



# APOHA

BUDDHIST NOMINALISM  
AND HUMAN COGNITION

EDITED BY

Mark Siderits

Tom Tillemans

Arindam Chakrabarti

Columbia University Press New York



Columbia University Press

*Publishers Since 1893*

New York Chichester, West Sussex

Copyright © 2011 Columbia University Press

All rights reserved

Library of Congress Cataloging-in-Publication Data

*Apoha : Buddhist nominalism and human cognition* / edited by Mark Siderits, Tom Tillemans, and Arindam Chakrabarti.

p. cm

includes bibliographical references and index.

ISBN 978-0-231-15360-7 (cloth : alk. paper) — ISBN 978-0-231-15361-4 (pbk. : alk. paper) — ISBN 978-0-231-52738-5 (ebook)

I. Buddhist logic. 2. Nominalism. 3. Negation (Logic) I. Siderits, Mark, 1946- II. Tillemans, Tom J. F. III. Chakrabarti, Arindam.

BC25.A65 2011

181'.043—dc22

2010049385



Columbia University Press books are printed on permanent and durable acid-free paper.

This book is printed on paper with recycled content.

Printed in the United States of America

# Contents

## *Preface* vii

Introduction	1
• ARINDAM CHAKRABARTI AND MARK SIDERITS	
1. How to Talk About Ineffable Things: Dignāga and Dharmakīrti on Apoha	50
• TOM TILLEMANS	
2. Dignāga's Apoha Theory: Its Presuppositions and Main Theoretical Implications	64
• OLE PIND	
3. Key Features of Dharmakīrti's Apoha Theory	84
• JOHN D. DUNN	
4. Dharmakīrti's Discussion of Circularity	109
• PASCALE HUGON	
5. Apoha Theory as an Approach to Understanding Human Cognition	125
• SHÖRYŪ KATSURA	
6. The Apoha Theory as Referred to in the <i>Nyāyamañjarī</i>	134
• MASAOKI HATTORI	
7. Constructing the Content of Awareness Events	149
• PARIMAL G. PATIL	
8. The Apoha Theory of Meaning: A Critical Account	170
• PRABAL KUMAR SEN	

9. Apoha as a Naturalized Account of Concept Formation	207
• GEORGES DREYFUS	
10. Apoha, Feature-Placing, and Sensory Content	228
• JONARDON GANERI	
11. Funes and Categorization in an Abstraction-Free World	247
• AMITA CHATTERJEE	
12. Apoha Semantics: Some Simpleminded Questions and Doubts	258
• BOB HALE	
13. Classical Semantics and Apoha Semantics	273
• BRENDAN S. GILLON	
14. Śrughna by Dusk	283
• MARK SIDERITS	

*Bibliography* 305

*List of Contributors* 321

*Index* 325

"Without Brackets: A Minimally Annotated Translation of Ratnakīrti's *Demonstration of Exclusion*," translated by Parimal G. Patil, is available on Apoha's page at <http://cup.columbia.edu/>.

## Preface

This is a book about the Buddhist nominalist theory of apoha. The basic idea of the apoha theory is that a general term like "cow" refers to all those things that are not non-cows. This is one of those philosophical ideas that might seem too smart by half. Students of Indian and Buddhist philosophy have often wondered whether the theory could be any more than a facile logical trick, one that ultimately fails to solve the problem it addresses: explaining our ability to use general terms without supposing that universals and other equally odd abstract entities exist. For four days in May 2006, fourteen scholars gathered at Crêt Bérard, a retreat in the hills above Lausanne, Switzerland, to try to answer this question. While the basic idea of the apoha theory is simple, its history and development in India and Tibet are quite complex. Among the scholars attending the conference were experts in various facets of that history. But the gathering also included scholars working in such diverse areas as philosophy of language, linguistics, and cognitive science. The aim was to try to arrive at a better understanding of what the apoha theorists were actually saying and then see if their view turns out to be a promising approach to the study of human cognition.

The conference at Crêt Bérard spurred much excitement and a great deal of subsequent discussion. The papers collected here represent the fruits of that discussion. While all the contributors save one were at the conference, none of the papers in this volume was presented there. Some of these papers are descendants of conference presentations, refined in the light

of much discussion and debate. But other papers were only written after the conference, reflecting new insights that grew out of that discussion and debate. This is not your usual volume of conference proceedings; this is collaborative research on a problem that until now has received only sporadic attention from individual scholars. Our hope is that by combining philosophical and textual-historical approaches we have allowed the ideas in this Buddhist material to find their rightful place in a contemporary discussion.

There are many who contributed to the success of the project. The conference would not have been possible without the generous support of the Elisabeth de Boer Fund of the University of Lausanne. Thanks are also due to Thomas Doctor and Heidrun Köppl, who worked tirelessly to ensure that everything ran smoothly, not only during the conference but at arrival and departure as well. We also wish to acknowledge the efforts of the staff of Crêt Bérard, who made all the conference participants feel welcome in their charming surroundings. Thanks are due as well to the editors of *Acta Asiatica* for granting permission to reprint (a slightly amended version of) an article by Masaaki Hattori that first appeared in the pages of their publication. Some of the editorial work involved in assembling this volume was made possible thanks to HK research support provided by Seoul National University and support from the National Institute of Advanced Study, Bangalore. Wendy Lochner, our editor at Columbia University Press, was an early supporter of this project. She deserves our thanks not just for her help in bringing this volume to fruition but also for her efforts to advance comparative philosophical research in general. We also wish to express our gratitude to Christine Mortlock, Michael Haskell, and Robert Demke for their help in the production process.

"Without Brackets: A Minimally Annotated Translation of Ratnakīrti's *Demonstration of Exclusion*," translated by Parimal G. Patil, is available on Apoha's page at <http://cup.columbia.edu/>.

§. ΑΡΘΗΑ. §

# Introduction

• Arindam Chakrabarti and Mark Siderits •

**T**his is a book about the apoha theory of Buddhist nominalism.<sup>1</sup> The apoha theory is first and foremost an approach to the problem of universals—the problem of the one over many. That problem is one of explaining how it is possible, when we see a pot, to think of it as a pot and call it by the name “pot,” a name that applies to many other particular pots. What is the one thing, being-a-pot, that this particular shares with many other particulars? Is there really such a thing in the world, over and above the individual pots, or is it just a mental construction of some sort? To hold the first alternative is to be a realist about universals, to hold the second is to be a nominalist. The apoha theory is a distinctive Buddhist approach to being a nominalist.

To fully appreciate the apoha theory one must understand the problem of universals. The usual practice, in introducing the problem of universals, is to start with Plato and Aristotle. In this introduction, we start with classical Indian philosophy instead. The problem of universals has played important roles in both the Western and the Indian traditions, and we suspect that the terrain may have been more fully explored on the Indian side. The first part of this introduction is a brief survey of classical Indian approaches to the problem. Part 2 then takes up some modern and contemporary Western approaches to the related problem of what concepts are, and how we acquire and manifest our mastery of them. Part 3 uses the discussions of the first and second parts to construct a taxonomy of possible approaches



to the problem of universals and locate the individual papers on the apoha theory that make up the rest of the book in that taxonomy.

## PART 1: ARGUMENTS FOR AND AGAINST UNIVERSALS IN CLASSICAL INDIAN PHILOSOPHY

That the word for an ontological type or a category, in Sanskrit, is *padārtha*, literally, “meaning of a word,” deserves serious attention and reflection. Somewhat like Frege and his commentator Dummett, but nearly two thousand years before them, Indian philosophers of language saw clearly that a theory of meaning ultimately boils down to metaphysics. Apoha, historically, is Dignāga’s answer to the question, “What does a word mean?” By the time it is elaborated by Dharmakīrti in his autocommentary on the chapter on Inference-for-oneself, this theory of meaning becomes partly a complex metaphysics of relationless, propertyless bare particulars, and partly a psychology of imagined generalities concocted out of habitual practice-guiding exclusions. These bare particulars serve as the ultimate referents and constituents of a lived world where actions are undertaken on the basis of word-signified demarcations-from-the-other. To get the thrust of the underlying nominalist attack against mind-independent universals, we must first survey the semantic roots of Indian realisms concerning universals.

From the time that Patañjali wrote his “Great Commentary” to Pāṇini’s grammar, semantic analysis of case roles and verbs embroiled the Sanskrit grammarians in ontology. In Pāṇini’s grammar and its early commentaries (between fourth and second centuries B.C.E.), three crucial technical terms for a universal—*sāmānya*, *jāti*, and *ākṛti*—were already explicitly in use. The device of adding a *-tva* or *-tā* (roughly equivalent to the English “-ness”) to any nominal root, “X,” yields, as meaning, the property of being (an) X, shared by all Xs. From “substance” (*dravya*) one can thus mechanically abstract “substanceness” (*dravya-tva*), from “real” (*sat*), “reality” (*sattā*), from “humans,” “humanity.” With this device in place, it was natural to make the distinction between an individual substance and the property that makes it what it is, its abstract essence. But to parse our talk of concrete cows rather than of the bovine essence, the Grammarians drew the distinction between talking about one particular cow and talking about any cow or a cow in general (*Vyākaraṇa-Mahābhāṣya* on Pāṇini’s *Aṣṭadh*, sūtra 1.2.58 and 1.2.64). The distinction between the general and the particular also came up for discussion in the context of the logic of pluralization. What allows one to say “trees” or “men” instead of using the word for a tree or man as

many times as the number of trees or men one referred to? Because the direct meaning of a common noun is the shared universal property of the referents, one could eliminate all but one remaining (*ekaśeṣa*) occurrence of that word, when speaking generally of all instances or any instance. One could also issue universalizable moral imperatives such as "a cow ought not to be killed," which, Patañjali jokes, is not obeyed by simply sparing the life of an individual cow.

*Jāti* (a word which, in modern Indian vernaculars, has come to mean a class, caste, or even a nation, and which is the Sanskrit counterpart of the Latin *genera*) is used by Pāṇini for a shared property of all the particulars of one natural kind, which serves also to distinguish any one of them from things of other kinds. The particulars are called *vyakti*—a word that etymologically suggests a distinct concrete tokening of common and uncommon properties. The problem with this generalist theory of meaning—defended by Vāṇyāyana (perhaps third century B.C.E.)—was that when, in a descriptive or prescriptive sentence, the action denoted by the verb has to hook up with what the noun means, what is meant by the noun has to be a particular. For, after all, no one can bind cowness with a rope, cut the tree essence, or have lunch with humanity.

Thus, in Indian semantics the dispute between those who insisted that a word primarily means a universal and their rivals who held that it must be particular substances that are the first meanings of words is at least twenty-two centuries old. The word often used for a universal by Patañjali was *ākṛti* (literally, "shape"), which is more reminiscent of form than of a property. In answer to the basic question "what is a word?" Patañjali considers the option, "is it that which remains nondistinct among distinct individuals, untorn when individuals are torn down?" and answers, "no, that is not the word, that is only the universal form (*ākṛti*)."

The need to switch to imperishable universals as meanings was felt both by the Grammarians and the Mīmāṃsā school of Vedic hermeneutics, for whom the authority of authorless sentences of the Vedas rested on their being eternal. The relation between words and objects was said to be "entrenched" and permanent. If perishable particular horses, cows, humans, and plants were the meanings of words, how could they be the eternally connected meanings of these beginningless Vedic words? The word *gauḥ* (cow) is therefore best taken to be eternally connected to the timeless bovine essence.

The first clear recognition of the need to postulate universals might also have come as much from reflecting on the generality or repeatability of the audible words themselves as from the theory of meaning. That

there could be many pronunciations or distinguishable phonations of the same word was seen to be an unquestionable example of the one-in-many. That naturally went hand in hand with the idea of the real word type existing timelessly, independent of its temporal, perishable token utterances. Later, in the philosophy of Bhartṛhari (sixth century C.E., sometimes called a "linguistic nondualist"), word universals and meaning universals and our natural tendency to superimpose the former on the latter were elaborately discussed. Out of these discussions there eventually emerged the Jaina notion of a vertical universal (*urdhvatā-sāmānya*), as against the more common property universal, which was termed "horizontal universal" (*tiryak-sāmānya*). As in Descartes's example of the piece of wax, a single substance that assumes different forms at different times (first hard and white, then soft and colorless), the vertical universal is a case of one over many where the one is a single substance, while the many are different forms. In the case of the horizontal universal, by contrast, the one is a form while the many are the distinct substances (or other particulars) in which that form occurs. Here we see clear recognition of the fact that the problem of universals is fundamentally a problem of explaining sameness in difference.

### The Hot Topics for Debate

Between the fifth and fifteenth centuries C.E., the debate between mainstream Nyāya-Vaiśeṣika and Mīmāṃsā realists and dissident Buddhist nominalists raged around the existence of eternal essences. The major points of disputation were:

1. Must we explain the use of a common noun or the experience of community across a plurality of particulars by postulating a single real property inherent in each of those particulars? (Vaiśeṣika and Mīmāṃsā said yes with some caveats, Buddhists said no.)
2. Is the property totally distinct from the individuals that exemplify it? (Vaiśeṣika said yes, and Bhāṭṭa Mīmāṃsā said yes and no.)
3. Does a universal exist only in all its own instances or are universals omnipresent? (A trick question of the Buddhist nominalist, answered cautiously by the Vaiśeṣika.)
4. Do universals have any role in causation? (Vaiśeṣika said that they can cause our awareness of them, while for Buddhists anything that is eternal must be causally inert, hence nonexistent. For Udayana [eleventh-

century Nyāya-Vaiśeṣika], nomic relations of necessary concomitance are ontologically founded upon the universals inhering in causes and effects.)

5. Can the work that is done by universals be done by relations of resemblance between particulars? (Vaiśeṣika said no, Jainism and Mādhyama Vedānta said yes.)

### Classical Nyāya-Vaiśeṣika Realism About Universals

Universals come to occupy a crucial role as the fourth type of real in the scheme of six basic categories of reals or “things-meant-by-words” (*padārthas*; notice again the semantic orientation) listed in the *Vaiśeṣikasūtras* of Kaṇāda. In that canonical scheme, after three types of unrepeatables—substances, particular qualities, and motions—come common properties. Although substances, qualities, and motions are entities of different types, they share one common property: they are all real. What is this realness that is common to all substances, qualities, and motions? Realness is a generic essence present in many substances, many qualities, and many motions. It is a universal, the highest one. Then there are less general features as well, the substancehood shared by all substances, the qualityhood common to all qualities, and the motionhood inherent in all motions. These second-tier universals are called “common-uncommon” since they function as defining properties belonging to all the members of the class to be defined and lacked by all others.

The *Vaiśeṣikasūtra*’s word for universal is *sāmānya* (the phonetic resemblance with “sameness” may not be entirely accidental), meaning “what is common.” The word for an individuator or particularity is *viśeṣa*, which means uncommon feature or specialty, the difference maker. Floweriness might be a common property, shared by roses, jasmines, and sunflowers. But the same property would be a difference maker when you compare a rose with fruits, seeds, stones, and animals, since none of these except the rose has floweriness. Hence Kaṇāda’s aphorism: “Universal and particularity depend upon understanding” (VS 1.2.3). Commentators hasten to point out that this formulation does not mean that universals are subjective or invented by our ways of understanding the world. All it means is that we find out by the verdict of our understanding whether some property is a pure universal or also a demarcator, as shown above.

Four broad arguments are proposed by these staunch realists for proving the existence of universals.

1. The evidence of sense perception is the strongest of all. Unless it leads to logical inconsistency, we must admit some common recurrent entity in each of those many things that sense perception shows to be of the same kind. This class character, the basis for our sense of sameness, is a universal for it is the same one found in many.
2. The argument from the meaning of general words, which runs as follows. A learnable common noun such as "bird" can denote an unlimited number of particulars of enormous variety. How the same word with the same meaning can correctly apply to so many diverse particulars calls for explanation. The explanation must lie in a distinction between reference (*śakya*) and sense (*śakyatāvachedaka*). Thanks to the existence of an objective universal, for example birdness, which serves as the shared sense, the same word can distributively refer to all birds or any bird. This does not boil down to one or the other of the two early extreme views that the bare particular or the pure universal is the primary meaning of a word. It is the balanced view that the meaning of a word is a particular possessing a general property, something serving as the common mode of presentation of its unlimited number of referents.
3. Then we have the argument from lawlike causal connections. Fire is a substance, but when it causes burning, its causal efficacy is not determined by its simply being a substance, for then any substance would burn. To explain what makes fire—and not any other substance—the cause of burning, we need to postulate fireness as the property that limits the causality of fire toward this effect. With the advent of extremely technical New Nyāya (around the thirteenth century) the need to have limiters (*avachedaka*) of causehood and effecthood became the standard ground for ontological commitment to universals.
4. Admission of universals also helped Nyāya solve the problem of justifying the inductive leap from observation of a few cases to a universal generalization covering all cases of a concomitance (e.g., where there is smoke there is fire). The common property observed in a few instances (the smokiness I perceive in the kitchen and at the bonfire) can, as it were, put us in direct perceptual touch with all the other instances where it also inheres, not in their individual details but in a generic way. Here the universal itself is supposed to play the role of the operative connection (linking bridge) between the sense organ and the apparently unobserved instances of that universal.

With all these supporting arguments for the universal's existence, the precise definition offered by Nyāya-Vaiśeṣika came down to this: "A uni-

versal is that which, being eternal, is inherent in many." Not any quality inhering in a substance is a universal. A wish inheres in a soul, but it is a short-lived episode, not eternal, hence not a universal. Colors are not universals in this system because they are unrepeatable tropes clinging to the particular surfaces. All colors share the universal colorhood. But two apples of exactly the same shade of red have two distinct red colors in them, just as each of them would have a distinct falling motion if they both fell. A universal must subsist wholly in each of its instances by the special relation of inherence. A universal must be wholly inherent in each of its instances. The word "inherent" must be taken seriously. A single string may be running through many flowers, but it—more precisely different segments of it—are only in contact with the flowers. The whole string cannot inhere in any one of them. This "single thread" analogy, therefore, is not entirely a happy one. We shall see later how Buddhist logicians refused to accept this idea of a single property running through many distinct objects. After asserting that "each entity is self-confined; they do not mix themselves with others, each of them being intrinsically unconnected to any other." Dharmakīrti gives a cartoonlike analogy: "Even if they are cognized together, by a cognition projecting itself as generality embracing, they would not be like the idols of spirits linked by a single string attached to their necks, there being nothing single [neither a property nor a relation] among these discrete particulars."<sup>2</sup>

What then is inherence? According to the orthodox realists, it is a kind of being-in, the converse of which is an intimate "having." Humanity inheres in me, just in case I have humanity. Now, *having* can be of many kinds. Things *have* qualities and motions. Wholes *have* parts. I *have* a pen in my hand. Rich people *have* big houses. The logical structure of each of these relations, of characterization, constitution, contact, and ownership, however, is utterly different. All four are more or less aptly reportable by the use of the preposition "in" or "of": the taste is in the apple; the room is or consists in the walls, roof, and floor; the pen is in between the fingers; and the house is of the rich merchant. Yet one initial grouping could be made to clarify their distinct structures. The taste and the room cannot exist without the apple or the room parts. The taste cannot float about on its own, minus the apple.<sup>3</sup> The room cannot stand independently of the walls. But that very pen can easily exist untouched by the hand, and that house can change hands. So the first two relations hold between pairs that are "incapable of standing apart from one another" (*a-yutasiddha*), whereas the other two relations hold between pairs that are "capable of standing apart from one another" (*yutasiddha*). However tightly my ring is stuck to my finger, it is

not inherent in it as inseparably as fingerness is inherent in my fingers. It is no physical glue but a metaphysical inseparability that joins the goatness to the goat, ties up the running and the black color of the goat to the goat, and binds the goat to its body parts. The kind of being-inseparably-in that connects the universal to its instances has to be distinguished from the way a berry lies in a bowl. For the sake of economy—the principle of not multiplying entities beyond necessity—the mainstream Nyāya-Vaiśeṣika metaphysicians posit only a single such relation as enough to link up innumerable pairs of universals and particulars, qualities and substances, and wholes and parts. For systemic reasons, this relation is supposed to be eternal as well. And this is inherence (*saṃavāya*). Even other universal-friendly realists, such as the Bhāṭṭa Mīmāṃsakas, give the Vaiśeṣika much grief over this peculiar theory of exemplification. The Bhāṭṭas themselves take the relation between a universal and its own exemplifier to be identity-in-difference. The Buddhist logician finds both inherence and identity-in-difference equally unpalatable.

Though you cannot experience Vaiśeṣika universals by themselves, they are ontologically independent of their instances. Even when all cows are destroyed in the world, cowness will still be around, for otherwise the possibility of a fresh cow coming to be remains inexplicable.

### Real Universals (*jāti*) and Titular Properties (*upādhi*): On Being a Cook

Though all universals are common features, not all common features corresponding to multiply applicable descriptions are, strictly speaking, universals. Even hardcore realists about universals feel the need for population control in the world of universals. Being a brahmin (a member of the priestly intellectual class) is taken to be a natural kind by Nyāya-Vaiśeṣika in the face of vehement opposition by anticaste Buddhists and Jains. But being a cook or being a tailor would not be considered a natural or real universal, even though it is a common feature of cooks or tailors. Nyāya-Vaiśeṣika philosophers suggest six tests that an alleged (semantically suggested) property must pass to count as a genuine universal. These tests or hurdles are called “universal blockers.”

1. If a property has only a single exemplifier then it is not a universal. “Being the Statue of Liberty” is not a universal and neither is timehood because there is no more than one Statue of Liberty or one time.

2. If two properties have exactly the same extension, for example, the property of being a homo sapiens and the property of humanity, they cannot be two distinct universals.
3. The domains of two universals can be either completely disjoint or one of them completely included in the other. They cannot be partially intersecting and partially excluding each other. Thus, being material and having a limited size cannot both be universals in Vaiśeṣika ontology, because while many things have both properties, open space is supposed to be material yet not limited in size, while the internal sense organ is supposed to be limited in size but immaterial.<sup>4</sup>
4. A regress-generating property is not a universal. Universalhood is not a universal, although all universals seem to have that property in common, for one could then multiply universals endlessly. Universals do not have further universals in them.
5. When the very nature of a characteristic is merely to distinguish its bearer, for example, an earth atom, from another particular of that kind, such an ultimate individuator should not be brought under a general category of individuatorhood, for that militates against its necessarily unique nature. Failing this test, the alleged generality "indivuatorhood" (*viśeṣatva*) does not qualify as a universal within Vaiśeṣika atomism.
6. The feature must bear inherence and no other relation to its bearer. Inherencehood is not a universal because, were it one, it would have to be related by inherence to inherence, which would be absurd. An absence cannot be a universal. Nor could the negativity common to all absences be a universal. Even though every rabbit is hornless, neither the absence of horn itself nor the absenceness of the absence resides in rabbits or absences by inherence.

Besides these, compound properties such as being a sturdy black cow or being either a cow or a buffalo are ruled out because universals are supposed to be simple.

What happens to the properties that get disqualified by a universal blocker? They are thrown into the mixed heap of titular, surplus, or imposed properties (*upādhi*). They might still be of much theoretical and practical use. Not only nonnatural generalities like being a New Yorker or being a carburetor, but even is-ness, knowability, and positive presence (shared by items of all six categories—substance, quality, motion, universal, inherence, and final individuator—but not found in absences) are merely



titular properties. Knowability and existence (is-ness) are (intensionally) distinct properties, in spite of being coextensive, because they are not universals.

### How Are Universals Known?

We need philosophical reasoning to grasp such deep universals as substancehood because many concrete instances of substancehood, such as time, atoms, and other people's souls, are not objects of perception. If the instances are perceptible, the universals must be directly perceptible as well. We see floweriness in a flower, just as we see its hue and smell its fragrance. According to Nyāya epistemology, to see Black Beauty as a horse we must first see its horseness (which is a perceived universal, though it is not perceived to be a universal).

But many strong arguments could be given against the perceptibility of universals. Let us examine several. First, if properties were perceived, we would perceive them even at the time of encountering the first exemplifier, but we do not. Hence, properties are abstracted, not seen. Both premises of this argument, of course, could be questioned. For the empirical knowledge of a common property to dawn gradually, a recognition must take place in the second, third, and subsequent sightings of the instances. To be faithful to the form of that recognition, "I have seen this sort of animal before," is to admit that even in the first instance that sortal property was seen.

Here is another antiperception argument. If properties were objects of perception, they would be causes of perception, but they are not. Therefore, they are not perceived. Again, both premises are rejected by Nyāya realists. Potness need not itself reflect light back into the retina for it to be causally relevant to the visual perception of potness. As long as the pot in which it inheres is in contact with the seeing eyes, it has a causally operative connection with the appropriate sense organ. If, of course, we define perception as prelinguistic and nonconceptual (as some Buddhists do), and universals are taken to be word-generated concepts, then to use that definition as an argument for the imperceptibility of universals would be crudely question begging.

Someone with Fregean sensibilities might propose another quick argument against the perceptibility of universals. Universals are not objects but functions. Therefore, they are not objects of perception. But there is a clear shift in the meaning of "object" between the premise and conclusion of this argument. In the West there is a basic resistance to admitting the sense perception of universals because universals are supposed to belong to the

intelligible realm. In *The Problems of Philosophy*, Russell claimed that we have direct acquaintance with universals, but that acquaintance was not meant to be sensory. Only David Armstrong, whose view about universals comes very close to Nyāya-Vaiśeṣika realism, seems to have warmed up to the idea of perceiving universals.

### Attacks from the Buddhist Nominalist

The Vaiśeṣikas' first argument for the existence of universals depends upon the generalization, "In every case, the sense of commonness or similarity felt by word users must be spawned by an objective universal." Surely this generalization is riddled with counterexamples. We have just seen how people feel a sense of similarity across the many cooks; yet Nyāya-Vaiśeṣika realists refuse to admit cookness as a universal. There is no good reason to posit these weird entities and every reason to eliminate them. So claimed the Yogācāra-Sautrāntika Buddhists. "It does not come there [from another place], it was not there already, nor is it produced afresh, and it has no parts, and even when it is elsewhere it does not leave the previous locus. Amazing indeed is this volley of follies!" (PV 1.152).<sup>5</sup> With this oft-quoted remark, Dharmakīrti summarizes his battery of objections against the Nyāya-Vaiśeṣika theory of universals. How can a universal remain the same while existing in distinct things and places? Does it scatter itself into parts or does it live in its entirety in each instance? When the locus moves, does it move? If cowness is everywhere, how can it be absent in a horse? If it is only where its instances are now, then how does it travel to a new place when a new cow is born there? It does not pervade the place where an individual is located, for then the place itself would be its instance, yet how can it manage to inhere in the individual that occupies that place? If the particular instance is needed as a manifestor of the ubiquitous universal, why can't we perceive the cow—its manifestor—independently of noticing the universal cowness? A lamp reveals the preexistent pot in a room, but you don't need to see the pot first before you notice the lamp!

Most of these difficulties, the realists retorted, suffer from a category mistake. They assume that a universal is just another kind of superparticular. But a universal is not a spatiotemporal thing, and that is why multiple location without divisibility is not a problem for it. In spite of such robust responses, Buddhist antirealism about universals became more and more trenchant in the second millennium, until we have such caustic remarks directed at the Vaiśeṣika realists as those of Paṇḍit Aśoka: "One can clearly see five fingers in one's own hand. One who commits himself to a sixth

general entity fingerhood, side by side with the five fingers, might as well postulate horns on top of his head" (SD 101-2).

### APOHA NOMINALISM IN A NUTSHELL: THE BUDDHIST EXCLUSIONIST ACCOUNT OF CONCEPT FORMATION

Buddhist logicians have an error theory about universals and permanent substances that they reduce to mental or physical particulars or simply eliminate. There are nothing but momentary quality particulars in the world. But the human mind, afflicted by perpetuation wishes and language-generated, deeply ingrained myths, has a tendency to cluster some of them together first in the fictional form of enduring substantial things (i.e., the mind constructs what the Jainas call vertical universals) and then further classify these "things" into types. This illusion of commonality, of course, has some pragmatic value because, except in thoughtless contemplative experience, our working cognitions of the world mostly take the form of predictive judgments or explanatory inferences on the basis of these apparently general features and their mutual connections. When a particular cow (which, in its turn, is a fictional cow shape superimposed on certain packets of quality tokens) is seen to be other than all other animals, the original indeterminate (concept-free) perceptual content somehow causally triggers this difference-obliterating tendency. The particular cow image is made to "fit" this linguistic and imaginative exclusion from the complementary class of horses, rabbits, pillars, and such things. The specificity of the particular cow—its numerical detailed differences from other cows—is ignored; instead, this mere exclusion from non-cows is foisted onto the perceptual content as a predicate. This exclusion masquerades as the universal cowness. To take Dharmakīrti's example, the universal "antipyreticness" is a useful figment of the imagination. In the external world, there is no single shared intrinsic property of different medicinal plants that all work as fever reducers except that they are other than those things that fail to relieve fever. Antipyreticness is an erroneous reification of this mere exclusion (apoha). This, in a nutshell, is the apoha nominalism of the Yogācāra Buddhist logicians, which is the topic of this collection of papers.

### Milder Nominalisms: Resemblance Theories

In the midst of this great battle between realists and nominalists, the Jaina syncretists stepped in with their typical reconciliatory message: that every

object of knowledge has an alternatively more-than-one (*anekānta*) nature. particularity and generality being just two of these. We cannot doubt that things do objectively resemble one another. These resemblances are real relations. But both the things and their mutual resemblances are particulars. Nothing has the burden of being strictly repeatable.

The Jainas reject the Buddhist version of nominalism, more or less on the same grounds that Kumārila Bhaṭṭa, the great Mīmāṃsaka, rejected it. Positive predicates, Kumārila had objected, cannot all be given a negative meaning. Since these exclusions are nonentities invented by erroneous imagination, to say that all our words mean them is to turn all words into empty terms. Indeed, since all exclusions are equally hollow in content, distinguishing one from another would be like trying to distinguish two nonexistents, one fictional fat man in the doorway from another bogus bald man in the doorway (to give Kumārila a Quinean example). Only those denials make sense which have something positive to deny. Since all descriptions capture only negations, this theory, ironically, strips our negations of all meaning, since there is nothing left to deny.

Although they use the Buddhist criticisms to do away with Vaiśeṣika realism about self-standing universals, the Jainas bring extremely pertinent charges against Buddhist nominalism. According to Dharmakīrti's "error theory," the projection of an external object-in-general meant by the word is possible because of a superimposition of the internally constructed exclusionary image form on external, nameless, uncategorized, positive particulars. But for a superimposition or false identification of the inner with the allegedly outer to happen, both the locus and the content of the error must be grasped by the same (error-committing) episode of awareness. Unless the rope is actually encountered (not as a rope) and the snake is recalled by the same piece of cognition, mistaking a rope for a snake is not possible. But what kind of perception will grasp both? Not a nonconceptual pure sensation, for it does not make any claims, hence makes no mistakes; and not the concept-laden ascertaining perception, for it never actually has access to any external or internal *particular*. So the apoha story is untellable under the assumptions of Buddhist epistemology. In its place, the Jainas propose a resemblance-based theory of perspective-dependent generality.

Prabhācandra anticipates the Russellian objection that at least all these resemblance relations would ultimately need a shared similarity universal. His answer is that, just as a Vaiśeṣika final individuator (*viśeṣa*) does not need another distinguisher, one resemblance does not need a higher-level resemblance or universal to explain why all those resemblances

are similar. While accounting for the similarity between ground-level particulars, the similarities also account for their own similarity to each other.<sup>6</sup>

### Contrasts with Western Metaphysics of Forms and Properties

It should be clear by now that there is no core theory of universals shared by all Indian philosophers. But we can discern five broad features that distinguish Indian theories of universals from their Western counterparts.

1. Even the strongest realist position, that of Nyāya-Vaiśeṣika, falls short of the ante rem realism of Plato's theory of ideas. Unlike Plato, Indian realists about universals were equally realists about the perceptible particulars of the external world. Earthly particulars were never thought to be less real copies of thinkable universals, even by those who believed in universals.
2. Even if we concede that Nyāya's universals were close to Aristotle's universal properties, which are immanent in concrete particulars, Aristotle could never agree that universals are themselves directly perceived by the same senses that grasp the corresponding particulars, which is the standard Nyāya position.
3. The peculiar form that nominalism took in the Indian Buddhist theory of word meanings as exclusions does not have any parallel in the West. We find an interestingly different counterpart of the Jaina and Mādhyamika theories of resemblance in Nelson Goodman, but apoha nominalism remains a unique contribution of Indian Buddhism.
4. Most Western realist accounts of universals take colors, smells, tangible textures, and such qualities, as well as relations such as "being larger than," as paradigm examples of universal properties. In Indian realist thought, these would count not as universals but as particulars. The distinction between such particular qualities (*guṇa*) and universal properties (*jāti*) has been sacrosanct. It is only very recently that the idea of quality particulars (or "tropes") has gained ground in Western analytic metaphysics. Neither are relations treated as genuine universals by any classical Indian realist.
5. Finally, the controversial and complex theory of inherence as a single concrete connector joining not only universals and their instances but also particular qualities to substances and, most puzzlingly, wholes to their parts, is totally foreign to Western realists.

## PART 2: CONTEMPORARY WESTERN THEORIES OF CONCEPTS ON THE MARKET

Since the Sanskrit technical term *vikalpa*—central to apoha theory—is sometimes translated as “concept” (or “conceptual construction”), and with at least one (highly contentious) interpretation of apoha theory construing it as a version of conceptualism, it would be good to preface our investigation of the apoha theory with a quick survey of the available theories of concepts (and concept formation, and concept possession) in modern and contemporary Western philosophy. Unfortunately, like most frequently used words in a disputatious philosophical tradition, the word “concept” is so full of ambiguities that while all these theories are called “theories of concepts,” it is far from clear that they are trying to explicate the same concept of a concept. Some of these theories are concerned with what it is that we can do when we possess a concept and how we manifest such possession; others are concerned with what kind of mental representation a concept is; and yet others are interested mainly in the psychosocial story of how we acquire concepts. But can these theories even have different concepts of a *concept*, if they do not have, at some level, the same concept of a concept? One suspects there is a paradox lurking here. Still, our prephilosophical idea of concepts can be unpacked with the help of the following characterizations. Something that is meant by a predicate (“... is wise,” “... is square,” “... is a metal”), by a pluralizable common noun (“dog,” or for that matter “concept” itself), and most directly by an abstract noun (“substanteness,” “existence,” “justice”), that which renders general propositions possible, that by which our thoughts are constituted, minimally, is called “a concept.” Let us now look at an ahistorically ordered list of available theories of concepts so understood.

1. Classical definitionism: Beginning with Socrates, and fully matured in Aristotle’s and Aquinas’s doctrine of essences, the dominant Western account of concepts takes them as defining essences captured by necessary and sufficient conditions for counting as something. Being wet and being earth together constitute the concept of mud, because all and only mud is wet earth. The definition of a triangle captures the precise concept of a triangle. This is closely connected to the notions of a kind and a category, which are the types in which entities can be classified, divided, and defined according to definitions. Murphy (2004, 15) identifies three main claims of the classical theory: (a) that concepts are mentally represented as definitions, (b) that every object either does or does not fall under a concept,

unless it is an intrinsically vague or fuzzy concept like that of "big," and (c) that any item that satisfies the definition is as good an instance of a concept (an object falling under that concept) as any other, that all dogs, for example, are equally dogs and there is no typical/atypical distinction within members of the class of dogs. "The definition is the concept according to the classical theory."

This theory is widely seen to have been discredited by Wittgenstein's famous attack against it in *Philosophical Investigations*, showing our failure to find any properties had by all and only games. Definitions of all but the arbitrarily stipulated concepts of mathematics and logic are nearly impossible to agree upon. Once we start noticing borderline cases, most empirical and everyday concepts turn out to be vague, and the law of excluded middle (claim [b] above) seems to fail to apply to something even as basic as the concept of "alive." In addition to such "conceptual" warnings against the Aristotelian (or Vaiśeṣika) search for exact definitions, there is strong empirical evidence, from how people actually think, that concepts cannot be definitions. Different competing definitions of the same concepts seem to be entertained by different groups of people, who might for example debate whether a tomato is a fruit or a vegetable, whether flavor is a matter of taste or smell or both, whether tapestry is an art or a craft. There is no telling whether a low three-legged seat with a very small back is a chair or a stool, whether a fetus is a distinct person or not. There is no definition, yet there is a concept. So concepts are not definitions. Further, if concept *F* were the set of defining properties that all and only *F*s have in common, then analytic judgments explicating the concept of the subject by the predicate would be sharply distinguishable from synthetic ("ampliative") judgments where the concept of the predicate is added anew to the concept of the subject. But, as Quine has shown, the analytic-synthetic distinction cannot be sharply drawn. Hence concepts are not definitions.

For some, the final blow to the classical theory came when the empirical work of E. Rosch, C. B. Mervis, and others established the falsity of the third claim (c): some dogs, some chairs, some tomatoes are more typical than others and are more closely related to the common concepts of those items. A Chihuahua is not as typical a dog as a German shepherd, a white tomato the size of a cherry is not what one thinks of when one calls up the concept of a tomato, a penguin is not as typical a bird as a sparrow. These variations within the definitionally secured domain of a concept show what is called a "typicality effect." Typicality effects take the concept theorist away from the search for objective essences (recognition-transcendent satisfaction conditions) toward what people actually have in mind when they use a

general word or a predicate, from what the word means to what individual users tend to mean by it.

2. The British Empiricist idea/image theory: Diametrically opposed to the classical common essence theory of concepts is the extreme empiricist theory of Hume that concepts are ideas, which are faint copies of sensory impressions. My concept of a dog is a smudged memory image of many impressions of individual dogs that I have received through my senses. Concepts, under this theory, are creatures of the individual mind. The difficulty with this view is one that Berkeley pointed out before Hume: since an image is a mental particular, however lacking in detail an idea-image may be, it cannot be a general idea. Thus, if one took the idea theory of concepts seriously, then by virtue of the reduction of concepts to particular mental contents, conceptualism about universals would be reduced to nominalism. If there is nothing but featureless particulars anywhere in the world, ideas in people's minds would also be featureless particulars.

The idea theory of Hume, therefore, does not give us any explanation of abstract general concepts, insisting instead that there aren't any. It also misconstrues the categorical distinction between particular sensations and general concepts as a distinction of degree of vivacity, as if a particular visual image of a cat, when its contour gets smudged and its colors paler, turns into the concept of a cat-in-general.<sup>7</sup> This is as dubious as the suggestion that an adjective or verb is a half-forgotten proper name. There are moreover Wittgensteinian reasons against such a theory of concepts (or meanings of words): it makes meanings and constituents of thought completely private. You and I do claim to have, sometimes, the same concept of a tree, but we can never have the same idea of a tree, let alone an identical mental particular called up by the word "tree." That makes the Humean version of the idea theory of concepts unattractive.

3. Wittgenstein's family resemblance theory: After rejecting the classical theory of definable essences corresponding to every meaningful general word, Wittgenstein's positive agenda was to motivate us to observe and describe the different ways in which a common noun or an adjective is used, without trying to insist, *a priori*, on some single thread of meaning running through all those uses. To master a word, and hence the concept-in-use attached to it, is to gradually develop command over a network of disjunctively woven criteria, an arbitrarily chosen subset of which can be used to identify the objects falling under it. The reason hopscotch is called a game may have nothing in common with the reason computer chess is called a game, though there could be intermediate examples of games such as cricket which have features in common with both hopscotch and chess.



These reasons do not form any unitary core of necessary conditions but neither are they subjective or mental such that each user has a private idea why they call some activity a game.

The problem with this powerful theory is that while its negative force, rejection of the classical Platonic and empiricist theories of concepts, is clear, its positive account of what it is to have the same concept as others, or how any of us picks up these loosely woven sets of criteria from diverse contexts of use, is far from clear. While Wittgenstein crucially distinguishes between following a meaning rule correctly and merely seeming to, he does not, for example, tell us how a misapplication of a concept is detected. How do we know whether someone is using the concept *game* or the concept *entertainment*, if the overlaps between the two show closer internal connection than the widely divergent cases of a game, which are held together by the concept *game*?

4. Kantian theory: Wittgenstein was not the first Western thinker to associate concepts with rule following. Already in Kant's *Critique of Pure Reason* concepts were defined as rules for synthesizing experiences. But what does Kant mean by a rule? First, it is a schematic recipe for pattern recognition, for running through and holding together certain bunches of sensory inputs as if they represent an object outside the sensation. Second, it is like a major premise of a syllogism, ready to subsume an instance or a subcategory under a more general predicate, to draw a "judgment" as a conclusion. Longuenesse (1998, 50) helps us here: "The two meanings of 'rule,' as rule for sensible synthesis [the concept as schema] and as discursive rule . . . or major premise of a possible syllogism . . . are indeed linked. Because one has generated a schema, one can obtain a discursive rule by reflection and apply this rule to appearances." Sensibility as a faculty is mere susceptibility to external stimuli. Mere sensing is not enough to give us an object outside the experience, let alone to form a general judgment about the object. Only understanding can organize the received sensory data into representations of this or that object of this or that category or sort. Understanding is the faculty that employs pure concepts and, with the help of imagination, generates empirical concepts. So, to have a concept is to know how to combine or synthesize an array of sensory stimuli such that the unified awareness could be intentionally directed toward an object, out there, from which the awareness distinguishes itself. Such is the concept of a substance, of any old tree or any water pot, which organizes sensations, memories, expectations, and permanent possibilities of perceptions such that the seeing or touching counts as an experience of a substance, of a tree, or of a pot.

About the actual psychological process of concept formation, Kant has an agnostic attitude: "The concept 'dog' signifies a rule according to which my imagination can delineate a figure of a four-footed animal in a general manner, without limitation to any single determinate figure such as experience, or any possible image that I can represent *in concreto*, actually presents. This schematism of our understanding, in its application to appearances and their mere form, is an art concealed in the depths of the human soul, whose real modes of activity nature is hardly likely ever to allow us to discover, and to have open to our gaze" (B 181).

Despite its obscurity, Kant's "rule of synthesis" account contains a valuable lesson that is uncannily similar to Dharmakīrti's account of the passage from direct perception of particulars to action-prompting ascertainment of objects. Though it is the senses that put us in contact with the external world, it is concepts that enable us to make claims about common objective targets of perception and practical activity. Through their generality, concepts make reidentification possible, and so become object makers.

5. Fregean function theories: Using his distinction between sense and reference, Gottlob Frege held that a concept is the reference of a predicate expression. More recently, Christopher Peacocke (1992b) has claimed that concepts are the senses of such expressions. Both theories share the view that a concept is like a predicate expression in that it serves as a function that takes us from one sort of object to another sort of object. According to Frege's semantic theory, the expression "    is red," for instance, plays the role of mapping objects onto truth-values. So when the blank indicated by "   " is filled by an expression that refers to a stoplight, it yields the truth-value True, while inserting an expression that refers to grass yields the truth-value False. The reference of such an expression is the function itself. For Frege this is as much a part of the world as are stoplights and grass. Otherwise, he believes, we could not explain how different speakers of the same language, or speakers of distinct languages, could all be said to grasp the same concept when they come to know that stoplights are red.

Peacocke's view is that a concept is not the function itself but its mode of presentation, the way in which it is grasped by those who are said to possess the concept. Like Frege, Peacocke thinks of a concept as something the possession of which enables the subject to represent an object in determinate ways, ways that may be true or false. Peacocke's view is also like Frege's in making concepts objective constituents of the extramental world. Both are clearly heirs of the Kantian insight that possession of a concept

involves something like mastery of a rule, something that confers the ability to group together disparate mental presentations and to make inferences. The chief difficulty of both, from the perspective of our present concerns, is that in making concepts inhabitants of a mysterious "third realm" that is neither physical nor mental, they leave it unclear how concepts could play a causal role in human cognition. No Indian philosopher, whether a realist like the Naiyāyikas or a Buddhist nominalist, would accept such entities. While they might represent an elegant solution to any number of problems in formal semantics, they would strike such thinkers as ontological overkill. Frege's version also suffers from the paradoxical consequence that the expression "the concept *red*" does not refer to a concept. This is so because the reference of a definite description must be an object, something "complete" or saturated in nature, while concepts, as functions, are by nature "incomplete" or unsaturated. While Frege himself embraced this paradox, many see it as indicating a serious deficiency in the view.

6. Similarity-based theories: Bolstering Wittgenstein's family resemblance theory with empirical research, Eleanor Rosch and others came up with the influential thesis that instances fall under a concept, not in a yes/no fashion (as classical definition theory predicted), but in a probabilistic way, depending upon how many of the typical features are available in an instance. To be a dog, an animal need not have all the necessary conditions of being a dog—there being no such set of defining conditions at all—but only a sufficient number of conditions which are associated with the complex mental representation of a typical dog. Since the concept *robin* has many more of the structural elements of the superordinate concept *bird* than the concept *chicken* has, and the concept *penguin* has even less than *chicken*, robins are more readily recognized to be birds than chickens or penguins. Categorization, say, of household goods as furniture, thus turns out to be a feature-matching process shading off from the more typical instances of chairs and tables all the way to hat stands and ottomans.

The resulting prototype theory claims that a concept is constituted by a network of similarities with the best instances of a type. A concept, instead of being a fixed definition of a universal essence, may be a set of weighted feature representations most users of a word have in mind—a set which could be added to or subtracted from continuously. Picking up the concept *egg* and the concept *eggplant* from white, oval-shaped and dark purple "best examples," respectively, a child may fit into a community of users of the words "egg" and "eggplant" by becoming ready to also include purple eggs and white eggplants, as color moves away from the core set of criterial attributes to the periphery, through other more salient bases of resemblance.

Of all the resemblance-based theories of concepts, prototype theory is the most empirically well corroborated and statistically computed.

A recent variant on this basic approach is Jesse Prinz's (2002) proxytype theory, which purports to provide a theoretical framework for explaining how the ability to use concepts might be realized in the brain. Taking as a model D. Marr's computationalist approach to the neurophysiology of visual processing, Prinz sees concept formation as a matter of the formation of networks of stored images that exhibit dynamic interactivity and context sensitivity and yet are ultimately traceable to repeated occurrences of sensory stimulation. The result might be called neo-Lockean, but unlike Locke's theory there is built into the theory a way of accounting for the flexibility with which we deploy such concepts as *dog*. Not only can it explain the typicality effects that prototype theory is designed to handle, but it can also account for variations across contexts of concept application. For on Prinz's theory a context determines which subset of the set of features in the network stands proxy for the concept as a whole.

Prototype theory, proxytype theory, and the closely related exemplar theory all do a far better job of accounting for the actual practice of concept application than classical definition theory and its offshoots. But such approaches are not without their own shortcomings. One is the "pet-fish" problem. A dog or a cat is a typical pet; a middle-sized trout or a tuna is a typical fish. And no one can deny that the concept *pet fish* is composed of these two concepts. Yet, the typical pet fish, a guppy or a goldfish, does not remotely resemble a dog or a cat, a trout or a tuna, and has features not computable through resemblance from the intersection of the weighted group of features of the constituent prototypes. An adequate theory of concepts must account for their compositionality, the fact that we form complex concepts by somehow putting together their simpler constituents. Proxytype theory might claim the ability to handle this problem, but it is not clear how this will work. And in any event, all such approaches have one major flaw from the perspective of the Buddhist nominalist. They all take as given the ability to respond to similarity of stimulation. For the Buddhist nominalist, a key task for any theory of concepts is to explain how, in a world of pure particulars, certain stimulations come to seem similar to subjects. A theory of concepts cannot then take such an ability as a given.

7. Theory theory of concepts: The fact that people possess concepts that come with no feature- or exemplar-based stereotype and the fact that radically dissimilar stereotypes are associated by different users of the same concept throw into doubt the very basis of similarity-based theories of concepts. Mastering and deploying a concept may not be a matter of

feature matching or probabilistic frequency measurement in a space of perceptible resemblance to paradigm instances. Having and applying a concept might be more like theoretical problem solving. Concepts would then be minitheories for the categories they range over, theories that yield beliefs about causal connections and deep intrinsic properties of things falling under them. Gopnik and Meltzoff (1996) identify three characteristics of such a "theory": (1) Structurally, a theory is a system of abstract entities and laws, postulating perception-transcendent causal relations based on counterfactual reasoning. (2) Functionally, a theory enables the user to make predictions, give explanations of observable behavior, and give reasons for further taxonomies. (3) Dynamically, a theory tests itself with imagined counterevidence and initially defends itself by taming all recalcitrance, but is ultimately open to revision and theory change. Our concept of a natural kind like *whale* or a psychological or clinical phenomenon such as *depression* or *cancer* might be such a theory. A theory theory of concepts is, in a sense, an attempt to retrieve the realist insights behind the classical essentialist concept of a concept, noting that concept users have more faith in hidden essences underlying causal mechanisms than in observable superficial features or family resemblances.

One major problem with reducing concepts to theories is that theories themselves consist of, and hence presuppose the notion of, concepts. A theory theory of concepts would thus lead to either an infinite regress or circularity, if it identifies concepts with theories of one sort or other. Theory theorists might meet this charge by embracing concept holism: interdependent theories are determined by the entire system of beliefs held true by an individual. My concept *cancer* would be determined, not by its constituent elements or criteria, but by all the other theories and the inferential role of this cancer theory in producing my entire system of beliefs directly or indirectly involving cancer. This would entail the repugnant consequence that I cannot share my concept of cancer with anyone else who holds different theories and beliefs. This sort of holism would mean that any disagreement between two people on any topic entails that they are using different concepts. A second difficulty is that, like prototype theory, the theory theory has difficulty explaining compositionality (the pet-fish problem): combinations of concepts seem simple and straightforward compared to combinations of theories, something for which there is no good account.<sup>8</sup>

8. Fodor's informational atomism: Back to Hume? Over several decades Jerry Fodor has developed a provocatively "retro/remix" sort of theory of concepts building on what he now confesses to be Humean (psychological) billiard balls, setting his face against such major twentieth-century philos-

ophers of thought as Wittgenstein, Quine, Ryle, Davidson, Dummett, Putnam, and Peacocke, whom he lumps together as "conceptual pragmatists." According to Fodor's theory, our language of thought is made up of word-like mental particulars called "concepts," which are structureless, simple cognitive blobs directly referring to properties and entities in the world. Concepts, like Hume's "ideas," are mental representations that are not determined by our ability to do anything. Our possession of them need not be explained in terms of our knowing how to engage in verbal and nonverbal behavior. Most of these concepts are, according to Fodor, innate. And it is here that the archempiricist (copy of the sense impression) *idea* *idea* is cobbled together with the archrationalist Cartesian innatism and the semantic hook up between concepts and the world starts to look like an unexplained brute fact. Dog (an information-rich but unanalyzable concept) just ranges over real world dogs in virtue of being the mental particular that it is. "Cognitive processes are constituted by causal interactions among mental representations, that is, among semantically evaluable mental particulars" (Fodor 2006, 135). The information atoms that are Fodor's concepts, in turn, pick out external objects and properties and real relations. Fodor's concepts are avowedly mental particulars. This makes his view vulnerable to the concerns about communicability lying behind the private language argument.

9. The ability theory: We have seen that any variety of imagistic theory of concepts, as well as the holistic theory theory, leads to privacy and unsharability of concepts across individuals. Wittgenstein, while proving that no private system of rules can count as a language, also showed that mental representations are useless as explanations of what makes a word mean what it means. Mental-representation accounts of concept possession take concepts to be word analogues in a language of thought. (Dummett [1993b, 97] calls this "the code conception of language.") But if understanding is translating a string of words into a string of concepts, the same problems that arose with regard to the meanings of words would arise with regard to the meanings of concepts. Besides, saying, in an associationist way, that we understand the Sanskrit word *gauḥ* when the concept *cow* comes to mind in hearing those Sanskrit syllables is no help. "There is really no sense to speaking of a concept's coming to someone's mind. All we can think of is some image coming into mind which we take as in some way representing the concept, and this gets us no further forward, since we still have to ask in what his associating that concept with that image consists" (Dummett: 1993b, 98). Rejecting all Lockean representational theories of concepts, Dummett has proposed that concepts are best looked upon as recognitional capacities that, when exercised, constitute the understanding of sentences manifested in communicative practice. To have the concept of a mountain

lion is to be able to recognize one when one sees it, or at least to be able to justify accepting and rejecting statements using the word "mountain lion" as true or false. Since such abilities could be socially shared and publicly compared with norms, to identify concepts with such recognitional abilities would enable us to account for interpersonal concept sharing.

Of course, critics reject this ability theory on at least three powerful grounds. First, a concept (recall Frege's idea of the meaning of a predicate) is supposed to be applicable to or true of objects. But my ability to recognize marigolds is not true of those flowers. Second, abilities fail the compositionality test. Those with the ability to recognize pets and recognize fish do not necessarily recognize pet fish. Third, Fodor lambastes the recognitional-capacity account of concept possession as a kind of "concept pragmatism" that leads ultimately to "a kind of flirting with idealism" (2004, 31). In fact, a better charge against the ability view would be that it ignores the "inner story" and leaves no room for the distinction between acting socially as if one gets a concept and really getting (i.e., episodically having the right mental representation which is) a concept. While Dummett's embracing of meaning-theoretic antirealism tends to confirm the idealism alert, it is extremely hard to swallow the warning against idealism from Fodor, who is proposing to go back to Hume's private ideas theory of concepts (which led straight to phenomenalism). Against any broadly dispositional practical capacity or implicit knowledge theory of concept possession, Fodor says bluntly "concepts are mental particulars" (1998, 3). It should follow that word meanings are mental particulars. So the meaning of my words can never be the meaning of your words, because my mental particulars are not your mental particulars.

One thing we learn from a Dummettian ability theory's emphasis on the ability to manifest concept possession is this: Just as a theory of meaning should be a theory of understanding, similarly a theory of concepts should be a theory of concept possession and concept use. There is no point characterizing meanings or concepts in such a fashion (e.g., as involving some synaptic connections in the brain) that one cannot give any usable criterion for distinguishing someone who knows the meaning or possesses the concept from someone who does not. It remains to be seen whether Buddhist apoha theorists can provide this.

### PART 3

The reader could be forgiven for sensing a disconnect between the contents of parts 1 and 2. For the most part the modern theories of concepts

discussed in part 2 skirt the metaphysical issues that so exercised classical Indian philosophers. It might be thought that the downfall of classical definitionism explains this: that Indian realism requires real essences, so that in the wake of Wittgenstein's family-resemblance account, the field is left open to nominalism. But a sophisticated realism like that of Nyāya can agree that many of our concepts exhibit indeterminacy of application, typicality effects and the like, and still insist that some privileged set of concepts must carve nature at its joints (since otherwise there is no explaining the efficacy of concept-guided practice). And to say that nature has joints is to say that certain particulars by nature share something in common.

Perhaps it is thought that developments in modern logic—specifically the replacement of syllogistic logic by predicate calculus—have shown that belief in real universals rested on an improper understanding of the logic of our language. But even if we agree with Frege that predicate expressions are best thought of as functions or rules that take us from objects to truth-values, there is still the question how such rules come to be grasped.<sup>9</sup> The answer Quine (1953, 68) gives—by “inductive generalization”—marks him as what Armstrong (1978, 16) calls an “ostrich nominalist,” someone who refuses to countenance universals but sees no need to give a reductive analysis of our seeming commitment to them. For one could learn to use “red” by induction from ostended instances only if the instances pointed to by the teacher were taken by both teacher and pupil to resemble one another. And if the felt resemblance has no basis in reality, it is a major miracle that the lesson can be shared among all the speakers of a language. But if the basis of felt resemblance among the instances is some sort of real resemblance, then it requires explaining why this does not yield commitment to universals. For, argues the realist, any two things resemble one another in some respect or other, so the instances used in teaching would have to resemble one another in some particular regard—for example, in redness.

So the metaphysical concerns that fueled the classical Indian debate and shaped the apoha theory have not been resolved in recent Western philosophy, they have just been set to one side. Still, the issues involved in the recent discussions are not without importance to our understanding and appreciation of the apoha theory. An apoha-theoretic nominalist owes us something the realist is under no obligation to provide. The realist can claim that we learn to use “red” from the ostended instances because when we perceive those particulars we also perceive the redness that inheres in each. A nominalist who eschews all talk of real resemblances (e.g., by claiming that the perceiving of resemblance is due to conceptualization) must



instead posit concepts as devices that mediate between input from particulars and pattern-exploiting output (e.g., "learning by induction"). Then there will be questions to answer.

1. Are (at least some) concepts innate (Descartes, Fodor), or is concept possession something acquired only through experience? If the former, what explains the apparent congruence between innate concepts and the world? If the latter, how does this learning take place in the absence of real resemblances?
2. Are concepts mental particulars (Locke, Hume, Fodor), or are they to be understood instead as rules or schemata (Kant, Frege, similarity-based theories, theory theory)? If the former, how are we able to express our thoughts in a public language, and how does a particular have application to a many? If the latter, how are such rules mastered and how do we tell when such mastery has been achieved?
3. Are concepts the sorts of things that can enter into causal interaction with other things (Hume, Kant, prototype theory, etc.), or not (Frege, Peacocke). If the latter, then what explains the difference that concept possession makes to our worldly success? If the former, then if it is also true that only particulars are causally efficacious, how does a concept have application to a many?
4. How is the compositionality of concepts to be explained? How is it that from our grasp of such concepts as *pet*, *fish*, *blue*, and *lotus*, we are able to understand the expressions "pet fish" and "blue lotus"?

Even the most fully developed formulations of apoha theory fail to give clear-cut answers to all these questions. But since we wish to know not only what the apoha theorists said, but also whether what they said is philosophically significant, we shall want to work out, where possible, how they might answer them. The essays in this volume explore the apoha theory not merely as an historical artifact, but as a novel approach to an important philosophical problem. If this approach has real philosophical merit, it should yield viable answers to questions we might put to it concerning the concepts (*vikalpa*) that play a central role in the theory. At the same time, we will want to assess the theory's adequacy as a solution to the metaphysical puzzles that fueled the Indian debate over universals. Of these, perhaps the following are the most pressing:

5. Does the resort to exclusions in place of real universals succeed in explaining how particulars appear to naturally fall into kinds or classes

without in the end bringing in real resemblances? If so, how? If not, is there some account of why real resemblances need not be backed up by real universals?

6. Does the apoha theory really claim that the meanings of words are all negative in nature? If so, then how does it account for the fact that we take the meanings of words like "pot" and "yellow" to be positive? And how is it possible to arrive at anything positive entirely on the basis of negation? Does negation not require the existence of something positive to be negated? If the something positive is the particular, then won't the exclusion of what is distinct from that particular just give us back the particular itself and not something general? If the something positive is another concept, then insofar as concepts are really negations that are only taken as positive entities because of ignorance, how does this avoid circularity or infinite regress?



The papers in this collection attempt to answer these and related questions in two distinct ways: by examination of the historical record, looking at the works of individual apoha theorists and their critics, and by standing far enough back from the text-historical details to extract a somewhat idealized "apoha theory" and subjecting it to philosophical scrutiny. (Needless to say, each does some of both; but a given paper will inevitably put greater emphasis on one than on the other.) The first paper, by Tom Tillemans, draws a distinction between two types of apoha theory that is used extensively in other historically oriented papers. This is the distinction between "top-down" and "bottom-up" approaches. Tillemans points out that the Buddhist philosophical tradition within which the apoha theory developed (Yogācāra-Sautrāntika, the school of "Buddhist logic") draws a sharp distinction between the real particulars that make up the world and the general concepts we employ in thinking and talking about them. Given this radical scheme-content dichotomy, the question arises how speech and thought successfully mesh with the world. A top-down approach seeks to answer this question by starting with the resources of logic and language and showing how these can be used to pick out pure particulars when the latter lack all hint of generality or shared natures. A bottom-up strategy, by contrast, tries to bridge the gap by showing how the causally efficacious pure particulars could generate felt resemblances and thus give rise to general concepts. The top-down approach Tillemans identifies with Dignāga,

the founder of the Yogācāra-Sautrāntika movement. The bottom-up approach he associates with Dharmakīrti, the “commentator” who significantly reshaped Dignāga’s system.

While other authors see the top-down and bottom-up approaches as complementary, Tillemans does not. He claims that while Dignāga thought the nominalist could solve the problem of one over many by the ingenious use of two negations, Dharmakīrti essentially abandoned this approach in favor of a purely causal story that only pays lip service to the founder’s views about exclusion. On this view, Dignāga’s answer to (5) is that the realist’s universal can be replaced by a conceptually constructed exclusion of the other, so that nominalism can avoid commitment to both universals and real resemblances. But Tillemans is skeptical that this solution will work independently of the widely held predicate nominalism, according to which there are no common reals, only common terms. So he denies that the resort to exclusions succeeds by itself in allowing the nominalist to circumvent universals or real resemblances.

Tillemans is less skeptical about the bottom-up approach he finds in Dharmakīrti. On this approach, the particular with which one is in perceptual contact causes a mental image that one then mistakes for a class character. It is, in other words, just a brute fact that two distinct particulars can cause the same judgment (e.g., “this is blue”). This brute fact does include a subjective component—it is due to “beginningless ignorance” that we make the same judgment about what are actually quite different entities. This supposedly absolves Dharmakīrti from the charge of smuggling universals into his account. Still Tillemans concedes that this will make Dharmakīrti look like an ostrich nominalist to some. His reply is that Dharmakīrti might better be thought of as a “happy nominalist,” someone who feels no need to explain the utility of our judgments of sameness.

In distinguishing between the approaches of Dignāga and Dharmakīrti, Tillemans is principally interested in how these play out in their answers to (5) and only alludes in passing to some of our other questions, such as (4), the question about compositionality. (He takes Dignāga’s approach to fail this test, but is silent on whether or how Dharmakīrti’s causal approach fares any better.) But resemblances between his interpretation of Dharmakīrti and Hume would suggest that some answer to a question like (2), the question how a public language is possible if meanings are mental particulars, is called for. Other papers that follow Tillemans in his views about Dharmakīrti will have more to say about this.

Since the *apoha* theory seems to have begun with Dignāga, it is appropriate that the first detailed historical study in this collection be on the formu-

lation of the theory presented in the fifth chapter of his *Pramāṇasamuccaya*, the chapter entitled "The Examination of Apoha." Ole Pind's contribution situates Dignāga's views squarely within the context of a debate with the Nyāya-Vaiśeṣika school over the nature of knowledge derived from verbal testimony. In classical Indian epistemology, a major topic of dispute is the number and nature of the means of knowledge (*pramāṇas*). Nyāya holds that when the testimony of a trustworthy source produces a true cognition in the hearer, this constitutes a distinct kind of cognitive instrument that is not to be assimilated to inference or any other sort of reliable epistemic process. Dignāga disagrees, holding that verbally acquired knowledge is just the product of a particular kind of inference. This disagreement stems in large part from the fact that while Nyāya is realist about universals, Dignāga and his school deny that there are universals, on the grounds that being eternal they would be causally inert. So where Nyāya could claim that a kind term (e.g., "cow") denotes a particular inhered in by a universal (e.g., a cow inhered in by cowness), Dignāga must devise some other account of how general terms refer and how sentences can bring about cognition of facts. His basic answer to the question how words refer is that just as inference can only put us in touch with facts indirectly, so words cannot denote particulars, only objects in general, and that this can only be achieved through exclusion.

This position is grounded in the view to which Tillemans alluded in claiming that Dignāga's school employs a strict scheme-content dichotomy: the view that perception takes as object the unique (and hence unconceptualizable) particular, while inference (which now includes any cognition that is verbally expressible) has as object a conceptual construction that is routinely but wrongly taken to be the same mind-independent real that is cognized in perception. Nyāya, by contrast, maintains that it is one and the same object (e.g., a fire) that is cognized both by perception and by inference. It is Dignāga's strict separation of these two objects of knowledge, plus his rejection of real universals, that results in his having to give some account (according to Tillemans, a top-down account) of how scheme and content mesh. As Pind makes clear, Dignāga approaches this task by focusing on how the meaning of a word can be learned.

Dignāga argues that in the absence of universals, the only positive entity that remains to serve as the object of linguistic cognition, the particular, cannot be the referent, since the particulars in the extension of a kind term are potentially infinite in number. But Dignāga also observes that such an account of word meaning leaves unexplained an interesting phenomenon: that there is uncertainty as we descend the taxonomic tree

from determinable to determination, but not when we ascend the same tree. Suppose we know of some particular that it is a tree. This leaves room for doubt as to whether it is an elm, an oak, or a jackfruit tree; but from the knowledge that it is a tree, there is certainty that it is solid, a substance, and an existent. Dignāga is here employing the Nyāya taxonomy of universals, which starts with *existenceness* and descends through a variety of levels of determination to such particulars as a mango tree, an occurrence of scarlet, and the downward motion of a cow's tail. Now, any particular that is a mango tree is also a tree, a solid, a substance, and an existent. Hence, if the meaning of "solid" were the particulars that it denotes, then when we know of a particular that it is a solid, this should convey the information not only that it is a substance and an existent (ascending the taxonomic tree) but also that it is a tree and a mango tree (descending the taxonomic tree). In fact, it conveys the former but not the latter information. Hence, the meaning of a general term cannot be the particulars that are its extension.

Pind takes the point here to be that the grasp of word meaning must be, like what happens in inference, the grasp of something abstract, of a type and not of tokens. Awareness of the connection between word and object is not, properly speaking, inductive. But if there are no universals, what can this abstract object be? The suggestion is that since it cannot be anything positive, it must be a mere absence or lack. And here is a second place where it is clear Dignāga must have had Nyāya's taxonomic tree in mind. The idea is that since there can be no positive nature that is shared by all the mango trees, the abstract object that the word denotes must be what is picked out by the expression "that which is other than what is excluded by 'mango tree.'" The thought is simply that the determinable "tree" constitutes a field of determinations in which the nature of each determination ("elm," "oak," "mango," etc.) is fixed by contrast with the others in the field. The result is a mental construction (all mere absences being such) that can nonetheless be (mis)taken for something positive and quite real.

This approach may answer some questions, but it raises others. Pind takes Dignāga to have been chiefly concerned to answer question (6), the question whether word meanings are indeed wholly negative in nature (he answers in the affirmative), and thereby also to answer (5), how it is that particulars appear to fall into classes (through the generation of mentally constructed general natures). Dignāga saw confirming evidence for these answers in the fact that this approach also helps answer question (4) about compositionality. But although Dignāga uses the notion of a concept in his account, it is not clear just what he takes a concept to be. He appears to answer (2) by denying that concepts are themselves real particulars, but

then one wonders how he would answer (3), the question about causal efficacy. And then there are the further questions stemming from his official answer to (6), questions having to do with the fact that since Dignāga's purely negative account of word meaning seems logically equivalent to the positive account of the realist, this approach to avoiding universals seems to lead to either circularity or infinite regress. It is just this charge that his critics took up and from which many of the later apoha theorists sought an escape.

John Dunne's contribution to this volume examines Dharmakīrti's revised formulation of the apoha theory and assesses its success as a nominalism. Where Dignāga sought to account for our awareness of a world that is structured in the form of a taxonomic tree, Dharmakīrti tries to explain how a world of pure particulars could cause us to form the false but nonetheless useful judgment that things naturally form kinds. So Dharmakīrti's enterprise is appropriately labeled "bottom-up" by contrast with Dignāga's "top-down" approach. The question is whether these are complementary: can Dharmakīrti be seen as filling in important blanks in Dignāga's account, or is he rather deploying a rival theory under the cover of a strategic use of Dignāga's terminology? As Dunne explains it, Dharmakīrti's chief concern is to describe a process whereby a mental image that is copied from a perceptual cognition could come to be taken as resembling other perceptions in such a way as to give rise to the sense that their objects form a kind. Dharmakīrti's first answer is that this depends on the objects each being taken to perform some function in which we take an interest. Since the objects are in fact unique particulars, it is actually false that they share the common nature of performing that function. Each fulfills it in its own specific way; it is only our interest in that function that makes us overlook these mutual differences and judge the objects to be alike in this respect. All they actually share is their common difference from those things that do not perform that function. And this difference being a mere negation, there is no temptation to take it for a real universal or resemblance. So it begins to look like Dharmakīrti is after all trying to explain how, in the context of a given set of human interests, general natures could be constructed in an entirely negative way, thereby providing a psychological model of Dignāga's semantic theory.

Dunne considers the objection that this account appears to presuppose that distinct particulars can all perform a function that we take to be the same. If the many particulars that we call "fire" do not in fact share the common power to cause heat, why should it prove useful for us to perceive them as all alike in this respect? Dharmakīrti's answer is that it is just the

ultimate nature of each particular to cause an effect that we will judge to be the same. Dunne construes this appeal to the ultimate nature of things as a kind of concession that beyond a certain point the apoha approach must give way to the happy nominalist's way with the realist's one over many argument. This would lend credence to the claim that Dharmakīrti's apoha theory is importantly different from Dignāga's. But this might equally be taken as no more than Dharmakīrti's way of pointing out that since the ultimate nature of the particular is inexpressible, the question of what ultimately explains our judging one particular to resemble another is ill formed. For if it is true that all concepts involve exclusion, that no exclusion is ultimately real, and that all explanation involves conceptualization, then any attempt to explain our so judging that appeals to how things are apart from all conceptual construction is necessarily illegitimate. All this means is that our assessment of the apoha theory can only be based on its adequacy to what Dignāga and Dharmakīrti would call the conventional truth—how things are in a world that is in some sense already conceptually constructed.

On Dunne's account, Dharmakīrti gives the beginnings of answers to (2), the question whether concepts are mental particulars or rules, and (3), how concepts causally interact with other things. Dunne describes Dharmakīrtian concepts as "Janus-faced" insofar as they are at one and the same time mental particulars (copied images) and entities that apply to multiple instances (abstract types). If such a strategy can be made to work, it would effectively solve the dilemma posed by (3), the dilemma that abstractions lack causal efficacy while particulars lack generality. But can it be made to work? The idea seems to be that the mental image can be taken as resembling a multiplicity of perceptual images by virtue of its being indistinct or obscure. One might, though, wonder whether there could be such a thing as a visual image of a tiger with an indeterminate number of stripes. Dunne also suggests that the image's serving as a concept has to do with our taking it as resembling other images, and this would bring Dharmakīrti closer to thinking of concepts as rules or schemata. To the extent that exclusion can be seen as a mental operation, this way of thinking of concepts conforms well with Dignāga's negative semantics. But then it is unclear what role the mental image is to play in Dharmakīrti's theory of concepts. If mastery of a concept is the mastery of a kind of rule-governed behavior (namely, behavior that overlooks the differences among a class of particulars), then the particular mental image that is the other face of the Dharmakīrtian concept starts to look like something of an idle cog.

Dunne also represents Dharmakīrti as addressing (1), the question whether concepts are innate or acquired through experience. The answer

comes in the theory of the "imprint" (*vāsanā*), which Dharmakīrti uses to explain how we come to see certain things as resembling. Now to call this mechanism an imprint is to suggest that the disposition is one we acquired through prior experience. But Dharmakīrti is aware that not all concept acquisition can come through language-mediated learning, since prelinguistic infants and nonhuman animals are able to respond differentially to certain classes of stimuli. And so Dharmakīrti speaks of certain imprints as stemming from "beginningless ignorance." Dunne takes Dharmakīrti to have just two such imprints in his theory, namely, the general disposition to find similarity and the disposition to mistake representations for objects (i.e., the error of implicitly accepting a naïve realist view of perception). But it would seem Dharmakīrti needs far more specific resemblance-perceiving dispositions to make his account work: the infant must be predisposed to perceive distinct nipples as "the same," for instance. And the claim that such dispositions result from "beginningless ignorance" need not be construed as calling them innate, at least not in the sense that contrasts with calling them acquired. Buddhist tradition holds that the series of rebirths prior to the present life has no beginning. On this view, a given disposition might be present in each life in the series (and hence "innate" in one sense) and yet acquired through experiences had in a prior life (hence in some sense "learned"). In the Buddhist context it is not clear that question (1) has a clear sense.

Question (6) poses a dilemma for those who take seriously the claim that the meaning of a word is something wholly negative in nature, an exclusion. Negation requires a negandum, and presumably this must be something that is positive in nature. But if it is the real particular, then the two negations involved in the *apoha* formula "not non-cow" return us to the particular with which we began and not a class character. The alternative is to take "non-cow" to denote something positive that is shared among all the things not called cows. But the class of non-cows is so heterogeneous that it is hard to imagine their sharing anything except the property of not being cows, so either the account is circular or it leads to an infinite regress. This was the challenge Dharmakīrti's opponents raised for his formulation of the *apoha* theory. It, and Dharmakīrti's response, are the subject of Pascale Hugon's contribution to this volume.

As Hugon describes it, Dharmakīrti's response seems puzzling. He chooses the second horn of the dilemma, but asserts that the realist is faced with the same difficulty. For the alleged interdependence of the two classes *cow* and *non-cow* means that one lacks independent access to the meaning of "cow" just as much as to "non-cow." The realist has a ready response to this challenge though: they can claim that the presence of the perceptible



cowness in the cow and its absence from the non-cow give us a way of telling whether something belongs to the class of cows. But this then licenses Dharmakīrti to respond that the judgment of similarity performs the same function in the apoha theory. And since such a judgment relies not on real resemblances but instead on the interests of the subject, the apoha theory avoids commitment to entities that Dharmakīrti claims are ontologically problematic. So the *tu quoque* response is actually part of a complex and subtle dialectical strategy designed to bring out the extent to which apoha semantics mirrors that of the realist. In this respect, it might be observed, Dharmakīrti sounds very much like Dignāga, who quite cheerfully acknowledges that the “exclusion of the other” behaves just like a universal, all that is missing from the apoha theory being the realist’s ontological baggage. The difference with Dignāga lies in Dharmakīrti’s notion of the judgment of similarity. Hugon is herself silent on the question whether this is consistent with Dignāga’s account and whether it is rationally defensible.

The subject of Shōryū Katsura’s contribution is the account of the three meanings of the expression “exclusion of the other” (*anyāpoha*), which was first developed by the commentator Śākyabuddhi and further elaborated by Dharmottara. This account became increasingly important as the tradition begun by Dignāga sought to clarify and extend his and Dharmakīrti’s insights. The first thing that might be meant by *anyāpoha* is the unique particular that is the object of perception. This is “excluded from the other” in the sense that its being unique is just its being distinct from everything else. A second possible use of the expression is to denote the universal or object-in-general that is the intentional object of conceptual cognition (including inference and all other forms of linguistically mediated cognition). This involves “exclusion of the other” in the sense that all thought of kind membership involves differentiation. Dharmottara adds (no doubt with Dharmakīrti’s psychological model in mind) that the particular is the direct object of perception and the indirect object of conceptual cognition, while the object-in-general is the direct object of conceptual cognition and the indirect object of perception. This leads to the question how there can be coordination between these two forms of cognition. The answer is that the mental image copied from a perception is of the nature of the “exclusion of the other” in the sense that, due to the activation of impressions from past experience, it is such as to be incompatible with representations of things taken as dissimilar. And it is the excluding nature of this mental image that is the third sense of *anyāpoha*. This third sense is clearly related to the “Janus-faced” nature of Dharmakīrti’s concepts to which Dunne refers.

In assessing the apoha theory as a theory of human cognition, Katsura points out that even after Dharmakīrti’s emendations and Śākyabuddhi’s

clarifications, there remains the problem of explaining how things can appear to naturally fall into kinds (Dignāga's use of the taxonomic tree) or resemblance classes (Dharmakīrti's judgment of similarity). He suggests that in the end the apoha theorist must invoke human conventions. But in defense of this he cites the "Ugly Duckling" theorem of the theoretical physicist Satoshi Watanabe. This theorem concerns the intuitive idea that natural kinds are constructed out of resemblance classes: that, for example, two lotuses belong to a single kind because they share more predicates in common than do a lotus and a mugwort plant. But as Watanabe proved, this is false; there being infinitely many predicates that any two particulars (including a lotus and a mugwort plant) share, this way of understanding the sense that a lotus resembles another lotus more than it does any other entity cannot be made to work. So, the suggestion is, natural kinds must involve factors pertaining to the human subject; resemblance nominalism is a nonstarter.

Hugon showed that a deeper understanding of the apoha theory can be obtained by starting with its critics—in the case of her contribution, Kumārila and Uddyotakara. In his contribution to this volume, Masaaki Hattori investigates the deep and subtle critique of apoha developed by another Naiyāyika, Jayanta. One objection of particular interest is that the apoha theory is unable to account for the fact that in the expression "blue lotus," the denotations of the two words must share a common locus and be in the qualifier-qualified relation. This is said to be a problem for the apoha theorist because the Buddhist denies that absences are real entities, so that if what words denote are exclusions or absences, they can have no locus and enter into no relations. If on the other hand one holds that absences are reals, there will be the difficulty that two absences must have separate loci and likewise cannot be in the qualifier-qualified relation. The problem here is the one raised by question (4), only put in ontological form. Dignāga and Dharmakīrti both answer it by in effect making the particular the indirect object of the individual word. So the particular denoted by a use of "lotus" is differentiated from all those particulars that are nonlotuses, but among these are some that are nonblue; the use of "blue" in this use context serves to differentiate the denoted particular from those that are lotuses but nonblue.

Question (6) also receives much attention from Jayanta. For the realist about absences, an absence requires an absentee or object of negation, something positive that gives content to the negation. This leads to an objection first formulated by Kumārila, that the difference between the object of negation for "cow" (namely, the class of lions, tigers, elephants, squirrels, horses, zebras, etc.) and the object of negation for "horse" (namely,

the class of lions, tigers, elephants, squirrels, cows, zebras, etc.) is vanishingly small. For each class differs from the other just in containing one kind not contained by the other, and there are potentially infinitely many kinds of animals. Here we have an interesting variant on Watanabe's Ugly Duckling theorem, only used against the apoha theory.

Jayanta's discussion of the apoha theory includes what Hattori takes to be references to the views of Dharmottara, and Hattori outlines the contributions of this important thinker to the later tradition. Chief among these is the view that the denotation of a word can be neither existent nor nonexistent. This odd-sounding view is the consequence of what would become a kind of master argument for the apoha theory in the hands of Jñānaśrīmitra and Ratnakīrti: that since a word may be used in a particular context of use to denote either an existent or a nonexistent, the denotation of a word type can be neither and consequently must be a mere mental construction. Dharmottara's positive view is that word meaning is something mentally constructed and superimposed on the reals. Jayanta takes this to be significantly different from Dharmakīrti's view, but it is not entirely clear that the views cannot be reconciled.

Parimal Patil's first contribution to this volume is a discussion of what was probably the last Indian formulation of the apoha theory, that of Jñānaśrīmitra and Ratnakīrti. (His second contribution, available at the companion website for this volume, is a translation of Ratnakīrti's *Apoha-siddhi*.) One of their aims, he claims, was to deploy the apoha theory to develop a general theory of mental content—of the intentional object of such mental states as perception, inference, and verbal cognition. He sees the theory as involving a kind of paradigmaticism (like that of prototype theory), according to which one's command of the concept associated with the word "cow" involves the ability to recall a particular mental image on the basis of which one forms the appropriate exclusion class (*non-cow*) and its complement (the class of things like the paradigm, and hence cows). But, Patil claims, the success of the theory depends on both the failure of all the competing realist and nominalist theories and a number of questionable assumptions.

Among the latter is the assumption that there will be intersubjective agreement among the exclusion classes formed by the utterance of "cow" for English speakers, given that according to the theory these are formed through recall of a particular mental image. Since the image that occurs to each speaker will be distinct from that of any other speaker, Patil is here raising the "private language" horn of the dilemma in question (2). Now Dharmakīrti's response to this challenge (as Dunne and Hattori both dis-

cuss) is to invoke some common interest (such as the desire for milk) that all members of the exclusion class fail to satisfy. (Remember that failure can be construed as a lack or absence, something that Buddhists claim involves no ontological commitment.) Patil raises the question why this functionally determined exclusion class should agree across subjects, and he suggests that a modern and naturalistically inclined apoha theorist who appeals to natural selection to answer the question will quickly exhaust their explanatory resources and be reduced to saying that is just how it is. Presumably this is because a selectionist explanation will make reference to similarities across individual members of a species—for instance, the fact that all humans are mammals and so must share a taste for milk. And any such similarities that are appealed to in the explanation will themselves require further explanation. So the modern apoha theorist must either dogmatically assert that this is how things are or else embark on an infinite regress. Patil clearly finds the latter alternative uninviting, but one might wonder whether Jñānaśrīmitra necessarily agrees. As Patil discusses at the end of his paper, Jñānaśrīmitra does not think that any theory (including the apoha theory) can be ultimately true, but he nonetheless holds that a given theory might be more useful than its alternatives and so stand as a better formulation of the conventional truth than the other options. Moreover, Jñānaśrī also holds that there is a potentially infinite hierarchy of such theories, each standing as a better approximation to the ultimate truth than its predecessors.<sup>10</sup> So here too, as in the appeal to “beginningless ignorance,” there may be more tolerance for infinite regresses than one might expect.

Patil also raises a number of questions concerning the account of mental content he finds implicit in Jñānaśrīmitra and Ratnakīrti. Among these is the question whether a contemporary apoha theorist would want to take on board Jñānaśrī's claim that perceptual and inferential/verbal cognitions have distinct phenomenal characters. He takes this to be the claim that while perception and inferential/verbal cognition of, for example, a fire have the same conceptual content, they differ by virtue of their qualitative character as experiences, their “what-it-is-like-ness.” But it is not entirely clear that any apoha theorist says this. Dharmakīrti does say that a perception and the immediately following mental image that forms the basis for conceptual cognition differ in that the first is vivid while the second is indistinct, and this sounds like a difference in phenomenal character. But this is, according to all apoha theorists, a difference between a cognition without conceptual content and one with conceptual content—between a perception properly so called and what is sometimes called a perceptual

judgment. Now there is an interesting problem in the neighborhood, one that stems from the fact that on Dharmakīrti's formulation there must also be a phenomenal difference between the perceptual judgment and a cognition that is entirely inferential in nature. When your sense of vision comes in contact with the fire, there is first the mental image,  $M_1$ , the vivid and unconceptualized presentation of color and shape, and then a moment later the perceptual judgment that employs an indistinct copy,  $M_2$ , of this presentation to form the concept *fire*. When you later tell me about your experience, your use of the word "fire" arouses in me a mental image,  $M_3$ , that I use in working out what the fact is that you are reporting. Now we are told that while  $M_1$  and  $M_2$  differ in terms of relative distinctness, the latter is routinely mistaken for the former, so they must resemble one another in phenomenal character. On the other hand, no one ever mistakes  $M_3$  for either  $M_1$  or  $M_2$ , yet it still must resemble them in phenomenal character. So there does seem to be a puzzle here. But current cognitive science tells us that subjects are able to make more fine-grained discriminations on the basis of a presented image than on the basis of a recalled image (Metzinger 2003, 43–62). Since  $M_2$  and  $M_3$  are recalled images, with the former being the product of immediate recall, it is possible that differences in phenomenal character among the images are the result of such functional differences. Patil is of course right to wonder whether this is of strictly semantic significance. But an adequate theory of concepts must account for more than just the semantic facts.

Patil also makes much of Jñānaśrī's claim that there can be no adequate account of mental content insofar as nothing that will serve as representational content (something that is necessarily conceptual in nature) can fully capture what is immediately given in raw experience. One might wonder, though, whether this is not just another way of making the point that Dunne has in mind when he refers to the "Janus-faced" nature of Dharmakīrti's concepts. One point it is not always easy to keep in mind is that the representationalist theory of perception was not the consensus position in classical Indian philosophy. Modern theories of concepts tend to presuppose such a view and with it the idea that mental content is necessarily "in the head." Externalist views of content and a direct realist view of perception are still minority views. In the Indian context, on the other hand, non-Buddhist philosophers consistently held the direct realist view, and even among Buddhist philosophers representationalism was controversial (Dhammajoti 2007, 136). In this context it seems plausible that the insistence one finds in Dharmakīrti and his successors that ordinary people routinely mistake the object of inferential cognition for the

object of perception might reflect an older tradition of representationalist insistence that perceptual judgments cognize the external object only indirectly through a kind of automatic inference. It is that older tradition that seems to lie behind Dignāga's insistence that there are two distinct kinds of objects of cognition. Perhaps it is likewise at work in Jñānaśrī's claim that no single account can be given of cognitive content.

P. K. Sen's paper begins with a survey of some of the more important critiques of apoha theory by such thinkers as Uddyotakara, Kumārila, Jayanta, and Vācaspati. He then develops his own set of objections to the theory, along lines not unlike those of its Nyāya critics. Unlike Pind, Sen does not interpret Dignāga as having started with an argument against the existence of universals. This makes an important difference to how we assess his critics such as Uddyotakara and Kumārila, many of whose objections otherwise seem to miss the point. If all Dignāga were doing was developing an exclusion-based semantics that is formally equivalent to the realist semantics of Nyāya or Mīmāṃsā, and there were no independent reason to think that the meaning of a kind term (something that is distributed across a range of particulars) could not involve real universals, then we would be inclined to agree with Uddyotakara, Kumārila, and the like that the realist theory is preferable to the more cumbersome and counterintuitive apoha theory.

As Sen's discussion brings out, once the Buddhist argument against real universals was on the table, the debate over the apoha theory centered largely on issues of ontology. Uddyotakara's objections, for instance, all presuppose that whatever serves as the semantic value of an expression must be some real thing; he quite fails to grasp that apoha theorists are struggling to articulate an alternative vision according to which abstract objects are mental constructs. Of course Sen himself sees this, and so he neatly turns the tables on the apoha theorist, who sees the twin problems of causal efficacy and instantiation as the Achilles's heel of a realist semantics. To the apoha nominalist who asks how an eternal universal can cause anything and how a single universal can be present in many locations simultaneously, Sen asks how something imaginary can do either one. He concedes that Nyāya realism comes at an ontological price, but he insists that the Buddhist nominalist alternative has its costs as well. And he warns against dismissing the realist view based on a caricatured understanding of realism. Sen points out, for instance, that Nyāya does not posit a real universal for every general term; their semantics makes do quite well with a basic stock, through the combination of which other meanings are "cooked up." And the apoha theory is, he says, vulnerable to the problem

of typicality effects common to prototype theory, proxytype theory, and other exemplar-based theories of concepts. If giving milk and serving as a draught animal are the functions that a Dharmakīrtian cites as the basis for the construction of the exclusion "not non-cow," then what are we to make of buffaloes and of aged and infirm cows? While Sen does not point this out, a sophisticated Nyāya theory of universals does contain resources for handling this problem: although the basic stock of real universals might be inflexible in their application, the problems associated with typicality effects might arise only with kind terms whose semantic values involve combinations of that basic stock, with flexibility built into the principles of combination. But then the apoha theorist might use a similar strategy to answer Sen's buffalo objection.

Through his examination of accounts of the apoha theory in critical Nyāya and Mīmāṃsā texts, Sen often manages to cast the theory in an interesting light. One instance of this is Śrīdhara's account of Dharmottara's arguments for apoha. One such argument is, according to Śrīdhara, that since it is agreed that goats, horses, camels, and squirrels have nothing in common, they can be considered as constituting a class only under the expression "non-cows." But according to the apoha theorist, since there are no real universals such as cowness, the things called "cows" likewise have nothing in common. So it stands to reason that these too should be taken as constituting a class only under the expression "not non-cow." This interesting and quite sensible-sounding argument is one that is more readily apparent in Śrīdhara's description than in any number of more "friendly" sources.

Georges Dreyfus's paper concerns Dharmakīrti's formulation of the apoha theory, only this time as seen through the eyes of some of his Tibetan commentators. Dreyfus holds that central to Dharmakīrti's overall theory is the role played by concept formation. But he begins by examining whether Dharmakīrti may be thought of as a resemblance theorist, someone who holds that we judge particulars to belong to a kind because they all resemble one another. Dharmakīrti does say that we naturally judge certain particulars to resemble others. Dreyfus points out, though, that he carefully refrains from asserting that such judgments are based on real resemblances among the particulars, an assertion that would lead to a failed nominalism. Instead, he seeks to explain such judgments in terms of the fact that the particulars involved perform a similar function (in the case of the stock example of the medicinal herbs, that of reducing fever). This makes it clear that Dharmakīrti wishes to bring out the interest-relative nature of perceived resemblance. It also leaves room for appeal to either

karmic imprints or the forces of natural selection to account for a widely shared primitive similarity space, something that can then be employed to explain how we might construct more complex, language-mediated concepts. Crucial to this account, according to Dreyfus and some Tibetan commentators, is once again the "Janus-faced" nature of Dharmakīrtian concepts, which are at once mental particulars with unique and inexpressible phenomenal content and exclusions that are multiply instantiable. As such they are meant to perform the crucial function of mediating between perception and thought, thereby explaining how thinking can be constrained by reality despite its being incapable of capturing reality.

Dreyfus is nonetheless concerned about what he takes to be the seeming paradox at the heart of Dharmakīrti's system, that the (presumably correct) judgment "that is a cow" is as much in error as "that is Santa Claus" or "that is a garland of sky-flowers." And of course this worry is not misplaced; non-Buddhist philosophers regularly beat up on the apoha theorists on just these grounds. To dispel the paradox, Dreyfus recommends the distinction certain Tibetan commentators drew between precritical application and critical examination. In the former context the judgment "this is a cow" is deemed correct. In the latter it is erroneous, not only because there is no such thing as the property of being a cow but also because the subject of predication, the "this," is likewise a fictional construction. Critical examination nonetheless helps us see how such constructions as *cowness* and the spatiotemporally extended substance denoted by "this" are causally connected to real particulars, thereby explaining how a judgment that could not possibly be veridical can nonetheless contribute to successful practice. This, it might be added, could be usefully compared to our present understanding of color perception. While we perceive physical objects as colored, we know that mind-independent reality contains no such thing as color. It is nonetheless possible to explain how the mind-independent nature of physical objects causes us to perceive color and through this causal link also explain why judgments of color should have pragmatic value.

The gap between sensation and thought is also the subject of Jonardon Ganeri's essay. But Ganeri's contribution is among those that stand back from the historical details of the apoha theory and attempt to evaluate it on strictly philosophical grounds. In this case the assessment centers on the question whether the apoha theory has anything to offer to current efforts to close the gap between conceptual and nonconceptual representation. Taking Austen Clark's work on sentience as a model of such efforts, he spells out several places where the notion of exclusion or apoha might do important work. One such place arises because there is no straightforward



correlation between stimuli and the qualities presented in sensation. In the case of color vision, for instance, any number of distinct combinations of wavelengths will all produce the experience of seeing red. What this is taken to show is that presented qualities are sorted into kinds on the basis of their place in a discrimination ordering. This, Ganeri points out, is tantamount to claiming that being red is just a matter of being excluded from what would be deemed nonred. To this it might be added that this is just the insight that seems to have been behind Dignāga's claim that determination under a determinable is fixed by contrastive relations.

A second place where Ganeri sees an opening for the apoha theory is in explaining the ability of sentient creatures to assign perceived features to places. This ability is crucial to the "bundling" of qualities that bundle theorists like modern empiricists and Buddhist reductionists employ to bridge the gap between sensation and thought. If space were given as just another quality alongside red and sweet, it could not be used to bundle red and sweet together in the construction of quality clusters (an ability that is crucial in turn to the construction of substances such as tomatoes). Ganeri suggests that the key difference lies in the fact that a quality like red has an incompatibility range—the instantiation of red at a particular place is incompatible with the instantiation of such other qualities as blue and yellow—while there is nothing comparable for places. So the ability to do something akin to referring to a particular place, an ability that a sentient creature of any degree of complexity must have, is likewise to be grounded in a kind of discriminatory exclusion. Ganeri worries, however, that Dharmakīrti's idealism might get in the way here. If, as most scholars now believe, Dharmakīrti's final position is that of Yogācāra, then he must deny the existence of space and consequently must explain the appearance of space as a matter of spatial qualities being given as part of the flux of sensory impressions. But this last bit need not follow. There is much evidence from cognitive science that the spatial organization of sensory data proceeds on the basis of action simulation on the part of the sentient creature: "here" is distinguished from "there," for example, on the basis of what can be done with and without locomotion. And it might be possible for the idealist to simply appropriate this way of individuating apparent places, given that actions come with their own presented qualities.

A third place where Ganeri thinks exclusion might help explain how sensory capacity could give rise to something like our conceptual scheme is in the ability to see the world as containing enduring, reidentifiable physical objects. This ability involves resources that are clearly unavailable at the level of pure sentience. How is the gap crossed? Ganeri points

out that a creature with the ability to assign features to places will also be able to detect places that lack the absence of a feature. And, he points out, "no absence of pot here" is transformable, through the canceling of the two negations, into "this is a pot." The latter judgment clearly involves the concept of a reidentifiable particular. The suggestion is thus that our scheme of enduring substances can be constructed out of perceptual contents through the judicious application of exclusion. Having thus indicated places where the notion of exclusion might play a useful role in a modern gap-closing exercise, Ganeri goes on to explore ways in which specific details of Dharmakīrti's account can be tied to such an account. On this view the apoha theory is more than an interesting historical artifact.

Amita Chatterjee is similarly sanguine about the prospects for a cognitive science account of understanding that is informed by insights from the apoha theory. One current debate in cognitive science concerns whether cognition should be understood as a computational process (akin to the operations of a digital computer) or instead should be seen in terms of the dynamic interaction of the perception and action systems of an ecologically situated organism. One question for dynamicists or noncomputation-alists is whether any theory of cognition can dispense entirely with mental representations (these being something that computational accounts seem well suited to account for). Chatterjee suggests that Dharmakīrti's formulation of the apoha theory might be used to sketch a noncomputational theory of mental representation, thereby enhancing the plausibility of the moderate form of noncomputationalism, which agrees with computationalism that mental representations play a role in higher-order cognition. This immediately suggests an answer to question (2), the question whether concepts are mental particulars or rules. A computational approach to representations sees them as rules or schemata (namely, the algorithms involved in computation), so if the apoha theory suggests a noncomputational approach to representations, then this option is ruled out. But the dynamicism of noncomputational approaches likewise rules out the possibility that a representation is a kind of mental particular. On a noncomputational approach, there is nothing that a concept actually is; our notion of a concept merely picks out one facet of a complex dynamic process.

This likewise suggests an answer to question (1), whether concepts are innate or learned. Chatterjee makes use of the notion of a Gibsonian affordance, which is a perceivable possibility for action. Affordances are relational properties that obtain between an organism and its environment, in this respect resembling the interests that Dharmakīrti claims shape the

formation of an exclusion class (such as the interest in load-bearing animals that helps shape the concept cow as what is other than those things that fail to satisfy this interest). Now some affordances are sufficiently stable features of the organism-environment relation that natural selection can operate to foster the appearance of resemblances among the particulars involved. This process yields concepts that are innate for the species, the equivalent of Dharmakīrti's imprints due to "beginningless ignorance." But other affordances are more ephemeral. These will be the source of learned concepts, some of which are then shared through the medium of language. This account thus appears to resolve the old debate between rationalists and empiricists in favor of rationalism, since it holds that at least some concepts are innate. But as the notion of an affordance makes clear, the innateness at work here holds only at the level of the individual organism, not that of the species. From the perspective of the species, all concepts are acquired through interaction with the environment. Completely dissolved is rationalism's mystery of a preestablished harmony between mind and world.

The deepest source of affinity between the dynamicist approach and the apoha theory lies, for Chatterjee, in their shared conviction that "Objects in the world are created in stages by dynamic interaction between organisms and the world." What Dharmakīrti's formulation offers is a way of spelling out how these stages lead from the raw, unconceptualized perception of unique particulars, through the (protoconceptualized) perception of primitive resemblances such as of color, to full-fledged perceptual judgments about enduring substances such as pots and cows. Because this is a stage-wise process, representations are required for the account to work. Dharmakīrti offers a means of understanding in a dynamicist way how one set of interactions might come to be treated as perceptual input—serving as representations—at the next higher level. On this understanding of Dharmakīrti's bottom-up approach, apoha is not a logical operation we perform on sensory input, but a node in the interactive interplay of organism and environment.

A final point worth remarking on is Chatterjee's suggestion that the apoha theory might be made to work without the scheme-content dichotomy that was built into it from its inception in Dignāga's thought. This would require restricting the theory to what Buddhists call the conventional truth. Such an apoha theory would then be, like dynamicist theories of cognition, fully compatible with Putnam's internal realism. And it would thereby avoid all the difficulties that come with a strict separation

between scheme and content, while at the same time avoiding the problem of relativism. To this it might be added that Madhyamaka appropriations of Dharmakīrti's thought aim to accomplish just this. Mādhyamikas routinely deny that there is such a thing as the ultimate truth; yet Dharmakīrti's formulation of the apoha theory makes full use of the distinction between conventional and ultimate truth. So when a Mādhyamika such as Śāntarakṣita claims that the apoha theory gives the best available account of human cognition, this is tantamount to saying that the ultimate-conventional distinction holds only conventionally.

The next two papers, by Bob Hale and Brendan Gillon, take up a characterization of the apoha theory (first developed in Siderits 1982) that includes the claim that the characteristic apoha expression "not non-cow" involves two distinct kinds of negation: verbally bound negation ("that is *not* non-cow") and nominally bound negation ("that is not *non-cow*"). On this understanding of apoha, the use of two distinct negations is meant to answer the nongeneralization difficulty first raised by Kumāṛila (see Hugon, this volume): how, in a world of particulars, can reference to a real yield something general in nature? If the two negations involved in the apoha expression are of the same sort, then it seems they should cancel out, in which case we return to the particular with which we began. The suggestion is that the combination of two different types of negation does not obey the classical principle of double negation, so the generalization required for the apoha theory to work is logically feasible.

In his critical assessment of this strategy, Hale begins by raising what he calls the problem of compositionality, but this is a different problem than that of "blue lotus" and "pet fish" discussed under the rubric of question (4). He takes the central claim of the apoha theory to be that "*x* is *P*" is to be analyzed as "*x* is not non-*P*," which leads to the difficulty that since the latter involves a complex expression in which "*P*" is a constituent, the composition principle (the principle that the meaning of a complex expression is a function of the meanings of its constituents) must be violated. This is in fact the circularity objection of Kumāṛila and Uddyotakara discussed by Hugon. In exploring possible responses to this objection, Hale describes what he calls a nominalistic compositional semantics, which has as primitive expressions only logically proper names, the basic logical operators (including verbally bound negation), plus a distinctness predicate that functions like nominally bound negation for proper names. Hale grants that this strategy might be made to work if we equip the language user with something like Dharmakīrti's "indistinct" mental images. But he points out

that then it is the mental images that do all the nominalistic heavy lifting (functioning as they do much like Locke's "abstract ideas"), in which case the resort to the two-negations strategy seems unnecessary.

Hale goes on to point out that the appeal to interests that is at the heart of Dharmakīrti's causal story appears to require that interests be repeatable kinds. In this case the nominalist is back to the usual dilemma of either building in real resemblances at ground level or else embarking on an infinite regress. In the end, Hale thinks the best the apoha nominalist might be able to do is simply claim that while "x is not non-P" is not the correct analysis of "x is P," it is ontologically less committed than "x is P" and so can be used to replace "x is P." But, it could be pointed out, this takes the apoha theory back to the most minimal characterization of Dignāga's formulation. What other resources the apoha theorist might summon to try to answer the many objections to which Dharmakīrti was responding, Hale does not say.

Brendan Gillon's contribution sets out to sketch a formal semantics for apoha nominalism interpreted in accordance with the two-negations idea. His result is negative: with the two negations interpreted as set-theoretic internal and external negation, no combination of the two yields the desired result that shows how cognition can proceed from the particular to the general. Other authors claim there is no evidence that any apoha theorist had the two-negations strategy in mind. Gillon is silent on this historical question. His conclusion is that if they had meant the apoha theory to work in this way, there is nothing in modern semantics that would vindicate their hypothesis.

Mark Siderits takes stock of his understanding of the apoha theory in light of the criticisms of others on both historical and theoretical grounds. To do so he spells out in considerable detail what he takes the bottom-up approach of Dharmakīrti and his followers to amount to when we stand back from the historical details and perform a rational reconstruction of the theory. This yields a response to question (2) that, like that of Chatterjee, rejects the dichotomy of concepts as either mental particulars or schemata. Thus, in response to the criticism that Dharmakīrti only gives us what Dummett calls a "code conception" of language (in which words serve as code for communicating ideas, which are private states of the speaker), the apoha theorist would say the concept of a concept is the reification of a dynamic process in which mental particulars play a role and which can be analyzed post hoc in terms of rule-governed behavior. This dynamic process is consistent over time for a cognizer, as well as across cognizers, because it leads to action that satisfies the interests of cognizers.

Siderits also suggests a response the apoha theorist might make to the common charge at the heart of question (5), that the theory inevitably smuggles in universals in the guise of real resemblances. This response involves what might be called a strategy of endless deferral, in effect conceding that the theory involves an infinite regress but denying that it is vicious. But it also tackles the question how there can be causation without universals, given that causal laws state relations between *kinds* of entities. The proposed response suggests that causation can be treated by the Buddhist nominalist as something to be reduced to more fundamental entities, in effect adding causation to the long list of things the Buddhist calls mere conceptual fictions. But this move will no doubt appear unsettling to those who see causation as crucial to the bottom-up account: how is the account to work if there really are no such things as causal laws?

Also addressed is question (6), the question whether the two negations involved in the apoha theory are any more than a gimmick. Siderits concedes that the two-negations strategy does not by itself explain how one can come to exhibit mastery of a concept, be it one that is "innate" (such as *yellow*) or "acquired through experience" (such as *pot*). Apoha theorists of a Dharmakīrtian stripe explain such mastery in terms of the interest-guided employment of mental images. He suggests, though, that the appeal to two distinct negations does respond to a very different question: not the epistemological question of how one masters a concept, but the metaphysical question of how there can be exploitable patterns in a world of unique particulars. The suggestion is that apoha theorists hit upon the idea of deploying a combination of verbally bound and nominally bound negations as a way of showing how Nyāya's taxonomic tree could be retained in a nominalistic universe. Of course Siderits must now concede that there is no historical evidence directly supporting his claim. But he does point to an episode in later Nyāya history that includes all the major features of his model of apoha, including a two-negation strategy linking particulars and universals and an objection based on the nongeneralization problem.

Finally, Parimal Patil's translation of Ratnakīrti's *Aphasiddhi*, or "Demonstration of Exclusion," is available at the companion website for this volume, [www.cup.columbia.edu/apoha-translation](http://www.cup.columbia.edu/apoha-translation). Patil has chosen to give his translation the title "Without Brackets." In most translations of Indian philosophical texts, one encounters a great deal of material enclosed in square brackets. This material is there for a good reason. Classical Indian philosophers placed such value on conciseness of expression that their writings are all but unintelligible to those not well versed in the tradition in which they worked. Consequently, a literal translation of the

text is likely to be of little or no use unless supplemented by a considerable amount of background material, and many scholars have elected to use square brackets to indicate that supplementary material is being inserted. But this device interferes with comprehension of the argument, particularly for nonspecialists. Patil (like some other scholars) has chosen to avoid their use while still providing the material necessary to make the text intelligible. It will still prove challenging. But the hope is that those with little or no prior exposure to Indian philosophical texts and traditions will still be able to follow the argument after having read the essays that precede it in this volume.

## Notes

1. The Sanskrit term *apoha* means "exclusion." The *apoha* theory is the theory that the meaning of a general term is "the exclusion of what is other." We will be using the word in *italicized* form, just as the word *karma*, which is Sanskrit in origin, is now used without italics.
2. "sarva eva hi bhāvā svarūpa-vyavasthitayaḥ. te nātmānaṃ pareṇa miśrayanti. . . na hi sambandhināpi anyenānye samānā nāma tadvanto nāma syuḥ. bhūtavat kaṇṭhe guṇena" (PVSV ad PV I.40–41, ed. Gnoli, 24.25–25.6). Kaṇṭhakagomin explains that the reference here is to the popular religious practice of stringing together idols of the planets and other divinities.
3. At least not on the commonsense view. Buddhists maintain that strictly speaking what we call an apple is just a bundle of quality particulars such as the sweet taste, the red color, and so on. For them the commonsense view that there is no sweet taste without the apple that has it is only conventionally true.
4. Whether such "cross-cutting" disqualifies both the properties or only one of them, and whether the neat ontological hierarchy that is presupposed by this universal blocker is integral to a realist metaphysics, have been the subject of much contemporary debate (see Shastri 1964; Mukhopadhyaya 1984; and Ganeri 2001).
5. PV 1.152 (Gnoli): "na yāti na ca tatrāsīd asti paścān na cāṃśavat / jahāti pūrvam nādhāram aho vyasanasaṃtatiḥ" (Gnoli 1960).
6. See *Nyāya-kumuda-candra*, vol. 2, 560–1 (Kumar 1939, 1941). Versions of this theory were adopted by followers of Rāmānuja (Qualified Monist Vedānta) as well as by Mādhdva (Dualist Vedānta) logicians. Vyāsātīrtha of the latter school clarified how a single resemblance can reside, as it were, with one leg in the ressembler and with another leg simultaneously in numerous other similar particulars. The category of resemblance admitted by these philosophers is very different from the resemblance admitted by Prābhākara Mīmāṃsakas, for the latter were realists about universals, while the Jainas and the Mādhdvas rejected, as logically redundant, both universals and inherence. The only difference between Prābhākara and Vaiśeṣika, as regards universals, centers on their conceptions of inherence.

7. This, by the way, is almost suggested by the Buddhist author Arcāṭa in his HBT commenting on Dharmakīrti's HB: that a *vikalpa* is nothing but an unclear copy of a sensory image. At HBTĀ 287, Durvekamiśra, in his gloss on Arcāṭa on Dharmakīrti, says explicitly that the false impression of being determinate is given by the faint and unclear concept thanks to the nonconceptual inner perception which takes place at the same time, implying thereby that the conceptual awareness borrows its seeming clarity from the nonconceptual awareness.
8. Compositionality, incidentally, was also a concern for Dignāga, whose discussion of the expression "blue lotus" is motivated by anxiety to distinguish coreferential terms from same-sense-bearing synonyms. Apoha theory could not strictly speaking be a theory theory of concepts, since it does not appeal to ontologically intrinsic essences at all; but the functional and dynamic aspects of a set of concomitance rules and a pragmatic-predictive capacity are features that apoha theorists also ascribe to concept employment.
9. For Frege as a major source of modern nominalist bias, see Bergmann 1958.
10. For more on Jñānaśrīmitra's contextualist semantics, see McCrea and Patil 2006.