How Analytic Philosophy Has Failed Cognitive Science

April 13, 2008

Contents

I. Introduction

We analytic philosophers have signally failed our colleagues in cognitive science. We have done that by not sharing central lessons about the nature of concepts, concept-use, and conceptual content that have been entrusted to our care and feeding for more than a century.

I take it that analytic philosophy began with the birth of the new logic that Gottlob Frege introduced in his seminal 1879 *Begriffsschrift*. The idea, taken up and championed to begin with by Bertrand Russell, was that the fundamental insights and tools Frege made available there, and developed and deployed through the 1890s, could be applied throughout philosophy to advance our understanding of understanding and of thought in general, by advancing our understanding of concepts - including the particular concepts with which the philosophical tradition had wrestled since its inception. For Frege brought about a revolution not just in *logic*, but in *semantics*. He made possible for the first time a *mathematical* characterization of meaning and conceptual content, and so of the structure of sapience itself. Henceforth it was to be the business of the new movement of analytic philosophy to explore and amplify those ideas, to exploit and apply them wherever they could do the most good. Those ideas are the cultural birthright, heritage, and responsibility of analytic philosophers. But we have not done right by them. For we have failed to communicate some of the most basic of those ideas, failed to explain their significance, failed to make them available in forms usable by those working in allied disciplines who are also professionally concerned to understand the nature of thought, minds, and reason.

Contemporary cognitive science is a house with many mansions. The provinces I mean particularly to be addressing are cognitive psychology, developmental psychology, animal psychology (especially primatology), and artificial intelligence. (To be sure, this is not all of cognitive science. But the points I will be making in this paper are not of similarly immediate significance for such other subfields as neurophysiology, linguistics, perceptual psychology, learning theory, and the study of the mechanisms of memory.) Cognitive psychology aims at reverse-engineering the human mind: figuring out how we do what we do, what more basic abilities are recruited and deployed (and how) so as to result in the higher cognitive abilities we actually display. Developmental psychology investigates the sequence of stages by which those abilities emerge from more primitive versions as individual humans mature. Animal psychology, as I am construing it, is a sort of combination of cognitive psychology of non-human intelligences and a phylogenetic version of ontogenetic human developmental psychology. By contrast to all these empirical inquiries into actual cognition, artificial intelligence swings free of questions about how any actual organisms do what they do, and asks instead what constellation of abilities of the sort we know how to implement in artifacts might in principle yield sapience.

Each of these disciplines is in its own way concerned with the empirical question of how the trick of cognition is or might be done. Philosophers are concerned with the normative question of what counts as

www.alhassanain.org/english

doing it - with what understanding, particularly discursive, conceptual understanding consists in, rather than how creatures with a particular contingent constitution, history, and armamentarium of basic abilities come to exhibit it. I think Frege taught us three fundamental lessons about the structure of concepts, and hence about all possible abilities that deserve to count as concept-using abilities.[1] The conclusion we should draw from his discoveries is that concept-use is intrinsically stratified. It exhibits at least four basic layers, with each capacity to deploy concepts in a more sophisticated sense of 'concept' structurally presupposing the capacities to use concepts in all of the more primitive senses. The three lessons that generate the structural hierarchy oblige us to distinguish between:

concepts that only *label* and concepts that *describe*,

the *content* of concepts and the *force* of applying them, and

concepts expressible already by *simple* predicates and concepts expressible only by *complex* predicates.

AI researchers and cognitive, developmental, and animal psychologists need to take account of the different grades of conceptual content made visible by these distinctions, both in order to be clear about the topic they are investigating (if they are to tell us how the trick is done, they must be clear about exactly which trick it is) and because the empirical and in-principle possibilities are constrained by the way the abilities to deploy concepts in these various senses structurally presuppose the others that appear earlier in the sequence. This is a point they have long appreciated on the side of basic *syntactic* complexity. But the at least equally important - and I would argue more conceptually fundamental - hierarchy of *semantic* complexity has been largely ignored.

II. First Distinction: From Labeling to Describing

The Early Modern philosophical tradition was built around a *classificatory* theory of consciousness and (hence) of concepts, in part the result of what its scholastic predecessors had made of their central notion of Aristotelian forms. The paradigmatic cognitive act is understood as classifying: taking something particular *as* being of some general kind. Concepts are identified with those general kinds.

This conception was enshrined in the order of logical explanation (originating in Aristotle's *Prior Analytics*) that was common to *everyone* thinking about concepts and consciousness in the period leading up to Kant. At its base is a doctrine of *terms* or *concepts*, particular and general. The next layer, erected on that base, is a doctrine of *judgments*, describing the kinds of classificatory relations that are possible among such terms. For instance, besides classifying Socrates as human, humans can be classified as mortal. Finally, in terms of those metaclassifications grouping judgments into kinds according to the sorts of terms they relate, a doctrine of *consequences* or *syllogisms* is propounded, classifying valid inferences into kinds, depending on which classes of classificatory judgments their premises and conclusions fall under.

It is the master-idea of classification that gives this traditional order of explanation its distinctive shape. That idea defines its base, the relation between its layers, and the theoretical aspiration that animates the whole line of thought: finding suitable ways of classifying terms and judgments (classifiers and classifications) so as to be able to classify inferences as good or bad solely in virtue of the kinds of classifications they involve. The fundamental metaconceptual role it plays in structuring philosophical thought about thought evidently made understanding the concept of classifying itself a particularly urgent philosophical task. Besides asking what differentiates various kinds of classifying, we can ask what they have in common. What is it one must *do* in order thereby to count as *classifying* something as being of some kind?

In the most general sense, one classifies something simply by responding to it differentially. Stimuli are grouped into kinds by the response-kinds they tend to elicit. In this sense, a chunk of iron classifies its environments into kinds by rusting in some of them and not others, increasing or decreasing its temperature, shattering or remaining intact. As is evident from this example, if classifying is just exercising a reliable differential responsive disposition, it is a ubiquitous feature of the inanimate world. For that very reason, classifying in this generic sense is not an attractive candidate for identification with conceptual, cognitive, or conscious activity. It doesn't draw the right line between thinking and all sorts of thoughtless activities. Pan-psychism is too high a price to pay for cognitive naturalism.

That need not mean that taking *differential responsiveness* as the genus of which *conceptual classification* is a species is a bad idea, however. A favorite idea of the classical British empiricists was to require that the classifying response be entering a *sentient* state. The intrinsic characters of these sentient states are supposed to sort them immediately into repeatable kinds. These are called on to function as the *particular* terms in the base

level of the neo-Aristotelian logical hierarchy. *General* terms or concepts are then thought of as sentient state-kinds derived from the particular sentient state-kinds by a process of *abstraction*: grouping the base-level sentient state-repeatables into higher-level sentient state-repeatables by some sort of perceived *similarity*. This abstractive grouping by similarity is itself a kind of classification. The result is a path from one sort of consciousness, sentience, to a conception of another sort of consciousness, sapience, or conceptual consciousness.

A standing felt difficulty with this empiricist strategy is the problem of giving a suitably naturalistic account of the notion of sentient awareness on which it relies. Recent information-theoretic accounts of representation (under which heading I include not just Fred Dretske's theory, which actually goes by that name, but others such as Jerry Fodor's asymmetric counterfactual dependence and nomological locking models[2]) develop the same basic differential responsiveness version of the classic classificatory idea in wholly naturalistic modal terms. They focus on the information conveyed about stimuli - the way they are grouped into repeatables - by their reliably eliciting a response of one rather than another repeatable response-kind from some system. In this setting, unpalatable pan-psychism can be avoided not, as with traditional empiricism, by insisting that the responses be sentient states, but for instance by restricting attention to flexible systems, capable in principle of coming to encode many different groupings of stimuli, with a process of *learning* determining what classificatory dispositions each one actually acquires. (The classical American pragmatists' program for a naturalistic empiricism had at its core the idea that the structure common to evolutionary development and individual learning is a Test-Operate-Test-Exit negative feedback process of acquiring practical habits, including discriminative ones.[3])

Classification as the exercise of reliable differential responsive dispositions (however acquired) is not by itself yet a good candidate for *conceptual* classification, in the basic sense in which applying a concept to something is *describing* it. Why not? Suppose one were given a wand, and told that the light on the handle would go on if and only if what the wand was pointed at had the property of being *grivey*. One might then determine empirically that speakers are grivey, but microphones not, doorknobs are but windowshades are not, cats are and dogs are not, and so on. One is then in a position reliably, perhaps even infallibly, to apply the *label* 'grivey'. Is one also in a position to *describe* things *as* grivey? Ought what one is doing to qualify as applying the *concept* grivey to things? Intuitively, the trouble is that one does not know what one has found out when one has found out that something is grivey, does not know what one is taking it to be when one takes it to be grivey, does not know what one is describing it *as*. The label is, we want to say, uninformative.

What more is required? Wilfrid Sellars gives this succinct, and I believe correct, answer:

It is only because the expressions in terms of which we describe objects, even such basic expressions as words for the perceptible characteristics of

molar objects, locate these objects in a space of implications, that they describe at all, rather than merely label.[4]

The reason 'grivey' is merely a *label*, that it classifies without informing, is that nothing *follows* from so classifying an object. If I discover that all the boxes in the attic I am charged with cleaning out have been labeled with red, yellow, or green stickers, all I learn is that those labeled with the same color share *some* property. To learn what they *mean* is to learn, for instance, that the owner put a red label on boxes to be discarded, green on those to be retained, and yellow on those that needed further sorting and decision. Once I know what *follows* from affixing one rather than another label, I can understand them not as *mere* labels, but as *descriptions* of the boxes to which they are applied. Description is classification with *consequences*, either immediately practical ("to be discarded/examined/kept") or for further classifications.

Michael Dummett argues generally that to be understood as conceptually contentful, expressions must have not only *circumstances* of appropriate application, but also appropriate *consequences* of application.[5] That is, one must look not only *upstream*, to the circumstances (inferential and non-inferential) in which it is appropriate to apply the expression, but also *downstream* to the consequences (inferential and non-inferential) of doing so, in order to grasp the content it expresses. One-sided theories of meaning, which seize on one aspect to the exclusion of the other, are bound to be defective, for they omit aspects of the use that are essential to meaning. For instance, expressions can have the same circumstances of application, and different consequences of application. When they do, they will have different descriptive content.

1] I will write a book about Hegel,

and

2] I foresee that I will write a book about Hegel,

say different things about the world, describe it as being different ways. The first describes my future activity and accomplishment, the second my present aspiration. Yet the circumstances under which it is appropriate or warranted to assert them - the situations to which I ought reliably to respond by endorsing them - are the same (or at least, can be made so by light regimentation of a prediction-expressing use of 'foresee'). Here, to say that they have different descriptive content can be put by saying that they have different truth conditions. (That they have the same assertibility conditions just shows how assertibility theories of meaning, as one-sided in Dummett's sense, go wrong.) But that same fact shows up in the different positions they occupy in the "space of implications." For from the former it follows that I will not be immediately struck by lightning, that I will write some book, and, indeed, that I will write a book about Hegel. None of these is in the same sense a consequence of the second claim.

We might train a parrot reliably to respond differentially to the visible presence of red things by squawking "That's red." It would not yet be *describing* things as red, would not be applying the concept red to them, because the noise it makes has no significance for it. It does not know that it follows from something's being red that it is colored, that it cannot be

wholly green, and so on. Ignorant as it is of those inferential consequences, the parrot does not grasp the concept (any more than we express a concept by 'grivey'). The lesson is that even observational concepts, whose principal circumstances of appropriate application are non-inferential (a matter of reliable dispositions to respond differentially to non-linguistic stimuli) must have inferential consequences in order to make possible description, as opposed to the sort of classification effected by non-conceptual labels.

The rationalist idea that the inferential significance of a state or expression is essential to its *conceptual* contentfulness is one of the central insights of Frege's 1879 *Begriffsschrift* ("concept writing") - the founding document of modern logic and semantics - and is appealed to by him in the opening paragraphs to define his topic:

...there are two ways in which the content of two judgments may differ; it may, or it may not, be the case that all inferences that can be drawn from the first judgment when combined with certain other ones can always also be drawn from the second when combined with the same other judgments…I call that part of the content that is the same in both the conceptual content [begriffliche Inhalt]. [6]

Here, then, is the first lesson that analytic philosophy ought to have taught cognitive science: there is a fundamental conceptual distinction between classification in the sense of *labeling* and classification in the sense of *describing*, and it consists in the *inferential* consequences of the classification: its capacity to serve as a premise in inferences (practical or theoretical) to further conclusions. (Indeed, there are descriptive concepts that are purely *theoretical* - such as gene and quark - in the sense that in addition to their inferential consequences of application, they have *only* inferential *circumstances* of application.) There is probably no point in fighting over the minimal circumstances of application of the concepts concept and conceptual. Those who wish to lower the bar sufficiently are welcome to consider purely classificatory labels as a kind of concept (perhaps so as not to be beastly to the beasts, or disqualify human infants, bits of our brains, or even some relatively complex computer programs wholly from engaging in conceptually articulated activities). But *if* they do so, they must *not* combine those circumstances of application with the consequences of application appropriate to genuinely *descriptive* concepts - those that *do* come with inferential significances downstream from their application.

Notice that this distinction between labeling and describing is untouched by two sorts of elaborations of the notion of labeling that have often been taken to be of great significance in thinking about concepts from the classical classificatory point of view. One does not cross the boundary from labeling to describing just because the reliable capacity to respond differentially is *learned*, and in that sense flexible, rather than *innate*, and in that sense rigid. And one is likewise developing the classical model in an orthogonal direction insofar as one focuses on the metacapacity to learn to distinguish arbitrary Boolean combinations of microfeatures one can already reliably discriminate. From the point of view of the distinction between

labeling and describing, that is not yet the capacity to form *concepts*, but only the mastery of *compound* labels. That sort of structural articulation upstream has no *semantic* import at the level of description until and unless it is accorded a corresponding inferential significance downstream.

III: Ingredient vs. Free-Standing Content: Semantically Separating Content from Force

Once our attention has been directed at the significance of applying a classifying concept - downstream, at the consequences of applying it, rather than just upstream, at the repeatable it discriminates, the grouping it institutes - so that *mere* classification is properly distinguished from *descriptive* classification, the necessity of distinguishing different *kinds* of consequence becomes apparent. One distinction in the vicinity, which has already been mentioned in passing, is that between *practical* and *theoretical* (or, better, *cognitive*) consequences of application of a concept. The significance of classifying an object by responding to it one way rather than another may be to make it appropriate to *do* something else with or to it - to keep it, examine it, or throw it away, to flee or pursue or consume it, for example. This is still a matter of inference; in this case, it is *practical* inferences that are at issue. But an initial classification may also contribute to further classifications: that what is in my hand falls under both the classifications raspberry and red makes it appropriate to classify it also as ripe - which in turn has practical consequences of application (such as, under the right circumstances "falling to without further ado and eating it up," as Hegel says in another connection) that neither of the other classifications has individually. Important as the distinction between practical and cognitive inferential consequences is, in the present context there is reason to emphasize a different one.

Discursive intentional phenomena (and their associated concepts), such as assertion, inference, judgment, experience, representation, perception, action, endorsement, and imagination typically involve what Sellars calls "the notorious 'ing'/'ed' ambiguity." For under these headings we may be talking about the *act* of asserting, inferring, judging, experiencing, representing, perceiving, doing, endorsing, and imagining, or we may be talking about the *content* that is asserted, inferred, judged, experienced, represented, perceived, done, endorsed, or imagined. 'Description' is one of these ambiguous terms (as is 'classification'). We ought to be aware of the distinction between the act of describing (or classifying), applying a concept, on the one hand, and the content of the description (classification, concept) - *how* things are described (classified, conceived) - on the other. And the distinction is not merely of theoretical importance for those of us thinking systematically about concept use. A distinctive level of conceptual sophistication is achieved by concept users that themselves distinguish between the contents of their concepts and their activity of applying them. So one thing we might want to know about a system being studied, a non-human animal, a prelinguistic human, an artifact we are building, is whether *it* distinguishes between the *concept* it applies and what it *does* by applying it.

We can see a basic version of the distinction between semantic content and pragmatic force as in play wherever *different* kinds of practical significance can be invested in the *same* descriptive content (different sorts of speech act or mental act performed using that content). Thus if a creature can not only say or think that the door is shut, but also ask or wonder

whether the door is shut, or order or request that it be shut, we can see it as distinguishing in practice between the content being expressed and the pragmatic force being attached to it. In effect, it can use descriptive contents to do more than merely describe. But this sort of practical distinguishing of pragmatic from semantic components matters for the semantic hierarchy I am describing only when it is incorporated or reflected in the *concepts* (that is, the *contents*) a creature can deploy. The capacity to attach different sorts of pragmatic force to the same semantic content is not sufficient for *this* advance in structural semantic complexity. (Whether it is a necessary condition is a question I will not address - though I am inclined to think that in principle the answer is 'No'.)

For the inferential consequences of applying a classificatory concept, when doing that is describing and not merely labeling, can be either *semantic* consequences, which turn on the *content* of the concept being applied, or *pragmatic* consequences, which turn on the *act* one is performing in applying it. Suppose John issues an observation report: "The traffic light is red." You may infer that it is operating and illuminated, and that traffic ought to stop in the direction it governs. You may also infer that John has a visually unobstructed line of sight to the light, notices what color it is, and believes that it is red. Unlike the former inferences, these are not inferences from what John *said*, from the *content* of his utterance, from the concepts he has applied. They are inferences from his *saying* it, from the pragmatic force or significance of his *uttering* it, from the fact of his *applying* those concepts. For what he has *said*, that the traffic light is red, could be true even if John had not been in a position to notice it or form any beliefs about it. Nothing about John follows just from the color of the traffic light.[7]

It can be controversial whether a particular consequence follows from how something is described or from describing it that way, that is, whether that consequence is part of the descriptive content of an expression, the concept applied, or stems rather from the force of using the expression, from applying the concept. A famous example is expressivist theories of evaluative terms such as 'good'. In their most extreme form, they claim that these terms have no descriptive content. *All* their consequences stem from what one is doing in using them: commending, endorsing, or approving. In his lapidary article "Ascriptivism,"[8] Peter Geach asks what the rules governing this move are. He offers the archaic term 'macarize', meaning to characterize someone as happy. Should we say that in apparently describing someone as happy we are not really describing anyone, but rather performing the distinctive speech act of macarizing? But why not then discern distinctive speech acts for *any* apparently descriptive term?

What is wanted is a criterion for distinguishing semantic from pragmatic consequences, those that stem from the content of the concept being applied from those that stem from what we are doing in applying that concept (using an expression to perform a speech act). Geach finds one in Frege, who in turn was developing a point made already by Kant.[9] The logical tradition Kant inherited was built around the classificatory theory of consciousness we began by considering. Judgment was understood as

classification or predication: paradigmatically, *of* something particular *as* something general. But we have put ourselves in a position to ask: is this intended as a model of judgeable contents are constructed, or of what one is doing in judging? Kant saw, as Frege would see after him, that the phenomenon of *compound* judgments shows that it can*not* play both roles. For consider the hypothetical or conditional judgment

3] If Frege is correct, then conceptual content depends on inferential consequences.

In asserting this sentence (endorsing its content), have I predicated correctness of Frege (classified him as correct)? Have I described him as correct? Have I applied the concept of correctness? If so, then predicating or classifying (or describing) is not judging. For in asserting the conditional I have *not* judged or asserted that Frege is correct. I have at most built up a judgeable content, the antecedent of the conditional, by predication. For embedding a declarative descriptive sentence as an unasserted component in a compound asserted sentence strips off the pragmatic force its free-standing, unembedded occurrence would otherwise have had. It now contributes *only* its *content* to the *content* of the compound sentence, to which alone the pragmatic force of a speech act is attached.

This means that embedding simpler sentences as components of compound sentences - paradigmatically, embedding them as antecedents of conditionals - is the way to discriminate consequences that derive from the *content* of a sentence from consequences that derive from the *act* of asserting or endorsing it. We can tell that 'happy' *does* express descriptive content, and is *not* simply an indicator that some utterance has the pragmatic force or significance of macarizing, because we *can* say things like:

4] If she is happy, then John should be glad.

For in asserting that, one does *not* macarize anyone. So the consequence, that John should be glad, must be due to the descriptive content of the antecedent, not to its force.

Similarly, Geach argues that the fact that we can say things like:

5] If being trustworthy is good, then you have reason to be trustworthy,

shows that 'good' *does* have descriptive content.[10] Notice that this same test appropriately discriminates the different descriptive contents of the claims:

6] Labeling is not describing,

and

7] I believe that labeling is not describing.

For the two do not behave the same way as antecedents of conditionals. The stuttering inference

8] If labeling is not describing, then labeling is not describing,

is as solid an inference as one could ask for. The corresponding conditional

9] If I believe that labeling is not describing, then labeling is not describing,

requires a good deal more faith to endorse. And in the same way, the embedding test distinguishes [1] and [2] above. In each case it tells us, properly, that different descriptive contents are involved.

What all this means is that any user of descriptive concepts who can also form compound sentences, paradigmatically conditionals, is in a position to distinguish what pertains to the semantic *content* of those descriptive concepts from what pertains to the *act* or pragmatic *force* of describing by applying those concepts. This capacity is a new, higher, more sophisticated level of concept use. It can be achieved *only* by looking at compound sentences in which other descriptive sentences can occur as unasserted components. For instance, it is only in such a context that one can distinguish *denial* (a kind of speech act or attitude) from *negation* (a kind of content). One who asserts [6] has *both* denied that labeling is describing, *and* negated a description. But one who asserts conditionals such as [8] and [9] has negated descriptions, but has *not* denied anything.

The modern philosophical tradition up to Frege took it for granted that there was an special attitude on could adopt towards a descriptive conceptual content, a kind of minimal force one could invest it with, that must be possible independently of and antecedent to being able to endorse that content in a judgment. This is the attitude of merely *entertaining* the description. The picture (for instance, in Descartes) was that *first* one entertained descriptive thoughts (judgeables), and *then*, by an in-principle subsequent act of will, accepted or rejected it. Frege rejects this picture. The principal - and in principle fundamental - pragmatic attitude (and hence speech act) is judging or endorsing.[11] The capacity merely to entertain a proposition (judgeable content, description) is a late-coming capacity - one that is parasitic on the capacity to endorse such contents. In fact, for Frege, the capacity to entertain (without endorsement) the proposition that *p* is just the capacity to endorse *conditionals* in which that proposition occurs as antecedent or consequent. For that is to explore its descriptive content, its inferential circumstances and consequences of application, what it follows from and what follows from it, what would make it true and what would be true if it were true, without endorsing it. This is a new kind of distanced attitude toward one's concepts and their contents - one that becomes possible only in virtue of the capacity to form compound sentences of the kind of which conditionals are the paradigm. It is a new level of cognitive achievement - not in the sense of a new kind of empirical knowledge (though conditionals can indeed codify new empirical discoveries), but of a new kind of semantic self-consciousness.

Conditionals make possible a new sort of hypothetical thought. (Supposing that postulating a distinct attitude of supposing would enable one to do *this* work, the work of conditionals, would be making the same mistake as thinking that denial can do the work of negation.) Descriptive concepts bring empirical properties into view. Embedding those concepts in conditionals brings the contents of those concepts into view. Creatures that can do that are functioning at a higher cognitive and conceptual level than those who can only apply descriptive concepts, just as those who can do that are functioning at a higher cognitive and conceptual level than those who can only classify things by reliable responsive discrimination (that is, labeling). That fact sets a question for the different branches of cognitive

www.alhassanain.org/english

science I mentioned in my introduction. Can chimps, or African grey parrots, or other non-human animals not just use concepts to describe things, but also semantically discriminate the contents of those concepts from the force of applying them, by using them not just in describing, but in conditionals, in which their contents are merely entertained and explored? At what age, and along with what other capacities, do human children learn to do so? What is required for a computer to demonstrate this level of cognitive functioning?

Conditionals are special, because they make *inferences* explicit - that is, put them into endorsable, judgeable, assertible, which is to say propositional form. And it is their role in inferences, we saw, that distinguishes descriptive concepts from mere classifying labels. But conditionals are an instance of a more general phenomenon. For we can think of them as operators, which apply to sentences to yield further sentences. As such, they bring into view a new notion of conceptual content: a new principle of assimilation, hence classification, of such contents. For we begin with the idea of sameness of content that derives from sameness of pragmatic force, attitude, or speech act. But the Frege-Geach argument shows that we can also individuate conceptual contents more finely, not just in terms of their role in free-standing utterances, but also accordingly as substituting one for another as arguments of operators (paradigmatically the conditional) does or does not yield compound sentences with the same free-standing pragmatic significance or force. Dummett calls these notions "free-standing" and "ingredient" content (or sense), respectively. Thus we might think that

10] It is nice here,

and

11] It is nice where I am,

express the same attitude, perform the same speech act, have the same pragmatic force or significance. They not only have the same circumstances of application, but the same consequences of application (and hence role as antecedents of conditionals). But we can see that they have different *ingredient* contents by seeing that they behave differently as arguments when we apply another operator to them. To use an example of Dummett's,

12] It is *always* nice here,

and

13] It is *always* nice where I am,

have very different circumstances and consequences of application, different pragmatic significances, and *do* behave differently as the antecedents of conditionals. But this difference in content, this sense of "different content" in which they patently do have different contents, is one that shows up *only* in the context of compounding operators, which apply to sentences and yield further sentences. The capacity to deploy such operators to form new conceptual (descriptive) contents from old ones accordingly ushers in a new level of cognitive and conceptual functioning.

Creatures that can not merely label, but describe are *rational*, in the minimal sense that they are able to treat one classification as providing a *reason* for or against another. If they can use conditionals, they can

distinguish inferences that depend on the *content* of the concept they are applying from those that depend on what they are *doing* in classifying something as falling under that concept. But the capacity to use conditionals gives them more than just that ability. For conditionals let them *say* what is a reason for what, say *that* an inference is a good one. And for anyone who can do that, the capacity not just to *deny* that a classification is appropriate, but to use a *negation* operator to form new classificatory contents means brings with it the capacity to say that two classifications (classifiers, concepts) are incompatible: that one provides a reason to withhold the other. Creatures that can use this sort of sentential compounding operator are not just *rational*, but *logical* creatures. They are capable of a distinctive kind of *conceptual self-consciousness.* For they can describe the rational relations that make their classifications into descriptions in the first place, hence be conscious or aware of them in the sense in which descriptive concepts allow them to be aware of empirical features of their world.

IV. Simple versus Complex Predicates

There is still a higher level of structural complexity of concepts and concept use. I have claimed that Frege should be credited with appreciating both of the points I have made so far: that descriptive conceptual classification beyond mere discriminative labeling depends on the inferential significance of the concepts, and that semantically distinguishing the inferential significance of the contents of concepts from that of the force of applying them depends on forming sentential compounds (paradigmatically conditionals) in which other sentences appear as components. In each of these insights Frege had predecessors. Leibniz (in his *New Essay on the Human Understanding*) had already argued the first point, against Locke. (The move from thinking of concepts exclusively as reliably differentially elicited labels to thinking of them as having to stand in the sort of inferential relations to one another necessary for them to have genuine descriptive content is characteristic of the advance from empiricism to rationalism.) And Kant, we have seen, appreciated how attention to compound sentences (including "hypotheticals") requires substantially amending the traditional classificatory theory of conceptual consciousness. The final distinction I will discuss, that between *simple* and *complex* predicates, and the corresponding kinds of concepts they express, is Frege's alone. No-one before him (and embarrassingly few even of his admirers after him) grasped this idea.

Frege's most famous achievement is transforming traditional logic by giving us a systematic way to express and control the inferential roles of *quantificationally complex* sentences. Frege could, as the whole logical tradition from Aristotle down to his time (fixated as it was on syllogisms) could not, handle iterated quantifiers. So he could, for instance, explain why

14] If someone is loved by everyone, then everyone loves someone,

is true (a conditional that codifies a correct inference), but

15] If everyone loves someone, then someone is loved by everyone,

is not. What is less appreciated is that in order to specify the inferences involving arbitrarily nested quantifiers ('some' and 'every'), he needed to introduce a new kind of predicate, and hence discern a structurally new kind of *concept*.

Our first grip on the notion of a predicate is as a *component* of sentences. In artificial languages we combine, for instance, a two-place predicate 'P' with two individual constants 'a' and 'b' to form the sentence 'Pab'. Logically minded philosophers of language use this model to think about the corresponding sentences of natural languages, understanding

16] Kant admired Rousseau,

as formed by applying the two-place predicate 'admired' to the singular terms 'Kant' and 'Rousseau'. The kind of inferences that are made explicit by *quantified conditionals* - inferences that essentially depend on the contents of the predicates involved - though, require us also to distinguish a

www.alhassanain.org/english

one-place predicate, related to but distinct from this two-place one, that is exhibited by

17] Rousseau admired Rousseau,

and

18] Kant admired Kant,

but *not* by [16].

19] Someone admired himself,

that is, something of the form ∃x[Pxx], follows from [17] and [18], but not from [16]. The property of being a self-admirer differs from that of being an admirer and from that of being admired (even though it entails both).

But there is no *part* of the sentences [17] and [18] that they share with each other that they don't share also with [16]. Looking just at the sub-sentential expressions out of which the sentences are built does not reveal the respect of similarity that distinguishes self-admiration from admiration in general - a respect of similarity that is crucial to understanding why the conditional

20] If someone admires himself then someone admires someone,

(∃x[Pxx] **à** ∃x∃y[Pxy]) expresses a good inference, while

21] If someone admires someone then someone admires himself,

(∃x∃y[Pxy] **à** ∃x[Pxx]) does not. For what [17] and [18] share that distinguishes them from [16] is not a *component*, but a *pattern*. More specifically, it is a pattern of cross-identification of the singular terms that two-place predicate applies to.

The repeatable expression-kind 'admires' is a *simple* predicate. It occurs as a component in sentences built up by concatenating it appropriately with a pair of singular terms. 'x admires x' is a *complex* predicate.[12] A number of different complex predicates are associated with any multi-place simple predicate. So the three-place simple predicate used to form the sentence

22] John enjoys music recorded by Mark and books recommended by Bob,

generates not only a three-place complex predicate of the form Rxyz, but also two-place complex predicates of the form Rxxy, Rxyy, and Rxyx, as well as the one-place complex predicate Rxxx. The complex predicates can be thought of as patterns that can be exhibited by sentences formed using the simple predicate, or as equivalence classes of such sentences. Thus the complex self-admiration predicate can be thought of either as the *pattern*, rather than the *part*, that is common to all the sentences {"Rousseau admired Rousseau," "Kant admired Kant," "Caesar admired Caesar," "Brutus admired Brutus," "Napoleon admired Napoleon,"…}, or just as that set itself. Any member of such an equivalence class of sentences sharing a complex predicate can be turned into any other by a sequence of *substitutions* of all occurrences of one singular term by occurrences of another.

Substitution is a kind of *decomposition* of sentences (including compound ones formed using sentential operators such as conditionals). After sentences have been built up using simple components (singular

terms, simple predicates, sentential operators), they can be assembled into equivalence classes (patterns can be discerned among them) by regarding some of the elements as systematically replaceable by others. This is the same procedure of noting invariance under substitution that we saw applies to the notion of free-standing content to give rise to that of ingredient content, when the operators apply only to whole sentences. Frege called what is invariant under substitution of some sentential components for others a '*function*'. A function can be applied to some arguments to yield a value, but it is not a *part* of the value it yields. (One can apply the function *capital of* to Sweden to yield the value Stockholm, but neither Sweden nor *capital of* is part of Stockholm.) He tied himself in some metaphysical knots trying to find a clear way of contrasting functions with *things* (objects). But two points emerge clearly. First, discerning the substitutional relations among different sentences sharing the same simple predicate is crucial for characterizing a wide range of inferential patterns. Second, those inferential patterns articulate the contents of a whole new class of concepts.

Sentential compounding already provided the means to build new concepts out of old ones. The Boolean connectives - conjunction, disjunction, negation, and the conditional definable in terms of them (A **à** B if and only if ~(A&~B)) - permit the combination of predicates in all the ways representable by Venn diagrams, corresponding to the intersection, union, complementation, and inclusion of sets (concept extensions, represented by regions), and so the expression of new concepts formed from old ones by these operations. But there is a crucial class of new concepts formable from the old ones that are *not* generable by such procedures. One cannot, for instance, form the concept of a C such that for every A there is a B that stands to that C in the relation R. This is the complex one-place predicate logicians would represent as having the form {x: Cx & $\forall y \in A \exists z \in B[Rxz]$}. As Frege says, such a concept cannot, as the Boolean ones can, be formed simply by putting together pieces of the boundaries of the concepts A,B, and C. The correlations of elements of these sets that concepts like these, those expressed by complex predicates, depend on, and so the inferences they are involved in, cannot be represented in Venn diagrams.

Frege showed further that it is just concepts like these that even the simplest mathematics works with. The concept of a natural number is the concept of a set every element of which has a successor. That is, for every number, there is another related to it as a successor ($\forall x \exists y$[Successor(x,y)). The decisive advance that Frege's new quantificational logic made over traditional logic is a *semantic, expressive* advance. His logical notation can, as the traditional logic could not, form *complex* predicates, and so both express a vitally important kind of concept, and logically codify the inferences that articulate its descriptive content.

Complex concepts can be thought of as formed by a four-stage process.

First, put together simple predicates and singular terms, to form a set of sentences, say {Rab,Sbc,Tacd}.

Then apply sentential compounding operators to form more complex sentences, say {Rab **à** Sbc, Sbc&Tacd}.

Then substitute variables for some of the singular terms (individual constants), to form complex predicates, say {Rax **à** Sxy, Sxy&Tayz}.

Finally, apply quantifiers to bind some of these variables, to form new complex predicates, for instance the one-place predicates (in y and z) {∃x[Rax **à** Sxy], ∀x∃y[Sxy&Tayz]}.

If one likes, this process can now be repeated, with the complex predicates just formed playing the role that simple predicates originally played at the first stage, yielding the new sentences {∃x[Rax **à** Sxd], ∀x∃y[Sxy&Taya]}. They can then be conjoined, and the individual constant *a* substituted for to yield the further one-place complex predicate (in z) ∃x[Rzx **à** Sxd]&∀x∃y[Sxy&Tzyz]. We can use these procedures to build to the sky, repeating these stages of concept construction as often as we like. Frege's rules tell us how to compute the inferential roles of the concepts formed at each stage, on the basis of the inferential roles of the raw materials, and the operations applied at that stage. This is the heaven of concept formation he opened up for us.

V. Conclusion

The result of all these considerations, which have been in play since the dawn of analytic philosophy, well over a century ago, is a four-stage *semantic* hierarchy of ever more demanding senses of "concept" and "concept use." At the bottom are concepts as reliably differentially applied, possibly learned, *labels* or classifications. Crudely behaviorist psychological theories (such as B. F. Skinner's) attempted to do all their explanatory work with responsive discriminations of this sort. At the next level, concepts as *descriptions* emerge when merely classifying concepts come to stand in *inferential*, *evidential*, *justificatory* relations to one another - when the propriety of one sort of classification has the practical significance of making others appropriate or inappropriate, in the sense of serving as *reasons* for them. Concepts of this sort may still all have observational uses, even though they are distinguished from labels by also having inferential ones.[13] Already at this level, the possibility exists of empirical descriptive concepts that can *only* be properly applied as the result of inferences from the applicability of others. These are *theoretical* concepts: a particularly sophisticated species of the genus of descriptive concepts.

At this second level, conceptual content first takes a distinctive *propositional* form; applications of this sort of concept are accordingly appropriately expressed using *declarative sentences*. For the propositional contents such sentences express just are whatever can play the role of premise and conclusion in *inferences*. And it is precisely being able to play those roles that distinguishes applications of descriptive concepts from applications of merely classificatory ones. Building on the capacity to use inferentially articulated descriptive concepts to make propositionally contentful judgments or claims, the capacity to form sentential *compounds* - paradigmatically *conditionals*, which make endorsements of material inferences relating descriptive concept applications propositionally explicit, and *negations*, which make endorsements of material incompatibilities relating descriptive concept applications propositionally explicit - brings with it the capacity to deploy a further, more sophisticated, kind of conceptual content: *ingredient* (as opposed to free-standing) content. Conceptual content of this sort is to be understood in terms of the contribution it makes to the content of *compound* judgments in which it occurs, and only thereby, indirectly, to the force or pragmatic significance of endorsing that content.

Ingredient conceptual content, then, is what can be *negated*, or *conditionalized*. The distinctive sort of definiteness and determinateness characteristic of this sort of conceptual content becomes vivid when it is contrasted with contents that cannot appear in such sentential compounds. My young son once complained about a park sign consisting of the silhouette of what looked like a Scottish terrier, surrounded by a red circle, with a slash through it. Familiar with the force of prohibition associated with signs of this general form, he wanted to know: "Does this mean 'No Scotties allowed'? Or 'No dogs allowed'? Or 'No animals allowed'? Or 'No pets allowed'"? Indeed. A creature that can understand a claim like "If the

red light is on, then there is a biscuit in the drawer," without disagreeing when the light is not on, or immediately looking for the biscuit regardless of how it is with the light, has learned to distinguish between the content of descriptive concepts and the force of applying them, and as a result can entertain and explore those concepts and their connections with each other without necessarily applying them in the sense of endorsing their applicability to anything present. The capacity in this way to free oneself from the bonds of the here-and-now is a distinctive kind of conceptual achievement

The first step was from merely *discriminating* classification to *rational* classification ('rational' because inferentially articulated, according to which classifications provide reasons for others). The second step is to *synthetic logical* concept formation, in which concepts are formed by logical compounding operators, paradigmatically conditionals and negation. The final step is to *analytical* concept formation, in which the sentential compounds formed at the third stage are *decomposed* by noting invariants under substitution. This is actually the same method that gave us the notion of ingredient content at the third stage of concept formation. For that metaconcept arises when we realize that two sentences that have the same pragmatic potential as free-standing, force-bearing rational classifications can nonetheless make different contributions to the content (and hence the force) of compound sentences in which they occur as unendorsed components - that is, when we notice that substituting one for the other may change the free-standing significance of asserting the compound sentence containing them. To form *complex* concepts, we must apply the same methodology to sub-sentential expressions, paradigmatically singular terms, that have multiple occurrences in those same logically *compound* sentences. Systematically assimilating sentences into various equivalence classes accordingly as they can be regarded as substitutional variants of one another is a distinctive kind of *analysis* of those compound sentences, as involving the application of concepts that were not *components* out of which they were originally constructed. Concepts formed by this sort of analysis are substantially and in principle more expressively powerful than those available at earlier stages in the hierarchy of conceptual complexity. (They are, for instance, indispensible for even the simplest mathematics.)

This hierarchy is not a *psychological* one, but a *logical* and *semantic* one. Concepts at the higher levels of complexity presuppose those at lower levels not because creatures of a certain kind cannot in practice, as a matter of fact, deploy the more complex kinds unless they can deploy the simpler ones, but because in principle it is impossible to do so. Nothing could count as grasping or deploying the kinds of concepts that populate the upper reaches of the hierarchy without also grasping or deploying those drawn from its lower levels. The dependencies involved are not empirical, but (meta)conceptual and normative. The Fregean considerations that enforce the distinctions between and sequential arrangement of concept-kinds do not arise from studying how concept-users actually work, but from investigation of what concept use fundamentally is. They concern not how the trick (of concept use) is done, but what counts as doing it - a normative, rather than

www.alhassanain.org/english

an empirical issue. That is why it is philosophers who first came across this semantic hierarchical metaconceptual structure of concept-kinds.

But cognitive scientists need to know about it. For it is part of the job of the disciplines that cognitive science comprises to examine - each from its own distinctive point of view - all four grades of conceptual activity: the use of more complex and sophisticated kinds of concepts, no less than that of the simpler and less articulated sorts. The move from merely classificatory to genuinely descriptive concepts, for instance, marks a giant step forward in the phylogenetic development of sapience. I do not think we yet know what non-human creatures are capable of taking that step. Human children clearly do cross that boundary, but when, and by what means? Can non-human primates learn to use conditionals? Has anyone ever tried to teach them? The only reason to focus on that capacity, out of all the many linguistic constructions one might investigate empirically in this regard, is an appreciation of the kind of semantic self-consciousness about the rational relations among classifications (which marks the move from classification to rational description) that they make possible. Computer scientists have, to be sure, expended some significant effort in thinking about varieties of possible implementation of sentential compounding - for instance in exploring what connectionist or parallel distributed processing systems can do. But they have not in the same way appreciated the significance of the question of whether, to what extent, and how such "vehicleless" representational architectures can capture the full range of concepts expressed by complex predicates. (Their lack of syntactically compositional explicit symbolic representations prohibits the standard way of expressing these concepts, for that way proceeds precisely by substitutional *de*composition of such explicit symbolic representations.) These are merely examples of potentially important questions raised by the hierarchy of conceptual complexity that cognitive scientists have by and large not been moved so much as to ask.

Why not? I think it is pretty clear that the answer is *ignorance*. Specifically, it is ignorance of the considerations, put forward already by Frege, that draw the bright metaconceptual lines between different grades of concepts, and arrange them in a strict presuppositional semantic hierarchy. Any adequately trained cognitive scientist - even those working in disciplines far removed from computational linguistics - can be presumed to have at least passing familiarity with the similarly four-membered Chomsky hierarchy that lines up kinds of grammar, automaton, and *syntactic* complexity of languages in an array from most basic (finite state automata computing regular languages specifiable by the simplest sort of grammatical rules) to most sophisticated (two-stack pushdown automata computing recursively enumerable language specifiable by unrestricted grammatical rules). But the at least equally significant *semantic* distinctions I have been retailing have not similarly become a part of the common wisdom and theoretical toolbox of cognitive science - even though they have been available for a half-century longer.

The cost of that ignorance, in questions not asked, theoretical constraints not appreciated, promising avenues of empirical research not pursued, is

great. Failure to appreciate the distinctions and relations among fundamentally different kinds of concepts has led, I think, to a standing tendency systematically to overestimate the extent to which one has constructed (in AI) or discerned in development (whether by human children or non-human primates) or reverse-engineered (in psychology) what *we* users of the fanciest sorts of concepts do. That underlying ignorance is culpable. But it is not the cognitive scientists themselves who are culpable for their ignorance. The ideas in question are those that originally launched the whole enterprise of analytic philosophy. I think it is fair to say that as we philosophers have explored these ideas, we have gotten clearer about them in many respects. For one reason or another, though, we have not shared the insights we have achieved. We are culpable for having kept this treasure trove to ourselves. It is high time to be more generous in sharing these ideas.

END

Notes

1 It ought to be uncontroversial that the last two of the three lessons are due to Frege. Whether he is responsible also for the first is more contentious. Further, I think both it and a version of the second can be found already in Kant. (As I argue in my 2006 Woodbridge Lectures, "Animating Ideas of Idealism: A Semantic Sonata in Kant and Hegel," forthcoming in the *Journal of Philosophy*.) But my aims here are not principally hermeneutical or exegetical - those issues don't affect the question of what we philosophers ought to be teaching cognitive scientists - so I will not be concerned to justify these attributions.

2 Dretske, Fred: *Knowledge and the Flow of Information* (MIT Press - Bradford, 1981), Fodor, Jerry: *A Theory of Content* (MIT Press - Bradford, 1990).

3 I sketch this program in the opening section of "The Pragmatist Enlightenment (and its Problematic Semantics)" *European Journal of Philosophy*, Vol 12 No 1, April 2004, pp. 1-16.

4 Pp. 306-307 (§107) in: Wilfrid Sellars: "Counterfactuals, Dispositions, and Causal Modalities" In *Minnesota Studies in the Philosophy of Science, Volume II: Concepts, Theories, and the Mind-Body Problem*, ed. Herbert Feigl, Michael Scriven, and Grover Maxwell (Minneapolis: University of Minnesota Press, 1958), p.225-308.

5 I discuss this view of Dummett's (from his *Frege: Philosophy of Language* second edition [Harvard University Press 1993], originally published in 1974), at greater length in Chapter Two of *Making It Explicit* [Harvard University Press, 1994], and Chapter One of *Articulating Reasons* [Harvard University Press, 2000].

6 Frege, *Begriffsschrift* (hereafter *BGS*), section 3. The passage continues: "In my formalized language [Begriffsschrift]...only that part of judgments which affects the possible inferences is taken into consideration. Whatever is needed for a correct inference is fully expressed; what is not needed is...not."

7 One might think that a similar distinction could be made concerning a parrot that merely reliably responsively discriminated red things by squawking "That's red." For when he does that, one might infer that there was something red there (since he is reliable), and one might also infer that the light was good and his line of sight unobstructed. So both sorts of inference seem possible in this case. But it would be a mistake to describe the situation in these terms. The squawk is a label, not a description. We infer from the parrot's producing it that there is something red, because the two sorts of events are reliably correlated, just as we would from the activation of a photocell tuned to detect the right electromagnetic frequencies. By contrast, John offers *testimony*. What he says is usable as a premise in our own inferences, not just the fact that his saying it is reliably correlated with the situation he (but not the parrot) reports (though they both respond to it).

8 *The Philosophical Review*, Vol. 69, No. 2, 221-225. Apr., 1960.

9 I discuss this point further in the first lecture of "Animating Ideas of Idealism" [op.cit.].

10 Of course, contemporary expressivists such as Gibbard and Blackburn (who are distinguished from emotivist predecessors such as C.L. Stevenson precisely by their appreciation of the force of the Frege-Geach argument) argue that it need not follow that the right way to understand that descriptive content is not by tracing it back to the attitudes of endorsement or approval that *are* expressed by the use of the expression in free-standing, unembedded assertions.

11 In the first essay of "Animating Ideas of Idealism" [op.cit.] I discuss the line of thought that led Kant to give pride of place to judgment and judging.

12 This point, and the terminology of 'simple' and 'complex' predicates, is due to Dummett, in the second chapter of his monumental *Frege's Philosophy of Language* [op.cit.].

13 A key part of the higher *inferential* grade of conceptuality (which includes the former, but transforms it) is that it is *multipremise material* inferences that one learns to draw as conclusions (=responses) now to Boolean combinations of the relatively enduring states that result from one's own responses.