Erkenntnis
https://doi.org/10.1007/s10670-023-00694-3

ORIGINAL RESEARCH

Compressing Graphs: a Model for the Content of Understanding

Felipe Morales Carbonell[1]

Received: 18 January 2023 / Accepted: 24 April 2023

Abstract
In this paper, I sketch a new model for the format of the content of understanding states, Compressible Graph Maximalism (CGM). In this model, the format of the content of understanding is graphical, and compressible. It thus combines ideas from approaches that stress the link between understanding and holistic structure (like as reported by Grimm (in: Ammon SGCBS (ed) Explaining Understanding: New Essays in Epistemollogy and the Philosophy of Science, Routledge, New York, 2016)), and approaches that emphasize the connection between understanding and compression (like Wilkenfeld (Synthese 190(6):997–1016, 2018)). I argue that the combination of these ideas has several attractive features, and I defend the idea against some challenges.

Keywords Understanding · Structure · Compression · Mental representation · Graph theory

> It was an experience of water and interconnection [...] I was trailing my hand in the water and I thought about how the water was moving around my fingers, opening on one side and closing on the other. And that changing system of relationships where everything was kind of a similar and kind of the same, and yet different. That was so difficult to visualize and express, and just generalizing that to the entire universe, that the world is a system of ever-changing relationships and structures struck me as a vast truth. Which it is. [...] And writing is the process of reducing a tapestry of interconnections to a narrow sequence. And this is in a sense illicit. This is a wrongful compression of what should spread out.
>
> Ted Nelson, in Werner Herzog's *Lo and Behold*

✉ Felipe Morales Carbonell
ef.em.carbonell@gmail.com

[1] Universidad de Chile, Isabel Riquelme Sur 1769 Dpto. 43, Maipú, Santiago, Chile

Published online: 20 May 2023

1 Introduction

In this paper, I offer a model of the content of understanding states, *Graph Maximalism* (GM), where the content of understanding states can be represented in terms of graphs, and a variant of this model, *Compressible Graph Maximalism* (CGM), that introduces a component of compression to GM.[1] This model is intended as concrete way to combine the ideas that a) understanding involves some relation to structure, broadly speaking, and b) understanding can be fruitfully characterized in terms of compression.

My goal is to show that the combination of these ideas is fruitful. In fact, my argument in defense of the view will be mainly abductive: I will defend the view on the basis of its theoretical promise. I am mostly interested in showing that a 'package view' like CGM should be endorsed by theoreticians who defend structure-based or compression-based accounts of understanding. Although I present the view in opposition to what we may call propositionalism (the view that the format of the content of understanding is merely propositional), the spirit of the project is not on the main polemical–here, I try to be ecumenical even when I feel a preference for how to deal with certain points of detail (in many cases, defending those preferences would take us away from the main point, so I avoid committing to them altogether).

I should also make it clear at the outset that I do not intend to propose a full theory of understanding. That would require, among other things (but crucially, as attested by recent work on understanding), an account of the gradability of understanding and of when understanding attributions are appropriate. This is not something I will attempt to provide here (although I will give some suggestions on this point in Sect. 6.6). My focus will rather be in building an account of the *format* of the content of understanding, something that most philosophical accounts omit to give or treat only summarily. An account of the format of understanding provides at least a partial account of the *measures* of understanding, and that in turn gives material to construct an attribution story. So even though I will not attempt to give an account of appropriate understanding attribution, I will touch on some of the issues that bear on the point.

I will first (Sects. 2 and 3) go back to the ideas that understanding bears relation to structure and compression, and give a brief overview of why they are attractive by summarizing the views of some of their proponents. This is mainly intended as set-up, and readers who are familiar with these ideas can skip over to Sect. 4, where I sketch Graph Maximalism. For modelling, Graph Maximalism uses some tools from graph theory, so for readers who are not well acquainted I begin by offering a brief introduction of the main concepts that are relevant here. Then, I proceed by outlining the basic idea (4.2), and then a fuller version (4.3). Once this is done, I outline the advantages I think the model has: it inherits the intuitive appeal of the

[1] In this paper, I will defend the idea that understanding states in general have graphical contents, that is, that for any state of understanding u, the format of the content of u is graphical. It is possible to qualify this further, but for reasons that will be made clearer later, I think the general version of the view is preferable.

understanding-and-structure accounts, it can help in the explanation of abilities related to representation manipulation and cognitive control, if offers a potentially unified format for the representations involved in understanding, and it yields a powerful story of the measures of understanding (4.4). In Sect. 5 I extend GM with the idea that the relevant graphs can be compressed, give an overview of some ways in which this can be implemented, and show the epistemological significance of adopting a model of this kind. Finally, in Sect. 6 I address an assorted array of challenges to the approach.

2 Understanding and Structure

The idea that understanding has something to do with structure, broadly construed, is a commonplace in the literature on understanding, although the specifics vary widely. To orient ourselves, it is useful to distinguish between the *target* and the *content* of understanding. The target of understanding is what is understood; when we say that Ann understands how cheese is made, the expression 'how cheese is made' refers to the target of her understanding.[2] Ann's understanding plausibly (although not unquestionably) involves some form of representation of the target; and this representation will have content, in a given format (propositional, pictorial, graphical, or else).[3] Maybe Ann has something like a 'mental picture' of a procedure to make cheese: first, you do this, and then this other thing, and so on. What is important for us now is that we can talk of the structure of both the target and the content.

One may think, concerning the relationship between understanding and structure, that understanding targets are or must be structured in certain ways. One instance of this is Wittgenstein's famous idea that there is a 'kind of understanding which consists in "seeing connections"' (2009, 122). 'Connections' here refers to what we see in the target.[4] Various accounts of understanding emphasize the point that understanding involves some form of grasp of explanatory structure.[5]

In other proposals, what matters is rather that the content is structured in certain ways. For example, in Elgin (2009), understanding

[2] In fact, here we should distinguish between the *topic* of understanding and the *target proper*. 'How cheese is made' can be taken as a subject matter, a partition of logical space in various ways in which cheese is made, or as a particular cell in that partition, that is, the particular way in which cheese is made in the world that is relevant to our assessment of Anne's understanding.

[3] I will use representation-talk throughout, but I am not really committed to the existence of representations in the classical sense. It is enough for me that something can serve a certain functional profile, namely, one that is similar to the one of classical representations (that it can be said to be about something, that it is information-bearing, that it can be at least partially communicated). In any case, at this stage we are simply reviewing the literature on the connection between understanding and structure, and it is patent that there is a strongly representationalist strand to it.

[4] Things quickly become more complicated, since Wittgenstein allows for understanding to be had by *inventing* connections.

[5] Hazlett (2018) offers an account of the role of structure for understanding where 'structure' stands for something objective as well as joint-carving, and is required for understanding as apt theorizing: according to the view, it is not possible to understand as apt theorizing if there is no objective structure.

> [...] is in the first instance a cognitive relation to comprehensive, coherent sets of cognitive commitments. (323)

For understanding subjects, Elgin also requires that they are able to see the connections between the elements of these bodies of information; this information must also be encoded in the content of understanding.[6] Gardiner (2012) offers an argument about the value of integration that can be seen in this light as well.[7]

Both dimensions can also be conjoined, and this perhaps is the most common view. Elgin herself requires that understanding is grounded in fact in a range of interesting cases. To give a more explicit example, Kvanvig (2003) claims that understanding

> [...] requires the grasping of explanatory and other coherence-making relationships in a large and comprehensive body of information. (192)

In Kvanvig's proposal the target is a cluster of different relationships, at the same time that the content is structured. He says nothing of the format of the content, but a view like this can easily accommodate a plurality of them. This is what we find in Zagzebski (2001), which is rather characteristic of this line of thought:

> [...] understanding is not directed toward a discrete object, but involves seeing the relation of parts to other parts and perhaps even the relation of part to a whole. It follows that the object of understanding is not a discrete proposition. One's mental representation of what one understands is likely to include such things as maps, graphs, diagrams, and three-dimensional models in addition to, or even in place of, the acceptance of a series of propositions. (241)

Another author who takes a mixed view of this kind is Stephen Grimm (2016, and elsewhere). His view is interesting first because he requires the target of understanding to be structured:

> the differences among these various objects of understanding can be (and has been) overstated, and the reason is that in all of these cases understanding seems to arise from a grasp of what we might call dependency relations. Although when it comes to more complex structures (the House of Representatives, for example) more of these relations are grasped than when it comes to

[6] A different account that stresses the importance of the structure of content is Meynell's 2020 pictorial account.

[7] Gardiner (2012) aims to attack veritism (the view that epistemic value is concerned only with truth). Her argument asks us to imagine to robots, Alpha and Beta, who *ex hypothesi* have the same propositional information about some topic, but who differ in that Alpha stores this information as a list of propositions whereas Beta stores them in a format that includes explicit connections, in 'something akin to hyperlinks'. Gardiner argues that Beta's epistemic state is more valuable than Alpha's, because of the integration that this representational format affords. In fact, Gardiner argues (taking some hints from Stroud (1979)) that Alpha's state could not have the same epistemic value as Beta's, because of its lack of integration: the relation between the relevant facts simply cannot be represented by any number of propositions. I will return to this point in Sect. 4.4.

understanding particular states of affairs, this does not amount to a difference in kind but instead to a difference in degree. (214)

In terms of the content, he also suggests that the content of understanding is structured. Drawing from the work of Gopnik et al. (2004), he suggests that the content of understanding takes the form of mental maps. These mental maps should be taken to be ‘mobile’, in the sense of being able to adapt and change as the variables represented on the map take different values (ibid, 216). In Grimm’s account, understanding is manifested in the ability to see what could happen to the target in the respects of interest if it underwent a variety of changes.

3 Understanding and Compression

The second general idea we are interested here is that understanding is somehow connected with compression. Whereas, as we saw above, the idea that understanding is related to structure is widespread, the connection with compression has only been recently been made explicit, in the work of Wilkenfeld (2018). The idea has, however, antecedents in some earlier views on the connection between explanation and learning and compression (see, for example, Grünwald (2007) and Li and Vitanyi (2019), and in some views on mathematical learning (see Thurston (1990) and Gray and Tall (2007)).

Roughly speaking, representations of smaller size (by some measure) of some target compress it. In informational terms, a compressed representation R relative to a source representation R’ encodes the same information in less bits. The possibility of compression depends on the availability of redundancy and patterns in the source; this is why, for example, truly random sequences are incompressible–they are never redundant. Using this relative notion of representation-to-representation compression, we can define a notion of object-to-representation as follows: a representation R compresses an object O if it represents it and the size of R is smaller than the size of another given representation or of a maximally lengthy non-redundant representation of O (a representation that in this sense is ‘as good as it gets’).

Given this conception of compression, it is easy to see the appeal of linking understanding and compression. Complexity and size can be seen as impediments to understanding. A large assortment of facts without a way to systematize them seems to offer no easy way to gain understanding. The ability to compress depends on the ability to detect patterns in the available structure. Compression, then, manifests the realization of such capacities.

Wilkenfeld’s theory is that understanding *is* compression. The core of his account is his proposal for comparative judgements of understanding:

> Understanding-as-Compression A person p_1 understands object o in context C more than another person p_2 in C to the extent that p_1 has a representation/process pair that can generate more information of a kind that is useful in C about o (including at least some higher order information about which information is relevant in C) from an accurate, more minimal description length [representation.]

Wilkenfeld makes use of a Minimal Description Lenght (MDL) framework. The idea is that given a data set D and a set of hypotheses H, finding regularity in the data set consists in finding the set of hypotheses in H that compresses the data set the most.[8] In particular, in Wilkenfeld's proposal the goal is to minimize 'the sum of the length of the description of a general hypothesis about how the data is structured combined with a description of the specific data in terms of that hypotheses' (5), so that it is possible to generate the most information from the less information possible.

4 Graph Maximalism

For those who agree at least in part with both the understanding-and-structure and understanding-and-compression ideas, it may be desirable to find a way to combine them. In his paper on understanding as compression, Wilkenfeld (2018) tried to keep these approaches somewhat apart, plausibly in part because he wanted to stake out a place for his account as a distinct viable theoretical option. I think, however, this might give a wrong idea of the viability of a combined account. Wilkenfeld's criticism of Grimm's account, who serves as an exemplar of the kind of mixed structuralist account of understanding we are interested in here, does not touch on the focus on structure as much as the idiosyncratic focus of Grimm's theory on the idea that understanding is a species of knowledge. In fact, Wilkenfeld admits that possessing a mechanism for being able to check the counterfactual consequences of changes to relevant variables concerning a topic 'will be a good way to encode some information, as one will be able to remember general connections between variables along with lists of deviations rather than having to remember an indefinite number or pairwise (or more) relations' (18). What he is saying here is that Grimm's idea that the content of understanding takes the shape of mental maps is in effect in line with an MDL model.

So, what gives if we drop the understanding-as-knowledge idea and retain both the idea that the content of understanding is structured and the idea that understanding involves compression?

Rather than producing a combination of Grimm and Wilkenfeld's accounts, I will proceed by constructing a theory of the *format* of the content of understanding from the ground up. By *format* I mean the type of representational structure of the content's vehicle. The format problem is whether the vehicle of this content is propositional (the traditional view), pictorial (like Meynell seems to suggest), graphical (as I will suggest here), or heterogeneous. Starting from scratch from this angle will allow us, I hope, to produce a clearer and more general picture of the problem space. It is important that I emphasize that I will not deal in full with the more substantive issue of what is in fact represented in understanding. While I will make suggestions

[8] A fuller overview of MDL is beyond my purposes here, so the reader is advised to refer to Grünwald (2007) for details. Later, in Sect. 5.4, I will describe in some detail a compression scheme for graphs in line with MDL ideas.

as to what that is or may be, I don't want to defend a particular view, at least not here. My hesitancy to do so also stems from the impression that a model where the content of understanding is heterogeneous is more plausible than one where it isn't (but keep in mind that this is about heterogeneity in the content, not necessarily in the format of the content). The need to account for this possibility will be a constraint for the construction of my account.

In summary of what comes ahead, I will propose a 'graphical' account of the format of the content of understanding. According to this view, which I will call Graph Maximalism, the content of understanding has graphical structure, that is, it can be modelled with graphs. Before moving on to sketching the basic view (4.2), and then a fuller version of the idea (4.3), I will first give a brief summary of the graph-theoretical concepts I will use throughout (4.1). After I do all this, I will argue that the view has many desirable features as part of a full account of understanding (4.4).

4.1 Graph Theory

A *graph* in the sense intended here is an object composed of vertices and edges.[9] Formally, it is a pair $\langle V, E\rangle$, where V is a nonempty set of vertices, and E is a set of pairs of items in V.[10] For our purposes here, we are interested in *general graphs*, where it is allowed for an edge to go from and arrive at the same edge, forming a loop, and where it is allowed for multiple edges to connect the same vertices.[11] We will also allow for vertices and edges to have various kinds of information attached (Fig. 1).

A vertex in a graph can be *adjacent* to another if there is an edge between them. There is a *walk* between two vertices if there is a sequence of adjacent vertices between them (the edge between to adjacent vertices is already a walk). A *path* is a walk with no cycles. A vertex can be adjacent to more than one other vertex. The number of vertices that a vertex is adjacent to is its *degree*. When vertices are connected by directed edges, we say that a vertex's *outdegree* is the number of distinct edges that begin at the vertex, and that its *indegree* is the number of edges that arrive at the vertex.

In a *disconnected* graph, there is at least a pair of vertices that are not connected by any path (for example, if in G_1 the edge between c and d was missing, there would not be a path between c and d, and the graph would be disconnected). In general, the *connectivity* of a graph is a measure of how easy it is for it to become disconnected after a series of manipulations. The existence of any given path between two vertices in the graph depends on the connectivity of the graph. A graph where each vertex is

[9] A good introduction to graph theory can be found in Hartsfield and Ringel (2003). For a broader view, see Bondy (2008), and for an algorithmic approach, see Jungnickel (2005).

[10] I will allow edges to be both directed or undirected. The main difference is that directed edges are ordered pairs of vertices.

[11] This is sometimes called a pseudograph, because it lacks some properties of graphs that disallow these things.

connected to every other and thus there is a path for any combination of vertices is called a *complete* graph.

A graph G is a *subgraph* of a graph F iff every vertex of G is a vertex of F and every edge of G is an edge of F.

Besides the pair-of-sets representation given above, graphs can be represented in various ways. An *adjacency list* of a graph G is a set of pairs $\langle V, A\rangle$, where V is a vertex in G and A is a set of vertices that are adjacent to V in G. An *incidence matrix* is a $V \times E$ matrix where each cell represents how many times a vertex and edge are incident. An *adjacency matrix* is a $V \times V$ matrix where each cell represents the number of edges adjoining a pair of vertices (Fig. 2).

In Sect. 4.4 I will make use of two more graph-theoretical notions that are of a more specialized use. First, the *degree matrix* G_D of a graph G will be a diagonal $V \times V$ matrix that represents how many vertices are adjacent to each vertex (compare with the adjacency matrix). Second, a *Laplacian matrix* L of a graph will be the difference of the degree matrix D and the adjacency matrix A, that is, $D - A$ (Fig. 3).

Note that the degree matrix is not suitable as a graph representation in the sense that adjacency or incidence matrices are: different graphs can share the same degree matrix.

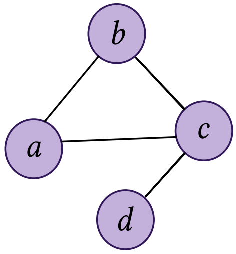
4.2 The Basic Idea

As I already pointed out, Graph Maximalism hinges on the idea that the format of the content of understanding states is graphical. What this means is that the structure of the content is the structure of a graph in the mathematical sense we just described: it contains items that play the role of vertices, and other items that connect them and play the role of edges. Accordingly, the attitudes directed towards the content of understanding are directed towards graphical structures. This vindicates in part Elgin's idea that understanding is in the first place a relationship to a comprehensive body of information; in our case, this body of information takes the shape of a graph.[12]

The graphical model contrasts with a propositional model, where the content of understanding is composed of propositions. A model like that may be implicit in some reductive accounts of understanding, where understanding is taken to consist in knowledge of some sort, for example. However, it is more fair to say that in fact most accounts have never actually addressed the issue of the format of the content of understanding, so they have defaulted into something that is more traditional. As

[12] There are actually two ways in which we can take this idea. On the one, we can say that the content of an understanding state *is* a kind of graph. On the other, we can say that it suffices that the content can be *represented* by means of a graph. As a reviewer points out, the former is *prima facie* a far stronger claim. For my purposes here, the weaker claim suffices. However, perhaps it would be better to say the following: the content of an understanding state is a *concrete* graph, an object (of some kind) with graphical structure (so, it exemplifies what we may call an *abstract* graph, which is the subject matter of graph theory), similarly to how Fig. 1 (the diagram inscribed in the page) is a concrete graph that exemplifies an abstract graphical structure. In these cases we recognize that these objects have this structure because we seem to be able to represent some of their features in a graphical manner.

Fig. 1 $G_1 = \langle V = \{a, b, c, d\}, E = \{\{a, b\}, \{b, c\}, \{a, c\}, \{c, d\}\}\rangle$



we saw in Sect. 2, there have been some exceptions like Zagzebski and Elgin, but they have not been fully explicit on their own proposals either, falling back into a form of undiferentiated pluralism regarding the format. While I think that a form of pluralism is perhaps plausible (a point I alluded to before), I think the hypothesis that the format of the content of understanding is uniform should be reconsidered outside the narrow confines of the propositional model.[13] For my purposes here, I will assume as a working hypothesis that a graphical format can be the single format for the content of understanding. The graphical model should be evaluated in terms of its universality only once it is made more precise–which is what I am trying to do here. If the graphical model cannot do all the work that we want out of an account of the format of understanding, let us then adopt a pluralist view (and this may allow for propositional content). However, I hope to convince the reader that the graphical model can go a long way.

In its basic form, Graph Maximalism can be formulated as a necessary condition for understanding, as follows:

> Graph Maximalism (Basic) S understands X only if S's cognitive state U can be characterized in terms of a pair $\langle G, D\rangle$ of a graph G relevantly connected to X, and a dispositional profile D related to G, that includes the capacity to process it and manipulate it.

Put less formally, Graph Maximalism requires that in order for someone to understand, the content of their cognitive state must take the format of a graph, which the subject must be capable to process and manipulate. The idea that the subject 'has' the data in the graph available must be understood in a more or less thick sense, and an account of attribution would plausibly specify a sense that is thick enough for the subject to merit attribution (for example, it might then be required that the subject not only has the data, but is also able to use it competently for a range of tasks). Here we are satisfied with much less.

One important class of capacities that should belong to D is the ability to construct other representations (including other graphs) from G on demand, and much

[13] A defender of the propositional model could argue the same point, suggesting that propositional contents can be heterogeneous in ways that accommodate the intuitions of those critics of the propositional model. This has been in fact suggested in the case of so-called mental maps. Camp (2018) offers an attack on those attempts.

Adjacency matrix

	a	b	c	d
a	0	1	1	0
b	1	0	1	0
c	1	1	0	1
d	0	0	1	0

Adjacency list

a	b, c
b	a, c
c	a, b, d
d	c

Incidence matrix

	a,b	a,c	b,c	c,d
a	1	1	0	0
b	1	0	1	0
c	0	1	1	1
d	0	0	0	1

Fig. 2 Alternative representations of G_1

Fig. 3 Other graph matrices of G_1

Degree matrix

	a	b	c	d
a	2	0	0	0
b	0	2	0	0
c	0	0	3	0
d	0	0	0	1

Laplacian matrix

	a	b	c	d
a	2	-1	-1	0
b	-1	2	-1	0
c	-1	-1	3	-1
d	0	0	-1	1

of what follows hinges on these capacities (to anticipate a bit from Sect. 5, a graph is in some sense compressible when the subject has the capacity to produce a compressed representation of it). We may say that this is *latent* content of the state, and that a state $\langle G_1, D_1 \rangle$ with latent content L_1 is *l-equivalent* to the state $\langle G_1 \cup_{D_1} L_1, D_1 \rangle$, where $G_1 \cup_{D_1} L_1$ is whatever graph results from updating G_1 with the contents of L_1 (given the capabilities to update G_1 in D_1).

Note that the class of subjects who satisfy the condition that Graph Maximalism imposes on understanding is broader than the class of subjects who at some point of evaluation are attributable with understanding. Since their cognitive state may be the basis for understanding attributions, we may say that they are in an *understanding-like* state–a state that is like an arbitrary understanding state in some respects but that is not necessarily attributable with understanding. Graph Maximalism can thus be construed as a definition of what it is for a state to be understanding-like. This is important because a theory of understanding should offer an account of understanding-like states, but I will not press the point further here.[14]

[14] Cf. Elgin's (2009) criticism of Kvanvig (2003) concerning the issue of whether certain individuals should be attributed with understanding in an 'honorific' but not literal sense, which can be interpreted in terms of whether the subjects are attributable with understanding or merely with something understanding-like. From a broader perspective, we should acknowledge that 'mere' understanding-like states can also have epistemological significance.

4.3 The Full Picture

The basic picture I just gave of the GM model is insufficient in several respects.

First, it does not properly characterize the state of subjects in view of their potential coupling to their environment/context. It may be desirable to be able to distinguish between two subjects who have the same representation of a topic in worlds where the topic is realized differently. As I pointed out above, we need to distinguish between the target and the content of understanding. In a representational framework, we will want to measure how well the representation of the target matches the target. This is a property of the state that we can't recover purely from the intrinsic properties of the representation.

Accordingly, we will distinguish between what we may call the *internal graph*, which is the representation itself, and the target, and make both characterize the state of understanding subjects. For simplicity purposes, and given that we may suppose that the target of understanding is itself structured, we may talk of an *external graph* instead of a target. Otherwise, we might also call this the *blob* or *lump*.[15] For cases where representational accuracy is not relevant, we may neglect to deal with specifying an external graph.

The introduction of an external graph or a blob as target requires some observations. First, there is the problem that understanding subjects might simply lack the means to acquire information about such graph. But more importantly, the model seems to require that its users (namely, us) are able to discriminate what such graph could contain. We are allowed perhaps to assume that our own understanding offers a picture of whatever that is, but this is of course unsatisfactory.

Second, the idea that there would be only one of each internal and external graphs may be implausible. A subject's overall understanding state, the sum of all that they understand, will plausibly target multiple disconnected topics–such state should be at least characterized in terms of multiple external graphs, and it may require multiple internal graphs as well if the subject has no way to piece all those topics together.[16] Note that this will not only happen at this level: dealing with a delimited topic may also require the subject to target distinct structures in different representational clusters (for example, understanding a social system may involve the representation of causal structure but also of symbolic structure).[17] The same target can also be represented by different internal graphs (for example, when different representations may serve different purposes or when the subject has not settled between conflicting views). Finally, the same internal graph could represent different targets.

[15] Cf. Eklund (2008).

[16] The world might be 'dappled', as Cartwright (1999) suggests. See Strevens (2017) for a critical discussion.

[17] This suggests that the external graph might itself be a kind of composite, with a layered structure, or even that there could be a multitude of external graphs. In this case, different layers or graphs could be more relevant to different tasks.

Thirdly, and given the previous points, understanding states should also be characterized in part by the kind of mapping there is between internal and external maps.[18]

Given these adjustments to the view, Graph Maximalism should be stated more precisely as follows:

> Graph Maximalism (Full) S understands at t only if S is at t in a state U characterizable by a 4-tuple $\langle G_I, G_E, M, D\rangle$, where G_I is a non-empty set of internal graphs, G_E is a possibly empty set of external graphs, M is a mapping between elements in G_I and G_E, and D is a dispositional profile associated to possible manipulations of G_I.

While this will be the form of GM that I will work with in this paper, it still fails to accommodate a final complication with the format of the internal graph(s). While the same graph can have different drawings (concrete realizations), we might want to say that different drawings of the same graphs will offer different mental contents, and that their 'appearance' itself offers information that is relevant to their understanding state.[19] To some degree, this can be simulated by graphical states that store the topology of drawings, but it seems more plausible that the format of understanding states should embody this more directly. So the model should really deal with the drawings of graphs rather than the graphs properly speaking. This preserves many of the properties of the graphical model and extends it in the way required.[20]

4.4 Advantages of GM

A graphical model like GM has several nice features.

First, GM offers a concrete implementation of the idea that understanding has to do with structure. The full GM model has a way to account for both external and internal structure, and it stresses how understanding is not just a matter of having certain information available, but also having the capacities to process and manipulate it. To the degree that these ideas are attractive (and as we have seen in Sect. 2 this impression seems to be widespread), the GM model seems to be initially plausible.

Adding to this last point, a graphical model like GM should be able to explain at least partially the specific abilities that are characteristic of understanding, such

[18] In principle, we could ask that the mapping relation satisfies certain constraints, but I think this is more relevant to the issue of understanding attribution (it could be that a subject's state is only attributable with understanding if the mapping relation satisfies certain conditions). As I said before, I don't want to offer an account of understanding attribution here.

[19] With mathematical graphs, appearance is a matter of how the graph is embedded in some space (cf. Gross and Tucker (1987) on topological graph theory, with deals with this kind of object). In the present case, there seem to be two options: the internal graph may afford different conditions for mental access, or it may have an imaginistic component with spatial properties.

[20] The resulting format is not only graphical, but also in a sense pictorial, which allows us to capture some of the insights from Meynell (2020).

as representational manipulability (Wilkenfeld , 2013) and cognitive control (Hills , 2016). Whatever cognitive capacities understanding involves, they must supervene on the properties of the underlying representational format: the relevant subjects would not have these capacities if the format of their representations was not as it is. The scope of the capacities invoked in this supervenience claim has to be understood narrowly. For some of these capacities, what format they will be supported by will not matter for their possession; plainly, for example, the capacity to produce explanations can be had by someone whose understanding is based on propositions or on graphical structure. However, the format could matter for their specific performance or reliability. In the limiting case, there can be capacities that can only be supported by specific formats. Stroud (1979) and Gardiner (2012) have sketched arguments to this effect in favor of the idea that understanding couldn't have a purely propositional format. Stroud, in particular, wanted to show that the structural relations between beliefs cannot be captured by further beliefs; something else, perhaps dispositions, is necessary. This is compatible with understanding having a propositionally formated content linked to some appropriate dispositional profile. Gardiner goes further, formulating her argument in terms of the difference between a subject whose beliefs are stored as a list of propositions and another who also stores the links between beliefs.[21] The latter can do something that the former cannot do, which is storing information about the relations between their beliefs; Gardiner argues that this cannot be done by the former on the risk of regress.[22]

A graphical model also offers something that can act as a unified format for thought. Graphs are very flexible and can serve as a mechanism to emulate other formats as well. For example, sequences of propositions can be easily represented in the form of linear graphs.[23] In this sense at least, we could see the graphical model as not entirely in opposition to the propositional model, since what it captures can be taken as a special case of what can be captured by a more general graphical model (namely, it captures contents that can be represented as linear graphs). Arguments can be represented by trees. From an architectural point of view, a common format that can serve as input or output of a wide range of mechanisms is desirable. Besides, the supposition that there is such format also allows us to simplify our account of understanding. Whether this supposition is empirically supported is another question (I will return to this point in Sect. 6). Ultimately, while I think the idea that graphs can act as a unifying format for thought is attractive, I don't wish to commit to it, since I think pluralism about the format may have certain advantages as well. I return to this point in Sect. 5.4.

[21] See footnote 7.

[22] On a similar point, one may be reminded of some ingenious arguments in Kitcher and Varzi (2000) to the effect that pictures (and we may try the assumption that at least some pictures are graphs) are worth $2^{\aleph_0}$ sentences.

[23] Unordered collections of propositions can be represented by a disconnected graph, where every vertex/proposition has a degree of 0, or as complete graphs (from any proposition of the collection you can proceed to any other). Now, this lack of structure is arguably not even true of written text, since there can be an assortment of different relations between the sentences in a textual body. To that extent, the purely propositional model is also a bad representation of the expression of thoughs.

Most importantly, the graphical model seems to support a much richer account of the measures of understanding, and in turn a more specific attribution story. For one, the capacity to measure how well the internal graph matches the external graph is built in, and this gives us a measure of *accuracy*. Likewise, the *breadth* of the state will plausibly be measured in terms of how large the union of external graphs that are matched by the internal graphs is. *Depth* can be construed in terms of how *dense* the graph is. If (some) edges represent compatibility relations, *coherence* will require that the subgraph of the vertices that are adjacent to each other through compatibility edges is complete. And so on.

While it is not my intention to provide a full account of the different measures that are available in the graph model, it is important to point out that for some of these measures, the graph model is much more efficient than the alternatives, and that in some cases these measures can only be described if we make use of something like it. In these cases, the graphical model offers perspicuity whereas the ordinary ways to characterize these measures are much less descriptively rich or simply non-existent.

An example of this is *robustness*. A typical reconstruction of the notion is modal: roughly, a subject's understanding is robust if at large certain features of their understanding (such as coherence) are resistant to the introduction of new information, where resistance is understood in terms of counterfactuals to the effect that if certain events concerning understanding (in the case at hand, informational updates) happened, the relevant properties of understanding would remain relevantly the same. This is only a rough sketch of the point, but it will do for my purposes here.

In the graphical model, changes to the content can be represented as changes to graphs (the insertion or deletion of vertices, the addition or deletion of edges, and so on). We may be worried about the robustness of the structure of the graphs through changes of these sorts. For example, we may worry about the ease with which the graph (or, most likely, a subgraph of it that encodes consistency relationships) could become disconnected–a function of the (sub)graph's connectivity. Connectivity can be measured quite precisely in the graphical model. The most straightforward measure of connectivity is the number of edges or vertices that it would take to delete in order to get a disconnected graph. Alternatively, the degree of connectivity could be measured as a real value. For example, it is known that the second eigenvalue of the Laplacian matrix gives a continuous measure of how close the graph is from being disconnected (a value of 0 means it is disconnected, and small values that it is close to being disconnected)–this measure is called *algebraic connectivity*. There is nothing particularly significant about this measure in particular; there is an abundance of measures of connectivity in the literature.[24] My suggestion is that this body of knowledge can and should be applied to the study of epistemic states of a 'holistic' character.

[24] See Oellermann (1996) and Freitas et al. (2022) for surveys. Some connectivity measures are computationally expensive, so there is some interest in devising connectivity measures that are useful more broadly. Cf. Beineke et al. (2002).

5 Compressible Graph Maximalism

We now have a more or less complete picture of a graphical model of the format of understanding states. While this account captures the understanding-and-structure intuitions, it does not capture the understanding-and-compression intuitions. Fortunately, GM can be extended to accommodate them quite simply. In this section, I will offer an extension to GM that introduces two dimensions in which understanding states can exhibit compression. I will call the resulting model *Compressible Graph Maximalism* (CGM). First, I will stablish some additional distinctions that will be of use later, concerning the compression of graphs (Sect. 5.1). Then, I will argue that in a compressible graph model we can think of the issue of how the internal graph compresses the external graph (Sect. 5.2), and also of the issue of how the internal graph can itself be compressed (Sect. 5.3). Finally, I will discuss some implementations of the idea (Sect. 5.4).

5.1 Graph Compression and Summarization

As we pointed out, compression relies on the existence of structural patterns. The structural properties of graphs are ripe with possibilities for compression. We already saw different ways to represent graphs, as matrices or lists of several kinds. Computationally, these representations afford different tradeoffs in terms of size and algorithmic complexity. For example, adjacency matrices make it easy to check if two vertices are connected by an edge, but are fairly large because they have to encode a lot of unnecessary information about what connections fail to exist (all cells in the matrix with 0 values). On the contrary, adjacency lists make it easy to see what vertices are adjacent to a given vertex, and take up much less space, but make it harder to check if an edge exists.

These representations of graphs are lossless, since all the information about the graph can be recovered from them. For certain purposes, lossless representations are not necessary. In particular, sometimes we want to represent only the most significant features of a graph; in those cases, we can make use of lossy formats. More generally we will talk about making a *summary* of a graph, or *graph summarization*. It is important to note that what counts as a good graph summary is task-dependent: since a summary is intended to represent only features of interest, interest needs to be defined in context.[25] While summarization is a form of compression, in what follows I will prefer using the term ‘compression’ to refer kinds of compression that do not select features by a measure of interestingness (unless by the context it makes sense to refer to the general class).

Previous work on the role of compression for understanding has not made a distinction between compression and summarization. For many purposes, for understanding what matters is not just the size of the representations used, but their ability to represent useful features. Significantly, it is possible to generate information that

[25] Liu et al. (2018) give an excellent overview of summarization techniques and challenges.

is not explicitly represented from a summary just as it is possible to generate it from a merely compressed representation. What will differ in these cases may be the quality of the generated information, which may in turn impact performance of the tasks where understanding is required. But it is possible for the same *task-relevant* information to be generated from a summary and a merely compressed representation–in this case, if there is a difference in performance, it may depend on the properties of the particular mechanisms that generate this information.

5.2 Compressing the External Graph

How can we integrate compression into the GM model? Part of the answer is straightforward, since in a model like GM, there are only two points where compression can happen or where it can be introduced. On the one hand, we can see the internal graphs as compressed representation of their targets, the external graphs. On the other, we can see that the internal graphs can themselves be more or less compressed. In this section I will deal with the first point.

As we already saw, GM incorporates an external graph as part of the state that characterizes a subject's understanding. This was initially done so we can capture properties like accuracy in the model as a measure of the state. As it happens, this feature of the model also allows us to measure in principle how well an internal graph compresses an external graph that it is mapped to. In this sense, external compression is a constitutive feature of a process that generates internal graphs–internal graphs can be generated by compressing the external graph and storing it.[26] The external graph G_E is *e-compressible* to the degree that the subject can produce a compressed internal graph of it.

There are some complications. In general, we can suppose that many kinds of representation of extra-mental items involves some form of compression (informally, representations tend to not capture all available information about their extra-mental targets, so they are lossy with regards with a supposed ideal complete representation of their targets).[27] This, however, does not mean that it is in principle possible to assess if a representation is compressed in absolute terms–we do not generally have access to those ideal complete representations. So we cannot in general measure how much information was lost in the process of compressing the external graph–to do that we would need access to information that we do not have access to.[28] In practice, to assess that a given internal graph compresses an external graph,

[26] In the full GM model there may be multiple internal graphs, and each could be the product of a different operation where the external graph is compressed.

[27] Representations of representations could in principle be copies of their targets, so they won't necessarily involve compression.

[28] This can be generalized as an objection to understanding accounts that rely on ideal states in order to account for degrees of understanding (like in Khalifa (2017) and Kelp (2021))–in those accounts degrees correspond to the distance between a given state and some ideal. If we cannot in general say how a state approximates ideal conditions, we cannot say to what degree of understanding it corresponds to (although we could still make relative comparisons, and that is all it takes to give an account of outright attribution). Cf. Baumberger (2019).

we need to be able to judge that at least there is some representation of the external graph that is less compressed, and this additional representation needs to be produced through interaction with the external graph in some way that is comparable to the way by which the internal graph was produced.

This complication arises because if the internal graph is lossy there is in principle no way to recover the information that was lost. Even if we had an ideal decompressing mechanism that was able to extract as much information from the graph as possible, we would not in principle be able to dismiss the possibility that the target could have elements that the graph does not represent. One could have a mechanism that generates *new* information that could supplement in some ways that which can be extracted by such ideal decompressing mechanism, however (call the result of a mechanism like this an *inflated* representation of the target). On the basis of something like this we might be able to say that the target *might* be compressed by the graph: there is a possible object that is similar to the target of the graph that when compressed would take more space than the best representation of the target we can recover from the graph. But we cannot say with certainty that this object *is* the target.

While there is no way to circumvent this limitation, it can be defused somewhat. A subject could simply make use of an inflated representation as a measure of the relative compression of the internal graph–this inflated representation may itself be graphical, although this is not strictly speaking necessary. While the resulting measure of compression is not accurate, it may be a good enough approximation for most purposes, since it would allow for relative comparisons. There are various ways to accommodate this element into the GM model. For example, the inflated representation may replace the external graph in the characterization of the understanding state of the subject, or it may be added as an additional parameter to such characterization. It could also be enough that the subject has the capacity to produce an inflated representation on-demand, in which case the inflated representation could be taken as part of the virtual content of the state. I think this option may be preferable since what counts as a good representation of the target will depend on the context and thus may need to be updated dynamically.

In any case, compression of the external graphs can be a source of performance differential between subjects engaged in tasks that make use of their understanding of certain subject matters. How efficiently they have compressed the graph can make a difference to their competence in using the relevant information when prompted. The degree of a subject's understanding is in part a function of their compression of their external graphs. The contribution of external compression to the degree of understanding is not straightforward, however: it may be that a more compressed internal graph could be less efficient for a given task that one that is less compressed. Whether compression improves performance is a matter of tradeoffs.

5.3 Compressing the Internal Graph

Once we have an internal graph, we may ask whether it can be compressed further. Unlike external compression, which is a form of representation production (as we just saw), internal compression is a form of representation *manipulation*. A subject S's internal graph G_I is *i-compressible* iff S can produce a compressed graph from G_I, that is, if $C(G_I) \neq G_I$, where $C(G_I)$ represents an application of the compression manipulation on G_I.[29]

For internal compression, we need to consider two types of cases: first, cases where graphs are compressed in place, and second, cases where compressed graphs are added to the set of internal graphs. When the compression operation is lossless, this difference is not significant, but in-place lossy compression entails a net loss of available information extractable from the internal graph in situations where there is no redundancy. Suppose G_I was a set of two internal graphs G_{I1}, G_{I2} that losslessly compress the same information from an external graph G_E (this is the original state s_0). Compressing G_{I1} further losslessly causes no loss of overall available information (this is state s_1), which means we can go back to state s_0. Then, compressing G_{I2} losslessly does not cause loss of information either (this is state s_2), which means we can go back to state s_1 and thus to s_0. Compressing G_{I1} lossily does not reduce overall available information, since it was redundant (this is state s_4). But compressing G_{I2} lossily at this point makes the state *does* reduce the overall available information, and it becomes impossible to go back to the previous states (this is state s_5).[30]

Allow me to digress slightly with some general observations about what the GM model suggests about the nature of understanding, since they bear on the dynamics of internal compression. The content of a subject's understanding is subject to continuous change during their interaction with the world and during their own inner lives. In terms of the GM model, this means that internal graphs are constantly being updated, by adding, replacing or removing vertices and edges. Each update to this state offers a new opportunity for compression, which is itself a kind of update to the content of internal graphs. As new facts are learnt, new patterns and consequently redundancy may be discovered and made available.[31] This process is not necessarily explicit nor guided by the subject itself. It is not, however, a shadowy process behind the subject's cognitive life. Rather, it *is* the systematic operation of the subject's cognition as a whole. Part of the dynamics of understanding will have to do, then, with the coordination of the different components of this overall system. For example, since in the full GM model there may be a multitude of internal graphs mapped to the same targets, we may expect them to go out of sync as the system develops. For the system to remain consistent, some kind of monitoring is necessary.

Like in the case of compression of the external graph, compression of the internal graph can be a source of performance differential between subjects. A more efficiently compressed internal graph can make a difference in the competence of

[29] $C(G_I)$ does not need to be smaller than G_I. Because of the combinatory pigeonhole principle, if C is a lossless operation, compressing some internal graphs is bound to generate larger representations. In practice, this leads to the use of domain-specific compression algorithms. See Tate (2003).

[30] It is already impossible to go back to s_3 from s_4 if it is not known that G_{I1} and G_{I2} are supposed to represent the same target (note that this information is tracked in M in the full GM model). If it is known, one could simply replace G_{I1} for G_{I2}.

[31] In fact, it is possible that the subject or their cognitive system may learn new compression strategies as a result of this as well.

a subject to any task that requires understanding. So it is plausible that the degree of someone's understanding is a function of their internal compression as well. The same caveats that we raised about the contribution of compression to the degree of understanding in the external case apply here.

5.4 An Implementation

Until now, we have considered the question of the compression of understanding content abstractly. We should take a close look at the question of how this content could be compressed, if its format is graphical. There is an abundance of compression schemes for graphs, so I will limit to describe one scheme that is both simple and illustrative.[32]

Navlakha et al. (2008) describe a general graph representation that has some interesting compression features. In this model, a graph G is represented by a pair $\langle S, C\rangle$ where S is a graph $\langle V_S, E_S\rangle$, the *summary*, and C is a set of what we will call *edge corrections*. The vertices in the summary graph are themselves sets of vertices in G. When there is an edge between two vertices v'_S and v''_S in the summary, one can draw an edge between every vertex $w \in v'_S$ to every vertex $w' \in v''_S$. It may turn out that some of these edges do not exist in the original graph, or that some edges are still missing. These differences between the original graph and the summary are recorded as edge corrections: if some edge $\{v_1, v_2\}$ is missing, it will be recorded as $\langle +, \{v_1, v_2\}\rangle$, and if it should be removed, as $\langle -, \{v_1, v_2\}\rangle$.

To understand how this representation can be useful in a compression scheme, let's consider its application to the simple graph G_1 from Sect. 4.1. First, we can represent G_1 exactly by a pair R_1 where C is empty, and S is structurally identical to G_1,[33] that is,

$$\begin{aligned} R_1 = \langle S = \langle \{A = \{a\}, B = \{b\}, C = \{c\}, D = \{d\}\}, \\ \{\{A, B\}, \{B, C\}, \{A, C\}, \{C, D\}\}\rangle, \\ C = \emptyset\rangle \end{aligned}$$

If we summarize the graph as if G_1 was a complete graph (that is, one where each vertex is connected to every other), we would need to correct for a couple of unnecessary edges. We could then represent G_1 as

$$\begin{aligned} R_2 = \langle S = \langle \{A = \{a, b, c, d\}\}, \{\{A, A\}\}\rangle, \\ C = \{\langle -, \{a, d\}\rangle, \langle -, \{b, d\}\rangle\}\rangle \end{aligned}$$

which is more compact. While the differences are not so striking in such small graph, they can be significant in large ones (Fig. 4).

---

[32] Besta and Hoefler (2018) give a survey of lossless graph compression schemes, and Liu et al. (2018) one for graph summarization. A more limited survey can be found in Zhou (2015), from which I take the scheme I discuss here.

[33] Strictly speaking, while in G_1 the set of vertices is $\{a, b, c, d\}$, in R the set of vertices should be $\{\{a\}, \{b\}, \{c\}, \{d\}\}$. Similar points can be made about the sets of edges.

What this illustrates is that depending on how the graph is summarized, it is possible to achieve varying degrees of compression. In fact, the scheme corresponds quite closely to one idea from the MDL literature, which is that compression can be achieved by providing a general hypothesis about the data along with additional information about the specific case at hand in terms of the hypothesis.[34] The former role is played by the summary graph, and the latter by the edge correction set. Figuring out a compression algorithm for graphs that uses this type of representation involves figuring out how to produce a summary/edge correction pair with a small cost (understanding cost as the sum of the size of these components); finding the representation with the smallest cost is the same as computing what is called its MDL representation. An MDL representation is essentially lossless, but the scheme can accommodate lossy compression by, for example, removing edges from an MDL representation and checking if some constraints on error are satisfied. Navlakha et al. (2008) describe various algorithms for finding MDL representations, as well as others for approximate representations; the reader is advised to check them for the details.

Describing this model of graph representation, and how compression schemes could make use of it, is merely meant to illustrate some of the principles for graph compression. Concretely, what schemes could actually be implemented in human understanding is an empirical question. Furthermore, on the assumption that human representations are somewhat heterogeneous (Pearson & Kosslyin , 2015), it may seem implausible that human cognitive systems could make use of a single compression scheme. Different mechanisms could be specialized to different types of tasks; and in fact the particular format of the contents of understanding could vary even though its graphical character could remain constant.[35] For some purposes, simple graphs are sufficient; for others, general graphs are necessary; for yet another, topology matters; in some cases, tree-like graphs are enough; and so on. Here I am satisfied with outlining the basic idea of a graphical model of understanding that incorporates an element of compression.

6 Challenges

We have an overview of how the CGM model and some of its purported advantages. I will now dedicate some time to addressing some potential concerns with and challenges to the model, in no particular order: the worry that the model introduces extraneous properties to our theory of understanding (6.1), the concern that it is not clear what the model actually represents (6.2), the problem that it does not seem to handle perspective (6.3), the objection that it is too committed to representationalism

[34] If you remember, this was also the idea behind Wilkenfeld's MDL-based proposal.

[35] The graphical model is compatible with there being other non-graphical formats for representation–it is not necessary part of the model that all representation is uniformly graphical. Here I have assumed the weaker thesis that the vehicles of understanding could be uniformly graphical, but I would not think it strange if ultimately the picture would have to be more complicated even in this particular case.

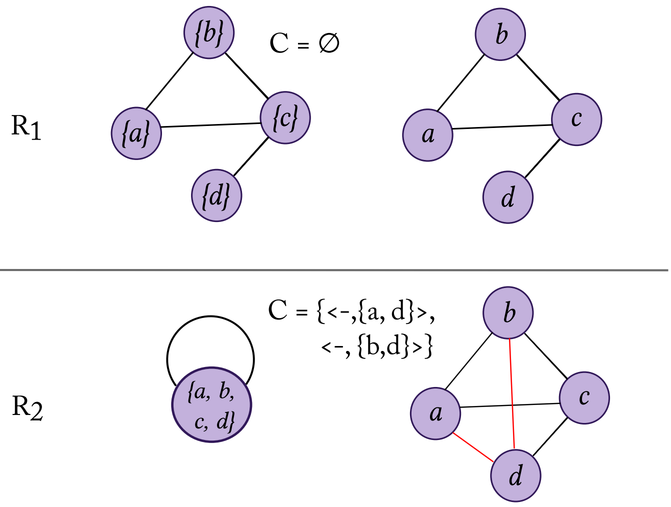


Fig. 4 R_1 and R_2, and the expansions of their summaries (corrected edges are marked in red)

when there are alternatives (6.4), the issue that it seems to focus too much on the internal state of subjects and does not account for the use of tools for understanding (6.5), the question of what kind of attribution story can be given for this model (6.6), and the problem that it may not be an accurate picture of human understanding (6.7).

6.1 The Model Introduces Extraneous Properties

Making the format of the content understanding explicitly take on the form of graphs allows us, as we have seen, to utilize all manner of graph-theoretical ideas to describe understanding states. In Sect. 4.4 this was touted as an advantage. However, this also introduces a mass of mathematical truths the instances of which could then be interpreted as truths about understanding states but which do not match at all our preconceptions of what understanding is like or are entirely disconnected to the issues that we are worried about related to understanding.[36] This 'excess theoretical load' can be taken as a sign that the model is too weak, as well as raise the worry that the model may simply end up being inaccurate due to overfitting.

Part of these worries can be weakened if we observe that models can be useful even if they have unneeded features, as long as modellers are explicit about the scope of the model, that is, what aspects of the target are intended to be represented by the model (Weisberg , 2013). One strategy, accordingly, would be to indicate that the GM model represents understanding states only in so far as it can be useful to do so. That it includes features that are not in fact used by the theory of understanding but which could perhaps be interpreted as so usable could then be presented as a further advantage, in that the model may still be useful if the requirements we make

[36] This is analogous to the phenomenon of junk-theorems. Cf. (Hamkins (2020), p. 18).

of the model changed. Since the model already has those features, we can use it also in an exploratory manner (cf. Gelfert (2016)); for example, we can use it to examine the question of how to best account for measures of understanding such as robustness, as I already sketched.

6.2 It is not Clear What the Graphical Model Actually Represents

The GM model that CGM builds on the idea that the content of understanding states is graphical. But what are the parts of the posited graphs supposed to be? That is, what are the vertices and edges that the model describes? In fact, we have already summarily touched upon this issue, but it is necessary to make some additional observations.

I think we should treat this issue liberally: edges and vertices could in principle be anything. The graphical format does not impose any substantive constraint on content. Internal graphs may encode, just to give a few examples, logical, causal, or spatial structure, and in each of these cases edges and vertices will represent different kinds of objects and relations. What is actually represented will depend on the kind of cognitive tasks that these representations could be embedded in. In this sense, it is plausible that if there are constraints on content, they are given as constraints on the inputs for graph-encoding and -manipulating mechanisms.

This liberal attitude towards the kind of things that the model may represent would allow us to incorporate some additional ideas from the literature. In Sect. 2 we saw how Grimm (2016) suggests that understanding is encoded in 'mobile maps', taking inspiration from how Gopnik et al. (2004) use Bayes nets to describe causal systems. Bayes nets are directed acyclic graphs. Vertices represent causal variables, and edges between two vertices $X$and Y represent that changes to the value of the X result in specific changes in the value of Y. This allows users of Bayes nets to saturate the nets with values to check for the behavior of the causal systems that they represent. Grimm suggests that this can happen more broadly with maps that encode all sort of dependency relations. The GM format can capture this idea straightforwardly. Furthermore, introducing the element of compression allows for causal inference: roughly, variables that make a difference should not be lost when compressed without loss of accuracy in the predictions of the net. Furthermore, it may be hypothesized that when some X causes Y, better compression is achieved when X is compressed first and then Y given X, than the other way around.[37]

6.3 The Model does not Account for Perspective

Related to the previous point, it may be observed that the model does not seem to take notice of perspective–which is an important element that a full account of understanding needs to talk about (cf. Camp (2019), Massimi (2022)).

[37] Cf. Budhathoki and Vreeken (2028), Pranay and Nagaraj (2021), Vreeken (2015) and Wieczorek and Roth (2016) for related work.

I don't have a settled view on how perspective matters to understanding, but I don't see an intrinsic incompatibility between the GM account and there being perspectival effects in understanding. As I pointed out earlier, in some cases we may want to deal explicitly with the topological features of graphs, and these may be used to encode how a target is seen from a given perspective. The visual metaphor is somewhat apt, because we can see pictures and maps (which are clearly able to encode perspective) as a special kind of topological graph where nodes have different surfaces with additional properties attached. So, if pictorial and cartographic content can encode perspective, so does graphical content.

6.4 Everything that the Model does can be Done in a Purely Dispositional Model

Since on the face of it the CGM model seems thoroughly representationalist, those who are uneasy with representationalism would object in principle to the approach taken here. There are two possible responses. On the one hand, one could bite on the representationalist bullet, and argue that the approach is not less legitimate than any other that is put in representationalist terms. On the other hand, one could weaken the notion of representation in use, and claim, for example, that whenever we talk about representations and content, we could also make do with a functionally equivalent story that involves neither (for example, one could appeal to a story about content like those provided by Hutto and Myin (2017) or Moyal-Sharrock (2021)). While some would still object to keeping the terminology of content and representation, pursuing this move may seem more easily defensible than abandoning the terminology altogether. There is no room here to settle the issue, but I think the GM model could be adopted in some form by both representationalists (straightforwardly) and non-representationalists (with some adjustments).

A different but related objection to the model as presented is that, given that a subject's understanding state is partially characterized by their dispositional state, and this can include several abilities to generate graphs on demand, perhaps it is not necessary to have an explicit set of internal graphs (possibly held in memory). My response is that relying on generating and regenerating graphs on demand is likely to be more expensive cognitively than holding on to it in storage, even if architecturally doing away with internal graphs is in principle feasible.[38]

A related worry with the approach is that the model is fundamentally static (it only depicts time-slices of understanding processes), and that for that reason it does not have the capacity to account for more dynamic marks of understanding such as ingenuity and creativity. I acknowledge that this is something that a full account of understanding needs to address, and I would admit that it would be a major deficiency of the model if it didn't allow us to at least partially make sense of this. While the model presented here is not intended to be a full account of understanding, we can sketch some reasons to think that the CGM's account has some explanatory potential in relation to this issue as well. During inquiry, one may end up in a

[38] This is independent of the issue of whether memory itself could be accounted for non-representationaly.

situation where there is no more relevant information to extract from the resources that one already possesses, or where this information cannot be compressed any further, without inquiry having settled (I use inquiry only as an example, I believe something similar can happen with other types of task). In this case, one often has to either gather new information or generate it, or attempt to see or rearrange the old information in new ways. Ingenuity can play a role here. In doing this one could gain the capacity to compress the content, because we may obtain more structural redundancy to exploit.[39] So it seems like making things more compressible matters to understanding in inquiry. For different tasks, like creating a puzzle or coming up with a good research question, one may want to introduce artificial difficulty for compression.

6.5 The Model Focuses too much on Internal State

Another issue concerning the implementation of the model is that it seems to focus merely on internal properties of understanding states. While nominally the states include external graphs, these are only used to assess the accuracy of the internal graphs. However, understanding may require active interaction with objects 'outside of the head', like instruments, props, and other symbolic systems (cf. Ylikoski (2014) and Toon (2015)). Internalizing the content of understanding states could also prevent us from making sense of collective understanding, if we wanted to allow for that (cf. Delariviere (2020)).

The worry is misplaced as a challenge to the graphical model. Internal graphs could in principle be in some cases arrangements of things–for example, they could be analogue models structured in ways that could be interpreted graphically.[40] It is also not necessary that internal graphs are owned by particular individuals–they could be distributed.

6.6 It is not Clear What Kind of Attribution Story can be Added to the Model

As I said before, it is not my goal to provide a full account of understanding here. However, some observations can be useful to show how what I propose here bears on the issues of the gradability of understanding and attribution. I hope this also clears things up for readers who may struggle to see how the graphical proposal can bear on a full account of understanding.

[39] It is worth noticing that compressibility can be increased either by introducing falsehoods or by introducing new truths. Indirectly, one could then argue that someone could further their understanding through falsehoods, but I think we need to keep those two ideas separate because the contribution of compressibility to the degree of understanding could be indirect–again, in this paper I am not attempting to give an account of the conditions for understanding attribution. On the point of understanding by falsehoods, see Rancourt (2015), Le Bihan (2017), Lawler (2021) and Elgin (2022), among many others.

[40] In this case, there may also be an internal graph for understanding the analogue model as a system, in addition to the model itself encoding graphical structure for understanding the target of the model.

A typical strategy for building an account of understanding attribution builds on an account of degrees of understanding.[41] It is well acknowledged that the possession of understanding is a matter of degrees. But what do these degrees measure? There are many proposals; in Sect. 3 we saw one from Wilkenfeld (2018). In that proposal the degree of understanding (in a given context) is a function of the capacity to extract information from a compressed representation. In principle, someone who adopts a compressible graphical model like the one proposed here could endorse something similar–in fact, Wilkenfeld's proposal could be lifted as is and used to give an attribution story. That, however, could be underutilizing the resources of the graphical part of the model. Given that this part offers a rich range of potential measures for understanding states, the best way to make use of the model would be to take the degree of understanding as a function of the structural properties of the graphs as well as of the compression level of the relevant representations; for example, someone could want to say that a state with a more connected internal graph has a greater degree of understanding than one that is less connected, all else being equal. In some cases, these properties could be more important to the appropriateness of someone's understanding than compression, as is suggested in the case of Ted Nelson's quote in the epigraph. It is important to note that comparisons between understanding states can track a wider range of properties than those that are expressed in 'better' or 'worse' judgements, or in outright attributions. As I already suggested in Sect. 4.4, the graphical model offers an abundance of potential measures for comparisons.

6.7 The Model is not True of Human Understanding

Finally, and most importantly, it may be objected that while the model captures a potential format for the content of understanding-like states, it is not clear that *human* understanding can be described as using this kind of format. Ultimately, this question is empirical. However, there is an argument to be made to the effect that there is some evidence in favor of the applicability to the human case.

First, there is traction to the idea that the format of representations (and the vehicles of mind more broadly) is not (purely) propositional. I will not give an overview of the literature here, but it is important to notice at least the literature on so-called S-representations and mental maps (Cummins , 1991; Ramsey , 2007; Gladziejewski , 2015; Rescorla , 2009; Blumson , 2011; Gladziejewski & Milkowski , 2017; Lee , 2018; Camp , 2018; Shea , 2018), the literature on connectionist models of the mind, the literature on causal maps and Bayes nets (Gopnik and Glymour (2002), Tenenbaum et al. (2011), and more recently, the literature on graph neural networks (GNNs) (cf. Zhou et al. (2022)). This body of research collectively provides a strong reason to look into models like GM for a description of human cognition.

[41] Cf. Baumberger (2019) for an overview of approaches. An important point of covergence in the literature is the importance of context to characterize the conditions in which attributions are appropriate. In Morales Carbonell (2022) I also give an overview and present my favored approach, which is a form of higher-order contextualism (where the parameters for evaluation are also selected by context).

Second, something similar can be said about the role that compression can play in cognition. The literature on predictive coding (Clark , 2013; Hohwy , 2014) suggest that compression is a core feature of cognition. There is also a growing literature on applications of the MDL principle to the problem of learning (Grünwald , 2007; Robinet et al. , 2011).

Consequently, there are independent reasons to think that both structured representation and compression are important to human cognition. This makes it plausible that CGM can serve as a model, not just for understanding-like states of some sort, but also for the kind of understanding-like states that are typical of humans.

7 Conclusion

In this paper, I have sketched a model for the format of the content of understanding states, Compressible Graph Maximalism or CGM. The characteristic features of CGM are that (1) the format of the content of understanding states is graphical, and (2) that this graphical content can be compressed. The model combines the features of two important approaches to the nature of understanding, and I have argued that this combination offers some advantages. I expect this approach should be of interest to both those who think that understanding involves grasp of structure, and to those that think that understanding is compression, since it offers a concrete way to think about the format of the representations involved. While some challenges can be raised against the view, I think it can withstand them. In particular, one may worry if the model is empirically adequate with regard to human cognition. While there are reasons to think that it is at least compatible with some recent accounts of cognition, and thus is no less plausible than them, it could also be helpful to think of it as a framework for thinking about the construction of systems that can understand. The current project is thus in line with recent work on naturalized epistemology of understanding (cf. Khalifa et al. (2022)).

Acknowledgements This research was funded by ANID (Chile), project 3220017. Earlier versions of this work were presented at the 4th SURe workshop (NY/online, April 2022) and at the CLPS seminar in Leuven (October 2022). I am thankful for the feedback given in those occasions, specially from Conor Mayo-Wilson, Letitia Meynell, Leander Vignero, Jan Heylen, Wouter Termont, Giulia Lorenzi, Kristine Grigoryan, and Lorenz Demey. I am also thankful for the feedback received from reviewers.

Declarations

Conflict of interest The author has no conflicts of interest to disclose.

References

Baumberger, C. (2019). Explicating objectual understanding: Taking degrees seriously. *Journal for General Philosophy of Science, 50*, 367–388.

Beineke, L., Oellerman, O., & Pippert, R. (2002). The average connectivity of graphs. *Discrete Mathematics, 252*, 31–45.

Besta, M., Hoefler, T. (2018). Survey and taxonomy of lossless graph compression and space-efficient graph representations https://arxiv.org/abs/arXiv:1806.01799v2 [cs.DS].

Blumson, B. (2011). Mental maps. *Philosophy and Phenomenological Research, 85*(2), 413–434.
Bondy, U. J. A., & Murty,. (2008). *Graph Theory*. Berlin: Springer-Verlag.
Budhathoki, K., & Vreeken, J. (2028). Origo: Causal inference by compression. *Knowledge and Information Systems, 56*, 285–307.
Camp, E. (2018). Why maps are not propositional. In Non-Propositional. Intentionality (Ed.), *Montague; AGM* (pp. 19–45). Oxford University Press.
Camp, E. (2019). Perspectives and frames in pursuit of ultimate understanding. In S. Grimm (Ed.), *Varieties of Understanding: New Perspectives from Philosophy, Psychology, and Theology* (pp. 17–45). Oxford University Press.
Cartwright N (1999) The dappled world: A study of the boundaries of science. Cambridge University Press.
Clark, A. (2013). Whatever next? predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(3), 181–204.
Cummins, R. (1991). *Meaning and Mental Representation*. The MIT Press.
Delariviere, S. (2020). Collective understanding–a conceptual defense for when groups should be regarded as epistemic agents with understanding. AVANT xi(2). https://doi.org/10.26913/avant.2020.02.01.
Eklund, M. (2008). The picture of reality as an amorphous lump. In: Zimmerman TSJHD (ed) Contemporary Debates in Metaphysics. Blackwell.
Elgin, C. (2009). Is understanding factive? In A. Haddock, A. Millar, & D. Pritchard (Eds.), *Epistemic Value* (pp. 331–339). Oxford University Press.
Elgin, C. (2022). Models as felicitous falsehoods. *Principia, 26*(1), 7–23.
Freitas, S., Yang, D., Kumar, S., et al. (2022). Graph vulnerability and robustness: A survey https://doi.org/10.1109/TKDE.2022.3163672, https://arxiv.org/abs/arXiv:2105.00419 [cs.SI].
Gardiner, G. (2012). Understanding, integration and epistemic value. *Acta Analytica, 27*, 163–181.
Gelfert, A. (2016). *How to Do Science with Models: A Philosophical Primer*. Springer.
Gladziejewski, P. (2015). Explaining cognitive phenomena with internal representations: A mechanistic perspective. *Studies in Logic, Grammar and Rethoric, 40*, 63–90.
Gladziejewski, P., & Milkowski, M. (2017). Structural representations: Causally relevant and different from detectors. *Biology & Philosophy, 32*, 337–355.
Gopnik, A., & Glymour, C. (2002). Causal maps and bayes nets: A cogntiive and computational account of theory-formation. In The Cognitive (Ed.), *Siegal PCSSM Basis of Science* (pp. 117–132). Cambridge University Press.
Gopnik, A., Glymour, C., Sobel, D., et al. (2004). A theory of causal learning in children: Causal maps and bayes nets. *Psychological Review, 111*, 3–32.
Gray, E., & Tall, D. (2007). Abstraction as a natural process of mental compression. *Mathematics Education Research Journal, 19*(2), 23–40.
Grimm, S. (2016). Understanding and transparency. In: Ammon SGCBS (Ed.) Explaining Understanding: New Essays in Epistemollogy and the Philosophy of Science. Routledge, pp. 212–229.
Gross, J., & Tucker, T. W. (1987). *Topological Graph Theory*. Dover Publications.
Grünwald, P. (2007). *The Minimum Description Lenght Principle*. The MIT Press.
Hamkins, J. D. (2020). *Lectures on the Philosophy of Mathematics*. The MIT Press.
Hartsfield, N., & Ringel, G. (2003). *Pearls in graph theory: a comprehensive introduction*. Dover Publications.
Hazlett, A. (2018). Understanding and structure. *Making Sense of the World: New Essays in the Philosophy of Understanding* (pp. 135–158). Oxford University Press.
Hills, A. (2016). Understanding why. *Nous, 50*(4), 661–688.
Hohwy, J. (2014). *The Predictive Mind*. Oxford University Press.
Hutto, D. D., & Myin, E. (2017). *Evolving Enactivism: Basic Minds Meet Content*. The MIT Press.
Jungnickel, D. (2005). *Graphs, Networks and Algorithms*. Springer-Verlag.
Kelp, C. (2021). *Inquiry, Knowledge and Understanding*. Oxford University Press.
Khalifa, K. (2017). *Understanding, Explanation, and Scientific Knowledge*. Oxford University Press.
Khalifa, K., Islam, F., Gamboa, J. P., et al. (2022). Integrating philosophy of understanding with the cognitive sciences. *Frontiers in Systems Neuroscience, 16*. https://doi.org/10.3389/fnsys.2022.764708.
Kitcher, P., & Varzi, A. (2000). Some pictures are worth $2^{\aleph_0}$ sentences. *Philosophy, 75*(3), 377–381.
Kvanvig, J. (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge University Press.

Lawler, I. (2021). Scientific understanding and felicitous legitimate falsehoods. *Synthese, 198*(7), 6859–6887.
Le Bihan, S. (2017). Elightening falsehoods: A modal view of scientific understanding. In S. G. C. Baumberger (Ed.), *Explaining Understanding: New Perspectives from Epistemology and Philosophy of Science* (pp. 111–135). Routledge.
Lee, J. (2018). Structural representation and the two problems of content. *Mind and Language, 34*(5), 606–626.
Li, M., & Vitanyi, P. (2019). *An Introduction to Kolmogorov Complexity and its Applications* (4th ed.). Springer.
Liu, Y., Safavi, T., Dighe, A., et al. (2018). Graph summarization methods and applications: A survey. *ACM Computing Surveys, 51*(3–62), 1–34.
Massimi, M. (2022). *Perspectival Realism*. Oxford University Press.
Meynell, L. (2020). Getting the picture: A new account of scientific understanding. *The Aesthetics of Science: Beauty, Imagination and Understanding* (pp. 36–62). Routledge.
Morales Carbonell, F. (2022). Undersanding attributions: Problems, options, and a proposal. *Theoria, 88*(3), 558–583. https://doi.org/10.1111/theo.12380
Moyal-Sharrock, D. (2021). From deed to word: Grapless and kink-free enactivism. *Synthese, 198*, 405–425.
Navlakha, S., Rastogi, R., & Shrivastava, N. (2008). Graph summarization with bounded error. In: SIGMOD'09: Proceedings of the 2008 ACM SIGMOD International COnference of Management of Data, ACM, pp. 419–432.
Oellermann, O. (1996). Connectivity and edge-connectivity in graphs: A survey. *Congresus Numerantium, 116*, 231–252.
Pearson, J., & Kosslyin, S. (2015). The heterogeneity of mental representation: Ending the imagery debate. In Proceedings of the National Academy of Sciences 112(33):10089–10092.
Pranay, S., & Nagaraj, N. (2021). Causal discovery using compression-complexity measures. *Journal of Biomedical Informatics* 117(103742).
Ramsey, W. (2007). *Representation Reconsidered*. Cambridge University Press.
Rancourt, B. (2015). Better understanding through falsehood. *Pacific Philosophical Quarterly*. https://doi.org/10.1111/papq.12134
Rescorla, M. (2009). Cogntiive maps and the language of thought. *British Journal of Philosophy of Science, 60*, 377–407.
Robinet, V., Lemaire, B., & Gordon, M. B. (2011). Mdlchunker: A mdl-based cognitive model of inductive learning. *Cognitive Science, 35*(7), 1352–1389.
Shea, N. (2018). *Representation in Cognitive Science*. Oxford University Press.
Strevens, M. (2017). Dappled science in a unified world. In H. Chao, J. Reiss, & S. Chen (Eds.), *Philosophy of Science in Practice: Nancy Cartwright and the Nature of Scientific Reasoning*. Berlin: Springer.
Stroud, B. (1979). Inference, belief and understanding. *Mind, 86*, 179–196.
Tate, S. R. (2003). Complexity measures. In K. Sayood (Ed.), *Lossless Compression Handbook*. Academic Press.
Tenenbaum, J., Kemp, C., Griffiths, T., et al. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science, 331*(6022), 1279–1285.
Thurston, W. P. (1990). Mathematical education. *Notices of the AMS, 37*, 844–850.
Toon, A. (2015). Where is the understanding. *Synthese, 192*(12), 3859–3875.
Vreeken, J. (2015). Causal inference by direction of information. In: Proceedings of the 2015 SIAM International Conference on Data Mining, Society for Industrial and Applied Mathematics, pp. 909–917.
Weisberg, M. (2013). *Simulation and Similarity: Using Models to Understand the World*. Oxford University Press.
Wieczorek, A., Roth, V. (2016). Causal compression https://arxiv.org/abs/arXiv:1611.00261 [stat.ML].
Wilkenfeld, D. A. (2013). Understanding as representation manipulability. *Synthese, 190*(6), 997–1016.
Wilkenfeld, D. A. (2018). Understanding as compression. *Philosophical Studies, 176*, 2807–2831. https://doi.org/10.1007/s11098-018-1152-1
Wittgenstein, L. (2009). Philosophical Investigations, 4th edn. Wiley-Blackwell.
Ylikoski, P. (2014). Agent-based simulation and sociological understanding. *Perspectives on Science, 22*, 318–335.
Zagzebski, L. (2001). Recovering understanding. *Knowledge, Truth and Duty: Essays on Epistemic Justification* (pp. 235–251). Responsibility and Virtue. Oxford University Press.

Zhou, F. (2015). Graph compression, Tech. rep. *Department of Computer Science and Helsinki Institute for Information Technonology HIIT*.

Zhou, Y., Zheng, H., Huang, X., et al. (2022). Graph neural networks: Taxonomy, advances, and trends. *ACM Transactions on Intelligent Systems and Technology, 13*(1), 1–54. https://doi.org/10.1145/3495161