

A Category of Monadic Expression*

Matthew McMillan
University of Virginia

October 2026

Abstract

A formal framework is suggested for Leibniz’s system of monadic expression using the language of mathematical categories, and this model is compared with the Structural Representation model of expression that has been developed by various authors. Leibniz’s distinctive theses are modelled using categorical notions, including the Identity of Indiscernibles, Universal Expression, Pre-established Harmony, and the theses that every monad is characterised by its point of view and that all points of view are expressions of God. A perspective on Leibniz’s distinction between natural and arbitrary expression is offered.

1. Overview

1.1 Monadic expression

Leibniz’s mature ontology consisted of an infinity of individual substances called ‘monads’. Monads are true simples, “absolutely destitute of parts”¹. Monads have qualities, and they are distinguished only by their qualities (the Principle of Identity of Indiscernibles, or PII)². Each monad *expresses* every other monad, representing the entire universe in

*A first version of this essay was written in January 2017 and has circulated informally since then. AI disclosure: in creating this version, Claude Opus 5.5 was used for purely technical support, the figure on p. 11, headings, recent literature discovery, and standardising references.

1 “A New System of the Nature and the Communication of Substances”, in: *Philosophical Papers and Letters*, ed. and trans. by L. E. Loemker, Dordrecht ²1969 (henceforth cited as Loemker), pp. 453–461, esp. p. 456; cf. “The Monadology”, Loemker, p. 643, § 1.

2 “Monadology”, Loemker, p. 643, §§ 8–9.

its qualities (Principle of Universal Expression, or PUE)³. This representation of multiplicity within unity Leibniz called *perception*⁴. A monad can express another monad in different ways and respects and with different degrees of distinctness; the combination of ways in which a monad expresses all the others is its *point of view*⁵. A monad is “like a world apart”, suffering no direct or causal influence from the others⁶; instead its state changes according to a rule of *appetition*⁷. Since the appetitive principle is internal, the fact that a monad expresses the universe at one time does not guarantee that it will for all time. The Principle of Pre-established Harmony, or PPH, says that God arranged the original states and principles of appetition of all monads so that PUE holds for all time⁸, and further that a monad’s qualities are completely characterised by its perceptions, and a monad’s appetition is completely characterised by its order of perceptions⁹.

Leibniz’s concept of expression is of fundamental importance in this system. His doctrines of PPH and PUE are defined in terms of expression. Without this concept, the structure and nature of his system would be reducible to a multiplication of the theory for a single monad. In fact, as Leibniz told de Volder, if hyperbolically, his whole philosophy follows from the PUE¹⁰.

1.2 Aims

In this essay I endeavour to explain and formalise Leibniz’s system of mutually expressing monads through the concept of a *category* from modern mathematics. Let us call this a *category of monadic expression*, or the CME account. I have two goals. The first goal is to understand what Leibniz means by expression. In this vein my proposal gives an alternative to proposals involving functions and homomorphisms already

3 “Discourse on Metaphysics”, Loemker, p. 313, § 16; “The Principles of Nature and of Grace, Based on Reason”, Loemker, p. 637, § 3; “Monadology”, Loemker, p. 648, § 56.

4 “A New Method for Learning and Teaching Jurisprudence”, Loemker, p. 91, note 16; “Correspondence with Arnauld”, Loemker, p. 344; “Principles of Nature and of Grace”, Loemker, p. 636.

5 “First Truths”, Loemker, p. 269: “these expressions vary in perfection as . . . perspectives of the same city seen from different points”; cf. pp. 71, 76, 143, 207, 211 in G. W. Leibniz: *Philosophical Essays*, ed. and trans. by R. Ariew and D. Garber, Indianapolis – Cambridge 1989 (henceforth cited as AG).

6 “Discourse on Metaphysics”, Loemker, p. 312; “New System”, Loemker, p. 457.

7 “Principles of Nature and of Grace”, Loemker, p. 636.

8 “Monadology”, Loemker, p. 648, § 59; cf. also AG, 148, 195.

9 “Principles of Nature and of Grace”, Loemker, p. 636; “Correspondence with De Volder”, Loemker, p. 537.

10 “Correspondence with De Volder”, Loemker, p. 531.

in the literature, and the contrasts suggest some new questions about the metaphysics of expression. The second goal is to demonstrate in one example how Leibniz's thought may be more naturally at home in formalisms of later twentieth- and early twenty-first-century mathematics than in those of the early twentieth-century, i.e. more in accord with synthetic, categorical, homotopical etc. foundations and frameworks than with analytic and set-theoretic ones. I do not mean to suggest that a categorical formalisation could give a complete or perfect account of what expression is for Leibniz¹¹.

The categorical monadology described here contrasts with what in secondary literature has come to be called the *Structural Representation* (SR) interpretation of expression, according to which one substance expresses another substance when the two have common structure¹². On the SR account, expression (or a given type of expression) is a *relation*. Expression is therefore an abstraction, and indeed one and the same relation of expression may obtain between multiple substances. Furthermore, on this account expression (of given type) is a *binary* relation. Therefore, mutual simultaneous expression is (reading naively off the formalism) a mere aggregation of relations obtaining between various substances expressing each other.

The categorical account provided below will differ from the SR account on each of these points. According to CME, no two monads can share the *same expression* of a third: an expression in CME is always 'of' a particular individual. (Expressions are individuated by the monads expressed 'in' and 'of', as well as other factors.) Further, mutual simultaneous expression is represented in CME in the *joint structure of expressions*. Many natural concepts of category theory, and some of Leibniz's distinctive doctrines, can be written in terms of the structure of mutual expression. For example, in category theory, what it is to be a 'universal object' is naturally described in terms of the joint expression of *all* other objects. Universality statements of this kind are not readily construed in terms of *relations* between substances. My contention is that Leibniz's doctrines, similar to universal properties in mathematics, are readily presented in categorical language that is designed to enable individuation of and quantification over all expressions of all individuals.

These points of difference between SR and CME, framed perhaps as questions about logical and ontological commitments, open new ground

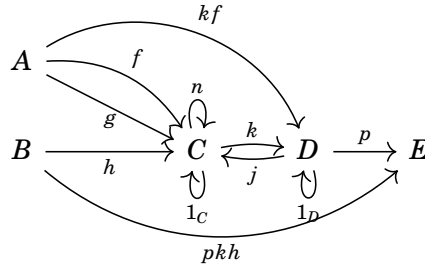
11 On limits of mathematical models of Leibnizean expression, see V. Debuiche: "L'expression leibnizienne et ses modèles mathématiques", in: *Journal of the History of Philosophy* 51/3 (2013), pp. 409–439; V. Debuiche: *Leibniz et l'expression* (= *Épistémè*), Aix-en-Provence 2021.

12 C. Swoyer: "Structural representation and surrogate reasoning", in: *Synthese* 87/3 (1991), pp. 449–508; S. Puryear: "Perception and representation in Leibniz", PhD thesis, University of Pittsburgh, 2006.

for the study of Leibniz's system.

2. A category of monadic expression

A category $\mathcal{C}$ is a class of objects and a class of arrows between the objects with the structural features that the arrows are *composable* and every object has a given *identity* arrow¹³. Composability of arrows means that each time the head of one arrow meets the tail of another in some vertex, a third arrow is provided by a determinate rule. This composition rule should be associative, for example $(pk)h = p(kh)$, and should respect the given identity arrows, for example $1_D k = k = k 1_C$. Here is an example, with some arrows omitted for clarity:



Given f and k as in the diagram, composability determines a unique arrow kf from A to D . The identity 1_C satisfies $1_C h = h$ and $k 1_C = k$. Given reciprocal arrows k and j , j is called the *inverse* of k if it is true that $jk = 1_C$ and $kj = 1_D$. When an arrow has an inverse, the objects it connects are said to be *isomorphic*.

Let us make an observation for later use about the individuation of arrows. It is possible to have distinct arrows between the same objects, so we could have $f \neq g$ and $n \neq 1_C$. It is not possible to have the same arrow between distinct pairs of objects, so we know $g \neq h$.

A category of monadic expression (CME) has monads for objects and expressions of one monad in another for arrows. Considering the diagram above, we might say any of: ' h is an expression of B in C ', ' B is expressed in C by h ', or $B \xrightarrow{h} C$.

In a striking passage Leibniz wrote:

"For the simplicity of a substance does not prevent the plurality of modifications which must necessarily be found together in the same

¹³ For general references see S. Mac Lane: *Categories for the working mathematician* (= Graduate Texts in Mathematics 5), New York ²1998; S. Awodey: *Category theory* (= Oxford Logic Guides 52), Oxford ²2010.

simple substance; and these modifications must consist of the variety of relations of correspondence which the substance has with things outside. In the same way there may be found, in one *center* or point, though it is perfectly simple, an infinity of angles formed by the lines which meet in it”¹⁴.

By considering a category as a graph, the CME model represents monadic expression quite literally as lines meeting in points. We do not attribute any significance to their *angles*, but the arrows’ mutual relationships do constitute the category structure.

When Leibniz says “these modifications must consist of the variety of relations”, we stress the *variety* and take it that this variety should be viewed as more than a mere aggregate, indeed as a structure or whole, for that is implied by the comparison to angles. This stress is, moreover, a requirement of the point he is making: had Leibniz intended a mere aggregate of relations of structural equivalence of various kinds, so that the “plurality of modifications” of one monad is expressed in a similarity of “plurality of modifications” of (various) others, then the variety of relations would not add anything to his statement viewed as an *explanation* of the variety of modifications.

3. Leibnizian thought in a CME

3.1 Distinctness

It is easy to explain *degrees of distinctness* of expression in the language of a CME. Perfectly distinct expression is represented by a *monomorphism*. (This concept generalises ‘injection’ from set-theoretical mathematics.) Thus, $C \xrightarrow{k} D$ is a perfectly distinct expression of C in D iff k ‘separates’ incoming expression:

$$\forall A, f, g, \text{ s.t. } A \xrightarrow{f} C \xrightarrow{k} D \text{ and } A \xrightarrow{g} C \xrightarrow{k} D, f \neq g \implies kf \neq kg.$$

In this way nothing is ‘lost’ by k — nothing, that is, which appears elsewhere in the structure of expression.

We might say an expression $C \xrightarrow{k} D$ is *strictly more distinct than* an expression $C \xrightarrow{\ell} D$ when the classes of expressions that k ‘separates’, in the above sense, strictly include the classes separated by ℓ . This defines a partial order of distinctness. Without introducing further structures, we cannot define a natural total order on distinctness.

¹⁴ “Principles of Nature and of Grace”, Loemker, p. 636.

3.2 Pre-established Harmony

The doctrine of *Pre-established Harmony* has two elements: (i) monads are in *perfect harmony* at any time, (ii) this constant harmony is *pre-established*, being grounded in an initial harmony together with a harmony of rules of appetition. We consider only (i). Let us adopt the following interpretation of (i): perfect harmony of monads means that the modifications of each monad are *completely characterisable* by the total structure of expression of the monad in (all) other monads.

To clarify this just a little: Leibniz considers in contrast to Harmony the possibility that each monad, instead of relating to the others by expressing them, lives in absolute independence, having qualities that are not expressed in other monads. An aggregate of such monads would be *disharmonious*, we might say, in that coperception of the aggregate would not be intelligible in terms of their mutual expression. The extent of a possible *partial* Harmony is the extent to which monads *are* characterised by their mutual expression, and (perfect, pre-established) Harmony obtains when the nature of each monad is exhausted by its *role* in the system of expression. This is one way of interpreting Leibniz's comparison of a monad to a *mirror* that reflects the universe: a monad, like a mirror, has no 'extra' material hidden inside; its nature is exhausted by the totality of how it is expressed in the world. The *situation* of a monad in the category, its *point of view* in the system of expression, is equivalent data to the monad's state.

The principle of Harmony is easily translated to the language of a CME. It becomes the assertion that the nature of every object is exhausted by its role in the category structure of the CME. This translation is apparently not a statement about the structure of the CME, rather about the interpretation of the CME; nonetheless it *does* have parallels in the use of categories in mathematics. Often in mathematics, a category is constructed by defining a class of abstract structures as the objects, and a class of structure-preserving morphisms as the arrows. In this situation the objects might, of course, have internal structure that is not recoverable from the compositional structure of the category itself. But there are other situations where the object structures can be recovered from the category structure. This is the case for the categories of groups and rings. For object structures defined using sets, a very minimal requirement to recover the structure is that one can recover the set; this condition is called *concretisability* of the category.

3.3 Identity of Indiscernibles

The principle of *Identity of Indiscernibles* could be represented in a CME by the absence of pairs of isomorphic yet distinct monads. A category in which isomorphic objects are necessarily identical is called a *skeletal* category. So we might argue that the PII is equivalent to the

proposition that the CME is a skeletal category¹⁵.

Although this is a reasonable proposal, it is not hard to complicate the story. Given two objects and an isomorphism between them, it is easy to set up a perfect correspondence between their ‘situations’ in the category structure, and thus isomorphic monads would be (arguably) indiscernible to other monad perceivers.

The converse is not so simple. It is not necessarily true that one can always find an isomorphism that engenders a given equivalence between the roles played by two objects in a category. Indeed, in order to write down a precise condition for this on the CME structure, we would need first to state precisely what it means for objects to have ‘equivalent situations’. We develop this topic just a little in §3.4¹⁶.

Another type of concern would be the following. The principle of *Identity of Indiscernibles* is strictly stronger than another principle which we might call Identity of Indifferents, the latter banning pairs of objects with identical qualities, and the former banning pairs merely *perceived* to have identical qualities (by any perceiver or a best perceiver)¹⁷. These variants might not agree in general, but if we assume Harmony, then they would seem to collapse to the same principle, since then a monad is completely characterised by the totality of its expression in the others.

3.4 Universal Expression

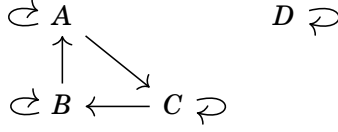
If Pre-established Harmony is the doctrine that there is nothing ‘extra’ in a monad which is not expressed in the world, then *Universal Expression* is the dual doctrine that there is nothing ‘extra’ in the world which is not expressed in every single monad. If the PPH concerns the expression of one in many, then the PUE concerns the expression of many in one.

This principle certainly does not hold for every category. Consider this one:

15 There is a large literature about identity and equivalence in mathematics and in philosophy. In categorical mathematics specifically, one can start with work on Voevodsky’s ‘univalent foundations’. See, for example, B. Ahrens, K. Kapulkin and M. Shulman: “Univalent Categories and the Rezk Completion”, in: *Mathematical Structures in Computer Science* 25/5 (2015), pp. 1010–1039.

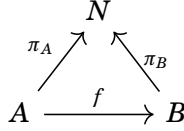
16 As with object-structure recovery discussed in connection with Harmony, this kind of question is associated with a complex history in category theory. It may be related to ‘concrete’ categories, ‘representable’ functors, and the Yoneda Lemma.

17 In distinguishing these I partly follow V. De Risi: *Geometry and monadology: Leibniz’s Analysis situs and philosophy of space* (= Science Networks. Historical Studies 33), Basel 2007, pp. 384–395.



The PUE cannot hold for this (simplistic) world: some of the world (specifically A, B, C) is not expressed by D .

Let us formalise a notion intended to model an aspect of a monad's 'situation'. Define a *view* at an object N in a category $\mathcal{C}$ as a collection of arrows to N (the 'legs'), one from each object in $\mathcal{C}$, such that all triangular diagrams that arise are commutative. For example, if $A \xrightarrow{f} B$ is any expression in the category, and $A \xrightarrow{\pi_A} N, B \xrightarrow{\pi_B} N$ are the legs of the view, this triangle must commute:



Written algebraically, this commutativity is the statement $\pi_A = \pi_B f$, which suggests an interpretation: 'the expression f is contained in the expression π_A of A in N '. Since f was arbitrary in the category, every expression is thus 'part of' an expression in the top object of the view. Thus we can interpret a view as an expression of the whole category, i.e. every other monad and expression, in the given one¹⁸. So the PUE would be associated with a requirement that views exist in the CME. Then the (full) situation of a monad might be described by the complete collection of views afforded in the CME.

3.5 God

Leibniz believed that God is a monad in the system of monads, but with a unique role in that system¹⁹. God's point of view, his situation, should be unique among views in its perfection: every other point of view should be subsumed by it. Leibniz illustrates this by comparing a generic perspective to a scenograph (a perspectival image of a building²⁰) and

¹⁸ This may become clearer with further technical details: a view is a natural transformation $\pi : \text{id}_{\mathcal{C}} \rightarrow \Delta_N$ from the identity functor to the constant functor on N . (The commuting triangles above are the naturality diagrams.) A natural transformation could itself be viewed as a kind of expression in the functor category.

¹⁹ This isn't uncontroversial, but we do not enter the debate here.

²⁰ Leibniz is certainly influenced by the emerging perspective geometry. See V. Debuiche and M. Brancato: "Scientia Perspectiva: Leibniz and Geometric Perspective", in: *Historia Mathematica* 63 (2023), pp. 47–69; V. Debuiche: "Perspective

God's perspective to an ichnograph (a floor plan used by architects)²¹. God's perspective is uniquely perfect. From his point of view, all others may be derived.

The mutually connected categorical concepts of *limit*²² and *universal object* are readily applied to illustrate, or model, how a single monad could be privileged in this way with regard to a system of expression modelled as a category. The limit we need, L , should have a 'universal view' in the mathematical sense of 'universal object' or categorical limit: it is a specific view, $\{A \xrightarrow{\pi_A} L\}_{A \in \mathcal{C}}$, expressing every other monad, but it is unique among views in having the following universal property: Any view $\{A \xrightarrow{\gamma_A} N\}_{A \in \mathcal{C}}$ is determined by composing the view at L with a particular (and determinate) expression of L in N . That is, there exists a unique expression $L \xrightarrow{n} N$ such that $\forall A \in \mathcal{C}, \gamma_A = n\pi_A$. Such an object (a colimit of the identity functor) is called a 'terminal object' (an object to which every other object has exactly one arrow). A very natural interpretation of this characterisation of L is that any point of view is an expression of God's point of view²³, and every expression of God determines a point of view. It is not necessarily the case that there are no arrows from a terminal object to other objects. In the category of sets, every singleton set is a terminal object, and arrows from a terminal object are equivalent to elements of the codomain set²⁴.

Another aspect of Leibniz's view about God as a monad is that God should be able to perceive every aspect of the entire universe in the qualities of each and every monad²⁵. The principle just presented says only that every other monad's point of view is represented in God. But if we are able to establish the PUE, then by conjoining that principle with this fact, we also see that the entire world is present to God in his perception of any single monad. But I admit that my defence of the PUE was perhaps the weakest of those given for the various doctrines²⁶.

in Leibniz's *Invention of Characteristica Geometrica*: The Problem of Desargues' Influence", in: *Historia Mathematica* 40/4 (2013), pp. 359–385.

21 B. Russell: *A Critical Exposition of the Philosophy of Leibniz*, London ²1937, p. 253; GP II, 438.

22 Strictly, a 'colimit'; I use 'limit' loosely for both limits and colimits.

23 "First Truths", Loemker, p. 269.

24 In this way it is possible for a categorical formalism to support expressions of God in finite substances, although there will be complications in the details. For work on the topic of finite monads expressing God (and consequences for the SR view), see S. Duncan: "Leibniz on the Expression of God", in: *Ergo* 2/4 (2015).

25 "Monadology", Loemker, p. 649, § 61.

26 Maunu also gives up this point in his attempt to explain PUE, and claims only that each monad expresses the *structure* of the universe; see A. Maunu: "Leibniz's Theory of Universal Expression Explicated", in: *Canadian Journal of Philosophy* 38/2 (2008), pp. 247–267, esp. p. 258. But I don't believe he can justify even *that*

4. Natural and arbitrary expression

Leibniz applied the concept of expression very widely, including to entities other than monads. Although our CME model is intended to illuminate Leibniz's principal doctrines of monadology, it is worth noticing that we can also apply categorical language to model expression for entities of other types. A central interpretive problem involving expression more generally is to understand Leibniz's distinction between *natural* and *arbitrary* expression. Illustrating the former is a cause and its total effect, and illustrating the latter are the characters of a written language²⁷.

Let us say that natural or direct expression ('expression_n') is expression as modelled above, i.e. as an arrow within a category. Arbitrary expression ('expression_a') is not direct expression *per se* but is similar to it, and arises when the structures of two entities belong to different systems: An entity A expresses_a another entity A' when A 's *role* in its system of expression_n parallels the *role* of A' in its own system, with this correspondence cashed out as an expression_n of A 's system in A 's system. That is to say, arbitrary expression is the natural expression of the system of expression of A' (or that part local to A') in the system of expression of A (or that part local to A); in this way it is an expression of higher order.

Using categorical language, we can model this with the concept of a *functor*, which we explain briefly now. Any category may be considered an object in the category *Cat* of all (small) categories, and an 'arrow' between categories in *Cat* is a functor. A functor $F : \mathcal{C} \rightarrow \mathcal{D}$ is a map from *both* objects and arrows of $\mathcal{C}$ to those of $\mathcal{D}$ preserving categorical structure: (1) arrow end-points; (2) composition, i.e. $F(gf) = F(g)F(f)$ (for all f, g); (3) identity, i.e. $F(1_A) = 1_{F(A)}$ (for all A). A functor models expression of one *system* of expression in another *system*.

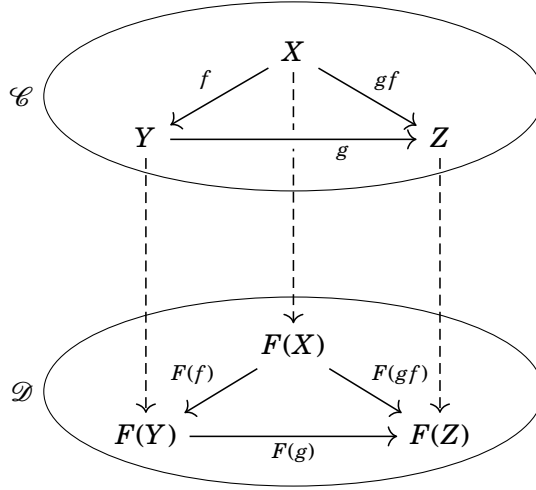
Now assume that A is an object of $\mathcal{C}$, A' of $\mathcal{C}'$, and there is a functor between these categories $F : \mathcal{C}' \rightarrow \mathcal{C}$ such that $F(A') = A$. This functor models expression of $\mathcal{C}'$ in $\mathcal{C}$. Then A' plays the role in the system $\mathcal{C}'$ that A plays in $\mathcal{C}$. So we can say that, through the functor F , A' is expressed_a by A .

Notice that A and A' may be very different kinds of entity since they are in different categories. But there is still a rule by which one can "move in thought"²⁸ from the expression to the expressed: that rule (al-

claim in his framework, since he misses the possibility of a disconnected 'D' as in the first figure of § 3.4.

27 "What Is an Idea?", Loemker, pp. 207–208.

28 Leibniz mentions such a rule in "What Is an Idea?", Loemker, p. 207. Brandom and Swyer emphasise the connection of expression to *reasoning* and *inference* that such a rule constitutes; see R. Brandom: "Leibniz and Degrees of Percep-



A functor $F : \mathcal{C} \rightarrow \mathcal{D}$.

be it backwards) is supplied by the functor F .

A useful illustration of this is an explanation of the way an algebraic equation like

$$x^2 + y^2 = 1^2 \quad (1)$$

expresses a circle²⁹. The fact that Leibniz believed the equation expresses a circle has baffled interpreters³⁰. Using categorical formalism we simply need to establish a functor between the category of algebraic equations and the category of geometric figures³¹. Such a functor would represent an expression_a of geometry in algebra, according to which Eqn. (1) in algebra corresponds to the circle in geometry.

A more general illustration on these lines might be the way a character of language expresses its referent. We would construct a functor (or

tion”, in: *Journal of the History of Philosophy* 19/4 (1981), pp. 447–479, esp. p. 450 ff.; Swoyer: “Structural representation”; C. Swoyer: “Leibnizian Expression”, in: *Journal of the History of Philosophy* 33/1 (1995), pp. 65–99. In CME we could describe this by letting arrows represent inference possibilities.

29 “What Is an Idea?”, Loemker, p. 207.

30 Swoyer conjectured that Leibniz meant the solution set of the equation (ordered pairs) expresses the circle (points in space); see Swoyer: “Leibnizian Expression”, p. 91. Puryear noticed that the equation must also express the solution set, so Swoyer just postponed the problem; Puryear, p. 22. But Puryear could not explain how this happens in his SR framework since no part of the algebraic structure of Eqn. (1) is shared by a circle.

31 The modern field of algebraic geometry involves a project like this.

some generalisation of this concept) from a categorical model of physical objects to a categorical model of language³². This functor would model linguistic expression and would include a map from objects to words³³. In both cases the symbol (word or equation) is arbitrary when considered singly, but it plays a non-arbitrary role in a larger structure.

5. Suppositions of a CME

5.1 Simplicity and individuation

Internal structure in an object of a category is not part of the structure of the abstract category itself. A mantra of category theory is that objects *are* how they *behave* vis-à-vis maps to other objects in the category. In this sense, the objects of a category are simple. Multiplicity, and structure, enters a category in two ways: the multiplicity of objects, and the multiplicity and structure of arrows. The ontology implied by such formalism fits naturally with Leibniz’s ontology. Monads do not have internal structure. There is a multiplicity of monads, and a multiplicity of qualities of monads, but these qualities are constituted by the ways in which monads express other monads; so there is a structured plurality of expressions. The insight of this essay begins with a recognition of this seemingly natural correspondence between monadology and the implied ontology of CME.

The CME model does imply further metaphysical commitments, or at least it suggests ones, that are less obviously compatible with monadology. The most basic of these concerns the individuation of expressions. According to CME, expressions are individuated *at least* by the data of the ordered pair of monads (the expressed and the expressing). It follows that *expression is not a relation*³⁴ on the CME model. From the point of

32 J. Lambek: “The Mathematics of Sentence Structure”, in: *The American Mathematical Monthly* 65/3 (1958), pp. 154–170; R. T. Oehrle, E. Bach and D. Wheeler (eds.): *Categorical Grammars and Natural Language Structures* (= Studies in Linguistics and Philosophy 32), Dordrecht 1988; D. J. Dougherty: “Closed categories and categorial grammar”, in: *Notre Dame Journal of Formal Logic* 34/1 (1993), pp. 36–49; M. Kracht: “Categorial Grammar and Formal Semantics”, in: *The mathematics of language* (= Studies in Generative Grammar 63), Berlin 2003, pp. 177–280.

33 Leibniz emphasises the “complex mutual relation [situm complexum] or order” that the characters of language have when conjoined or inflected; see “Dialogue”, Loemker, p. 184 (= GP VII, 192). Puryear takes this to be in *composite* characters, and applies the standard SR theory to explain their expression; Puryear, pp. 16–23, § 2.2. I think Leibniz means, rather, combination in *phrase* structure because he grants that “single words” might lack such *situs*.

34 Relations are *repeatable*; see F. MacBride: “Relations”, *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), ed. by E. N. Zalta, URL: <https://>

view of a monad B , each instance of expression in this monad is necessarily expression of some *particular* monad, such as A , modelled as an arrow $A \xrightarrow{h} B$. That is, it is the monad A which is expressed in B , and not some set of universals enjoyed by A which is expressed in B .

This feature of CME is not shared by the SR model (and related models) whereby B expresses A when a relation of a certain commonality of structure holds between A and B . By lumping ‘some common structure’ in with other qualities, we have for such models, more generally, the fact that B expresses A when B and A have some composite of qualities in common; i.e. that expression is a *relation of likeness*.

Let us briefly suggest an argument in favour of the metaphysics implied by CME. Leibniz believes that perception is the representation of multiplicity (the external world) in unity (the perceiver); the expression of many in one. This implies a composition of true unities in the representation³⁵. And of course these unities are expressions of the other monads. So an individual expression must be of a monad, and plausibly of a *particular* monad³⁶. We could not have one monad expressing two other monads by one and the same expression.

5.2 Composition and identity

The structural properties of a category, i.e. the axioms, directly suggest further metaphysical commitments about expression that should be

plato.stanford.edu/archives/win2016/entries/relations/ (visited on 25/12/2016).

35 I mean to mimic his argument for simple monads as necessary from the multiplicity of beings, but apply it here within perception. See also “Monadology”, Loemker, p. 649.

36 Maunu tries to work out roughly the *opposite* view, starting from a concern that Adam’s expression *husbandhood to Eve* makes Adam too dependent on Eve (a particular); see A. Maunu: “Leibnizian Soft Reduction of Extrinsic Denominations and Relations”, in: *Synthese* 139/1 (2004), pp. 143–164; Maunu: “Universal Expression”, p. 254. I think he misreads Leibniz, for Leibniz is *not* talking about a particular Adam, but a partial concept of Adam; see “Correspondence with Arnauld”, Loemker, p. 335. A particular Adam, thinks Leibniz, would express the particular Eve, equivalently Eve’s complete description (albeit confusedly), not an incomplete description of Eve. I believe I follow Strawson and Rutherford on this; see P. F. Strawson: *Individuals: An Essay in Descriptive Metaphysics*, London 1964, p. 128; D. Rutherford: *Leibniz and the rational order of nature*, Cambridge 1995, p. 185. Maunu’s project with “monadic-Ramsey-ascriptions” (conjoin all true statements involving monad A and replace other monad names with existentially quantified variables) seems circular to me. He avoids names for other monads but in no way writes relations in terms of incomplete concepts for them, since otherwise his deduction of PUE fails. Also, if it seems hard to believe that a monad expresses the infinite detail of another, as required to express its particularity, recall that the whole universe is supposed to be deducible from each monad’s complete concept.

assessed against Leibniz's philosophy. These include, principally, compositionality and identities.

Compositionality in the CME means that if A is expressed in B , and B in C , then *ipso facto* A is expressed in C (and by a determinate expression which factors through B). It means that *every* instance of a monad C expressing another monad B involves an instance of C expressing *every* monad A which is expressed in B .

Existence of identities in the CME means that every object expresses itself perfectly. This fact may seem trivial or suspicious. Formally speaking, the role of the identity arrow in an abstract category is to allow us to notate the situation of perfect expression of one object in *other* objects: one forms a loop of arrows ending at the object from which it started, and asserts equality of the composite arrow with the identity on that object. In a CME, this assertion would mean that the first arrow along the loop is a 'perfectly distinct expression' suffering no loss of information.

5.3 Plurality of expressions

Another supposition, or rather presupposition, that we have not squarely addressed is the fact that there may be a plurality of expressions of a given monad A in another B . This is very important in the abstract category structure; some expressions will express relatively little information, others will express more. In mathematics, some categories have a formal notion of a 'zero' arrow, a kind of trivial arrow that always exists yet plays an important role in the system with other arrows.

In more metaphysical terms, this supposition says that expressions are not *only* individuated by the expressed and the expressing monads. They are *also* individuated by some further features. Although Leibniz does not speak explicitly to this question (to my knowledge), it may be reasonable to suppose that he would affirm the plurality of expressions between two given monads because the variety of expressions is supposed to reflect the variety of qualities possessed by monads.

6. Discussion

6.1 Structural representation

The Structural Representation theory of Leibnizean expression has been developed by Kulstad, McRae, Swoyer, Puryear, and De Risi³⁷. Kul-

37 M. Kulstad: "Leibniz's Expression Thesis", PhD thesis, University of Michigan, 1975; R. F. McRae: *Leibniz: Perception, Apperception, and Thought*, Toronto 1976; M. A. Kulstad: "Leibniz's Conception of Expression", in: *Studia Leibnitiana* 9/1 (1977), pp. 55–76; Swoyer: "Structural representation"; Swoyer: "Leibnizian Expression"; Puryear; De Risi, p. 321 and pp. 429–436.

stad's idea was that expression can be modelled as a function between sets; McRae and then Swoyer added a requirement that the function should preserve some *structure* (the type of structure understood from context), and so expression is modelled by a structure-preserving function, which they identify as an isomorphism. Puryear applied SR to interpret Leibniz's expression concept for a variety of types of entity and structure. De Risi developed a more complex treatment of the *algebra of expression* with homomorphisms instead of isomorphisms, and used the kernel in Noether's homomorphism theorem to model the part of a monad's qualities which are not expressed distinctly³⁸. Undoubtedly Leibniz did understand something close to the modern mathematical concept of isomorphism, albeit without the language of functions, and used it to illustrate expression and representation³⁹. I should not be construed as arguing that the SR view is completely misguided. What, then, ought to be taken away from this proposal of a categorical model?

6.2 Comparison

The principal advantage of the CME model of expression over variants of the SR model is its suitability for application to the monadology. In CME one is committed to simple entities with a plurality of modifications that are manifested as a variety of expressions of other monads. In SR models the entities are necessarily not simple, since expression is defined as a relation of similarity of the internal structure of parts.

Similarly, the CME model also fares better with regard to its capacity to represent Leibniz's distinctive doctrines of the monadology. It is very hard to represent these doctrines in a serious way using an SR model, whereas we have seen fairly specific proposals for doing so in a CME model. The point is that CME treats the whole system of expression, whereas SR treats only the question of what it is for one entity to express another.

One conclusion we may draw, then, is the observation that Leibniz's own use of the concept of expression appears to be subject to a duality of interpretation⁴⁰. For some purposes, his discussion seems to focus on a single instance (one entity by another), and what expression means

38 This can be captured even more naturally in CME using a *factorisation system*; see J. Adámek, H. Herrlich and G. E. Strecker: *Abstract and Concrete Categories: The Joy of Cats* (= Pure and Applied Mathematics), New York 1990, ch. IV.

39 K. Stiegler: "Der Begriff des Isomorphismus und der Darstellung in der Metaphysik von Leibniz", in: *Akten des II. Internationalen Leibniz-Kongresses. Hannover, 17.–22. Juli 1972*, ed. by K. Müller, H. Schepers and W. Totok, vol. IV (= *Studia Leibnitiana, Supplementa XV*), Wiesbaden 1975, pp. 173–186.

40 See also L. E. Herrera Castillo: "La exigencia del cuerpo. Perspectivas sobre la expresión", in: *Logos. Anales del Seminario de Metafísica* 54/2 (2021), pp. 297–311.

in such a case; and here the entities are not monads. But the concept of expression is also critically important in his monadology, and here a categorical model seems much more suitable. So we find that a new interpretive question is opened. Did Leibniz recognise this duality? Did he recognise a tension stemming from the duality? Is there a Leibnizean reconciliation of the apparent tension? Or at least, a formal reconciliation that is compatible with Leibniz's philosophy?

6.3 Open questions

The CME model suggests new questions about Leibniz's system of monadic expression.

1. *Individuation and plurality for expression.* The CME model contrasts with the SR model on the question of whether expressions are universals, i.e. whether the same expression can obtain between multiple pairs of distinct monads. (Conversely: whether the expressing and expressed monads separate expressions.) Furthermore, the CME model implies that expressions are even *more* finely individuated than by the monad pairs they are associated with. Are these features compatible with Leibniz's philosophy? Are they determined by it?
2. *Compositionality and identity for expression.* The CME model implies a structure of composition of expressions with identities. Would Leibniz accept these structures? Are they implied by his philosophy? Are they truly compatible with it? I hope that exegetical work will bring some evidence to bear on these questions.
3. *Points of view.* How, more precisely, should we formalise *situations* or *points of view* in a CME? The sketch provided above is only an initial attempt. Both the PII and the PUE as we developed them would be strengthened significantly by a more detailed construction of the concept of situation, and by more rigorous proofs of the principles from that.

Perhaps of most interest is the fact that new questions arise naturally in the attempt to formalise Leibniz's concept of expression in the language of categories.