities can be coordinated across different minds.

We take no particular position on the nature of the mental, such as whether it is reducible to the physical. The mental is simply identified as whatever in humans plays the usual roles of intentionality, content, thought, perception and direction of action.

2. Directions for a theory of propositions

In this section we list several specifications which our theory of propositions will attempt to meet. Some are uncontroversial and some less so. We give brief reasons in their favour but cannot consider all the arguments of those who have taken different paths. Our aim is to motivate a causal-mentalist approach and to situate it in the spectrum of theories of propositions.
A theory of propositions will offer a clear answer as to the metaphysical nature of whatever is identified as propositions.

A theory should be clear as to whether it takes propositions to be Platonist abstract entities and if so which ones, or types of sentences and if so how the tokens are united into types (since a type is only identified by the properties that unite it), or mental entities and if so which ones exactly, or is a deflationary theory and if so what is left after deflation, and so on.

To defend a non-deflationary, or realist, theory of propositions is not necessarily to posit entities which are primary truth-bearers and to which thought or predication has some (external) relation. That way of speaking could concede too much to the Fregean Platonist view of propositions that once dominated the field. What the alternative might be, however, needs to become clear in the outcome. It will need to be explained how the entities to be posited accomplish the roles which propositions traditionally play.

It will explain both the public, objective aspect and the mental aspect of propositions.

Propositions are shareable and public in that two or more people can accept, believe, assert or communicate the same proposition. That is the point of them. Any account of the nature of propositions should attempt to explain those facts. Prima facie, a mentalist account of propositions could find this difficult, as mental entities are not public, nor do they appear to be shareable.

On the other hand, ‘any [view] according to which propositions represent things as being a certain way and so have truth conditions in virtue of their very natures and independently of minds and languages is in the end completely mysterious and so unacceptable’ (King 2009, 261; Pickel 2017). The problem of what propositions are does not arise from questions in physical or biological science. The roles played by propositions arise from humans attributing meaning to language, from language inducing thoughts, and from similar pieces of language reliably inducing similar thoughts and behaviours in different persons. Those explananda essentially involve the mental, so a theory of propositions must explain the relations of propositions to the mental. That is a prima facie difficulty for theories that propositions are to be identified with non-mental entities such as Platonist abstracta, sets of possible worlds, states of affairs or sentences. Our approach is thus
in accordance with the naturalist trends in cognitive science that have tended to supplant Fregean anti-psychologism in the last hundred years. We comment on Frege below.

- It will explain compositionality: how the nature and roles of a proposition are a function of (or at least relate to) the nature and roles of its terms

A well-formed sentence, like ‘John is tall’, surely expresses some thought. An ill-formed string of symbols, like ‘!@#$%’, does not. Something must account for this difference. Likewise, something must account for the fact that ‘John is tall’ resembles ‘John is short’ in one way and ‘Sue is tall’ in a different way. The structural features of propositions account for the difference (Duncan 2018).

To think of propositions as structured entities is to think of them as complex entities having an ordered relation among their parts or constituents. The ordered relation of the proposition’s constituents greatly matters not only in distinguishing between meaningful strings of symbols and nonsensical strings, but also in identifying whether two propositions with the same constituents are the same or different propositions. For example, the English sentence ‘John loves Mary’ expresses the same proposition as the Tagalog ‘Mahal ni John si Mary’, since structurally both have the same ordered set <loving, John, Mary>. But the sentence ‘Mary loves John’ expresses a different proposition from these two, since the proposition it expresses has a different structural ordering. Various accounts of the ordering relation are found in Schiffer (2012); King (2019); McGrath and Frank (2018); Hanks (2009).

While it is the majority position that propositions are structured, the great diversity in accounts of the nature of their parts opens a window onto what an account of propositions must explain (Briggs and Jago 2012; Chalmers 2012). Are propositions made up of symbols? Names? Abstracta? Possibilia? Intensions or meanings (King 2019)? Mental entities of some kind? Whichever is correct, surely any theory of the nature of propositions must include a theory of the nature of the parts and of the relation of parts to the whole.

- It will explain the apparently close relation between propositions and states of affairs
Propositions cannot be states of affairs such as ‘Snow’s being white’ because there are no false states of affairs. It certainly seems that true and false propositions do not differ in their nature as propositions but only in their relation to how the world is. States of affairs are ways the world actually is and lack the mental or interpretive aspect of propositions. Nevertheless, it is surely not a coincidence that true propositions like ‘Snow is white’ have a structure that mirrors that of the corresponding state of affairs (and the structure of a false proposition can mirror that of a possible but non-actual state of affairs). It is desirable to have some explanation of that.

- **It will accomplish those tasks without reliance on philosophical notions which threaten to be as obscure than the explanandum**

We aim for something more ambitious than accounts of propositions that take for granted concepts like ‘representation’, ‘content of concepts’, ‘reference’, ‘intentionality’, ‘correspondence’, ‘information’, ‘third realm’ or ‘object of’. The problem is that those notions span the mental and the extra-mental in ways as mysterious as propositions themselves.

Leading contemporary mentalist theories of propositions have not attempted to do that. Soames’s theory of propositions as ‘cognitive event types’ (Soames 2010) relies on a primitive notion of representation (Caplan 2016). Hanks’s theory grounded on primitive ‘acts of predication’ (Hanks 2015) and Davis’s theory grounded on ‘declarative thought-types’ (Davis 2005, 20–23) suffer this ‘ungroundedness’ problem.

We do not claim there is anything wrong with those notions, as philosophical interpretations of a causal story. Indeed, we will argue that propositions as we understand them do represent and their terms do refer. The objection is to taking those notions as primitive. We aim to give a free-standing causal account of the mental entities that play the role of propositions, with those entities arising in a way that does not rely on those notions. They can then be subject to an external philosophical analysis using notions such as representation and content.

- **Platonism will be a philosophy of last resort**

For Fregeans, the proposition expressed by a sentence is a structured relation of *senses*, where these senses correspond to the constituents of the
sentence expressing them. ‘Senses’ are not mental entities (a thesis dubbed ‘psychologism’ by Frege) but *abstracta*—the ‘realm-mates of Platonic properties and relations’ (Jubien 2001, 47; a clear account of the Fregean picture is found in Hanks 2015, 12–20). ‘Abstract’ here means ‘abstract object’ in the sense of full-blooded Platonism: non-spatial and causally inefficacious, as are numbers on a Platonist view (Rosen 2020). But Platonism in general labours under the weight of over two thousand years’ worth of substantial objections such as its incompatibility with naturalism and its difficulties with epistemic access. In the context of propositions in particular, Frege’s notion of their ‘graspability’ by the mind was always recognised as difficult and was a prime focus of attack by leaders of the (naturalist) ‘cognitive turn’ in analytic philosophy. (Sacchi 2006)

Platonism is a metaphysics of last resort—understandable if still problematic with apparently eternal objects like numbers, but surely even less attractive concerning entities like meanings that are so closely related to human utterances and intentions.

- **A naturalist, anti-Platonist theory of propositions will make essential reference to the causal origins of propositions**

  Just as biology is incomprehensible without evolution, and a naturalist philosophy of mathematics would require explanation of how mathematical knowledge can arise in minds, so a non-Platonist theory of propositions requires an account of how they come to be. If propositions are not *abstracta* or other entities ‘out there’ and they play some role in cognition, they must fit into a causal, scientific story. Thus, we regard purely philosophical mentalist accounts of propositions such as Davis (2005), Soames (2010) and Hanks (2015) as incomplete though valuable. Instead, we will connect our theory with causal theories of reference (Devitt and Sterelny 1999), symbol grounding in Artificial Intelligence (Harnad 1990) and perceptual symbols in cognitive science (Barsalou 1999). We will argue however that by keeping to the correct level of analysis we do not need to delve into details of neuroscience.

- **A naturalist, anti-Platonist theory of propositions will maintain continuity between human and higher animal belief**
Although we cannot defend it at length here, our approach accepts that (some) propositional attitudes are in the first instance pre-linguistic, and so exist in higher animals and human neonates. As Fodor writes, ‘You can, surely, believe that it’s raining even if you don’t speak any language at all. To say this is to say that at least some human cognitive psychology generalizes to infra-human organisms; if it didn’t, we would find the behavior of animals utterly bewildering, which we don’t.’ (Fodor 1978, 512; at length in Nelson 1983) The basic causal mechanisms we discuss will be shared with higher animals.

- Finally, the theory will then show how the entities identified as propositions fulfil the traditional roles of propositions as truth-bearers, objects of belief and other propositional attitudes, and relata of logical relations.

This is the main task of any theory that purports to account for the nature of propositions.

To summarise the discussion so far and prepare for the next section, we present a classification diagram of theories of propositions. (Fig 1) Our theory belongs in the shaded box at bottom right

![Classification of theories of propositions](image)

Figure 1: Classification of theories of propositions
The diagram will be referred to later in explaining the categories of the mental entities being discussed.

3. Propositions as mental entities?

So, why believe that propositions are mental entities? Soames notes that there is something unsatisfactory about the project of trying to place propositions ‘out there’ away from minds. He writes,

The key to solving [the problem] is to recognize the obvious fact that predication is something that agents do... Instead of explaining the intentionality of the cognitive activity of agents in terms of an imagined conceptually prior intentionality of the propositions they entertain, we must explain the intentionality of propositions in terms of the genuine conceptually prior intentionality of the cognitive activity of agents who entertain them (King, Soames, and Speaks 2014, 33).

This agent-, mind-dependent nature of propositions is something that Platonist and other theories such as physicalist and possible-worlds theories have not accounted for satisfactorily. Propositions are bearers of meanings, and there is no getting past the fact that meaning is something intentional; something with a social aspect, certainly, but mediated through individual minds.

The mentalist option should therefore be revisited, despite the prima facie problems it faces. The dialectic for our view is simple. Since propositions are not ‘out there’, then they must be at least partly ‘in here’, or mental. Furthermore, since propositional attitudes like belief are mental, it would simplify matters if propositions were as well—the Platonist problems with relations between minds and abstracta would be avoided if propositions were ‘inside’ the mind.

But there is an obvious problem. Can we explain how two people can believe the same proposition? If a proposition is just in one individual’s mind, as a token mental entity, then the proposition is not shareable. But the point was that propositions are shareable. Two people can both believe that Great Britain is a monarchy. Mental entities are private; propositions are, if not exactly on public display, at least interpersonally available.
On the other hand, while mental tokens are private, mental types can be shared. As such, these mental types can thus be treated as scientifically acceptable entities. To be more precise, the properties defining the type are shareable, while the tokens of the type are not. That possibility opens an opportunity for progress.

In order to maintain clarity about the nature of the mental entities to be discussed, we begin with several necessary distinctions, as laid out in the bottom right of Fig 1. A mental entity, such as a belief, may be either occurrent (a ‘datable mental act’, Textor 2017) or persistent (an underlying accessible mental structure that is activated as required, as when one is said to believe Pythagoras’s Theorem even when one is not thinking about it.) The nature of mental life is for occurrent mental acts to be caused by some process such as perception, then persistent traces to be laid down in a process of learning.

A mental entity may also be token or type: my occurrent enjoyment of warmth is of a type with yours, while my persistent love of summer is also of a type with yours. It is important that the occurrent/persistent distinction is not the same as the token/type distinction. (It is also possible to distinguish between type and property—a species of birds is a type defined by possessing the properties characteristic of that species (Wetzel 2018, sec. 3)—but that distinction does not play an important role here.)

Should a mentalist theory aim to analyse persistent or occurrent propositions? Existing mentalist theories have preferred occurrent propositions. Soames (2010) uses mental event types, Hanks (2015, 6) similarly takes propositions to be types of ‘mental and spoken actions’, Davis’s (2005, 20–23) ‘declarative thought-types’ are occurrent ones. We think, in accordance with the causal story to be told below, that long-term, persistent or dispositional, propositions (bottom right of Fig 1, shaded) are primary and occurrent thoughts or expressions of them are secondary and transitory. Propositions, like the terms such as names that are their constituents, should be persistent entities in the first instance which account for the unity of the event types that express them; so, their expressions in mental events such as predications are secondary.

However, we do not deny that the entities in the three other quadrants exist, nor that they can be rightly said to be true, or believed; one can
occasionally believe or speak truly. We will indeed defend a mentalist ana-
logue of Hanks’ claim that ‘linguistic types inherit their semantic properties
from their tokens.’ (Hanks 2011, 41) That is natural since we begin with a
causal story, and causes produce tokens—but ones with reliable properties
since by and large like causes produce like effects, and it is those properties
that account for their fulfilling the role of propositions.

4. Propositions as mental entities coordinated
by cognitive algorithms

Propositions should be a kind of thing satisfying barely compatible re-
quirements. They should be in some sense mentally dependent, but in order
to be shareable they should not be individual (token) thoughts. They should
be objective, yet neither denizens of the world of physical states of affairs
nor of a world of abstracta.

4.1 Propositions as types of persistent mental entities

One possibility for the kind of thing that could in principle satisfy those
requirements is a property of mental entities (and hence the type defined
by the property). Properties, although sometimes called ‘abstract’, are not
abstract objects in a Platonic realm, lacking causal power. On the contrary,
things act in virtue of the properties they have, as when we see something
as yellow because yellow things affect us in a certain way. One might well
take an Aristotelian (anti-Platonist) realist view of properties (Armstrong
1978), but here we do not commit ourselves to any particular metaphysics
of properties. We need only the scientific acceptability of causal properties
and their powers, as when we observe that Newton’s law of gravitation relates
(the properties) gravitation, mass and distance, or Weber’s law relates stim-
ulus (in general) to perception (in general). Properties as we discuss them will
be understood in that minimal scientific and naturalist sense, according to
which particulars have causal powers in virtue of the properties they have.

Strangely, Frege’s own arguments for his Platonist position do not rule
out this possibility. He argues concerning the ‘thought’ or what we would
call the content of a proposition:
Is a thought an idea? If the thought I express in the Pythagorean theorem can be recognized by others just as much as by me then it does not belong to the content of my consciousness, I am not its bearer ... If every thought requires a bearer, to the contents of whose consciousness it belongs, then it would be a thought of this bearer only and there would be no science common to many ... So the result seems to be: thoughts are neither things of the outer world nor ideas. A third realm must be recognized (Frege 1956, 301–2).

That is correct, but to conclude that if something is neither a physical object nor an occurrent idea, then it must be a Platonic abstract object neglects certain other possibilities. Properties of mental entities are another, and, as we have explained, they are not abstracta (as they have causal powers, like the properties of physical objects). Properties of mental entities satisfy Frege’s desiderata for ‘thoughts’, that they be apprehensible by many minds because they belong ‘neither to my inner world as an [occurrence, particular] idea nor yet to the outer world of material, perceptible things’ (Hanks 2015, 3–4).

My (occurrent) thought of cats and your thought of cats are numerically different mental entities, one in my mind and one in yours. But they can have properties in common. For example, they can occur at the same time, and they can both be partially caused by perceptual experience of cats. So, in principle, my thought and your thought could have a property (or stand in a relation) that makes them (tokens of) the same proposition. But what property, and how is that property acquired? It needs to be a property anchored in the ‘out there’ (like the causation of perception by real cats), that gives the two thoughts, yours and mine, in some way a sufficiently similar relation to external states of affairs. ‘Out there’ should mean, in naturalist fashion, out there in the real physical and social world, not out there in a Fregean Platonic ‘third realm’ causally detached from the real physical and social world.

A beginning on explaining how this is possible is made by a propositional theory of perception, such as that of Armstrong’s Perception and the Physical World. Armstrong writes:
Physical objects or happenings stimulate our sense organs. As a causal result of this we acquire immediate knowledge of their existence and their properties ... This knowledge is not necessarily verbalized knowledge, but it is always knowledge which it is logically possible to put verbally. It is propositional in form... The acquiring of immediate knowledge in this way is perception (Armstrong 1961, 191).

That is of no immediate help in analysing what propositions are: since it analyses perception in terms of propositions, it leaves propositions themselves unanalysed. But it does suggest that, if some independent account can be given of how perception creates mental entities, the entities so created can fulfil the role of propositions (at least, those propositions stating claims about perceived reality). In particular, since everyone’s perceptual apparatus is set up by biological causes to be similar, there will normally result a high degree of commonality between the results of my perception and yours (of the same thing). If the results of perception are propositions, then it will have been explained how the mental entities common to our believing a proposition do have properties in common: they are caused by the same perceptual object affecting the mind through a similar causal process.

4.2 Cluster analysis to generate atomic perceptual propositions

We begin by separating the case of propositions about immediate perceptual reality from others, such as inferred propositions and ones believed on the basis of testimony. That is in accordance with the long philosophical tradition from the Aristotelian ‘nothing in the intellect that was not first in the senses’ through Locke to Carnap’s Aufbau. It is also in accordance with much recent philosophy of language. As Devitt and Sterelny put it,

A causal theory of natural kind terms, like one for names, divides in two. First, there must be a theory of reference fixing, which explains how a term is linked to a referent in the first place ... Second, there must be a theory of reference borrowing, which explains the social transmission of a term to those having no contact with its referent (Devitt and Sterelny 1999, 88).
That applies to propositions as much as to the terms in them. There are
good reasons for that separation into two levels. The understandability of
'It’s sunny in California' is parasitic on the understandability of 'It's sunny
here'. If we can grasp the nature of immediate perceptual propositions, we
can hope to move on to others which stand in a logically dependent relation
to them (such as logical combinations of immediate perceptual propositions,
or merely possible ones, or reported ones).

The two-level approach is also suggested by the commonality of perceptu-
tial recognition to us and nonhuman primates. The ability to form pre-
symbols out of perceptual input is found in the higher primates, but appar-
ently not the compositionality and inference of human language (Zuber-
bühler 2018), nor reference to more distant or abstract entities. As Witt-
genstein neatly puts the two stages, ‘A dog believes his master is at the
doors. But can he also believe his master will come the day after tomorrow?’
(Wittgenstein 1953, 174; discussion in DeGrazia 1994) Therefore, it would
be desirable if the most basic account of what a proposition-like entity is
could rely only on what is possible for the primate mind, leaving more com-
plex operations peculiar to humans for a later stage.

Certainly, achieving an independent account of perceptual grounding
has proved very difficult, both at the in-principle level desired by philosophy
and at the more detailed level required by cognitive science (and even more
so, at the implementable level sought by Artificial Intelligence). The exten-
sive work in those areas has however produced some models that are useful
in demonstrating how a mental entity such as one whose properties play
the role of a term or proposition could be at once internal to the mind,
caused by an appropriate part of physical reality, and endowed with prop-
erties shareable by other minds. We will use the simple example of cluster
analysis as applied to the symbol grounding problem (Franklin 1996). This
is not simply an example or metaphor, but an exercise in abstract task
analysis to identify what must be done, minimally, to create an atomic
perceptual proposition.

Corresponding to the old philosophical problem of how words get their
meaning, cognitive scientists address the ‘symbol grounding problem’ (Har-
nad 1990). In an Artificial Intelligence system intended to perform tasks
like computer vision, how can the internal symbols which are to represent
pieces of external reality actually be attached to and caused by the perceptual input from that reality? In the vast flux of changing pixel values input to a vision system, how is the relevant part to be identified and labelled so that the system locates and tracks an object? The problem is not simply a technical one: here the main interest is in the task specification and in understanding the kind of algorithm that could possibly solve it, along with the nature of the result. (The distinction between task specification and algorithm is that a task specification states what output is to be created from what input—‘make a chocolate cake from these ingredients’; ‘sort a list of words into alphabetical order’—while a recipe or algorithm lists the steps to accomplish the task—a definite cake recipe for the cake; for sorting, an algorithm such as bubble sort which repeatedly goes through the list and swaps adjacent items if they are in the wrong order.)

Cluster analysis works like this. Its purpose is to take a heap of points as in Fig. 2 (that is, it is just given unlabelled points with their positions), and to conclude ‘There are two clusters, and these points are in cluster A and those are in cluster B.’ To the eye, it is easy, but finding and programming an algorithm that performs the task is non-trivial.

Figure 2: Points in two clusters

---

1 Figure adapted from: http://www.philender.com/courses/multivariate/notes2/cluster0.html.
The main application of cluster analysis is to ‘points’ that are not in geometrical space but in ‘feature space’ (as in Fig. 3, but typically the space has many dimensions). The axes represent features of objects and the degree to which objects possess them. So, a dot placed in the space represents an object, with its position in the space representing the degree to which it has each feature.2

![Figure 3: Clusters of points in feature space](http://2.bp.blogspot.com/_CWYkOgzhlyq0/TAJ4oFopasI/AAAAAAAAADO/NRrS4E3R1cs/s400/cluster_analysis_income_debt.gif)

For a vision system (artificial, animal or human) to recognise an object against background requires a form of cluster analysis: the pixels in the object, which are similar to one another in both position and colour, need to be ‘stuck together’ by the system’s software to create the ‘object’ cluster, while being separated from the ‘background’ cluster of pixels.

Then once one has an individual cat (say) cut out of the background and identified as a single object, one must perform cluster analysis again (in feature space) to recognise that cats are similar to one another and dogs are similar to one another (across a range of features) and there is not much in between. Hence, there are recognisable natural kinds: cat and dog.

---

2 For a technical survey, see (Jain, Murty, and Flynn 1999).
3 Figure adapted from: [http://2.bp.blogspot.com/_CWYkOgzhlyq0/TAJ4oFopasI/AAAAAAAAADO/NRrS4E3R1cs/s400/cluster_analysis_income_debt.gif](http://2.bp.blogspot.com/_CWYkOgzhlyq0/TAJ4oFopasI/AAAAAAAAADO/NRrS4E3R1cs/s400/cluster_analysis_income_debt.gif)
How similar inputs have to be (and similar in what ways) to count as in the same cluster is something that the clustering algorithm has to work out for itself. Cluster analysis is an exercise in discretisation—different points in a continuous input space end up classified into a single discrete cluster. Whether different perceivers can recognise sufficiently similar clusters given their somewhat different perceptual inputs is a matter for scientific investigation. The fact that communication using perceptually-generated concepts often succeeds suggests that they do. It is easy to think of the difficulties of classifying handwritten postcode digits, but software exists which does almost perfectly classify them into ten classes.

It is not merely that cluster analysis might be helpful for the problems of early perceptual grouping and of symbol grounding, but that the nature of the problems means that any solution to them must be some form of cluster analysis. Such problems all involve forming a discrete object out of a cluster, that is, a mass of neighbouring data points that are all close among themselves, but are all reasonably well separated from other data points.

A clustering algorithm is a perfectly comprehensible series of steps, a recipe implementable in software—whether software written by humans in an AI system or the mental implementation of such an algorithm in the human mind, however that is accomplished (and because we are working at the level of task specifications and algorithms, the theory can remain neutral on the mechanisms by which that is in fact accomplished). An implemented algorithm is a naturalistically acceptable entity, not an inhabitant of a Platonic world: it is just a series of regular scientifically-discoverable steps with inputs and outputs, like the process of photosynthesis. The algorithm does not contain any philosophical overhead such as ‘representation’ or ‘content’.

Nevertheless, its output consists of discrete items that it would be natural for the outside observer to recognise as uncannily resembling terms and propositions. The clusters identified by the algorithm in the flow of experience are labels of naturally grouped persistent objects. While the algorithm requires some structured input—the inflow of perceptual input must have some inhomogeneities actually present, and the feature dimensions are also given—the objects have been discovered in the data by the algorithm, not
presented to it beforehand. Nevertheless, we can find in its output items bearing a resemblance to all of the constituents of an atomic proposition \( Fa \). The ‘subject’ \( a \) is a cluster found in the points (the pixel positions and values, for example, clumped by the cluster algorithm to form natural discrete objects such as spot against background). The predicate \( F \) is a region of feature space; for example, the bottom left cluster in Fig. 3 lies in the low-income region of space. (That does require that the system has the capability to represent explicitly dimensions of feature space and their parts; we do not attempt to solve that problem here.) The outside observer is welcome to use the language of representation, so as to agree with Fodor’s ‘there are internal representations and propositional attitudes are relations that we bear to them.’ (Fodor 1978, 519) But representation is an interpretation of what has arisen from the algorithm, not something assumed or input at the start.

The proposition is the (persistent type) association between cluster and region, that is, the output of the clustering algorithm identifying the cluster and where it lies in feature space. One must distinguish between the points in the cluster actually being in that region (a state of affairs in the real world), and the clustering software’s explicit identification of the cluster as being in that region (an output of the software). It is the latter that is a proposition. As it is a type, produced similarly in different correctly-functioning minds, it can be common to different minds. In a pre-linguistic animal, there is no more to belief in a perceptual proposition than the algorithmic output (which should be sufficient for action on the basis of that belief). Humans may, as is their wont, add some conscious thoughts and linguistic expressions, but those are not essential to the belief in the proposition.

As with any implemented software, a system performing clustering can be analysed at three levels of abstraction. At the lowest level is the working implementation (code running on a machine, in the computer case; working neurobiology, in the human case). At the next level is the algorithm: the recipe or sequence of steps that is implemented in the code or neural activity. At the highest level is the task analysis or program specification, which describes what the algorithm is to do (in terms of transforming inputs to outputs) without laying down the steps it is to perform to do it. Different algorithms (such as different clustering algorithms) may perform the same
task (such as identifying natural clusters in data). For the purpose of identifying propositions, the important level of analysis is the highest, that of specifying the clusters (of objects and kinds) formed by the perceptual clustering. Thus, if Martians have a different algorithm for perception, but identify the same clusters in data, such as cats and mats, their proposition ‘the cat sat on the mat’ is the same proposition as ours, in virtue of their algorithm satisfying the same program specification as ours.4

The same—task analysis—level is appropriate for deciding whether two people believe the same proposition, such as ‘the cat sat on the mat,’ on the basis of their slightly differing perceptual experiences. The aim of perceptual algorithms is exactly to create persistent discrete unities out of the continuous flux of perceptual experience. It may succeed exactly, approximately, or not at all, and whether it does so the same way in two perceivers is a scientific question to be examined in the usual ways, such as by checking how similar are their inferences, answers and behaviour on the basis of the proposition believed.

Although task analysis or program specification is an obvious sense ‘more abstract’ than lower levels, it does not follow that it involves any naturalistically unacceptable abstracta. Just as photosynthesis has a more efficient implementation pathway in tropical plants than in temperate-zone ones but still effects the same biochemical transformation (without calling on any non-naturalist entities), so different cognitive algorithms may perform the same cognitive task, such as creating atomic perceptual propositions.

### 4.3 Properties of perceptual propositions generated by cluster analysis

It is not difficult to read off answers, in that model, to some of the troubling questions about propositions described earlier—at least for atomic propositions that summarise perceptual input.

As to the metaphysical nature of a proposition, it is an explicit internal (mental) persistent discrete entity, generated by internal algorithms acting

---

4 The separation of levels of analysis is less clear in the ‘perceptual symbol systems’ of Barsalou (1999), which otherwise attempts a similar causal analysis to the present one.
(in this case) on perceptual input. ‘Explicit’ does not mean ‘conscious’—the mind’s insight into its own workings is notoriously weak—but instead means a discrete output of mental processing, capable of entering into further processing, and in principle capable of being identified by the sciences of psychology and neuroscience.

The way in which the mental (persistent) tokens of propositions in different minds are coordinated and share properties follows from their being the outputs of identical (or very similar) algorithms run on similar data. Just as your cake and mine resemble each other (have similar persistent properties) if we use the same recipe on similar ingredients, so your token internal proposition that there is a cat before us resembles mine in virtue of our having run similar perception-processing clustering algorithms on similar visual input. It is true that if your experience of cats is only of white ones and mine only of black ones, there may be some mismatch between our concepts and thus potential for miscommunication. That is inevitable and algorithms can only do so much with the data they are given.

Compositionality has been explained above: the terms $a$ and $F$ in the proposition $Fa$ are themselves explicit as outputs of the clustering algorithm—they are exactly what the algorithm is designed to output. So is their relationship in the proposition. The fact that terms and propositions are discrete outputs is what allows them to become inputs in the higher levels of discrete processing: for example, to enter into explicit conjunctions and other logical composites and to be expressed in discretely structured natural languages. At least, it allows that in principle. There are some difficult questions as to how human cognition, but apparently animal cognition only minimally, achieves some explicit knowledge of the representational nature of its internal symbols. Those questions are about cognition and not directly about the nature of propositions and cannot be addressed here.

The relation between proposition and state of affairs is, in this simplest case though not necessarily elsewhere, the causal one between data input and algorithmic output. The cat’s being on the mat is a state of affairs.\(^5\) The perceptual clustering algorithm identifies discrete objects *cat* and *mat*

\(^5\) We take a realist approach to states of affairs as defended in (Armstrong 1997) and regard them as unproblematic entities in the present context, and as composed of particulars and universals as they seem to be.
in the visual experiential flux, tagged with their positions in space, computed from the visual directions of their parts. Some further processing, but of the same nature, is required to make explicit the spatial relation of cat and mat, inherited from the spatial relations of the perceived pixels belonging to each.

There is more than one way in which that simple causal relation of proposition and state of affairs can become more complicated. Firstly, there are many possible ways to go wrong between input and output; any algorithm implemented in real software can malfunction and, for example, output cluster labels when there are really no clusters in the data. In that case, there is no state of affairs corresponding to the output. Correspondence is thus a relation that may or may not hold between states of affairs and the output of the clustering software: there are certain clusters really in the data, and the algorithm does or does not successfully find them. No external notion of truth has been invoked—correspondence is defined in terms of the properties of the data and the software’s performance on the data. Again, the contrast between animal and human cognition is useful for understanding what has been claimed. An animal’s internal proposition can rightly be said to represent external reality, but that is a comment made by an outside human philosophical observer and is unknown to the animal. Humans have more insight into their cognition (we do not claim to have explained how) and their awareness of the representational nature of their symbols is useful not only philosophically but for such purposes as logical inference.

Secondly, there are other possibilities for lack of correspondence—mismatches between reality and the outputs of the software — when the discrete outputs of clustering algorithms are used as the inputs of recombinations. Human mental capacities, though apparently not cockroach mental capacities, include explicit recombination of the discrete chunks. ‘a is red’ and ‘b is blue’ permit the recombination ‘a is blue’; ‘X loves Y’ permits the recombination ‘Y loves X’. This ‘systematicity’ of thought was adduced in Fodor and Pylyshyn’s celebrated (1988) paper as being incompatible with a connectionist architecture for the mind; certainly, it points to the requirement that discrete recombinable entities should be found at a basic level in

---

6 Provided that ‘red’ and ‘blue’ are recognised as in the same category, something which itself needs to be a capacity of the software.
thought. The ability to ‘mix and match’ items is essential in exploring ‘what-if’ scenarios—thought experiments where humans imagine what would happen in situations that have not occurred yet, or ‘mental time travel’ (Suddendorf and Corballis 2007). It is the foundation of the human ability to plan. Recombination of (categorically compatible) terms yields something that might be the output of the clustering algorithm applied to some possible state of affairs. Or again, it might not, if the recombination yields an impossible state of affairs: there is no guarantee that recombination of mental items tracks what is really compossible in the outside world.

A third way in which the relation of proposition and state of affairs may become more complicated arises when the proposition is more logically complex than an atomic $Fa$ or $Gxy$. Those ways cannot be described in detail here. An example is that the relation of the truth-functionally complex proposition ‘$Fa$ and $Gy$’ to states of affairs can be explained as inherited from the relations of $Fa$ and of $Gy$ to states of affairs; that is the point of the connective ‘and’. Another story can be told about propositions with complexity like ‘X believes that $p’$, but clarification can be left until an account is given of how propositions fulfil roles such as objects of belief.

Not mentioned so far is any conscious mental content. It is not denied that there may be such qualia as feelings of aboutness or cognitive satisfaction when one entertains propositions. However, just as qualia of blue and red are not an essential part of the story of human perceptual discrimination between blue and red—which may take place without or before any such conscious sensation, even if such qualia exist—so the essential mental entity, the proposition, can perform its cognitive role whether or not accompanied by any qualia or awareness.

4.4 Propositions other than perceptual

It is of course not claimed that clustering algorithms applied to perception can perform an indefinitely large range of knowledge-generation tasks. Early perceptual grouping and object identification form a restricted range of knowledge, albeit a crucial and foundational one. Clustering is just a first and easily understood example, where the relation of input to output can be studied free of complications.
However, a good deal is known at least in principle about the generation of non-perceptual propositions from perceptual ones. In developmental psychology, for example, evidence on the production of more inferential knowledge, such as research on ‘Bayesian babies’, suggests that humans share inbuilt algorithms for inferential knowledge too (Denison, Reed, and Xu 2013). Something was said above about the human ability to speak of merely possible states of affairs through recombination. In the philosophy of language, much has been said about how the social nature of language (once the developing infant has inferred that a world of other minds exists out there) allows for one person to ‘catch’ propositions from another (Devitt and Sterelny 1999, 96–101). These abilities—apparently very minimal in animals—create a great variety of propositions. They require complex stories as to how they refer to reality (or fail to), but there is no reason to think they require a different account of the nature of propositions. If we understand the nature of perceptual propositions, recombinations of them and their parts are not particularly mysterious in principle.

Propositions involving reference to fictional and abstract entities also need their own story. Again, animals and probably human neonates seem unable to entertain such propositions, confirming the desirability of a two-stage approach that starts with perceptual propositions, and suggesting that such questions do not concern the nature of propositions but the nature of those entities. We agree with the argument of Moltmann (2013) that reference to abstract entities is both special and rare. No doubt the ability to refer to fictions and abstracta is a development of the ability to recombine. For example, the recognition that ‘a is red’ and ‘b is blue’ permit the recombination ‘a is blue’ can suggest detaching ‘blue’ as an entity in itself which can be the subject of propositions. If our theory satisfactorily covers the nature of propositions in animals and neonates, its extension to the range of sophisticated entities discussed by adult humans is a matter for special investigations on those topics.

As pointed out by McDaniel (2005), any realist theory of propositions also needs an account of propositions that have never been and will never be entertained by anyone. That is easy for Platonist theories—they all exist in the same way as entertained propositions—but harder for mentalist or linguistic theories. We argue that, as for the question of propositions about
abstract entities, it is a problem, but not one about the nature of propositions. It is a problem about the nature of uninstantiated properties. Just as the reality (or otherwise) of an uninstantiated shade of blue is a problem about the metaphysics of properties (Franklin 2015) rather than a problem in the science of colour, so unentertained propositions (on a mentalist view) is a problem about the uninstanitated in general rather than about propositions.

However, an analogy will explain why the problem of unentertained propositions should not be regarded as serious for a naturalist account. Let us consider the corpus of programs written in some computer language; for definiteness let us take an obsolete one such as FORTRAN V, in which a finite number of programs were written but which is no longer used. Suppose we are comparing two philosophies of the nature of computer programs. A Platonist one holds that computer programs (in all languages actual and possible) exist in a Platonic heaven, and an infinitesimal proportion of them are written down by some programmer. An alternative naturalist philosophy holds that computer programs are creations of programmers, even though it is an objective matter which of the symbol strings written down are well-formed FORTRAN V programs. A program is an actual string of symbols created by a person (or machine), and it is a well-formed FORTRAN V program if it follows certain rules. The Platonist will urge the “problem of unwritten FORTRAN V programs” as an objection to the naturalist theory. How serious is that objection?

It is not a serious objection. Many natural processes are generative of possibilities that are never realised. Darwinian evolution could generate many species other than those actually found in the history of life on earth, while organic chemistry could generate many compounds other than those it actually does. That is no reason to think that species or organic compounds are really Platonic entities. The Platonist does believe that any species or compound, actual or possible, reflects some Platonic archetype, but for the naturalist who rejects that on general metaphysical grounds, there is no further “problem of unrealised species/compounds/programs”. Those are just latent (and predictable) possibilities of the natural generative process. It is the same with propositions: given a naturalist story of how entertained ones are formed, such as has been presented here, the process
has natural generative possibilities for further propositions that may not ever be entertained. But that is not a reason to reject a naturalist process. The case of false propositions is considered below; false propositions, even perceptual ones, cannot be directly caused by perceivable reality.

5. Coordination of propositions between minds via game theory

Before turning to the question of how the mental entities that have been identified as propositions fulfil the roles initially laid down for propositions, such as being truth-bearers and objects of belief, one further issue needs to be addressed, namely the coordination of propositions between different believers (human believers, if not animal ones), and the public standing of propositions. Although it has been explained how (tokens of) propositions in different minds can share properties in virtue of being the output of similar algorithms on similar data, that still leaves propositions as (types of) private entities hidden in minds. How can they acquire sufficient public standing to allow communication? How can there be reliable coordination between propositions in different minds, so that standard communication cues like words and gestures reliably induce the same proposition (type) in different minds? That is a further task, which it seems that humans can do but animals cannot—cats can discretise perceptual input in a similar way to humans (that is a minimal requirement for recognising conspecifics and prey, which higher animals can do), but they do not appear to coordinate their internal propositions with other cats.

As a simple model for how that is possible, and possible for the sort of mental entities that have been identified as propositions by the theory being put forward here, consider the stability of strategies in game theory. In an Iterated Prisoner’s Dilemma game, the tit-for-tat strategy is a stable equilibrium. The game has two players, with a simultaneous choice at each step between two moves, ‘cooperate’ and ‘defect’. If both cooperate, the payoff is better for both than if both defect, but a temptation to defect results from the rule that if one player cooperates and the other defects, the defecting player is well rewarded and the cooperating player punished. The
tit-for-tat strategy (play the move that the other player played in the last round) permits cooperation to develop, to the benefit of both players, but avoids the perils of being played for a sucker time after time. The game is mentally dependent, in the sense that a round of play occurs as a result of the players’ intentions to play it (and of each player’s knowing that the other has such an intention). It is an objective, mind-independent fact, however, that tit-for-tat is the best way to play it—the way that leads reliably to the best outcome for each player in the medium to long term—so that rational players will tend towards that strategy. That strategy will then typically be observed to be implemented in the game (if the players are indeed rational): it will be observed that plays do agree with the move chosen by the other player in the previous round, and questioning may elicit agreement from the players that they are following that strategy. There is a mental type whose tokens in each player mean they are both playing tit-for-tat. Neither a pattern observed in the sequence of moves nor an intention to follow such a pattern is normally thought to require the existence of any Platonic entity such as an ‘abstract strategy’. There are no entities metaphysically more mysterious than instantiated patterns and the intentions of individuals.

The analogy between game-theoretic strategies and the interpersonal coordination of propositions is closer than it looks. To speak minimally and somewhat naively: propositions can typically be expressed in language, and one main purpose of doing so is to communicate. Communication is, among other things, a game-theoretic exercise. When a speaker enunciates a sign with the intention that the hearer should read it as a sign and take it in a certain way, the hearer’s move in the cooperative communication game involves his guessing the way in which the speaker intends the sign to be taken. It is an iterated game, so the history of cues that have worked is relevant to future moves. Thus, public language coordinates the induction of propositions across minds. (This is similar to the philosophy of communication of Lewis (1969), and Grice’s (1989, chap. 2) ‘cooperative’ principle of conversation, needed to allow a hearer to infer what a speaker intends from the speaker’s public utterance.) That does not imply, however, that the tokens of propositions themselves are public or somehow in the external world. Like the intention to play a round of a game, the token proposition
is a mental entity, but communication is possible because of the coordination of mental entities driven by the game of communication.

6. How propositions as algorithmically coordinated mental entities fulfil the traditional roles of propositions

It remains to be explained how this theory of propositions—mental entities sharing properties through their creation by common algorithms and coordinated by game-theoretic communication requirements—accounts for the roles of the proposition, as listed in the first section above.

Their role as truth-bearers follows from what has been said about correspondence. An implementation of an algorithm to find the real clusters in perceptual data can succeed or fail in doing so, whence the result can be called true or false. (Allowable) recombination of the discrete outputs of the algorithm—the terms of a potential proposition — creates an internal proposition that could be the output of the algorithm applied to some possible state of affairs; the recombination is true or false according as that state of affairs obtains or not.

Propositions as conceived here are the objects of belief because they are actually identified with (implicit, dispositional) beliefs. There is nothing to the implicit belief that $p$ over and above its having been installed in the mind by the belief-creating algorithm; which in the simplest case is the clustering algorithm applied to perception. (Of course, inferred beliefs need a further account of the process creating them out of immediate beliefs, but that does not bear on the nature of propositions itself.) Explicit, occurrent beliefs, such as consciously entertained or linguistically expressed ones, are created by the mysterious process by which the human mind can reflect on and bring to consciousness some of its contents. How that happens is again not directly relevant to the nature of propositions.

The story about the origin of propositions does not involve language in any essential way, so it has been explained in principle how propositions can be the meanings common to sentences of different languages. The clustering algorithm is independent of language—indeed, it is a precondition of the learnability of language. For an infant to associate the sound ‘cat’ with
experiences of cats, it must have solved four perceptual discretatisation or clustering problems first: (1) It has to cut individual cats out of the visual background, (2) it has to classify those objects as (potentially) a single kind, then (3) it has to segment the sound stream so that the syllable ‘cat’ is isolated, and finally, (4) it has to classify that occurrence of ‘cat’ with other occurrences of that syllable. Only after that can there be association between the syllable-type ‘cat’ and experience of the type cats. Thus, organising perceptual experience into discrete repeatable pieces has to take place before questions about the meaning of words can even arise. It is to be expected that linguistic expressions of simple terms and propositions would be intertranslatable between different languages, although there is scope for different languages to construct different complex concepts from combinations of the simples, resulting in some mismatch of meanings of statements in different languages. That is observed to be the case (Wierzbicka 1996).

It is harder to explain why propositions as conceived here, as coordinated mental entities, should be the relata of logical relations. It is not clear why mental entities, even when tied down and given objectivity by the causality of an algorithm or the coordination imposed by the cooperative game of communication, should be subject to the absolute, mind-independent necessity of relations such as logical consequence or logical contrariety. The clue lies in the connection between logical relations and truth. As an account of the truth of propositions has already been given, in terms of a certain kind of correspondence between propositions and states of affairs, logical relations should be explained in the same way. A particularly clear case is the close connection between the logical consequence relation between (certain) propositions and the inclusion relation between states of affairs. The state of affairs of this raven’s being black is a part of the state of affairs of this raven and its neighbour both being black. It is thus impossible for the latter state of affairs to obtain without the former also obtaining. If now a perceptual clustering algorithm (correctly) applies to those states of affairs and extracts the propositions that this raven is black and that both ravens are black, respectively, then the relation of logical implication holds between the latter and the former. It ‘mirrors’, so to speak, the inclusion relation between the states of affairs, meaning that the inclusion relation between the states of affairs is the truthmaker of the fact that it is
impossible for the second proposition to be true without the first also being true. That is, the latter proposition implies the former.

While that is the start of the story, it does not cover all cases. It does not explain why the proposition that this raven is white is implied by the proposition that this raven and its neighbour are both white, because those propositions are false and so do not correspond to any states of affairs, hence there is no inclusion relation between states of affairs to act as the truthmaker of their logical relation. Nevertheless, the potential for those propositions to result from actual states of affairs determines their logical relations. This raven's being white and these two ravens’ being white are possible states of affairs, the first of which would be a part of the second, if they obtained. Hence, it is impossible that the second should obtain without the first, whence the second proposition logically implies the first.

Naturally that will create a problem for propositions that purport to describe what are in fact impossible states of affairs. But again, it is not the business of the theory of propositions to explain the philosophical complexities of all the things that propositions could (or could not) be about.

7. Conclusion

There is every reason to agree with the tradition that there are such things as propositions, entities that fulfil roles such as objects of beliefs and relata of logical relations. But there is no reason why the requirement that propositions be objective and shareable should lead philosophy on a wild goose chase into a Platonic realm of abstracta. Objectivity is available to mental entities, via the cognitive algorithms and game-theoretic coordination strategies that we all share. Those algorithms induce in different minds (implicit) beliefs. The beliefs in different minds share properties that allow them to fill the roles of propositions.

Acknowledgments

A version of this paper was presented at the Propositions, Reference, and Meaning Workshop held in November 14-15, 2019 at the Institute of Philosophy, Slovak Academy of Sciences, Bratislava, Slovakia. We would like to thank the
organizers and participants of the workshop for their useful comments and suggestions.

References


