JONARDON GANERI

INDIAN LOGIC

1 ARGUMENTATION WITHIN DIALECTIC AND DEBATE: PRAGMATIC CRITERIA FOR GOOD ARGUMENTATION

1.1 Early dialogues: information-seeking, interrogation and cross-checking

The intellectual climate of ancient India was vibrant, and bristled with controversy. Debates were held on a great variety of matters, philosophical, scientific and theological. Quite soon, the debates became formal affairs, with reputations at stake and matters of importance in the balance. Already in the *Brhadāraṇyaka Upaniṣad* (c. 7^{th} century BCE), we find the sage Yājñavalkya being quizzed by the king's priestly entourage on tricky theological puzzles:

> Once when Janaka, the king of Videha, was formally seated, Yājñvalkya came up to him. Janaka asked him: 'Yājñvalkya, why have you come? Are you after cows, or discussion about subtle truths?' He replied: '*Both*, your majesty.' (BU 4.1.1).

What followed was a question-answer type dialogue in which Janaka interrogated the sage, not only to solicit information but to test Yājñavalkya's mettle. The sage had earlier granted Janaka a wish, and the wish he chose was the freedom to ask any question at will. Yājñavalkya was not to be released from this wish until he had fully satisfied Janaka's probing inquiry:

> [Janaka] 'Here, sire, I'll give you a thousand cows! But you'll have to tell me more than that to get yourself released!' At this point Yājñvalkya became alarmed, thinking: 'The king is really sharp! He has flushed me out of every cover.' (BU 4.3.33–4).

It is in fact a characteristic of the earliest recorded debates that they take the form of question-answer dialogues. As a form of debate, the goal of a question-answer dialogue is not restricted merely to one party soliciting information from another, for there are, as this dialogue shows, elements too of testing out one's opponent and cross-checking what he says. A particularly important early question-answer dialogue is the *Milinda-pañha*, or *Questions of King Milinda*. It records the encounter between a Buddhist monk Nāgasena and Milinda, also known as Menander, an Indo-Bactrian king who ruled in the part of India that had fallen under Greek influence at the time of Alexander's Indian campaign. The document dates from

around the first century CE, although Milinda's reign was 155–130 BCE. At the outset, Nāgasena insists that their dialogue is conducted as scholarly debate and not merely by royal declaration[1] —

> King Milinda said: Reverend Sir, will you discuss with me again?
>
> Nāgasena: If your Majesty will discuss (*vāda*) as a scholar, well, but if you will discuss as a king, no.
>
> Milinda: How is it then that scholars discuss?
>
> Nāgasena: When scholars talk a matter over one with another, then is there a winding up, an unravelling, one or other is convicted of error, and he then acknowledges his mistake; distinctions are drawn, and contra-distinctions; and yet thereby they are not angered. Thus do scholars, O King, discuss.
>
> Milinda: And how do kings discuss?
>
> Nāgasena: When a king, your Magesty, discusses a matter, and he advances a point, if any one differ from him on that point, he is apt to fine him, saying "Inflict such and such a punishment upon that fellow!" Thus, your Magesty, do kings discuss.
>
> Milinda: Very well. It is as a scholar, not as a king, that I will discuss. (MP 2.1.3).

Vāda, the type of dialogue Nāgasena depicts as that of the scholar, is one in which there are two parties. Each defends a position with regard to the matter in hand; there is an 'unravelling' (*nibbeṭhanam*; an unwinding, an explanation) and a disambiguation of the positions of both — a process of revealing commitments, presumptions and faulty argument; there is also a 'winding up' ending in the censure (*niggaho*; Skt. *nigraha*) of one party, a censure based on reasons he himself will acknowledge (*paṭikamman*; 're-action', rejoinder). This is a species of the persuasion dialogue, a 'conversational exchange where one party is trying to persuade the other part that some particular proposition is true, using arguments that show or prove

[1] A similar distinction, in the types of scientific debate held between physicians, will be drawn a little later by Caraka, a medical theorist, and an important source of information about ancient Indian logic. He says, in an echo of the *Meno* 7.5 c–d, that debate (*saṃbhāṣā*) among specialists is of two types — friendly (*sandhāya*) and hostile (*vigṛhya*). See *Caraka-Saṃhitā* 3.8.16–17 and Ernst Prets, 'Theories of Debate, Proof and Counter-Proof in the Early Indian Dialectical Tradition', in Balcerowicz, Piotr & Mejor, Marek eds., *On the Understanding of other cultures: Proceedings of the International Conference on Sanskrit and Related Studies to Commemorate the Centenary of the Birth of Stanislaw Schayer*, Warsaw 1999. (Warsaw: Oriental Institute, Warsaw University, 2000).

to the respondent that the thesis is true'[2]. Indeed, it would seem to be the species that has come to be known as the *critical discussion*, a persuasion dialogue in which the conflict is resolved 'only if somebody retracts his doubt because he has been convinced by the other party's argumentation or if he withdraws his standpoint because he has realized that his argumentation cannot stand up to the other party's criticism'[3]. Not every persuasion dialogue need end in one party recognising defeat, for an important function of the general persuasion dialogue is to be maieutic, helping each side to clarify the nature of their commitments and the presuppositions upon which their positions depend.[4] In the to-and-fro of such a dialogue, each party is allowed to retract earlier commitments, as it becomes clear what the consequences of such a commitment would be. This maieutic, clarificatory function of a dialogue is perhaps what Nāgasena intends when he speaks of an 'unravelling', and it seems clearer still in his characterisation of 'investigation' (*vikāra*) as a 'threshing-out':

> Milinda: What is the distinguishing characteristic, Nāgasena, of reflection (*vitakka*)?
>
> Nāgasena: The effecting of an aim.
>
> Milinda: Give me an illustration.
>
> Nāgasena: It is like the case of a carpenter, great king, who fixes in a joint a well-fashioned piece of wood. Thus it is that the effecting of an aim is the mark of reflection.
>
> Milinda: What is the distinguishing characteristic, Nāgasena, of investigation (*vikāra*)?
>
> Nāgasena: Threshing out again and again.
>
> Milinda: Give me an illustration.
>
> Nāgasena: It is like the case of the copper vessel, which, when it is beaten into shape, makes a sound again and again as it gradually gathers shape. The beating into shape is to be regarded as reflection and the sounding again and again as investigation. Thus it is, great king, that threshing out again and again is the mark of investigation.
>
> Milinda: Very good, Nāgasena. (MP 2.3.13–14).

So it is through reflection and argumentation that the parties to an investigation together thrash out a position. Nāgasena tells us very little about

[2]Douglas Walton, *The New Dialectic: Conversational Contexts of Argument* (Toronto: University of Toronto Press, 1998), p. 37.

[3]Frans van Eemeren & Rob Grootendorst, *Argumentation, Communication and Fallacies* (Hillsdale: Lawrence Erlbaum Associates,1992), p. 34.

[4]Walton (1988: 48).

the sort of argumentation that is appropriate, and we can learn little more about argument within persuasion dialogues from the *Questions of King Milinda* (although Milinda's repeated request to be given an illustration is suggestive of the importance that would later be attached to the citation of illustrative examples in good argumentation; see §1.3 below). And yet there is still something to learn. For the dialogue of the *Questions of King Milinda* is not, contrary to Nāgasena's initial statement, a straighforwardly scholarly debate, but proceeds instead with his being interrogated at the hands of Milinda. Ostensibly Milinda wishes to be informed as to the answer to a range of thorny ethical and metaphysical questions, but his questioning is not so innocent, and at times he seems intent on entrapping Nāgasena in false dichotomies and leading questions. So it is said of him:

> Master of words and sophistry (*vetaṇḍī*), clever and wise
> Milinda tried to test great Nāgasena's skill.
> Leaving him not, again and yet again,
> He questioned and cross-questioned him, until
> His own skill was proved foolishness. (MP 4.1.1).

Milinda here is significantly described as a 'master of sophistry' or *vetaṇḍī*, a practitioner of the dialogue form known as *viṭaṇḍā*, a 'refutation-only' type of dialogue in which the opponent defends no thesis of his own but is set only on refuting that of the proponent (see §1.4). The implication here is that such dialogues are essentially eristic. And it is, in particular, the eristic use of questioning that Milinda is a master of. Questions need not be innocent requests for information; they can also be disguised arguments. To reply to the question 'When did you stop cheating on your tax returns?' at all, affirmatively or negatively, is already to commit oneself to the 'premise' of the question, that one has indeed been cheating on one's tax returns. In the intellectual climate of ancient India, when interrogative dialogue was common-place, it was very well known that questions can be used to entrap the unwitting, and counter-strategies were invented to avoid entrapment. The Buddha himself was well aware that replying to a yes-no question can commit one to a proposition, whatever answer one gives, and his solution, famously, was to refuse to answer. Thus when asked a series of ten leading questions — is the soul is eternal? is it non-eternal? etc. — the Buddha declined to offer a reply. For any reply would commit him, against his wish, to the existence of souls. In the *Questions of King Milinda*, we see Nāgasena experimenting with a different technique to avoid entrapment. To some of Milinda's more devious yes-no questions, instead of refusing to reply at all, Nāgasena replies 'Both yes and no'! To others he replies 'Neither yes nor no'! For example:

> Milinda: He who is born, Nāgasena, does he remain the same or become another?

Nāgasena: Neither the same nor another.

Milinda: Give me an illustration.

Nāgasena: Now what do you think, O king? You were once a baby, a tender thing, and small in size, lying flat on your back. Was that the same as you who are now grown up?

Milinda: No. That child was one, I am another.

Nāgasena: If you are not that child, it will follow that you have had neither mother nor father, no! nor teacher. You cannot have been taught either learning, or behaviour, or wisdom. ... Suppose a man, O king, were to light a lamp, would it burn the night through?

Milinda: Yes, it might do so.

Nāgasena: Now, is it the same flame that burns in the first watch of the night, Sir, and in the second?

Milinda: No.

Nāgasena: Or the same that burns in the second watch and the third?

Milinda: No.

Nāgasena: Then there is one lamp in the first watch, and another in the second, and another in the third?

Milinda: No. The light comes from the same lamp all the night through.

Nāgasena: Just so, O king, is the continuity of a person or thing maintained. One comes into being, another passes away; and the rebirth is, as it were, simultaneous. Thus neither as the same nor as another does a man go on to the last phase of his self-consciousness. (MP 2.2.1)

The 'premise' of the question, that to change is to cease to be, is very effectively refuted with a 'neither yes nor no' reply. Nāgasena first makes Milinda acknowledge that, with this as the background premise, answering either 'yes' or 'no' leads to an absurdity. For if he is strictly identical to the child, then he must share that child's properties; and if he is different, then he cannot. Having exposed the false premise, Nāgasena, rejects it in favour of the view that persistence through time requires not strict identity but causal continuity. Here is a different kind of example:

Milinda: Does memory, Nāgasena, always arise subjectively, or is it stirred up by suggestion from outside?

> Nāgasena: Both the one and the other.
>
> Milinda: But does not that amount to all memory being subjective in origin, and never artificial?
>
> Nāgasena: If, O king, there were no artificial (imparted) memory, then artisans would have no need of practice, or art, or schooling, and teachers would be useless. But the contrary is the case.
>
> Milinda: Very good, Nāgasena. (MP 3.6.11).

Here the question's hidden premise is that memories are caused either wholly by what goes on in the mind or wholly by factors external to it, and the 'both yes and no' reply makes plain that what ought to be said is that memories are wholly caused either by what goes on in the mind or by factors external to it, but not caused wholly by one or the other. Again, subsidiary argumentation exposes the absurdity in replying with an unqualified 'yes' or an unqualified 'no'. It was perhaps in recognition of the tactical importance of such 'neither yes nor no' and 'both yest and no' replies that it became a common-place that there are *four possible ways* of responding to any question of the yes-no type, an idea that was systematised in the work of Nāgārjuna (§1.4). What we see very clearly in the *Questions of King Milinda* is a sophisticated early appreciation of the pragmatics of interogative dialogues.

1.2 On balance and fairness in the conduct of dialogue: The Kathāvatthu

The *Kathāvatthu* or *Points of Controversy* (circa third century BCE) is a book about method. It describes, for the benefit of adherents to various Buddhist schisms, the proper method to be followed in conducting a critical discussion into an issue of doctrinal conflict. Recent scholarship has largely focussed on the question of the extent to which there is, in the *Kathāvatthu*, an 'anticipation' of results in propositional logic.[5] For, while it is true that the formulation of arguments there is term logic rather than propositional, and true also that the propositional rules are nowhere formulated

[5] Aung, S.Z., *Points of Controversy, or, Subjects of Discourse: Being a Translation of the Kathāvatthu from the Abhidhammapiṭaka*, eds. S.Z. Aung and C.A.F. Rhys Davids. Pali Text Society, translation series no.5. London: Luzac & Co. 1915; reprint 1960; Schayer, St., "Altindische Antizipationen der Aussagenlogik", *Bulletin international de l'Academie Polonaise des Sciences et des Lettres*, classe de philologies: 90–96 (1933), translated in Jonardon Ganeri ed., *Indian Logic: A Reader* (London: Curzon, 2001); Bochenski, J. M., "The Indian Variety of Logic", in his *A History of Formal Logic*. Freiburg. Trans. I. Thomas, Notre Dame: University of Notre Dame Press (1961), pp. 416–447., reprinted in Jonardon Ganeri ed., *Indian Logic: A Reader*; Matilal, Bimal Krishna, *The Character of Logic in India*. Albany: State University of New York Press, 1998.

in the abstract, the codified argumentation clearly exploits manipulations that trade on the definition of material implication, on contraposition, and on at least one of *modus tollens*, *modus ponens* and *reductio ad absurdum*. The preoccupation with this question of anticipation, assumes, however, a methodology for the interpretation of Indian logic that suffers a number of serious disadvantages. For, first, in presupposing that the only matter of interest is the extent to which a given text displays recognition of principles of formal logic, the methodology fails to ask what it was that the authors themselves were trying to do, and in consequence, is closed to the possibility that these texts contribute to logical studies of a different kind. And second, in supposing that arguments have to be evaluated formally, the important idea that there are *informal* criteria for argument evaluation is neglected. In fact, the *Kathāvatthu* offers a particularly clear example of a text whose richness and interest lies elsewhere than in its anticipation of deductive principles and propositional laws. As a meticulous analysis of the argumentation properly to be used in the course of a dialogue of a specific type, its concern is with the pragmatic account of argument evaluation, the idea that arguments have to be evaluated as good or bad with regard to their contribution towards the goals of the dialogue within which they are embedded. The leading concern of the *Kathāvatthu* is with issues of balance and fairness in the conduct of a dialogue, and it recommends a strategy of argumentation which guarrantees that both parties to a point of controversy have their arguments properly weighed and considered. It is important, in the normative framework of the *Kathāvatthu*, that there is a distinction between the global aim of the dialogue as a whole — here to rehearse in an even-handed manner all the considerations that bear upon an issue of dispute, to clarify what is at stake even if no final resolution is achieved — and the local aim of each participant — to advocate the stance they adopt with regard to that issue by supplying arguments for it and attacking the arguments of the other parties.

A dialogue conducted in accordance with the prescribed method of the *Kathāvatthu* is called a *vādayutti*. The goal of a *vādayutti* is the reasoned examination (*yutti*; Skt. *yukti*) of a controversial point in and through a noneristic dialogue (*vāda*). The dialogue is highly structured, and is to be conducted in accordance with a prescribed format of argumentation. There is a given point at issue, for example, whether 'a person is known in the sense of a real and ultimate fact' (i.e. whether persons are conceived of as metaphysically irreducible), whether there are such things as morally good and bad actions, and so, in general, whether A is B. A dialogue is now divided into eight sub-dialogies or 'openings' (*aṭṭhamukha*). These correspond to eight attitudes it is possible to adopt with regard to the point at issue. So we have:

[1] Is A B?

[2] Is A not B?

[3] Is A B everywhere?

[4] Is A B always?

[5] Is A B in everything?

[6] Is A not B everywhere?

[7] Is A not B always?

[8] Is A not B in everything?

The introduction of an explicit quantification over times, places and objects serves to determine an *attitude* of proponent and respondent to the point of controversy. If the issue in question is, for example, whether lying is immoral, the clarification would be as to whether that proposition is to be maintained or denied, and in either case, whether absolutely, or only as relativised in some way to circumstances, times or agents. So an opening thesis here is by definition a point at issue together with an attitude towards it.

Each such'opening' now proceeds as an independent dialogue, and each is divided into five stages: the way forward (*anuloma*), the way back (*paṭikamma*), the refutation (*niggaha*), the application (*upanayana*), and the conclusion (*niggamana*). In the way forward, the proponent solicits from the respondent their endorsement of a thesis, and then tries to argue against it. In the way back, the respondent turns the tables, soliciting from the proponent their endorsement of the counter-thesis, and then trying argue against it. In the refutation, the respondent, continuing, seeks to refutes the argument that the proponent had advanced against the thesis. The application and conclusion repeat and reaffirm that the proponent's argument against the respondent's thesis is unsound, while the respondent's argument against the proponent's counter-thesis is sound.

It is significant to note that there is here no *pro-argumentation*, either by the respondent for the thesis or by the proponent for the counter-thesis. There is only *contra-argumentation*, and that in two varieties. The respondent, in the 'way back', supplies an argument against the proponent's counter-thesis, and in the refutation stage, against the proponent's alleged argument against the thesis. So we see here a sharp distinction between three types of argumentation — *pro argumention*, argumentation that adduces reasons in support of one's thesis, *counter* argumenation — argumentation that adduces reasons against the counter-thesis, and *defensive* argumentation, argumentation that defends against counter-arguments directed against one's thesis. The respondent, having been 'attacked' in the first phase, 'counter-attacks' in the second phase, 'defends' against the initial

attack in the third, and 'consolidates' the counter-attack and the defence in the fourth and fifth. The whole pattern of argumentation, it would seem, is best thought of as *presumptive*, that is, as an attempt to switch a burden of proof that is initially even distributed between the two parties. The respondent tries to put the burden of proof firmly onto the proponent, by arguing against the proponent while countering any argument against herself. The fact that the respondent does not offer any pro argumentation in direct support of the thesis means that the whole pattern of argumentation is technically *ab ignorantium*; that is, argumentation of the form "I am right because not proved wrong". But ab *ignorantium* reasoning is not always fallacious; indeed, it is often of critical importance in swinging the argument in one's favour in the course of a dialogue (see §1.5).

In the first stage, the way forward, the proponent elicits from the respondent an endorsement of a thesis, and then sets out to reason against it. Not any form of reasoning is allowed; indeed the *Kathāvatthu* prescribes a very specific method of counter-argumentation. Thus:

I. The Way Forward

Theravādin: Is the soul (*puggala*) known as a real and ultimate fact?
[1] Puggalavādin : Yes.
Theravādin: Is the soul known in the same way as a real and ultimate fact is known?
[2] Puggalavādin: No, that cannot be truly said.
Theravādin: Acknowledge your refutation (*niggaha*):
[3] If the soul be known as a real and ultimate fact, then indeed, good sir, you should also say, the soul is known in the same way as any other real and ultimate is known.
[4] That which you say here is false, namely, that we should say, "the soul is known as a real and ultimate fact", but we should not say, "the soul is known in the same way as any other real and ultimate fact is known."
[5] If the later statement cannot be admitted, then indeed the former statement should not be admitted either.
[6] In affirming the former, while denying the latter, you are wrong.

The respondent, here a *puggalavādin* or believer in the existence of personal souls, is asked to endorse the thesis. The proponent then attempts to draw out an implication of that thesis, an implication more over to which the *puggalavādin* will not be willing to give his consent. Here the thesis that persons are thought of as metaphysically irreducible elements of the world is held to imply that knowledge of persons is knowledge of the same kind as that of other types of thing. The *puggalavādin*, will perhaps want to draw an epistemological distinction between empirical knowledge of external objects and self-knowledge, and so will not endorse this derived proposition. And now the proponent, in a fresh wave of argumentation, demonstrates that it is inconsistent for the *puggalavādin* to endorse the thesis but not the

derived consequence. So a counter-argument has three components: the initial thesis or *thapanā* (Skt. *sthāpanā*), the derived implication or *pāpanā*, and the demonstration of inconsistency or *ropanā*.

It is in the *ropanā* that there seems to be an 'anticipation' of propositional logic. Of the four steps of the *ropanā*, the first, from [3] to [4], looks like an application of the definition of material implication or its term-logical equivalent:

$$(A \text{ is } B \rightarrow (C \text{ is } D) =_{\text{defn}} \neg((A \text{ is } B)\&\neg(C \text{ is } D)).$$

Notice here that an effect of soliciting from the respondent a 'no' in answer to the proponent's second question is that the negation is external and not internal. Thus, we have '$\neg(C \text{ is } D)$' rather than '$(C \text{ is } \neg D)$'. This what one needs in the correct definition of material implication.

The second step, from [4] to [5], looks like a derivation of the contraposed version of the conditional, a derivation that depends on the stated definition of the conditional. From that definition, and assuming that '&' is commutative, it follows that

$$(A \text{ is } B) \rightarrow (C \text{ is } D) \text{ iff } \neg(C \text{ is } D) \rightarrow \neg(A \text{ is } B).$$

The final step now is an application of *modus ponens*. So what we have is:

[1]	$(A \text{ is } B)$	premise
[2]	$\neg(C \text{ is } D)$	premise
[3]	$(A \text{ is } B) \rightarrow (C \text{ is } D)$	additional premise?
[4]	$\neg((A \text{ is } B)\&\neg(C \text{ is } D))$	3, defn. of $\rightarrow$
[5]	$\neg(C \text{ is } D) \rightarrow \quad \neg(A \text{ is } B)$	4, defn. of $\rightarrow$
[6]	$\neg(A \text{ is } B)$	2, 5, MP

This is how Matilal[6] reconstructs the *ropanā* stage of argumentation. Earlier, Bochenski[7] recommended a variant in which steps [3] and [4] "together constitute a kind of law of contraposition or rather a *modus tollendo tollens* in a term-logical version". Still another alternative is to see step [3] as a piece of enthymematic reasoning from the premise already given, rather than as the introduction of an additional premise. In other words, the 'if... then' in [3] is to be understood to signify the logical consequence relation rather than material implication. Then step [4] negates the premise in an application of reductio *ad absurdum*. That is:

[6] Matilal (1998: 33–37)
[7] Bochenski (1961: 423)

[1,2]	$(A \text{ is } B)$ & $\neg(C \text{ is } D)$	premise
[3]	$(C \text{ is } D)$	$1 + 2$, enthymematic derivation
[4]	$\neg((A \text{ is } B)$ & $\neg(C \text{ is } D))$	$1 + 2, 3$; RAA
[5]	$\neg(C \text{ is } D) \rightarrow \quad \neg(A \text{ is } B)$	4, defn. of $\rightarrow$
[6]	$\neg((A \text{ is } B)$ & $\neg(C \text{ is } D))$	5, defn. of $\rightarrow$

This reconstuction seems more in keeping with the overall pattern of argumentation — to take the respondent's thesis and derive from it consequences the respondent will not endorse, and thereby to argue against the thesis (and it preserves the repetition of the original). Here again we see that the form of argumentation in the *Kathāvatthu* is better understood if we bear in mind the function it is intended to serve within a dialogue context.

The same dialogue context is normative, in the sense that it gives the grounds for evaluating any actual instance of such argumentation as good or bad. It seems possible to understand the 'way forward' in terms of certain concepts from the theory of argumenation. Hamblin introduced the idea that each participant in a dialogue has a 'commitment store', a set of propositions to which they commit themselves in the course of the dialogue, primarily by asserting them directly.[8] In Hamblin's model, the commitments of each party are on public display, known to every participant in the dialogue. In order to represent the fact that this is very often *not* the case, Walton[9] employs a distinction between open or 'light-side' commitments, and veiled or 'dark-side' commitments. The veiled commitments of a participant are not on public view, and might not be known even to that participant themselves: but perhaps the participant trades on them in making certain kinds of dialogue move. Indeed, it is part of what Walton[10] calls the 'maieutic' role of dialogue to make explicit the veiled commitments of the participants, a process of clarification that is valuable even if it does not lead to the issue at stake being decided in favour of one party or the other.[11]

Something of this sort is what is being described in the initial stages of the 'way forward'. Steps [1] and [2] elicit from the respondent an explicit and open commitment to the propositions 'A is B' and '$\neg$ (C is D)'. ¿From the respective assertion and denial, these become parts of her explicit commitment store. But next, though the enthymematic argumentation that

[8] Hamblin, C. L., *Fallacies*. London: Methuen, 1970.

[9] Walton (1998: 50–51).

[10] Walton (1998: 58).

[11] The term 'maieutic', from *maieutikos* 'skill in midwifery, is taken from the *Theaetetus*, where Socrates describes himself as a midwife for beautiful boys - helping them to give birth to whatever ideas are in them, and test them for whether they are sound.

constitutes the *pāpanā* or stage [3], it is made clear that the respondent has a veiled commitment to the proposition 'C is D'. For this is shown to follow from propositions in the explicit commitment store of the respondent. Finally, the *ropanā* stage of reasoning reveals this newly explosed commitment to be inconsistent with the respondent's other explicit commitments. The overall effect is to force the respondent into a position where she must retract at least one of the propositions to which she has committed herself. Indeed, we can say that such a retraction is the primary goal of the way forward. The primary aim is not to *disprove* the thesis, but to force a retraction of commitment. So when we evaluate the argumentation used in this part of the dialogue, it is to be evaluated as good or bad with reference to how well it succeeds in forcing such a retraction, and not simply or only or even in terms of its deductive or inductive soundness. The strategic problem here is how to persuade the respondent to accept some proposition that is meant ultimately to be used to force a retraction, and the type of strategy being recommended is the one Walton calls that of "separating", where "two or more propositions are proved separately and then eventually put together in an argument structure that is used to prove one's own thesis or argue against an opponent's".[12] In setting out the reasoning in this way, the intention of the author of the *Kathāvatthu* is not to imply that precisely this sequence of arguments is sound. What is being shown is the *form* that any counter-argument should take. It is a description, in generic terms, of the strategic resources open to the proponent, and serves rather as a blue-print for any actual *vādayutti* dialogue.

At this point in the sub-dialogue that is the first opening, then, the burden of proof seems to lie squarely with the respondent, the *puggalavādin*, who is being pressured into the uncomfortable position of having to retract his stated thesis. The remaining four phases of the first opening are a summary of the strategic resources open to the respondent to recover his position, and indeed to turn the tables against the proponent. First, the way back. This is a phase of counter-attack, in which the respondent uses parallel reasoning to force the proponent too into a position of retraction with regard to the counter-thesis.

II. The Way Back

Puggalavādin: Is the soul not known as a real and ultimate fact?
[1] Theravādin: No, it is not known.
Puggalavādin: Is it not known in the same way as any real and ultimate fact is known?
[2] Theravādin: No, that cannot be truly said.
Puggalavādin: Acknowledge the rejoinder (*paṭikamma*):
[3] If the soul is not known as a real and ultimate fact, then indeed, good

[12] Walton (1998: 44).

sir, you should also say: it is not known in the same way as any other real and ultimate fact is known.
[4] That which you say is false, namely, that we should say "the soul is not known as a real and ultimate fact", but we should not say "it is not known in the same way as any other real and ultimate fact is known".
[5] If the latter statement cannot be admitted, then indeed the former statement should not be admitted either.
[6] In affirming the former while denying the latter, you are wrong.

At the end of the 'way back', if the respondent's arguments have gone well, the proponent has been pressed in the direction of retracting his commitment to the counter-thesis. If the respondent were to leave matters here, however, he would have failed in the global aim of the 'opening'. The aim of the opening is to shift the burden of proof decisively onto the proponent. After the second stage in the opening, however, the burden of proof is again symmetrically distributed among the parties to the dialogue — both are in a position of being pressed to retract their respective commitment. So, in the third phase, the respondent seeks, in a defensive move, to diffuse the argument of the proponent that is forcing this retraction. Again, the cited reasoning is schematic, it indicates a general strategy the details of which must be worked out differently in each specific case. The distinction being drawn is the one between *counter-argument*, and *defensive repost*, a distinction that makes sense only within the normative framework of a dialogical exchange.

The first opening in the *vādayutti* has rehearsed the best argumentation that available against someone whose attitude towards the point at issue is one of unqualified affirmation. Remember, however the global aim of a *vādayutti* — to be the form of dialogue most conducive to a balanced examination of the best arguments, both for and against. It is the function now of the second opening to rehearse the best argumentation against someone whose attitude towards the point at issue is one of unqualified denial, and of the subsequent openings to do likewise with respect to attitudes of qualified affirmation and denial. Even at the end of the dialogue, there may be no final resolution, but an important maieutic function has been served — the clarification of the commitments entailed by each position, of their best strategies and forms of argumentation. So, indeed, it is as a rich account of presumptive reasoning in dialogue, and not so much for its 'anticipations' of formal logic, that the *Kathāvatthu* makes a rewarding object of study.

1.3 Case-based reasoning, extrapolation and inference from sampling: The Nyāyasūtra

It was Henry Colebrooke[13] who first brought Indian logic to the attention of the English philosophical world, announcing in a famous lecture to the Royal Asiatic Society in 1824 his discovery of what he called the 'Hindu Syllogism'. Colebrooke's 'discovery' consisted in fact in a translation of an ancient Indian treatise called the *Nyāyasūtra*. It dates from around the 1^{st} or 2^{nd} century AD, and is said to be the work of Gautama Akṣapāda. Scholars are now inclined to regard it as the amalgamation of two earlier works on philosophical method, one a treatise on the rules and principles of debate, the other a discussion of more general issues in epistemology and metaphysics. In a section on the proper way for a debater to set out their argument, the *Nyāyasūtra* prescribes a five-step schema for well-formed argument, and it is this schema that Colebrooke identified as the Indian syllogism. We now know much much more than Colebrooke about the historical development of Indian logic. He, for instance, was unaware of the informal logic and anticipations of propositional calculus in the *Kathāvatthu* (§1.2), or the theories of the Buddhists Vasubandhu, Diṅnāga and Dharmakīrti on formal criteria for inference (§§2.1–5). And scholars had yet to learn the complexities of the later logical school of Navya-Nyāya (§§4.1–3), with its intriguing treatment of negation, logical consequence and quantification, and even, as Daniel Ingalls has shown in his pioneering book entitled *Materials for the Study of Navya-Nyāya Logic*, the formulation of De Morgan's Laws.[14] Nevertheless, in spite of Colebrooke's lack of acquaintance with the historical context, he and those who followed him were right to see the *Nyāyasūtra* as a treatise of fundamental importance in Indian logical thinking, and I would like to pick up and continue the thread of their discussion. I want to argue that the *Nyāyasūtra* begins a transformation in Indian thinking about logic. And this in two inter-related respects: in the beginnings of a shift of interest away from the place of *argumentation* within dialectic and debate and towards a greater concern with the more formal properties of sound *inference*; and in a parallel and correlated shift from *case-based* to *rule-governed* accounts of logical reasoning. I will discuss each of these in turn.

In the *Nyāyasūtra*, there is a more systematic discussion of the categories and methods of debate than in earlier debating manuals. Three kinds of debate are distinguished: good or honest debate (*vāda*), tricky or bad debate (*jalpa*) and a refutation-only debate (*viṭaṇḍā*):

[13] H. T. Colebrooke, "On the Philosophy of the Hindus: Part II - On the Nyāya and Vaiśeshika systems". *Transactions of the Royal Asiatic Society (1824)*, 1: 92–118; reprinted in Jonardon Ganeri ed., *Indian Logic: A Reader*.

[14] D. H. H. Ingalls, *Materials for the Study of Navya-Nyāya Logic* (Cambridge Mass.: Harvard University Press), 1951, pp. 65–67.

> Good debate (*vāda*) is one in which there is proof and refutation of thesis and antithesis based on proper evidence (*pramāṇa*) and presumptive argumentation (*tarka*), employing the five-step schema of argumentation, and without contradicting any background or assumed knowledge (*siddhānta*).
>
> Tricky debate (*jalpa*) is one in which, among the features mentioned before, proof and refutation exploit such means as quibbling (*chala*), false rejoinders (*jāti*), and any kind of clincher or defeat situation (*nigrahasthāna*).
>
> Refutation-only debate (*viṭaṇḍā*) is one in which no counter-thesis is proven. (NS 1.2.1–3).

Here is our first reference to the Indian five-step inference pattern. It is a schema for proper argumentation among disputants who are engaged in an honest, friendly, noneristic, and balanced debate (*vāda*). In the dialectical context in which such arguments are embedded, a proponent attempts to prove a thesis and to refute the counter-thesis of the opponent, both parties drawing upon a shared body of background knowledge and received belief, and using properly accredited methods for the acquisition and consideration of evidence. The aim of each participant in the dialogue is not victory but a fair assessment of the best arguments for and against the thesis. In Indian logic, *vāda* represents an *ideal* of fair-minded and respectful discourse. By contrast, in a tricky debate (*jalpa*), underhanded debating tactics are allowed, and the aim is to win at all costs and by any means necessary. The third kind of debate, the refutation-only debate (*viṭaṇḍā*), is the variety of dialogue preferred by the sceptics — to argue against a thesis without commitment to any counter-thesis. It is not entirely clear whether this is a type of good or tricky debate. We might conjecture, however, that if dialectic is a rough kin of *vāda*, and sophistic of *jalpa*, then the Socratic elenchus could be regarded as a species of *viṭaṇḍā*, which is not, therefore, an entirely disreputable method of debate.

The aim, in a good debate between friends, is the assessment of the best arguments for or against the thesis. And that leads to the question: how are arguments to be assessed or evaluated? Are the criteria for assessment formal, to do only with the form of the argument schema itself; or are they informal, pragmatic criteria, according to which arguments have to be evaluated as good or bad with regard to their contribution towards the goals of the dialogue within which they are embedded?

With this question in mind, let us look at the five-step proof pattern. The proper formulation of an argument is said to be in five parts: tentative statement of the thesis to be proved (*pratijñā*); citation of a reason (*hetu*); mention of an example (*udāharaṇa*); application of reason and example to the case in hand (*upanaya*); final assertion of the thesis (*nigamana*). An

unseen fire is inferred to be present on the mountain, on the basis of a plume of smoke; just as the two have been found associated in other places like the kitchen. The terms used here are defined in a series of admittedly rather gnomic utterances (NS 1.1.34–39):

1.1.32 'the parts [of an argument scheme] are thesis, reason, example, application and conclusion' (*pratijñāhetūdāharaṇopanayanigamanānyavayavāḥ*).

1.1.33 'the thesis is a statement of that which is to be proved' (*sādhyanirdeśaḥpratijñā*).

1.1.34 'the reason is that which proves what is to be proven in virtue of a similarity with the example' (*udāharaṇasādharmyāt sādhyasādhanaṃ hetuḥ*).

1.1.35 'again, in virtue of a dissimilarity' (*tathā vaidharmyāt*).

1.1.36 'the example is an illustration which, being similar to that which is to be proved, has its character' (*sādhyasādharmyāt taddharmabhāvī dṛṣṭānta udāharaṇam*).

1.1.37 'or else, being opposite to it, is contrary' (*tadviparyayād vā viparītam*).

1.1.38 'the application to that which is to be proved is a drawing in together (*upasaṃhāra*) "this is so" or "this is not so," depending on the example' (*udāharaṇāpekṣas tathety upasaṃhāro na tatheti vā sādhyasyopanayaḥ*).

1.1.39 'the conclusion is a restatement of the thesis as following from the statement of the reason' (*hetvapadeśāt pratijñāyāḥ punarvacanaṃ nigamanam*).

The basic idea is that an object is inferred to have one (unobserved) property on the grounds that it has another, observed, one — "there is fire on the mountain because there is smoke there". The most distinctive aspect of the schema, though, is the fundamental importance given to the citation of an example, a single case said either to be similar or else dissimilar to the case in hand. Suppose I want to persuade you that it is about to rain. I might reason as follows: "Look, it is going to rain (thesis). For see that large black cloud (reason). Last time you saw a large black cloud like that one (example), what happened? Well, its the same now (application). It is definitely going to rain (conclusion)."

Let us try to unpick the *Nyāyasūtra* definitions. Suppose we let 'F' denote the property that serves as the reason here (*hetu*), 'G' the property whose presence we are seeking to infer (*sādhya*), 'a' the new object about which we are trying to decide if it is G or not (*pakṣa*), and 'b' the cited example

(*udāharaṇa*). Then we seem to have a pair of schematic inferences, one based on similarity, the other on dissimilarity:

Five-step proof based on similarity

[thesis] Ga

[reason] Fa proves Ga, because b is similar to a.

[example] b has the 'character of a' because it is similar to a.

[application] a is the same as b with respect to G.

[conclusion] Ga

Five-step proof based on dissimilarity

[thesis] Ga

[reason] Fa proves Ga, because b is dissimilar to a.

[example] b does not have the 'character of a' because it is dissimilar to a.

[application] a is not the same as b with respect to G.

[conclusion] Ga

The counter-proof follows the same pattern, proving the counter-thesis ($\neg Ga$) by means of a different reason and example:

Counter-proof

[thesis] $\neg Ga$

[reason] $F'a$ proves Ga, because b is similar to a.

[example] c has the 'character of a' because it is similar to a.

[application] a is the same as c with respect to G.

[conclusion] $\neg Ga$

The five-step schema was interpreted in a particular way by Vātsyāyana, the first commentator on the *Nyāyasūtra*. His interpretation is largely responsible for shaping the direction Indian logic was later to take. At the same time, it was an interpretation that made the citation of an example essentially otiose. Vātsyāyana was, in effect, to transform Indian logic, away from what it had been earlier, namely a theory of inference from case to case on the basis of resemblance, and into a rule-governed account in which the citation of cases has no significant role.

Let us then consider first Vātsyāyana's interpretation. What Vātsyāyana says is that the similarity between a and b just consists in their sharing the reason property F. The basic pattern of inference is now: a is like b [both are F]; $Gb \therefore Ga$. *Or else*: a is unlike b [one is F and the other isn't]; $\neg Gb \therefore Ga$. Writing it out as before, what we have is:

Proof based on similarity

[thesis]	Ga	
[reason]	Fa	
[example]	Fb	b is similar to a [both are F].
[application]	Gb	
[conclusion]	Ga	

In a counterproof, a is demonstrated to be similar in some other respect to some other example, one that lacks the property G:

Counterproof

[thesis]	$\neg Ga$	
[reason]	$F'a$	
[example]	$F'c$	c is similar to a [both are F'].
[application]	$\neg Gc$	
[conclusion]	$\neg Ga$	

Thus, for example, a proof might be: the soul is eternal because it is uncreated, like space. And the counterproof might be: the soul is non-eternal because it is perceptible, like a pot.

The proposal is that if a resembles b, and b is G, then a can be inferred to be G too. But there is an obvious difficulty, which is that mere resemblance is an insufficient ground. Admittedly, the soul and space are both uncreated, but why should that give us any grounds for transferring the property of being eternal from one to the other? The *respect* in which the example and the case in hand resemble one another must be relevant to the property whose presence is being inferred. This is where the idea of a 'false proof' or 'false rejoinder' (*jāti*) comes in. Any argument that, while in the form of the five-step schema, fails this relevance requirement is called a 'false proof' and the *Nyāyasūtra* has a whole chapter (chapter 5) classifying and discussing them. A 'false rejoinder' is defined in this way:

> NS 1.2.18 'a *jāti* is an objection by means of similarity and dissimilarity' (*sādharymavaidharmyābhāṃ pratyavasthānaṃ jātiḥ*).

It appears to be admissible to transfer the property 'rain-maker' from one black cloud to another black cloud, but not from a black cloud to a white cloud. It appears to be admissible to transfer the property 'has a dewlap' from one cow to another cow, but not from one four-legged animal (a cow) to another (a horse). It is clear what now needs to be said. The argument is good if there exists a *general* relationship between the reason F and the property being proved G, such that the latter never occurs without the former.

It is the Buddhist logician Diṅnāga (480–540 CE) who seems to have been the first to make this explicit (see also §2.2). According to him, a reason must satisfy three conditions. Define a 'homologue' (*sapakṣa*) as an object other than a that possesses G, and a 'heterologue' (*vipakṣa*) as an object other than a that does not possess G. Then Diṅnāga's three conditions on a good reason are:

[1] F occurs in a.

[2] F occurs in some homologue.

[3] F occurs in no heterologue.

Condition [3] asserts, in effect, that F never occurs without G, and this, together with [1] that F occurs in a, implies of course that G occurs in a. In effect, the citation of an example in the original *Nyāyasūtra* formula has been transformed into a statement of a general relationship between F and G. There remains only a vestigial role for the example in condition [2], which seems to insist that there be an instance of F other than a which is also G. Diṅnāga is worried about the soundness of inferences based on a reason which is a property unique to the object in hand; for example, the inference "sound is eternal because it is audible". For if this is sound, then why not the counter-argument "sound is non-eternal because it is audible"? And yet there are many inferences like this that are sound, so it seems to be a mistake to exclude them all. In fact condition [2] soon came to be rephrased in a way that made it logically equivalent to [3], namely as saying that F occur only in homologues (*eva* 'only' is used here as a quantifier). In asking for the respect in which the example and the new case must resemble each other, for the presence of G in the example to be a reason for inferring that G is present in the new case, we are led to give the general relationship that any such respect must bear to G, and that in turn makes citation of an example otiose. The five-step schema becomes:

[thesis]	Ga
[reason]	because F
[example]	where there is F, there is G; for example, b.

[application] Fa

[conclusion] Ga

It is the five-step argument pattern so transformed that has suggested to Colebrooke and other writers on Indian logic a comparison with an Aristotelian syllogism in the first figure, *Barbara*. We simply re-write it in this form:

> All F are G.
> Fa.
> Therefore, Ga.

This assimilation seems forced in at least two respects. First, the conclusion of the *Nyāyasūtra* demonstration is a singular proposition. In Aristotle's system, on the other hand, it is always either a universal proposition with 'all' or 'no', or a particular proposition with 'some'. Second, and relatedly, the role of the 'minor term' is quite different: in the Indian schema, it indicates a locus for property-possession, while in Aristotle, the relation is 'belongs to'. Again, in reducing the Indian pattern to an Aristotelian syllogism, the role of the example, admittedly by now rather vestigial, is made to disappear altogether.

A rather better reformulation of the five-step schema is suggested by Stanisław Schayer,[15] who wants to see the Indian 'syllogism' as really a proof exploiting two rules of inference:

[thesis]	Ga	There is fire on a (= on this mountain).
[reason]	Fa	There is smoke on a.
['example']	$(x)(Fx \rightarrow Gx)$	For every locus x: if there is smoke in x then there is fire in x.
[application]	$Fa \rightarrow Ga$	This rule also applies for $x = a$.
[conclusion]	Ga	Because the rule applies to $x = a$ and the statement Ga is true, the statement Fa is true.

Two inference rules are in play here — a rule of substitution, allowing us to infer from '$(x)\zeta x$' to 'ζa', and a rule of separation, allowing us to infer from '$\theta \rightarrow \varphi$' and '$\theta$' to '$\varphi$'. Schayer thereby identifies the Indian syllogism with a proof in a natural deduction system:

THESIS. Ga because Fa.

[15] Schayer, St., "Altindische Antizipationen der Aussagenlogik", *Bulletin international de l'Academie Polonaise des Sciences et des Lettres*, classe de philologies: 90–96 (1933); translated in Jonardon Ganeri ed., *Indian Logic: A Reader*.

Proof.

1	(1)	Fa	Premise
2	(2)	$(x)(Fx \rightarrow Gx)$	Premise
3	(3)	$Fa \rightarrow Ga$	2, by $\forall$ Elimination
1,2	(4)	Ga	1 & 3, by $\rightarrow$ Elimination. ■

We have seen how the *Nyāyasūtra* model of good argumentation came to be transformed and developed by later writers in the Indian tradition in the direction of a formal, rule-governed theory of inference, and how writers in the West have interpreted what they have called the Indian 'syllogism'. I suggested at the beginning that we might try to interpret the Nyāya model along different lines altogether, seeing it an early atttempt at what is now called 'case-based reasoning'. Case-based reasoning begins with one or more prototypical exemplars of a category, and reasons that some new object belongs to the same category on the grounds that it resembles in some appropriate and context determined manner one of the exemplars. Models of case-based reasoning have been put forward for medical diagnostics and legal reasoning, and some have been implemented in artificial intelligence models. It has been shown, for example, that training a robot to solve problems by having it retrieve solutions to stored past cases, modifying them to fit the new circumstances, can have great efficiency gains over seeking solutions through the application of first principles. Perhaps something like this underlies a lot of the way we actually reason, and perhaps it was as an attempt to capture *this* type of reasoning that we should see the ancient logic of the *Nyāyasūtra* and indeed of the medical theorist Caraka.[16] In this model, a perceived association between symptoms in one case provides a reason for supposing there to be an analogous association in other, resembling cases. The physician observing a patient A who has, for example, eaten a certain kind of poisonous mushroom, sees a number of associated symptoms displayed, among them F and G, say. He or she now encounters a second patient B displaying a symptom at least superficially resembling F. The physician thinks back over her past case histories in search of cases with similar symptoms. She now seeks to establish if any of those past cases resembles B, and on inquiry into B's medical history, discovers that B too has consumed the same kind of poisonous mushroom. These are her grounds for inferring that B too will develop the symptom G, a symptom that had been found to be associated with F in A. A common etiology in the two cases leads to a common underlying disorder, one that manifests itsself in and explains associations between members of a symptom-cluster.

Can we find such a model of the informal logic of case-based reasoning in the *Nyāyasūtra*? Consider again NS 1.1.34. It said that 'the reason is

[16] *Caraka-Saṃhitā* 3.8.34: 'what is called "example" is that in which both the ignorant and the wise think the same and that which explicates what is to be explicated. As for instance, "fire is hot," "water is wet," "earth is hard," "the sun illuminates." Just as the sun illuminates, so knowledge of sāṃkhya philosophy illuminates'.

that which proves what is to be proved in virtue of a similarity with the example.' On our reading, what this says is that a similarity between the symptom F in the new case and a resembling symptom F' in the past-case or example is what grounds the inference. And NS 1.1.36 says that 'the example is something which, being similar to that which is to be proved, has its character'. Our reading is that the old case and the new share something in their circumstances, like having eaten the same kind of poisonous mushroom, in virtue of which they share a 'character', an underlying disorder that expains the clustering of symptoms. So the five-step demonstration is now:

[thesis]	Ga	
[reason]	Fa	F is similar to F' in b.
[example]	b	exhibits the same underlying structure as a, because it resembles a.
[application]		a is the same as b with respect to G.
[conclusion]	Ga	

Let us see if this pattern fits examples of good inference taken from a variety of early Indian logical texts. One pattern of inference, cited in the *Nyāyasūtra*, is causal-predictive: Seeing the ants carrying their eggs, one infers that it will rain; seeing a full and swiftly flowing river, one infers that it has been raining; seeing a cloud of smoke, one infers the existence of an unseen fire. Presumably the idea is that one has seen other ants carrying their eggs on a past occasion, and on that occasion it rained. The inference works if, or to the exent that, we have reasons for thinking that those ants and these share some unkown capacity, a capacity that links detection of the imment arrival of rain with the activity of moving their eggs. The pattern is similar in another kind of inference, inference from sampling: Inferring from the salty taste of one drop of sea water that the whole sea is salty; inferring that all the rice is cooked on tasting one grain. The assumption again is that both the sampled grain of rice and any new grain share some common underlying structure, a structure in virtue of which they exhibit the sydromes associated with being cooked, and a structure whose presence in both is indicated by their being in the same pan, heated for the same amount of time, and so forth.

I will make two final comments about these patterns of case-based reasoning. First, it is clear that background knowledge is essentially involved. As the *Nyāyasūtra* stresses in its definition of a good debate, both parties in a debate much be able to draw upon a commonly accepted body of information. Such knowledge gets implicated in judgements about which similarities are indicative of common underlying disorders, and which are not. Second, in such reasoning the example does not seem to be redundant or eliminable in favour of a general rule. For although there always will be a general law relating the underlying disorder with its cluster of symptoms,

the whole point of this pattern of reasoning is that the reasoner need not be in a position to *know* what the underlying disorder is, and so what form the general law takes. In conclusion, while the history of logic in India shows a strong tendency towards formalisation, the logic of ancient India tried to model informal patterns of reasoning from cases that are increasingly becoming recognised as widespread and representative of the way much actual reasoning takes place.

1.4 Refutation-only dialogue: vitaṇḍā

We have already seen how 'refutation-only' debate is defined in the *Nyāyasūtra*:

> Refutation-only debate (*vitaṇḍā*) is one in which no counter-thesis is proven. (NS 1.2.1–3).

For the Naiyāyika, to argue thus is to argue in a purely negative and destructive way. Here one has no goal other than to undermine one's opponent. People who use reason in this way are very like the sceptics and unbelievers of the epics. Vātsyāyana claims indeed that to use reason in this way is virtually self-defeating:

> A *vaitāṇḍika* is one who employs destructive criticism. If when questioned about the purpose [of so doing], he says 'this is my thesis' or 'this is my conclusion,' he surrenders his status as a *vaitāṇḍika*. If he says that he has a purpose, to make known the defects of the opponent, this too will is the same. For if he says that there is one who makes things known or one who knows, or that there is a thing by which things are made known or a thing made known, then he surrenders his status as a *vaitāṇḍika*.[17]

Viṭaṇḍā is the sceptic's use of argumentation, and it is a familiar move to attempt to argue that scepticism is self-defeating. In India, it is the Mādhyamika Buddhist Nāgārjuna (circa first century CE) who is most closely associated with the theoretical elaboration of refutation-only argumentation, through the method of 'four-limbed refutation' (*catuṣkoṭi*) and the allied technique of presumptive reasoning (*prasaṅga; tarka*). In the next section, I will show how this latter technique became a device for shifting the burden of proof onto one's opponent. First, I will examine the method of 'four-limbed' refutation in the context of Nāgārjuna's wider philosophical project.

Reasoning, for Nāgārjuna, is the means by which one 'steps back' from common sense ways of understanding to a more objective view of the world. Reason is a mode of critical evaluation of one's conceptual scheme. A more

[17] *Nyāyabhāṣya* 3, 15–20.

objective understanding is one in which one understands that things are not necessarily as they appear. It is a view from which one can see how and where one's earlier conceptions are misleading. One learns not to trust one's perceptions when a large object far away looks small, or a stick half submerged in water looks bent, and in learning this one exercises a mode of self-critical reason. So too a rational person learns not to trust their conceptions when they presuppose the existence of independent, self-standing objects. From the vantage point of an objective view, it is easy to see that one's old conceptions had false presuppositions. The real trick, however, is to be able to expose those presuppositions while still 'within' the old conception, and so to lever oneself up to a more objective view. This levering–up–from–within requires a new way of reasoning: Nāgārjuna's celebrated *prasaṅga-type* rationality. It is a self-critical rationality which exposes as false the existential presuppositions on which one's present conceptions are based. The same method can equally well be used to expose the false presuppositions on which one's dialectical opponents' views are based, and for this reason the whole technique is strongly maieutic, in the sense defined earlier.

A simple example will illustrate the kind of reasoning Nāgārjuna thinks is needed if one is to expose the presuppositions of one's conceptual scheme from within. A non-compound monadic concept 'F' has the following application-condition: it applies only to things which are F. It is therefore a concept whose application presupposes that there is a condition which divides the domain into two. For our purposes, the condition can be thought of either as 'belonging to the class of Fs' or 'possessing the property being-F'. Now take an arbitrary object, a, from some antecedently specified domain. There are apparently two possibilities for a: either it falls under the concept, or else it is not. That is, the two options are:

(I) F applies to a

(II) F does not apply to a.

Suppose that one can disprove *both* of these options. How one would try to do this will vary from case to case depending on the individual concept under scrutiny. But if one is able to disprove (I) and to disprove (II), then the concept in question can have no application-condition. The presupposition for the application of the concept, that there is a condition (class, property) effecting a division within the domain, fails. A later Mādhyamika master[18] expresses the idea exactly:

> When neither existence nor nonexistence presents itself before the mind, then, being without objective support (*nirālambana*) because there is no other way, [the mind] is still.

[18] Śāntideva, *Bodhicaryāvatāra* 9.34.

Sentences are used to make statements, but if the statement so made is neither true nor false, then, because there is no third truth-value, the statement must be judged to lack content.[19]

Nāgārjuna's developed strategy involves a generalization. A generalization is needed because many if not most of the concepts under scrutiny are *relational* rather than *monadic*; centrally: causes, sees, moves, desires. When a concept is relational, there are four rather than two ways for its application-condition to be satisfied (see Figure 1, page 26):

(I) R relates a only to itself

(II) R relates a only to things other than itself

(III) R relates a both to itself and to things other than itself

(IV) R relates a to nothing.

As an illustration of the four options, take R to the square-root relation $\surd$, and the domain of objects to be the set of real numbers. Then the four possibilities are exemplified by the numbers 0, 4, 1 and –1 respectively. For $\surd 0 = 0$, $\surd 4 = 2$ and also –2, $\surd 1 = 1$ and also –1, while finally –1 does not have a defined square root among the real numbers. The list of four options is what is called in Madhyamaka a *catuṣkoṭi*.

> Everything is thus, not thus, both thus and not thus, or neither thus nor not thus. That is the Buddha's [provisional] instruction. [*Mūlamadhyanaka-ka=arikā*, MK 18.8]
>
> Some say that suffering (*duḥkha*) is self-produced, or produced from another, or produced from both, or produced without a cause. [MK 12.1]
>
> Since every factor in existence (*dharma*) are empty, what is finite and what is infinite? What is both finite and infinite? What is neither finite nor infinite? [MK 25.22]

It is easy to see that the four options are mutually exclusive and jointly exhaustive. For the class of objects to which R relates a is either (IV) the empty set $\emptyset$ or, if not, then either (I) it is identical to $\{a\}$, or (II) it excludes $\{a\}$, or (III) it includes $\{a\}$. Not every relation exhibits all four options. (I) not exhibited if R is anti-reflexive. (II) is not exhibited if R is reflexive and bijective. (IV) is not exhibited if R is defined on every point in the domain. Note in particular that if R is the identity relation, then neither (III) nor (IV) are exhibited, not (III) because identity is transitive, and not (IV) because identity is reflexive. Indeed, options (III) and (IV)

[19] On presupposition and truth-value gaps, see P. F. Strawson, *Introduction to Logical Theory* (London: Methuen, 1952).

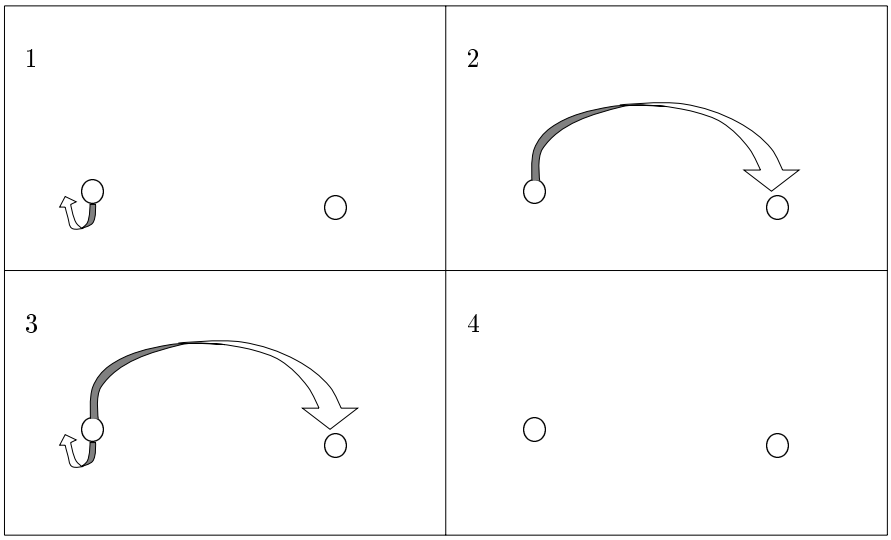


Figure 1. The Four Options

are not exhibited whenever R is an equivalence (transitive, symmetric, and reflexive) relation.

The next step in the 'refutation-only' strategy is to construct subsidiary 'disproofs', one for each of the four options. Although there is no predetermined procedure for constructing such disproofs, by far the most commonly used method is to show that the option in question has some unacceptable consequence (*prasaṅga*). I will examine this method in detail in §1.5. A major dispute for later Mādhyamikas was over what sort of reasoning is permissible in the four subsidiary disproofs, the proofs that lead to the rejection of each of the four options. It is a difficult question to answer, so difficult indeed that it led, at around 500 AD, to a fission within the school of Madhyamaka. The principal group (Prāsaṅgika, headed by Buddhapālita) insisted that only *prasaṅga*-type, 'presupposition-negating' reasoning is admissible. This faction is the more conservative and mainstream, in the sense that their teaching seems to be in keeping with Nāgārjuna's own method of reasoning. The important later Mādhyamika masters Candrakīrti and Śāntideva defended this view. A splinter faction, however, (Svātantrika, headed by Bhāvaviveka) allowed 'independent' inference or inductive demonstration into the disproofs. Perhaps this was done so that

the inferential methods developed by Diṅnāga (§2.2) could be deployed in establishing the Mādhyamika's doctrinal position. Clearly, the fewer restrictions one places on the type of reasoning one permits oneself to use, the greater are the prospects of successfully finding arguments to negate each of the four options. On the other hand, we have seen that the citation of paradigmatic examples is essential to this type of reasoning (§1.3), and it is hard to see how one could be entitled to cite examples in support of one's argument, when the very conception of those examples is in question.

The effect of the four subsidiary disproofs is to establish that none of the four options obtains:[20]

> Neither from itself nor from another, nor from both, nor without a cause, does anything whatever anywhere arise. [MK 1.1]
>
> One may not say that there is emptiness, nor that there is non-emptiness. Nor that both, nor that neither exists; the purpose for so saying is only one of provisional understanding. [MK 22.11]

The emptiness of the concept in question is now deduced as the final step in the process. For it is a presupposition of one of the four options obtaining that the concept does have an application-condition (a class of classes or relational property). If all four are disproved, then the presupposition itself cannot be true. When successful, the procedure proves that the concept in question is empty, null, *śūnya*. This is Nāgārjuna's celebrated and controversial '*prasaṅga*-type' rational inquiry, a sophisticated use of rationality to annul a conceptual scheme.

A statement is truth-apt if it is capable of being evaluated as either true or false. When Nāgārjuna rejects each of the four options, he is rejecting the claim that a statement of the form 'aRb' is truth-apt, since the four options exhaust the possible ways in which it might be evaluated as true. But if the statements belonging to a certain discourse are not truth-apt, then the discourse cannot be part of an objective description of the world (a joke is either funny or unfunny, but it cannot be evaluated as true or false.) The *prasaṅga* negates a presupposition for truth-aptness and so for objective reference.

Nāgārjuna applies the procedure in an attempt to annul each of the concepts that are the basic ingredients of our common-sense scheme. In each case, his method is to identify a relation and prove that none of the four options can obtain. On closer inspection, it turns out that his argumentation

[20] Further examples: MK 25.17, 25.18, 27.15–18. Interesting is the suggestion of Richard Robinson that the method of reasoning from the four options has two distinct functions, a positive therapeutic role as exhibited by the unnegated forms, and a destructive dialectical role, exhibited by the negated forms. Richard H. Robinson, *Early Mādhyamika in India and China*, (Madison, Milwauke and London: University of Winsconsin Press, 1967), pp. 39–58, esp. pp. 55–56.

falls into two basic patterns.[21] One pattern is applied to any concept involving the idea of an ordering or sequence, especially the concept of a causal relation, of a temporal relation and of a proof relation. The paradigm for this argument is Nāgārjuna's presentation of a paradox of origin (chapter 1), which serves as model for his analysis of causation (chapter 8), the finitude of the past and future (chapter 11), and suffering (chapter 12). The argument seeks to establish that a cause can be neither identical to, nor different from, the effect. If nothing within the domain is uncaused, then the four options for the realization of a causal relation are foreclosed.

The other pattern of argumentation in Nāgārjuna is essentially grammatical. When a relational concept is expressed by a transitive verb, the sentence has an Agent and a Patient (the relata of the relation). For example, "He sees the tree," "He goes to the market," "He builds a house." The idea of the grammatical argument is that one can exploit features of the deep case structure of such sentences in order to prove that the Patient can be neither identical to the Agent, nor include it, nor exclude it, and that there must be a Patient. Nāgārjuna uses this pattern of argumentation in constructing a paradox of motion (MK, chapter 2), and this chapter serves as a model for his analysis of perception (chapter 3), composition (chapter 7), fire (chapter 10), and of bondage and release (chapter 16). Indeed, the same pattern of argument seems to be applicable whenever one has a concept which involves a notion of a single process extended in time. What exactly these arguments show and how well they succeed is a matter of debate, but what we have seen is the elaboration of a sophisticated sceptical strategy of argumentation, based on the idea of 'refutation-only' dialogue.

*1.5 Presumptive argumentation (*tarka*) and burden-of-proof shifting*

Indian logicians developed a theory of what they call 'suppositional' or 'presumptive' argumentation (*tarka*). It is a theory about the burden of proof and the role of presumption, about the conditions under which even inconclusive evidence is sufficient for warranted belief. As we have already seen, it is a style of reasoning that is regarded as permissible within a well-conducted dialogue (*vāda*; see §1.3). In the canonical and early literature, *tarka* is virtually synonymous with reasoned thinking in general. The free-thinkers so derided in the epics were called *tārkikas* or 'followers of reason'. Even later on, when the fashion was to adorn introductory surveys of philosophy with such glorious names as *The Language of Reason* (*Tarkabhāṣā*, Mokṣākaragupta), *Immortal Reason* (*Tarkāmṛta*, Jagadīśa), *Reason's Moonlight* (*Tarkakaumudī*, Laugākṣi Bhāskara), it was usual to confer on a graduate of the medieval curriculum an honorific title like Master or Ford of

[21] On other patterns in Nāgārjuna's argumentation, see Robinson (1967: 48).

Reason (*tarkavāgīśa, tarkatīrtha*). Such a person is a master in the art of evidence and the management of doubt, knowing when to accept the burden of proof and also when and how to deflect it.

Extrapolative inference (*anumāna*, see §1.3) rests on the knowledge of universal generalisations, and it is the possibility of such knowledge that the most troubling forms of scepticism call into question. How can one be entitled to believe that something is true of every member of a domain without inspecting each member individually? How does one cope with the ineliminable possibility that an unperceived counterexample exists in some distant corner of the domain? The difficulty here is with the epistemology of negative existentials. The Buddhist Diṅnāga formulates the extrapolation relation as a 'no counterexample' relation. For him, x extrapolates y just in case there is no x without y (*y-avina x-abhāva*). The Navya-Nyāya logicians prefer a different negative existential condition, one derived from the reflexivity and transitivity of the extrapolation relation. Given transitivity, if x extrapolates y then, for any z, if y extrapolates z, so does x. The converse of this conditional holds too, given that the extrapolation relation is reflexive (proof: let $z = y$). So let us define an 'associate condition' (*upādhi*) as a property which is extrapolated by y but not x. Then x extrapolates y just in case there is no associate condition.[22] One can infer fire from smoke but not smoke from fire, for there is an associate condition, dampness-of-fuel, present wherever smoke is but not wherever fire is. Tinkering with the definition, though, does not affect the epistemological problem; it remains the one of proving a nonexistence claim.

Presumptive argumentation, *tarka*, is a device for appropriating a presumptive right — the right to presume that one's own position is correct even without conclusive evidence in its support. One is, let us imagine, in a state of doubt as to which of two hypotheses A and B is true. A and B are exclusive (at most one is true) but not necessarily contradictory (both might be false). Technically, they are in a state of 'opposition' (*virodha*).[23] The doubt would be expressed by an exclusive disjunction in the interrogative – Is it that A or that B? Uncertainty initiates inquiry, and at the beginning of any inquiry the burden of proof is symmetrically distributed among the alternative hypotheses. A piece of presumptive argumentation shifts the burden of proof by adducing a *prima facie* counterfactual argument against one side. The form of the argument is the same in all cases. It is that one alternative, supposed as true, would have a consequence in conflict with

[22] For a survey of the literature on this theory, see Karl Potter ed., "Indian Epistemology and Metaphysics", *Encyclopedia of Indian Philosophies*, Volume 2 (Princeton: Princeton University Press, 1977), pp. 203–206; Karl Potter and Sibajiban Bhattacharyya eds., "Indian Analytical Philosophy: Gaṅgeśa to Raghunātha", *Encyclopedia of Indian Philosophies*, Volume 6 (Princeton: Princeton University Press, 1993), pp. 187–192.

[23] Nandita Bandyopadhyay, "The Concept of Contradiction in Indian Logic and Epistemology," *Journal of Indian Philosophy* 16.3 (1988), pp. 225–246, fn. 1.

some set of broadly defined constraints on rational acceptability. The existence of such an argument gives one the right to presume that the other alternative is true, even though one has no conclusive proof of its truth, and even though the logical possibility of its being false remains open. In the psychologized language of the Nyāya logician, a suppositional argument is a 'blocker' (*bādhaka*) to belief in the supposed alternative, and an 'eliminator' (*nirvartaka*) of doubt. The Naiyāyika Vācaspati (9th century) comments:[24]

> Even if, following a doubt, there is a desire to know [the truth], the doubt still remains after the desire to know [has come about]. This is the situation intended for the application of presumptive argumentation. Of two theses, one should be admitted as known when the other is rejected by the reasoning called 'suppositional.' Thus doubt is suppressed by the application of presumptive argumentation to its subject matter... A means of knowing is engaged to decide a question, but when there is a doubt involving its opposite, the means of knowing fails [in fact] to engage. But the doubt concerning the opposite is not removed as such by the undesired consequence. What makes possible its removal is the means of knowing.

Vācaspati stresses that a thesis is not itself proved by a suppositional demonstration that the opposite has undesired consequences; one still needs evidence corroborating the thesis. But there is now a presumption in its favour, and the burden of proof lies squarely with the opponent. Presumptive argumentation 'supports' one's means of acquiring evidence but it not itself a source of evidence. It role is to change the standard of evidence required for proof in the specific context.

A radical sceptical hypothesis is a proposition inconsistent with ordinary belief but consistent with all available evidence for it. The aim of the radical sceptic is to undermine our confidence that our beliefs are justified, to introduce doubt. The Nyāya logicians' response to scepticism is not to deny that there is a gap between evidence and belief, or to deny the logical possibility of the sceptical hypothesis. It is to draw a distinction between two kinds of doubt, the *reasonable* and the *reasonless*. A doubt is reasonable only when both alternatives are consistent with all the evidence *and the burden of proof is symmetrically distributed* between them. One paradigmatic example is the case of seeing in the distance something that might be a person or might be a tree-stump. Udayana gives the epistemology of such a case: it is a case in which one has knowledge of common aspects but not of specific distinguishing features. What we can now see is that the example gets its force only on the assumption that there is a level epistemic playing field, with both hypotheses carrying the same *prima facie*

[24] *Nyāyavārttikatātparyaṭīkā*, below NS 1.1.40.

plausibility. Presumptive argumentation has the potential to break the impasse — imagine, for example, that the unidentified lump is just one of ten in an orderly row not there an hour ago. The perceptual evidence remains the same, but the burden of proof is on anyone who wants to maintain in this situation that the lump is a stump.

The other paradigm is knowledge of extrapolation relations. The problem here is that the thesis is one of such high generality that the burden of proof is already heavily against it! How can a few observations of smoke with fire ground a belief that there is fire *whenever* there is smoke? Suppositional argument has a different supportive role here. Its function is to square the scales, to neutralise the presumption *against* the belief in generality. It does so by finding *prima facie* undesirable consequences in the supposition that an associate condition or counterexample exists. Then sampling (observation only of confirmatory instances in the course of a suitably extensive search for counterexamples), though still weak evidence, can tilt the scale in its favour.

A presumptive argument moves from conjecture to unacceptable consequence. Modern writers often identify it with the medieval technique of *reductio ad absurdum*, but in fact its scope is wider. The 'unacceptable consequence' can be an out-and-out contradiction but need not be so. For we are not trying to prove that the supposition is false, but only to shift the burden of proof onto anyone who would maintain it. And for this it is enough simply to demonstrate that the supposition comes into conflict with some well-attested norm on rationality. Udayana, the first to offer any systematic discussion, does not even mention contradiction as a species of unacceptable consequence. He says[25] that presumptive argumentation is of five types –

1. self-dependence (*ātmāśraya*)
2. mutual dependence (*itaretarāśraya*)
3. cyclical dependence (*cakraka*)
4. lack of foundation (*anavasthā*)
5. undesirable consequence (*aniṣṭaprasaṅga*)

The last of these is really just the generic case, what distinguishes presumptive argumentation in general. The first four form a tight logical group. If the supposition is the proposition A, then the four types of unacceptable consequence are (1) proving A from A, (2) proving A from B, and B from A, (3) proving A from B, B from C, and C from A — or any higher number of intermediate proof steps eventually leading back to A, and (4) proving

[25] *Ātmatattvaviveka*, p. 863.

A from B, B from C, C from D, ..., without end. So what presumptive argumentation must show is that the supposition is ungrounded, its proof being either regressive or question-begging.

Two points are noteworthy about Udayana's list. First, rational unacceptability bears upon the proof adduced for the supposition, not the supposition itself. The underlying implication is that one has the right to presume that one's thesis is correct if one can find fault with the opponent's proof of the antithesis. Principles of this sort are familiar from discussion of the informal logic of arguments from ignorance in which one claims entitlement to assert A on the grounds that it is not known (or proved) that $\neg A$.[26] In general such a claim must be unfounded – it amounts to the universal appropriation of a presumptive right in all circumstances.

The second point to notice about Udayana's list, however, is that it is very narrow. Udayana places strict constraints on what will count as an unacceptable consequence, constraints which are more formal than broadly rational. Conflict with other well-attested belief is not mentioned, for instance. Udayana severely limits the scope of presumptive argumentation. His motive, perhaps, is to disarm the sceptic. For presumptive argumentation is the favoured kind of reasoning of the sceptic-dialecticians (and indeed the term Udayana uses is *prasaṅga*, the same term Nāgārjuna had used for his dialectical method). Sceptics typically will want to loosen the conditions on what constitutes an unacceptable consequence of a supposition, so that the scope for refutation is expanded. So what Udayana seems to be saying is that one does indeed have the right to presume that one's thesis is correct when the argument for the counter-thesis commits a fallacy of a particularly gross type — not mere conflict with other beliefs but formal lack of foundation. If the best argument for the antithesis is *that* bad, then one has a *prima facie* entitlement to one's thesis.

Śrīharṣa (c. AD 1140) is an Advaita dialectician, a poet and a sceptic.[27] He expands the notion of unacceptable consequence, noticing several additional types unmentioned by Udayana.[28] One is 'self-contradiction' (*vyāghāta*). It was Udayana himself[29] who analysed the notion of opposition as noncompossibility, and cited as examples the statements "My mother is

[26] Douglas Walton, *Arguments from Ignorance* (University Park, PA: The Pennsylvania State University Press, 1996).

[27] On Śrīharṣa: Phyllis Granoff, *Philosophy and Argument in Late Vedānta: Śrīhaṛṣa's Khaṇḍanakhaṇḍakhādya* (Dordrecht: Reidel Publishing Company, 1978); Stephen Phillips, *Classical Indian Metaphysics* (La Salle: Open Court, 1995), chapter 3.

[28] *Khaṇḍanakhaṇḍakhādya* IV, 19 (*aprasaṇgātmakatarkanirūpaṇa*, pp. 777–788, 1979 edition; section numbering follows this edition). Śrīharṛṣa the negative dialectician wants to criticise even the varieties of presumptive argumentation, although his own method depends upon it. So he says: "By us indeed were presumptive argumentations installed in place, and so we do not reject them with [such] counter-arguments. As it is said – 'it is wrong to cut down even a poisonous tree, having cultivated it oneself"' (p. 787).

[29] *Ātmatattvaviveka*, p. 533.

childless," "I am unable to speak", and "I do not know this jar to be a jar." In the first instance, the noncompossibility is in what the assertion states, in the second it is in the speech-act itself, while in the third the propositional attitude self-ascription is self-refuting (a case akin to the Cartesian impossibility of thinking that one is not thinking).

Another refutation-exacting circumstance is the one called 'recrimination' (*pratibandī*). This is a situation in which one's opponent accuses one of advancing a faulty proof, when his own proof suffers exactly the same fault! There is a disagreement about what this state of equifallaciousness does to the burden of proof. The practice of Naiyāyikas is to take the circumstance as tilting the balance against the opponent – the opponent discredits himself in pressing an accusation without seeing that it can be applied with equal force to his own argument. But Śrīharṣa quotes with approval Kumārila's assertion that "all things being equal, where the same fault afflicts both positions one should not be censured [and not the other]".[30]

Śrīharṣa, the sceptic, would like to see both parties refuted by this circumstance. The same point underlies his mention as an unacceptable consequence the circumstance of 'lack of differential evidence' (*vinigamanāviraha*), when thesis and antithesis are in the same evidential situation. Again, what we see is a jostling with the burden of proof. Here Śrīharṣa is saying that absence of differential evidence puts a burden of proof on both thesis and antithesis — doubt itself refutes. It is the sceptic's strategy always to seek to maximise the burden of proof, and so to deny that anyone ever has the right to presume their position to be correct. That is, as Stanisław Schayer observed a long time ago, a difference between the *tarka* of the Naiyāyika and the *prasaṅga* of a sceptic like Śrīharṣa or Nāgārjuna.[31] For the latter, the demonstration that a thesis has an allegedly false consequence does not commit the refuter to an endorsement of the antithesis. Nāgārjuna wants to maintain instead that thesis and antithesis share a false existential precommitment.

Simplicity (*laghutva*) is, Śrīharṣa considers and the Naiyāyikas agree, a *ceteris paribus* preference-condition. Of two evidentially equivalent and otherwise rationally acceptable theses, the simpler one is to be preferred. The burden of proof lies with someone who wishes to defend a more complex hypothesis when a simpler one is at hand. The Nyāya cosmological argument appeals to simplicity when it infers from the world as product to a *single* producer rather than a multiplicity of producers. Here too the role of the simplicity consideration is to affect the burden of proof, not itself to

[30] *Khaṇḍanakhaṇḍakhādya* II, 2 (*pratibandīlakṣaṇakhaṇḍana*, pp. 571–572). The full quotation is given in his commentary by Śaṃkara Miśra.

[31] Stanisław Schayer, "Studies on Indian Logic, Part II: Ancient Indian Anticipations of Propositional Logic," [1933], translated into English by Joerg Tuske in Jonardon Ganeri ed., *Indian Logic: A Reader*.

prove. Cohen and Nagel[32] make a related point when they diagnose as the 'fallacy of simplism' the mistake of thinking that "of any two hypotheses, the simpler is the *true* one." In any case, simplicity can be a product not of the content of a hypothesis but only of its mode of presentation — the distinction is made by the Naiyāyikas themselves.[33] And it is hard to see how it can be rational to prefer one hypothesis to another only because it is simpler in form.

We have assumed that the rival hypotheses are both empirically adequate, that is to say, they are both consistent with all known facts. Śrīharṣa mentions an unacceptable consequence involving empirical evidence (*utsarga*). It is an objection to the usual idea that if there is empirical evidence supporting one hypothesis but not the other, then the first is confirmed. Śrīharṣa's sceptical claim is that a hypothesis must be considered refuted unless it is conclusively proved; nonconclusive empirical evidence does nothing to affect this burden of proof. Likewise, he says, a hypothesis must be considered refuted if it is incapable of being proved or disproved — this at least seems to be the import of the unacceptable consequence he calls 'impertinence' (*anucitya*) or 'impudence' (*vaiyātya*).

Other varieties of suppositional refutation have been suggested along lines similar to the ones we have reviewed. Different authors propose different sets of criteria for rational nonacceptance. What we have seen is that there is, in the background, a jostling over the weight and place of the burden of proof. The sceptic presses in the direction of one extreme — that a thesis can be considered refuted unless definitively proven. The constructive epistemologist tries to press in the direction of the opposite extreme — that a thesis can be considered proved unless definitively disproved. The truth lies somewhere in between, and it is the role of presumptive argumentation to locate it.

2 BUDDHIST CONTRIBUTIONS IN INDIAN LOGIC: FORMAL CRITERIA FOR GOOD ARGUMENTATION

2.1 The doctrine of the triple condition (trairūpya)

The Buddhist logician Diṅnāga (c. 480–540 AD) recommends a fundamental restructuring of the early Nyāya analysis of reasoned extrapolation and inference. Recall that analysis. It is an inference from likeness and unalikeness. In the one case, some object is inferred to have the target property on the grounds that it is 'like' a paradigmatic example. The untasted grain of rice is inferred to be cooked on the grounds that it is in the same pan

[32] Morris R. Cohen and Ernest Nagel, *An Introduction to Logic and Scientific Method* (London: Routledge & Kegan Paul, 1934), p. 384.

[33] Bhimacarya Jhalakikar, *Nyāyakośa or Dictionary of Technical Terms of Indian Philosophy* (Poona: Bhandarkar Oriental Research Institute, 1928), s.v. *laghutvam*.

as a test grain which is found to be cooked. In the other case, the object is inferred to have the target property on the grounds that it is 'unlike' an example lacking the target property. Likeness and unalikeness are matters of sharing or not sharing some property, the reason-property or evidence grounding the inference. Examples are either 'positive' — having both the reason and the target property, or 'negative' — lacking both. Extrapolation is the process of extrapolating a property from one object to another on the basis of a likeness or unalikeness between them.

The difficulty is that not every such extrapolation is rational or warranted. The extrapolation of a property from one object to another is warranted only when the two objects are *relevantly* alike or *relevantly* unalike. That two objects are both blue does not warrant an extrapolation of solidity from one to the other; neither can we infer that they are different in respect to solidity because they are of different colours. What one needs, then, is a theory of relevant likeness or unalikeness, a theory, in other words, of the type of property (the reason property) two objects must share if one is to be licensed to extrapolate another property (the target property) from one to the other.

This is exactly what Diṅnāga gives in his celebrated theory of the 'reason with three characteristics' (*trairūpya*). Diṅnāga's thesis is that relevant likeness is an exclusion relation. Two objects are relevantly alike with respect to the extrapolation of a property S just in case they share a property excluded from what is other than S. In other words, a reason property H for the extrapolation of a target property S is a property no wider in extension than S (assuming that 'non' is such that $H \cap \text{non}S = \emptyset$ iff $H \subseteq S$). Here is the crucial passage in the *Pramāṇa-samuccaya*, or *Collection on Knowing*:

> The phrase [from II 1b] "through a reason that has three characteristics" must be explained.
>
> > [A proper reason must be] present in the site of inference and in what is like it and absent in what is not [II 5cd].
>
> The object of inference is a property-bearer qualified by a property. After observing [the reason] there, either through perception or through inference, one also establishes in a general manner [its] presence in some or all of the same class. Why is that? Because the restriction is such that [the reason] is present *only* in what is alike, there is no restriction that it is *only* present. But in that case nothing is accomplished by saying that [the reason] is "absent in what is not". This statement is made in order to determine that [the reason], absent in what is not [like the site of inference], is not in what is other than or incompatible

> with the object of inference. Here then is the reason with three characteristics from which we discern the reason-bearer.

Diṅnāga's important innovation is to take the notions of likeness and unalikeness in extrapolation to be relative to the *target* property rather than the reason property. Two objects are 'alike' if they both have, or both lack, the target property. Two objects are 'unalike' if one has and the other lacks the target property. We want to know if our object — the 'site' of the inference — has the target property or not. What we do know is that our object has some other property, the reason property. So what is the formal feature of that reason property, in virtue of which its presence in our object determines the presence or absence of the target property? The formal feature, Diṅnāga claims, is that the reason property is present only in what is alike and absent in whatever is unalike our object.

This can happen in one of two ways. It happens if the reason property is absent from everything not possessing the target property and present only in things possessing the target property. Then we can infer that our object too possesses the target property. It can also happen if the reason property is absent from everything possessing the target property and present only in things not possessing the target property. Then we can infer that our object does not possess the target property.

Call the class of objects which are like the site of the inference the 'likeness class', and the class of objects unlike the site the 'unlikeness class' (Diṅnāga's terms are *sapakṣa* and *vipakṣa*). Interpreters have traditionally taken the likeness class to be the class of objects which possess the target property, and the unlikeness class to be the class of objects which do not possess the target property. I read Diṅnāga differently. I take his use of the terms 'likeness' and 'unlikeness' here at face-value, and identify the likeness class with the class of things *in the same state* vis-à-vis the target property as the site of the inference. We do not know in advance what that state is, but neither do we need to. The pattern of distribution of the reason property tells us what we can infer – that the site has the target property, that it lacks it, or that we can infer nothing. My approach has several virtues, chief among which is that it preserves the central idea of likeness as a *relation* between objects rather than, as with the traditional interpretation, referring to a *property* of objects. I think it also avoids many of the exegetical problems that have arisen in the contemporary literature with regard to Diṅnāga's theory.

One of the traditional problems is whether the site of the inference is included in the likeness class or not.[34] If the likeness class is the class of objects possessing the target property, then to include it seems to beg the question the inference is trying to resolve: does the site have that property

[34] Tom F. Tillemans, "On sapakṣa," *Journal of Indian Philosophy* 18 (1990), pp. 53–80.

or not. But to exclude it implies that the union of the likeness and unlikeness classes does not exhaust the universe (the site cannot, for obvious reasons, be unlike itself). So one is left with two disjoint domains, and an apparently insuperable problem of induction – how can correlations between the reason property and the target property in one domain be *any* guide to their correlation in another, entirely disjoint, domain?[35]

If we take Diṅnāga's appeal to the idea of likeness at face-value, however, the problem simply does not arise. The site of the inference is in the likeness class on the assumption that likeness is a reflexive relation — but that begs no question, for we do not yet know whether the likeness class is the class of things which possess the target property, or the class of things which do not possess it. It is the class of things which are in the same state vis-à-vis the target property as the inferential site itself. We can, if needs be, refer to objects 'like the site but not identical to it;' or we can take likeness to be nonreflexive, and refer instead, if needs be, to 'the site *and* objects like it' — but this is a matter only of labelling, with no philosophical interest.

Another of the traditional problems with Diṅnāga's account is an alleged logical equivalence between the second and third conditions.[36] The second condition states that the reason property be present *only* in what is alike.[37] The third condition states that it be absent in what is not. But if it is present *only* in what is alike, it must be absent in what is not; and if it is absent in what is not alike, it must be present *only* in what is. Now it is clear that Diṅnāga's reason for inserting the particle *only* into his formula is to prevent a possible misunderstanding. The misunderstanding would be that of taking the second condition to assert that the reason property must be present in all like objects. That would be too strong a condition, ruling out any warranted inferences in which the reason property is strictly narrower than the target. On account of the meaning of the particle *only*, we can see that it is also one of the two readings of the statement:

> In what is alike, there is only the presence [of the reason]

where the particle *only* is inserted into the predicate position. Diṅnāga eliminates this unwanted reading of the second condition, but he does so in a disastrous way. He eliminates it by inserting the particle into the subject position:

> Only in what is alike, there is the presence [of the reason].

[35]Hans H. Herzberger, "Three Systems of Buddhist Logic," in B. K. Matilal and R. D. Evans eds., *Buddhist Logic and Epistemology: Studies in the Buddhist Analysis of Inference and Language* (Dordrecht: Reidel Publishing Company, 1982), pp. 59–76.

[36]Bimal Matilal, "Buddhist Logic and Epistemology," in Matilal and Evans (1982: 1–30); reprinted in Matilal, *The Character of Logic in India* (Albany: State University Of New York Press, 1998), chapter 4.

[37]There is some debate among scholars over whether it was Diṅnāga himself or his commenator Dharmakīrti who first inserts the particle *only* into the clauses.

The reason this is disastrous is that it makes the second condition logically equivalent to the third. Notice, however, that when *only* is in predicate position, there are still two readings. The reading one needs to isolate is the second of these two readings:

> In what is alike, there is *indeed* the presence [of the reason]

That is, the reason is present in *some* of what is alike.

Accordingly, the theory is this. The extrapolation of a property S to an object is grounded by the presence in that object of any property X such that X excludes non$S$ but not S. A reason property for S is any member of the class

$$\{X : X \cap S \neq \varnothing \ \& \ X \cap \ \text{non}S = \varnothing\}.$$

The clause 'but not S' (the second of Diṅnāga's three conditions) has a clear function now. It is there to rule out properties which exclude *both* non$S$ *and* S. Such properties are properties 'unique' to the particular object which is the site of the inference, and Diṅnāga does not accept as warranted any extrapolation based on them. I will look at his motives in the next section.

Reason properties are nonempty subsets of the properties whose extrapolation they ground. If two objects are 'alike' in sharing a property, and one has a second property of wider extension than the first, then so does the second. Inductive extrapolation, in effect, is grounded in the contraposed universal generalisation "where the reason, so the target." A difficult problem of induction remains – how can one come to know, or justifiably believe, that two properties stand in such a relation without surveying all their instances? Diṅnāga has no adequate answer to this problem (but see [Tuske, 1998; Peckaus, 2001]). Dharmakīrti, Diṅnāga's brilliant reinterpreter, does. His answer is that when the relation between the two properties is one of causal or metaphysical necessity, the observation of a few instances is sufficient to warrant our belief that it obtains (§2.3). Dṅnāga, however, is not interested in such questions. For him, the hard philosophical question is that of discovering the conditions for rational extrapolation. It is another issue whether those conditions can ever be known to obtain.

To sum up, Diṅnāga's three conditions on the reason are:

1.	Attachment	Presence in the site a attachment	(*pakṣadharmatā*)
2.	Association	Presence (only) in what is like	(*anvaya*)
3.	Dissociation	Absence in what is unlike	(*vyatireka*)

If we take these conditions to be independent, it follows that there are exactly seven kinds of extrapolative inferential fallacy — three ways for one of the conditions to fail, three ways for two conditions to fail, and one way for all three conditions to fail. So the new theory puts the concept of a fallacy on a more formal footing. A fallacy is no longer an interesting but

essentially *ad hoc* maxim on reasoned argument. It is now a formal failing of the putative reason to stand in the correct extrapolation-grounding relation. One way for the reason to fail is by not attaching to the site at all, thereby failing to ground any extrapolation of other properties to it. This is a failure of the first condition. Another way for the reason to fail is by 'straying' onto unlike objects, thereby falsifying the third condition. The presence of one property cannot prove the presence of another if it is sometimes present where the other one is not. (It can, however, prove the *absence* of the other if it is *only* present where the other is not — and then the absence of the first property is a proof of the *presence* of the second.) We might then think of the third condition as a 'no counter-example' condition, a counter-example to the extrapolation-warranting relation of subsumption being an object where the allegedly subsumed property is present along with the absence of its alleged subsumer. An extrapolation is grounded just as long as there are no counterexamples.

2.2 Diṅnāga's 'wheel of reasons' *(hetucakra)*

In addition to his *Pramāṇa-samuccaya*, Diṅnāga wrote another, very brief text on logic, the *Wheel of Reasons*, or *Hetucakranirṇaya*. Diṅnāga's aim here is to classify all the different types of argument which fit into the general schema $\langle p$ has s because it has $h\rangle$, and to give an example of each. It is here that he applies his theory of a triple-conditioned sign to show when an inference is sound or unsound, and the kinds of defect an inferential sign can suffer from. Hence, it leads to a classification of fallacious and non-fallacious inferences.

The 'wheel' or 'cycle' is in fact a 3 by 3 square, giving nine inference types. Diṅnāga derives the square as follows. A 'homologue' (*sapakṣa*) is defined as any object (excluding the locus of the inference) which is possesses the inferrable property, s. Now, a putative inferential sign, h, might be either (i) present in every homologue, (ii) present in only some of the homologues but not in others, or (iii) present in no homologue. Suppose we let 'sp' stand for the class of homologues. Then we can represent these three possibilities as 'sp+', 'sp±', and 'sp−' respectively. The same three possibilities are also available with respect to the class of heterologues (objects, excluding the locus, which do not possess the inferred property, s). We can denote these by 'vp+', 'vp±', and 'vp−' respectively. Thus, 'vp+' means that every member of vp (every heterologue) possesses the sign property, h, etc. Now since any putative inferential sign must either be present in all, some or no homologue, and also in either all, some or no heterologue, there are just nine possibilities (Figure 2):

Why does Diṅnāga say that only 2 and 8 are cases of a good inferential sign? Recall the three conditions on a good sign. The first is that the inferential sign must be present in the locus of inference. This is taken for

sp \ vp	+	-	±
+	1 deviating	2 goodK	3 deviating
-	4 contradictory	5 uniquely deviating	6 contradictory
±	7 deviating	8 good	9 deviating

Figure 2.

granted in the wheel. The second states that the inferential sign should be present in some (at least one) homologous case. In other words, a good sign is one for which either 'sp+' or 'sp±'. Thus the second condition rules out 4, 5 and 6. Similarly, the third condition states that the inferential sign should be absent from any heterologous case, i.e. that 'vp−'. This rules out 1, 4, 7 and 3, 6, 9. So only 2 and 8 represent inferential signs which meet all three conditions and generate good inferences. Note here that the third condition alone is sufficient to rule out every fallacious case except 5. Hence, seeing why Diṅnāga considers 'type-5' inferences to be unsound will reveal why he considered the second of the three conditions to be necessary (see below).

Diṅnāga next gives an illustration of each of the nine possibilities. They can be tabulated, as in Figure 3.

In each case, the locus of the inference is sound. Note that wherever possible, Diṅnāga cites both a 'positive confirming example', i.e. an object where both h and s are present, as well as a negative confirming example', i.e. an object where neither h nor s is present. Both support the inference. He also cites, where relevant, a 'counter-example', i.e. a case where h is present but s is absent. The existence of a counter-example undermines the inference. Let us look at four representative cases.

Case 2: A warranted inference. This inference reads: Sound is transitory, because it is created, e.g. a pot; space. Intuitively, this inference is sound, because the reason-property, createdness, is present only in places where the inferred property, transitoriness, is also present. Hence createdness is a

	s	h	positive example	negative example	counter-example
1	eternal	knowable	space	—	a pot
2	transitory	created	a pot	space	—
3	manmade	tansitory	a pot	space	lightning
4	eternal	crated	—	—	a pot
5	eternal	audible	—	a pot	—
6	eternal	manmade	—	lightning	a pot
7	natural	transitory	lightning	—	a pot
8	transitory	manmade	a pot	space	—
9	eternal	incorporeal	space	a pot	action

Figure 3.

good sign of transitoriness. The inference is supported first by an example where both are present, a pot, and second by an example where neither are present, space.

Case 3: 'deviating' (*asiddha*). This inference reads: Sound is manmade, because it is transitory, e.g. a pot; space. Intuitively, this inference is unsound, because the reason-property, transitoriness, is present in places where the inferred property, manmade, is absent. The counterexample cited is lightning — transitory but not manmade. Because we can find such a counter-example, the inferential sign is said to 'deviate' from the inferred property. Deviating inferences are ones which satisfy the second condition but fail the third.

Case 6: 'contradictory' (*viruddha*). The inference reads: Sound is eternal, because it is manmade, e.g. lightning. The sign here fails both conditions 2 and 3 — there is no case of a thing which is eternal and manmade, but there is a counter-example, for instance, a pot, which is manmade but non-eternal. Such an inference is called 'contradictory' because we can in fact infer to the contrary conclusion, namely that sound is non-eternal because it is manmade. We can do this because in the contrary inference, the homologous and heterologous domains are switched round.

Case 5: 'specific' (*asādhāraṇa*). Sound is eternal, because it is audible, e.g. a pot. The first point to notice is that there are no counter-examples to this inference, for there are no examples, outside the 'locus' domain of sounds, of an audible thing which is non-eternal. This is because there are no audible things other than sounds! Hence the third condition seems to be satisfied trivially. The characteristic of type-5 inferences is that the reason-property is 'unique' to the locus. According to Diṅnāga, such inferences are unsound, and the reason is that they fail the second condition - there is no homologue, i.e. an eternal thing other than sound, which is also audible.

But this just restates the characteristic feature of such inferences, it doesn't explain *why* they are unsound. Some modern authors argue that the significance of the second condition is more epistemological, than logical: the second condition implies that there must be a positive supporting example, and without such an example the inference, even if sound, carries no conviction. Diṅnāga might, however, have had a more formal or logical reason for rejecting type-5 inferences. The universal rule here is "Whatever is audible, apart from sound, is eternal". Now if a universal rule of the form '$(\forall x)(Fx \rightarrow Gx)$' is made true by there being no Fs, then so is the rule '$(\forall x)(Fx \rightarrow$ not-$Gx)$'. Hence, we could equally infer that sound is non-eternal because it is audible! This resembles the fault which the Nyāya called '*prakaraṇasama*' or 'indecisive'. Diṅnāga, it seems, wants to avoid this by saying that '$(\forall x)(Fx \rightarrow Gx)$' is true only if there is at least one F, which leads to the second condition.

Let us consider the argument from specifics further. I have said that an extrapolation-grounding property is a nonempty subproperty — a property narrower in extension than the property being extrapolated, and resident at least in the object to which that property is being extrapolated. The sweet smell of a lotus is a ground for extrapolating that it has a fragrance; its being a blue lotus is a ground for extrapolating its being a lotus. Extrapolation is a move from the specific to the general, from species to genus, from conjunction to conjunct. Extrapolation is a move upwards in the hierarchy of kinds. This model of extrapolation works well in most cases, but what happens at the extremes? The extreme in one direction is a most general property of all, a property possessed by everything. Existence or 'reality', if it is a property, is a property like this, and the theory entails that existence is always extrapolatable — the inference 'a is, because a is F' is always warranted. Diṅnāga's theory faces a minor technical difficulty here. Since everything exists, then everything is 'like' the site of the inference (in the same state as the site with respect to existence), and the unlikeness class is empty. So Diṅnāga has to be able to maintain that his third condition — absence of the reason property in every unlike object — is satisfied when there are no unlike objects. The universal quantifier must have no existential import. His innovative distinction between inference 'for oneself' (*svārthānumāna*) and inference 'for others' (*parārthānumāna*) is a help here. It is the distinction between the logical preconditions for warranted extrapolation and the debate-theoretic exigencies of persuasion. While it might be useful, even necessary, to be able to cite a supporting negative example if one's argument is to carry conviction and meet the public norms on believable inference, there is no corresponding requirement that the unlikeness class be nonempty if an extrapolation is to be warranted.

What happens at the other extreme? Extrapolation is a move from the more specific to the less specific, and the limit is the case when the reason property is entirely specific to the site of the inference. There is no doubt

but that Diṅnāga thinks that extrapolation breaks down at this limit. He calls such reason properties 'specific indeterminate' (*asādhāraṇānaikāntika*), and classifies them as bogus-reasons. Indeed it is the entire function of his second condition to rule out such properties. That is why the second condition insists that the reason property must be present in an object like the site. This condition is an addition to the first, that the reason property be present in the site — it demands that the reason be present in some other object like but not identical to the site. Diṅnāga's example in the *Collection on Knowing* [II 7d] is:

[Thesis] Sound is noneternal.

[Reason] Because it is audible.

In the *Wheel of Reasons* [5cd–7a], he gives another example:

[Thesis] Sound is eternal.

[Reason] Because it is audible.

What is the difference? In fact, the difference between these two examples holds the key to what Diṅnāga thinks is wrong. The property audibility, something specific to sound, does not determine whether sound is eternal or noneternal. In either case, audibility is absent from what is unlike sound (because it is unique to sound) but also from what is like sound (except for sound itself). This symmetry in the distribution of the reason property undermines its capacity to discriminate between truth and falsity. To put it another way, if we take the universal quantifier to range over everything except the site of the inference, sound, then it is true both that everything audible is eternal and that everything audible is noneternal – both are true only because there are no audibles in the range of the quantifier.

This seems to be Diṅnāga's point, but it is not very satisfactory. Sound *is* either eternal or noneternal, and so audibility *is* a subproperty of one or the other. One and only one of the above universal quantifications is true when the quantifier is *unrestricted*. In any case, just why is it that we should not reason from the specific properties of a thing? We do it all the time. Historical explanations are notoriously singular — unrepeated historical events are explained by specific features of their context. Diṅnāga, it seems, is like the follower of the deductive-nomological model in insisting on repeatability as a criterion of explanation. What about mundane cases like this one: the radio has stopped because I have unplugged it? Being unplugged by me is a property specific to the radio, and yet the form of the explanation seems unapproachable. Perhaps, however, what one should say is that the explanatory property is 'being unplugged', and not 'being unplugged by me', and the explanation rests on the generalisation 'whenever a radio is unplugged, it stops.' So then the restriction is not to any

property specific to the site, but only to those which are not merely tokens of some more general explanatory property. And yet there are still intuitively rational but specific inferences — that salt is soluble because it has a certain molecular structure, that helium is inert because it has a certain atomic number, flying creatures fly because they have wings. Why shouldn't the specific properties of a thing be implicated in inferences of its other properties?

What we see here is Diṅnāga's adherence to a strictly inductivist model of extrapolation. The specific property audibility does not ground an extrapolation of eternality or noneternality because there can be no inductive evidence for the extrapolation. Inductive evidence takes the form of objects in the likeness and unlikeness classes known to have or not to have the reason. One might think that one does have at least 'negative' evidence, for one knows that audibility is absent from any object in the unlikeness class. So why can one not infer from the fact that audibility is absent in unlike objects that it must be present in like objects? The answer is that one can indeed make that inference, but it does not get one very far. For we must recall again the way these classes are defined – as classes of objects like or unlike the site with respect to eternality. We do not know whether the site is eternal or noneternal, and in consequence we do not know whether unlike things are things which are noneternal or eternal. So while we have plenty of examples of eternal inaudibles and noneternal inaudibles, we still do not know which are the 'alike' ones and which the 'unalike'.

The explanation of salt's solubility by its specific molecular structure exemplifies a quite different model of explanation. It is a theoretical explanation resting on the postulates of physical chemistry. It is from theory, not from observation, that one infers that having an NaCl lattice structure is a subproperty of being soluble. Similarly, within the context of suitable theories about the nature of sound and secondary qualities, one might well be able to infer from sound's being audible to its being noneternal. Diṅnāga, in spite of his brilliance and originality, could not quite free himself from the old model of inference from sampling. His inclusion of the second condition was a concession to this old tradition. He should have dropped it. Later Buddhists, beginning with Dharmakīrti, did just that – they effectively dropped the second condition by adopting the reading of it that makes it logically equivalent to the third.

Diṅnāga's insistence that any acceptible inference should be accompanied by both positive and negative supporting examples provoked the Naiyāyika Uddyotakara to criticise and expand the Wheel. Uddyotakara points out that there are sound patterns of inference in which either the class of homologues or the class of heterologues is empty. These he calls the 'universally negative' (*kevala-vyatikekin*) and 'universally positive' (*kevalānvayin*) inferences. We now have a wheel with sixteen possible cases (Figure 4):

Here, 'o' means that the class (sp or vp) is empty. An example of a

sp \ vp	+	−	±	o
+		good		good
-				
±		good		good
0		good		

Figure 4.

sound 'universally positive' inference might be: "This exists because I can see it". There are no heterologues, because there are no things which do not exist, and so there are no negatively supporting examples. Nevertheless, we should recognise the acceptibility such an inference. Examples of 'universally negative' inferences are more difficult to find. The later Nyāya link such inferences with their theory of definition, considering such examples as "Cows are distinct from non-cows, because they have dewlap'. There are no objects which are distinct from non-cows except for cows, and hence no homologues. But the inference might have significance, for it tells us that the property of having dewlap serves to distinguish cows from non-cows, and hence can be used as a definition of cowhood[38].

2.3 Arguments from effect, essence and non-observation

Dharmakīrti (AD 600–660) offers a substantive account of the conditions under which the observation of a sample warrants extrapolation. His claim is that this is so if the reason property is one of three types: an 'effect' reason (*kārya-hetu*), a natural reason (*svabhāva-hetu*), or a reason based on nonobservation (*anupalabdhi-hetu*).[39]

In each case, the presence of the reason in some sense *necessitates* the presence of the target. An effect-reason is a property whose presence is *causally* necessitated by the presence of the target property – for example, inferring that the mountain has fire on it, because of smoke above it. The reason-target relation is a causal relation. Clearly one can, and later philosophers[40] indeed did, extend this to cover other species of causal inference,

[38]For further discussion, see B. K. Matilal, "Introducing Indian Logic", in Matilal (1998), reprinted in Jonardon Ganeri ed., (*Indian Logic: A Reader*

[39]Dharmakīrti, *Nyāyabindu* II 11–12.

[40]See Mokṣākaragupta's eleventh century *Tarkabhāṣā* or *Language of Reason*. Yuichi

such as cases when reason and target are both effects of a common cause. The generalisation 'night follows day' is true, not because day causes night but because both day and night are caused by the rotation of the earth. An example often cited is the inference of lemon-colour from lemon-taste, when both are products of the same cause, viz. the lemon itself. Still another example is the inference of ashes from smoke: ashes and smoke are both effects of fire. Such an inference has two steps. First, fire is inferred from smoke; second, ash is inferred from fire. The second step, in which we infer an effect from its cause, is possible only because ash is a necessary effect of fire.

A natural reason is one whose presence *metaphysically* necessitates that of the target property, for example the inference that something is a tree because it is a *śiṃśapā* (a species of tree). Dharmakīrti appears to regard the law "all *śiṃśapā*s are trees" as necessarily true, even if its truth has to be discovered by observation, and thus to anticipate the idea that there are *a posteriori* necessities.[41] He states, surprisingly, that the reason-target relation in such inferences is the relation of identity. Why? Perhaps his idea is that the two properties being-a-*śiṃśapā* and being-a-tree are token-identical, for the particular tree does not have two distinct properties, being-a-*śiṃśapā* and a separate property being-a-tree, any more than something which weighs one kilogramme has two properties, having-weight and having-weight-one-kilogramme. The properties as types are distinct, but their tokens in individual objects are identical. Trope-theoretically, the point can easily be understood. The very same trope is a member of two properties, one wider in extension than the other, just as the class of blue tropes is a subset of the class of colour tropes. But a blue object does not have two tropes – one from the class of blue tropes and one from the class of colour tropes. It is the self-same trope.

Is absence of evidence evidence of absence? According to Dharmakīrti, nonobservation sometimes proves absence: my failure to see an object, when all the conditions for its perception are met, is grounds for an inference that it is not here. The pattern of argument such inferences exemplify was known to the medievals as *argumentum ad ignorantiam*, or an 'argument from ignorance.' The pattern occurs whenever one infers that p on the grounds that there is no evidence that p is false. Dharmakīrti states that the argument depends on the object's being perceptible, i.e. that all the conditions for its perception (other than its actual presence) are met in the given situation. Douglas Walton, in a major study of arguments from ignorance,[42] claims that they depend for their validity on an implicit conditional premise — if p

Kajiyama, *An Introduction to Buddhist Philosophy: An Annotated Translation of the Tarkabhāṣā of Mokṣākaragupta*, Memoirs of the Faculty of Letters (Kyoto) 10 (1966), pp. 74–76.

[41] *Pramāṇavārttika* I, 39–42.

[42] Walton (1996).

were false, p would be known to be false. The characteristics of an argument from ignorance are then a 'lack-of-knowledge' premise — it is not known that not-p, and a 'search' premise — if p were false, it would be known that not-p. The underlying hidden premise mentioned by Dharmakīrti seems to be exactly the one Walton gives: if the object were here, one would see it. The necessity here is subjunctive. The argument has a presumptive status - one has a right to presume the conclusion to be true to the extent that one has searched for and failed to find counter-evidence. It is this idea that is strikingly absent in Diṅnāga. Warranted extrapolation depends not on the mere nonobservation of counterexamples, but on one's failing to find them in the course of a suitably extensive search.

In each of the three cases, the universal relation between reason and target is a relation not of coincidence but of necessity - causal, metaphysical or subjunctive. Dharmakīrti's solution to the problem of induction, then, is to claim that observation supports a generalisation only when that generalisation is lawlike or necessary. In this, I think he anticipates the idea that the distinction between lawlike and accidental generalisations is that only the former support the *counterfactual* 'if the reason property were instantiated here, so would be the target property'. In such a context, let us note, the observation of even a single positive example might sometimes be sufficient to warrant the extrapolation: I infer that any mango is sweet having tasted a single mango; I infer that any fire will burn having once been burnt.

Extrapolation is warranted when the reason-target is lawlike, but it does not follow that the extrapolator must know that it is lawlike. What Dharmakīrti has succeeded in doing is to describe the conditions under which extrapolation *works* — the conditions under which one's actions, were they to be in accordance with the extrapolation, would meet with success. It is a description of the type of circumstance in which extrapolation is rewarded (i.e. *true* — if, as it seems, Dharmakīrti has a pragmatic theory of truth[43]). As to how, when or whether one can know that one is in such a circumstance, that is another problem altogether and not one that Dharmakīrti has necessarily to address. For a general theory of rationality issues in conditions of the form 'in circumstances C, it is rational to do ϕ' or 'in circumstances C, it is rational to believe p'. And this is precisely the form Dharmakīrti's conditions take.

2.4 The Jaina reformulation of the triple condition

Diṅnāga had argued that there are three marks individually necessary and jointly sufficient for the warranted extrapolation from reason to target (§2.2).

[43]Shoryu Katsura, "Dharmakīrti's Concept of Truth," *Journal of Indian Philosophy* 12 (1984), pp. 213–235. Georges B. J. Dreyfus, *Recognizing Reality: Dharmakīrti's Philosophy and its Tibetan Interpretations* (Albany: State University of New York Press, 1997), chapter 17.

They are (1) that the reason be present in the site of the extrapolation, (2) that the reason be present (only) in what is similar to the target, and (3) that the reason be absent in what is dissimilar to the target. The second of these conditions is, arguably, equivalent to the third, which asserts that the reason property is absent when the target property is absent. That was supposed to capture the idea of a 'no counterexample' condition, according to which an extrapolation is warranted just in case there is nothing in which the reason is present but not the target. What happens to this account if one allows, as the Jaina logicians do, that a property and its absence be compossible in a single object?[44] What happens is that the three marks cease to be sufficient for warranted extrapolation. In particular, the third mark no longer captures the idea behind the 'no counterexample' condition. For now the absence of the reason property in a place where the target is absent does not preclude its *presence* there too! So the third mark can be satisfied and yet there still be counterexamples — cases of the presence of the reason together with the absence of the target.

The Jainas indeed claim that the three marks are neither necessary nor sufficient for warranted extrapolation. Their response is to substitute for the three marks a new, single, mark. It is clear that if the presence and absence of a property are compossible, then a distinction needs to be drawn between absence and nonpresence. The first is consistent with the presence of the property; the second is not. Early post-Diṅnāga Jainas like Akalaṅka and Siddhasena described the new mark in quasi-Buddhistic terms, as 'no presence without' (*a-vinā-bhāva*) — i.e. no presence of the reason without the target. Thus Akalaṅka:[45]

> An extrapolation is a cognition of what is signified from a sign known to have the single mark of no presence without the target (*sādhyāvinābhāva*). Its result is blocking and other cognitions.
>
> The relata of the causality and identity relations cannot be cognised without the suppositional knowledge (*tarka*) of their being impossible otherwise, [which is] the proof that this is the single mark even without those relations. Nor is a tree the own-nature (*svabhāva*) or the effect (*kārya*) of such things as shade. And there is no disagreement here.

There is an obvious reference to and criticism of Dharmakīrti here,[46] and also a mention of the important idea, which we have already discussed, that presumptive argumentation (*tarka*) is what gives us knowledge of the

[44] See B. K. Matilal, *The Central Philosophy of Jainism* (Ahmedabad: L. D. Institute of Indology, 1981).

[45] *Laghīyastraya*, verse 12.

[46] On Akalaṇka on Dharmakīrti: Nagin J. Shah, Akalaṇka's Criticism of Dharmakīrti's Philosophy (Ahmedabad: L. D. Institute, 1967), pp. 267–270.

universal generalisations grounding extrapolations. The crucial difference from the Buddhists is in the meaning of 'no presence'. For the Jainas, it has to stand for nonpresence and not for absence. That led them to reformulate the reason-target relation as a relation of necessitation. Siddhasena:

> The mark of a reason is 'being impossible otherwise' (*anyathānupannatva*) [Nyāyāvatāra 22].

Vādideva Sūri gives the developed Jaina formulation:

> A reason has a single mark, 'determined as impossible otherwise'. It does not have three marks, for fallacies are then still possible [Pramāṇanaya-tattālokālaṃkāra 3.11–12].

The idea is that the reason *cannot* be present if the target is not. It is impossible for the reason to be present otherwise than if the target is present. The presence of the reason necessitates the presence of the target.

I said that Diṅnāga's three marks are, for the Jainas, neither necessary nor sufficient. They are not sufficient because they permit extrapolation when the reason is both present and absent, and the target nonpresent. On what grounds are they thought not to be necessary? The theory of extrapolation as developed first by the early Naiyāyikas and then by Diṅnāga has a built-in simplifying assumption. The assumption is that extrapolation is always a matter of inferring from the presence of one property in an object to the presence of a second property *in that same object*. But that assumption excludes many intuitively warranted extrapolations. The main examples considered by the Jainas are: (i) the *śakaṭa* star-group will rise because *kṛttikā* star-group has risen; (ii) the sun is above the horizon because the earth is in light; (iii) there is a moon in the sky because there is a moon in the water.

These examples are said to prove that the first of Diṅnāga's three marks, that the reason property is present in the site, is not a necessary condition on warranted extrapolation. And yet, while it is certainly desirable to broaden the reach of the theory to cover new patterns of extrapolative inference, it is not very clear what these examples show. What is the underlying generalisation? What are the similar and dissimilar examples? In the first case, the extrapolation seems to be grounded in the universal generalisation 'whenever the *kṛttikā* arises, so too does the *śakaṭa*.' But then there is indeed a single site of extrapolation — the present time. The inference is: the *śakaṭa* will rise now because *kṛttikā* has now risen. A similar point could be made about the second example. There seems indeed to be an implicit temporal reference in both of the first two cases, an extrapolation grounded in a universal generalisation over times.

The third case is more convincing, yet here too one might try to discern a common site. For the true form of the extrapolation is: the moon is in

the sky because *it* is reflected in the water, an extrapolation grounded in a universal generalisation of the form 'objects cause their own reflections'. Certainly, however, there are patterns of extrapolation for which the 'single site' condition does not hold. If, for example, one can find a universal generalisation of the form '$\forall x \exists y(Fx \rightarrow Gy)$', then from '$\exists x Fx$' one can infer '$\exists x Gx$'. Perhaps *this* is the pattern of extrapolation the Jainas intend to exemplify with their example of a sky-moon and a water-moon. If so, it is represents an important criticism of a simplifying, but in the end also restricting, assumption in the classical theory of extrapolation.

3 JAINA CONTRIBUTIONS IN INDIAN LOGIC: THE LOGIC OF ASSERTION

3.1 Rationality and Consistency

What is the rational response when confronted with a set of propositions each of which we have some reason to accept, and yet which taken together form an inconsistent class? This was, in a nutshell, the problem addressed by the Jaina logicians of classical India, and the solution they gave is, I think, of great interest, both for what it tells us about the relationship between rationality and consistency, and for what we can learn about the logical basis of philosophical pluralism. The Jainas claim that we can continue to reason in spite of the presence of inconsistencies, and indeed construct a many-valued logical system tailored to the purpose. My aim in this chapter is to offer an interpretation of that system and to try to draw out some of its philosophical implications.

There was in classical India a great deal of philosophical activity. Over the years, certain questions came to be seen as fundamental, and were hotly contested. Are there universals? Do objects endure or perdure? Are there souls, and, if so, are they eternal or non-eternal entities? Do there exist wholes over and above collections of parts? Different groups of philosophers offered different answers to these and many other such questions, and each, moreover, was able to supply plausible arguments in favour of their position, or to offer a world-view from which their particular answers seemed true. The body of philosophical discourse collectively contained therefore, a mass of assertions and contradictory counter-assertions, behind each of which there lay a battery of plausible arguments. Such a situation is by no means unique to philosophical discourse. Consider, for instance, the current status of physical theory, which comprises two sub-theories, relativity and quantum mechanics, each of which is extremely well supported, and yet which are mutually inconsistent. The same problem is met with in computer science, where a central notion, that of putting a query to a data-base, runs into trouble when the data-base contains data which is in-

consistent because it is coming in from many different sources. For another example of the general phenomenon under discussion, consider the situation faced by an investigator using multiple-choice questionnaires, when the answers supplied in one context are in conflict with those supplied in another. Has the interrogee said 'yes' or 'no' to a given question, when they said 'yes' under one set of conditions but 'no' under another? Do their answers have any value at all, or should we simply discard the whole lot on account of its inconsistency? Perhaps the most apposite example of all is the case of a jury being presented with the evidence from a series of witnesses. Each witness, we might suppose, tells a consistent story, but the total evidence presented to the jury might itself well be inconsistent.

The situation the Jainas have in mind is one in which a globally inconsistent set of propositions, the totality of philosophical discourse, is divided into sub-sets, each of which is internally consistent. Any proposition might be supported by others from within the same subset. At the same time, the negation of that proposition might occur in a distinct, though possibly overlapping subset, and be supported by other propositions within it. Each such consistent sub-set of a globally inconsistent discourse, is what the Jainas call a "standpoint" (*naya*). A standpoint corresponds to a particular philosophical perspective.

Let us say that a proposition is *arguable* if it is assertible within some standpoint, i.e. if it is a member of a mutually supporting consistent set of propositions. The original problem posed was this: what is the rational reaction to a class of propositions, each of which is, in this sense, arguable, yet which is globally inconsistent? It seems that there are three broad types of response. The first, which I will dub *doctrinalism*, is to say that it will always be possible, in principle, to discover which of two inconsistent propositions is true, and which is false. Hence our reaction should be to reduce the inconsistent set to a consistent subset, by rejecting propositions which, on close examination, we find to be unwarranted. This is, of course, the ideal in philosophical debate, but it is a situation we are rarely if ever in. The problem was stipulated to be one such that we cannot decide, as impartial observers, which of the available standpoints, if any, is correct. If doctrinalism were the only option, then we would have no choice but to come down in favour of one or other of the standpoints, basing our selection, perhaps on historical, cultural, or sociological considerations, but not on logical ones.

A second response is that of *scepticism*. Here the idea is that the existence both of a reason to assert and a reason to reject a proposition itself constitutes a reason to deny that we can justifiably either assert or deny the proposition. A justification of a proposition can be defeated by an equally plausible justification of its negation. This sceptical reaction is at the same time a natural and philosophically interesting one, and indeed has been adopted by some philosophers, notably *Nāgārjuna* in India and the

Pyrrhonic sceptics as reported by Sextus Empiricus. Sextus, indeed states as the first of five arguments for scepticism, that philosophers have never been able to agree with one another, not even about the criteria we should use to settle controversies.

The third response is that of *pluralism*, and this is the response favoured by the Jainas. The pluralist finds some way conditionally to assent to each of the propositions, and she does so by recognising that the justification of a proposition is internal to a standpoint. In this way, the Jainas try "to establish a rapprochement between seemingly disagreeing philosophical schools"[47], thereby avoiding the dogmatism or "one-sidedness" from which such disagreements flow. Hence another name for their theory was *anekāntavāda*, the doctrine of "non-one-sidedness".[48]

In spite of appearances to the contrary, the sceptic and the pluralist have much in common. For although the sceptic rejects all the propositions while the pluralist endorses all of them, they both deny that we can solve the problem by privileging just one position, i.e. by adopting the position of the doctrinalist. (It seems, indeed, that scepticism and pluralism developed in tandem in India, both as critical reactions to the system-based philosophical institutions.) Note too that both are under pressure to revise classical logic. For the sceptic, the problem is with the law of excluded middle, the principle that for all p, either p or $\neg p$. The reason this is a problem for the sceptic is that she wishes to reject each proposition p without being forced to assent to its negation $\neg p$. The pluralist, on the other hand, has trouble with a different classical law, the law of non-contradiction, that for all p, it is not the case both that p and that $\neg p$, for she wishes to assent both to the proposition p and to its negation. While a comparative study of the two responses, sceptical and pluralist, would be of interest, I will here confine myself to developing the version of pluralism developed by the Jainas, and discussing the extent to which their system becomes paraconsistent. It is very often claimed that the Jainas 'embrace' inconsistency, but I will be arguing that this is not so, that we can understand their system by giving it a less strongly paraconsistent reading.

3.2 Jaina seven-valued logic

The Jaina philosophers support their pluralism by constructing a logic in which there are seven distinct semantic predicates (*bhaṅgī*), which, since they attach to sentences, we *might* think of as truth-values (for a slightly different interpretation, see Ganeri 2001, chapter 5). I will first set out the

[47] B.K. Matilal, *The Central Philosophy of Jainism*, Calcutta University Press, Calcutta 1977:61.

[48] For a good outline of these aspects of Jaina philosophical theory, see B.K. Matilal, *The Central Philosophy of Jainism*, and P. Dundas, *The Jains*, Routledge Press, London 1992.

system following the mode of description employed by the Jainas themselves, before attempting to reconstruct it in a modern idiom. I will follow here the twelfth century author Vādideva Sūri (1086–1169 A.D.), but similar descriptions are given by many others, including Prabhācandra, Malliṣena and Samantabhadra. This is what Vādideva Sūri says (*Pramāṇa-naya-tattvālokālaṅkāraḥ*, chapter 4, verses 15–21):[49]

> The seven predicate theory consists in the use of seven claims about sentences, each preceded by "arguably" or "conditionally" (*syāt*), [all] concerning a single object and its particular properties, composed of assertions and denials, either simultaneously or successively, and without contradiction. They are as follows:
>
> 1. Arguably, it (i.e. some object) exists (*syād asty eva*). The first predicate pertains to an assertion.
> 2. Arguably, it does not exist (*syān nāsty eva*). The second predicate pertains to a denial.
> 3. Arguably, it exists; arguably, it doesn't exist (*syād asty eva syān nāsty eva*). The third predicate pertains to successive assertion and denial.
> 4. Arguably, it is 'non-assertible' (*syād avaktavyam eva*). The fourth predicate pertains to a simultaneous assertion and denial.
> 5. Arguably, it exists; arguably it is non-assertible (*syād asty eva syād avaktavyam eva*). The fifth predicate pertains to an assertion and a simultaneous assertion and denial.
> 6. Arguably, it doesn't exist; arguably it is non-assertible (*syān nāsty eva syād avaktavyam eva*). The sixth predicate pertains to a denial and a simultaneous assertion and denial.
> 7. Arguably, it exists; arguably it doesn't exist; arguably it is non-assertible (*syād asty eva syān nāsty eva syād avaktavyam eva*). The seventh predicate pertains to a successive assertion and denial and a simultaneous assertion and denial.

The structure here is simple enough. There are three basic truth-values, true (t), false (f), and non-assertible (u). There is also some means of combining basic truth-values, to form four further compound values, which we can designate tf, tu, fu and tfu. There is a hint too that the third basic value is itself somehow a product of the first two, although by some other means of combination - hence the talk of simultaneous and successive assertion and

[49] Vādideva Sūri: 1967, *Pramāṇa-naya-tattvālokālaṃkāra*, ed. and transl. H. S. Battacharya, Jain Sahitya Vikas Mandal, Bombay.

denial. Thus, in Jaina seven valued logic, all the truth-values are thought to be combinations in some way or another of the two classical values.

There is, however, a clear risk that the seven values in this system will collapse trivially into three. For if the fifth value, tu, means simply "true and true-and-false", how is it distinct from the fourth value, u, "true-and-false"? No reconstruction of the Jaina system can be correct if it does not show how each of the seven values is distinct. The way forward is to pay due attention to the role of the conditionalising operator "arguably" (*syāt*). The literal meaning of "*syāt*" is "perhaps it is", the optative form of the verb "to be". The Jaina logicians do not, however, use it in quite its literal sense, which would imply that no assertion is not made categorically, but only as a possibility-claims. Instead, they use it to mean "from a certain standpoint" or "within a particular philosophical perspective". This is the Jaina pluralism: assertions are made categorically, but only from within a particular framework of supporting assertions. If we let the symbol "∇" represent "*syāt*", then the Jaina logic is a logic of sentences of the form "∇p", a logic of conditionally justified assertions. As we will see, it resembles other logics of assertion, especially the ones developed by Jaśkowski[50] and Rescher[51].

The first three of the seven predications now read as follows:

1. $|p| = t$ iff ∇p.

In other words, p is true iff it is arguable that p. We are to interpret this as saying that there is some standpoint within which p is justifiably asserted. We can thus write it as

1. $|p| = t$ iff $\exists\sigma\ \sigma : p$,

where "$\sigma : p$" means that p is arguable from the standpoint σ. For the second value we may similarly write,

2. $|p| = f$ iff $\nabla \quad \neg p$.

That is,

$$|p| = f \text{ iff } \exists\sigma\ \sigma : \neg p.$$

The third value is taken by those propositions whose status is controversial, in the sense that they can be asserted from some standpoints but their negations from others. These are the propositions which the Jainas are most concerned to accommodate. Thus

3. $|p| = tf$ iff $|p| = t$ & $|p| = f$.

[50] Jaśkowski, S.: 1948, "Propositional calculus for contradictory deductive systems"; English translation in *Studia Logica* **24**: 143 - 157 (1969).

[51] Rescher, N.: 1968, *Topics in Philosophical Logic*, Reidel, Dordrecht.

I.e.

$$|p| = tf \text{ iff } \nabla p \ \& \ \nabla \neg p,$$

or again

$$|p| = tf \text{ iff } \exists\sigma\ \sigma : p \ \& \ \exists\sigma\ \sigma : \neg p.$$

This way of introducing a new truth-value, by combining two others, may seem a little odd. I think, however, that we can see the idea behind it if we approach matters from another direction. Let us suppose that every standpoint is such that for any given proposition, either the proposition or its negation is assertible from within that standpoint. Later, I will argue that the Jainas did not want to make this assumption, and that this is what lies behind their introduction of the new truth-value "non-assertible". But for the moment let us make the assumption, which is tantamount to supposing that every standpoint is '*optimal*', in the sense that for any arbitrary proposition, it either supplies grounds for accepting it, or else grounds for denying it. There are no propositions about which an optimal standpoint is simply indifferent. Now, with respect to the totality of actual optimal standpoints, a proposition can be in just one of three states: either it is a member of every optimal standpoint, or its negation is a member of every such standpoint, or else it is a member of some, and its negation of the rest. If we number these three states, 1, 2 and 3, and call the totality of all actual standpoints, Σ, then the value of any proposition with respect to Σ is either 1, 2 or 3. The values 1, 2 and 3 are in fact the values of a three-valued logic, which we can designate **M3**. There is a correspondence between this logic and the system introduced by the Jainas (**J3**, say). The idea, roughly is that a proposition has the value 'true' iff it either has the value 1 or 3, it has the value 'false' iff it either has the value 2 or 3, and it has the value 'tf' iff it has the value 3. Hence the three values introduced by the Jainas represent, albeit indirectly, the three possible values a proposition may take with respect to the totality of optimal standpoints.

Before elaborating this point further, we must find an interpretation for the Jainas' fourth value "non-assertible". Bharucha and Kamat offer the following analysis of the fourth value:

> The fourth predication consists of affirmative and negative statements made simultaneously. Since an object X is incapable of being expressed in terms of existence and non-existence at the same time, even allowing for Syād, it is termed 'indescribable'. Hence we assign to the fourth predication ... the indeterminate truth-value I and denote the statement corresponding to the fourth predication as $(p\&\neg p)$.[52]

Bharucha and Kamat's interpretation is equivalent to

[52] Bharucha, F. and Kamat, R. V.: 1984, "*Syādvāda* theory of Jainism in terms of deviant logic", *Indian Philosophical Quarterly*, **9**: 181 – 187; 183.

4. $|p| = u$ iff $\nabla(p\&\neg p)$,

that is

$$|p| = u \text{ iff } \exists\sigma\ \sigma : (p\&\neg p).$$

Thus, for Bharucha and Kamat, the Jaina system is paraconsistent because it allows for standpoints in which contradictions are justifiably assertible. This seems to me to identify the paraconsistent element in the Jaina theory in quite the wrong place. For while there may be certain sentences, such as the Liar, which can justifiably be both asserted and denied, this cannot be the case for the wide variety of sentences which the Jainas have in mind, sentences like "There exist universals" and so on. Even aside from such worries, the current proposal has a technical defect. For what now is the fifth truth-value, tu? If Bharucha and Kamat are right then it means that there is some standpoint from which 'p' can be asserted, and some from which '$p\&\neg p$' can be asserted. But this is logically equivalent to u itself. The Bharucha and Kamat formulation fails to show how we get to a seven-valued logic.

Another proposed interpretation is due to Matilal. Taking at face-value the Jainas' elaboration of the fourth value as meaning "simultaneously both true and false", he says

> the direct and unequivocal challenge to the notion of contradiction in standard logic comes when it is claimed that the same proposition is both true and false at the same time in the same sense. This is exactly accomplished by the introduction of the [fourth] value - "Inexpressible", which can also be rendered as paradoxical.[53]

Matilal's intended interpretation seems thus to be

4. $|p| = u$ iff $\nabla(p, \neg p)$,

$$\text{i.e. } |p| = u \text{ iff } \exists\sigma(\sigma : p\&\sigma : \neg p).$$

Matilal's interpretation is a little weaker than Bharucha and Kamat, for he does not explicitly state that the conjunction '$p\&\neg p$' is asserted, only that both conjuncts are. Admittedly, the difference between Matilal and Bharucha and Kamat is very slight, and indeed only exists if we can somehow make out the claim that both a proposition and its negation are assertible without it being the case that their conjunction is. For example, we might think that the standpoint of physical theory can be consistently

[53] Matilal, B. K.: 1991, "*Anekānta*: both yes and no?", *Journal of Indian Council of Philosophical Research*, **8**: 1 – 12; 10.

extended by including the assertion that gods exists, and also by including the assertion that gods do not exist. It would not follow that one could from any standpoint assert the conjunction of these claims. Yet whether there is such a difference between Matilal's position and that of Bharucha and Kamat is rather immaterial, since Matilal's proposal clearly suffers from the precisely the same technical defect as theirs, namely the lack of distinctness between the fourth and fifth values.

Tere is another interpretation, one which gives an intuitive sense to the truth-value "non-assertible", sustains the distinctness of each of the seven values, but does not require us to abandon the assumption that standpoints are internally consistent. Recall that we earlier introduced the idea of an optimal standpoint, by means of the assumption that for every proposition, either it or its negation is justifiably assertible from within the standpoint. Suppose we now retract that assumption, and allow for the existence of standpoints which are just neutral about the truth or falsity of some propositions. We can then introduce a new value as follows:

4. $|p| = u \iff \exists\sigma(\neg(\sigma : p)\&\neg(\sigma : \neg p))$.

Neither the proposition nor its negation is assertible from the standpoint. For example, neither the proposition that happiness is a virtue nor its negation receives any justification from the standpoint of physical theory. We have, in effect, rejected a commutativity rule, that if it not the case that 'p' is assertible from a standpoint σ then '$\neg p$' is assertible from σ and vice versa $[\neg(\sigma : p) \iff (\sigma : \neg p)]$. Our new truth-value, u, is quite naturally called "non-assertible", and it is clear that the fifth value, tu, the conjunction of t with u, is not equivalent simply with u. The degree to which the Jaina system is paraconsistent is, on this interpretation, restricted to the sense in which a proposition can be tf, i.e. both true and false because assertible from one standpoint but deniable from another. It does not follow that there are standpoints from which contradictions can be asserted.

Why have so many writers on Jaina logic have felt that Jaina logic is paraconsistent in the much stronger sense. The reason for this belief is the account which some of the Jainas themselves give of the meaning of their third basic truth-value, "non-assertible". As we saw in the passage from Vādideva Sūri, some of them say that a proposition is non-assertible iff it is arguably both true and false *simultaneously*, as distinct from the truth value tf, which is *successively* arguably true and arguably false. We are interpreting the Jaina distinction between successive and simultaneous combination of truth-values in terms of a scope distinction with the operator "arguably". One reads "arguably $(t\&f)$", the other "(arguably t) & (arguably f)". If this were the correct analysis of the fourth truth-value, then Jaina logic would indeed be strongly paraconsistent, for it would be committed to the assumption that there are philosophical positions in which

contradictions are rationally assertible. Yet while such an interpretation is, on the face of it, the most natural way of reading Vādideva Sūri's elaboration of the distinction between the third and fourth values, it if far from clear that the Jaina pluralism really commits them to paraconsistency in this strong form. Their goal is, to be sure, to reconcile or synthesise mutually opposing philosophical positions, but they have no reason to suppose that a single philosophical standpoint can itself be inconsistent. Internal consistency was, in classical India, the essential attribute of a philosophical theory, and a universally acknowledged way to undermine the position of one's philosophical opponent was to show that their theory contradicted itself. The Jainas were as sensitive as anyone else to allegations that they were inconsistent, and strenuously denied such allegations when made. I have shown that it is possible to reconstruct Jaina seven-valued logic in a way which does not commit them to a strongly paraconsistent position.

The interpretation I give to the value "non-assertible" is quite intuitive, although it does not mean "both true and false *simultaneously*". My interpretation, moreover, is supported by at least one Jaina logician, Prabhācandra. Prabhācandra, who belongs to the first part of the ninth century C.E., is one of the few Jainas directly to address the question of why there should be just seven values. What he has to say is very interesting:

> (Opponent:) Just as the values 'true' and 'false', taken successively, form a new truth-value 'true-false', so do the values 'true' and 'true-false'. Therefore, the claim that there are seven truth-values is wrong.
>
> (Reply:) No: the successive combination of 'true' and 'true-false' does not form a new truth-value, because it is impossible to have 'true' twice. ... In the same way, the successive combination of 'false' and 'true-false' does not form a new truth-value.
>
> (Opponent:) How then does the combination of the first and the fourth, or the second and the fourth, or the third and the fourth, form a new value?
>
> (Reply:) It is because, in the fourth value "non-assertible", there is no grasp of truth or falsity. In fact, the word "non-assertible" does not denote the simultaneous combination of truth and falsity. What then? What is meant by the truth-value "non-assertible" is that it is impossible to say which of 'true' and 'false' it is.[54]

This passage seems to support the interpretation offered above. When talking about the "law of non-contradiction" in a deductive system, we must dis-

[54] Prabhācandra: 1941, *Prameyakamalamārtaṇḍa*, ed. M. K. Shastri, Nirnayasagar Press, Bombay; p. 683 line 7 ff.

tinguish between two quite different theses: (a) the thesis that "$\neg(p\&\neg p)$" is a theorem in the system, and (b) the thesis that it is not the case that both 'p' and '$\neg p$' are theorems. The Jainas are committed to the first of these theses, but reject the second. This is the sense in which it is correct to say that the Jainas reject the "law of non-contradiction".

I showed earlier that when we restrict ourselves to optimal standpoints, the total discourse falls into just one of three possible states with respect to each system. The Jainas have a *seven*-valued logic because, if we allow for the existence of non-optimal standpoints, standpoints which are just neutral with respect to some propositions, then, for each proposition, p say, the total discourse has exactly seven possible states. They are as follows:

1. p is a member of every standpoint in Σ.
2. $\neg$p is a member of every standpoint in Σ.
3. p is a member of some standpoints, and $\neg p$ is a member of the rest.
4. p is a member of some standpoints, the rest being neutral.
5. $\neg p$ is a member of some standpoints, the rest being neutral.
6. p is neutral with respect to every standpoint.
7. p is a member of some standpoints, $\neg p$ is a member of some other standpoints, and the rest are neutral.

Although Jainas do not define the states in this way, but rather via the possible combinations of the three primitive values, t, f and u, it is not difficult to see that the two sets map onto one another, just as they did before. Thus t = (1, 3, 4, 7), f = (2, 3 5, 7), tf = (3, 7), and so on.

Using many-valued logics in this way, it should be noted, does not involve any radical departure from classical logic. The Jainas stress their commitment to bivalence, when they try to show, as Vādideva Sūri did above, that the seven values in their system are all products of combining two basic values. This reflects, I think, a commitment to bivalence concerning the truth-values of propositions themselves. The underlying logic *within* each standpoint is classical, and it is further assumed that each standpoint or participant is internally consistent. The sometimes-made suggestion[55] that sense can be made of many-valued logics if we interpret the assignment of non-classical values to propositions via the assignment of classical values to related items is reflected here in the fact that the truth-value of any proposition p (i.e. $|p|$) has two values, the status of p with respect to standpoint σ ('$|p|_\sigma$') derivatively has three values, and the status of p with respect to a discourse Σ ('$|p|_\Sigma$'), as we have just seen, has seven.

[55] Haack, S.: 1974, *Deviant Logic*, Cambridge University Press, Cambridge; 64.

Consider again the earlier example of a jury faced with conflicting evidence from a variety of witnesses. The Jainas wouldn't here tell us 'who dun it', for they don't tell us the truth-value of any given proposition. What they give us is the means to discover patterns in the evidence, and how to reason from them. For example, if one proposition is agreed on by all the witnesses, and another is agreed on by some but not others, use of the Jaina system will assign different values to the two propositions. The Jainas, as pluralists, do not try to judge which of the witnesses is lying and which is telling the truth; their role is more like that of the court recorder, to present the totality of evidence in a maximally perspicuous form, one which still permits deduction from the totality of evidence.

So far so good. But there is another worry now, one which strikes at the very idea of using a many-valued logic as the basis for a logic of discourse. For, when we come to try and construct truth-tables for the logical constants in such a logic, we discover that the logic is not *truth-functional*. That is to say, the truth-value of a complex proposition such as '$p\&q$', is not a function solely of the truth-values of the constituent propositions 'p' and 'q'. To see this, and to begin to find a solution, I shall need briefly to describe the work of the Polish logician, Jaśkowski, who was the founder of discursive logics in the West, and whose work, in motivation at least, provides the nearest contemporary parallel to the Jaina theory.

3.3 Jaśkowski and the Jainas

Philosophical discourse is globally inconsistent, since there are many propositions to which some philosophers assent while others dissent. The Jainas therefore develop a logic of assertions-made-from-within-a-particular-standpoint, and note that an assertion can be both arguably true, i.e. justified by being a member of a consistent philosophical position, and at the same time be arguably false, if its negation is a member of some other consistent philosophical standpoint. This move is quite similar to that of the founder of inconsistent logics, Jaśkowski, who developed a "discussive logic" in which a proposition is said to be 'discussively true' iff it is asserted by some member of the discourse.

Jaśkowski motivates his paper "Propositional Calculus for Contradictory Deductive Systems" with two observations. The first is that

> any vagueness of the term a can result in a contradiction of sentences, because with reference to the same object X we may say that "X is a" and also "X is not a", according to the meanings of the term a adopted for the moment,

the second is that

> the evolution of the empirical sciences is marked by periods in which the theorists are unable to explain the results of experi-

> ments by a homogeneous and consistent theory, but use different hypotheses, which are not always consistent with one another, to explain the various groups of phenomena.[56]

He then introduces an important distinction between two properties of deductive systems. A deductive system is said to be *contradictory* if it includes pairs of theorems A and $\neg A$ which contradict each other. It is *over-complete*, on the other hand, if every well-formed formula is a theorem of the system. In classical logic, these two properties are conflated; hence the slogan "anything follows from a contradiction". The problem to which Jaśkowski addresses himself, therefore, is that of constructing a non-classical system which is contradictory but not over-complete. In classical logic, given two contradictory theses A, $\neg A$, we may deduce first that $A\&\neg A$, using the &-introduction or Adjunction Rule, $A, B \rightarrow A\&B$. Then, since $A\&\neg A$ iff $B\&\neg B$ for any arbitrary A and B, and since $B\&\neg B \rightarrow B$ from &-elimination or Simplification, $A\&B \rightarrow A$, it follows that B. More clearly:

1.	$A, \neg A$	
2.	$A\&\neg A$,	from 1 by Adjunction.
3.	$A\&\neg A$ iff $B\&\neg B$,	for any arbitrary A and B.
4.	$B\&\neg B \rightarrow B$,	by Simplification.
5.	$A\&\neg A \rightarrow B$,	from 3 and 4.
6.	B,	from 2 and 5 by Modus Ponens. ■

To get an inconsistent (contradictory but not over-complete) system, at least one step in this sequence must be broken. In Jaśkowski's new system, 'discursive logic', it is the Adjunction Rule which no longer holds. Jaśkowski considers the system in which many different participants makes assertions, each thereby contributing information to a single discourse. The best example, perhaps, is one already given, the evidence presented to a jury by witnesses at a trial. Jaśkowski then introduces the notion of discursive assertion, such that a sentence is discursively asserted if it is asserted by one of the participants in the discourse, and he notes that the operator "it is asserted by someone that..." is a modal operator for the semantics of which it should be possible to use an existing modal logic. Thus

> A is a theorem of **D2** iff $\Diamond A$,

where **D2** is Jaśkowski's two-valued discursive logic, and "$\Diamond$" is the operator "someone asserts that...". For some reason, Jaśkowski chooses a strong modal system, **S5**, to give the semantics of this operator, but this is surely a mistake. The reason is that the **S5** modal principle '$A \rightarrow \Diamond A$' does not seem to hold for a discursive system, since there will be truths which no-one

[56] Jaśkowski, S.: 1948, "Propositional calculus for contradictory deductive systems"; English translation in *Studia Logica*, **24**: 143 - 157 (1969); 144.

asserts. It would not be difficult, however, to use a weaker modal system than **S5**, for example $\mathbf{S2}^0$ or $\mathbf{S3}^0$, which lack the above principle, as the basis for **D2**. (The characteristic axiom of $\mathbf{S4}^0$, '$\Diamond\Diamond A \rightarrow \Diamond A$', does not seem to hold in a discursive system: it can be assertible from some standpoint that there is another standpoint in which p is assertible without there being such a standpoint). The point to note is that, in most modal systems, the Adjunction Rule fails, since it does not follow that the conjunction $A\&B$ is possible, even if A is possible and B is separately possible. And this too, is what we would expect from the discursive operator, for one participant may assert A, and another B, without there being anyone who asserts the conjunction. Jaśkowski therefore arrives at a system which is contradictory, since both A and $\neg$A can be theses, but, because it is non-adjunctive, is not over-complete.

3.4 The Logical Structure of the Jaina System

The parallels in motivation between Jaśkowski's discursive logic, and the Jaina system are unmistakable. There is, however, an important difference, to which I alluded earlier. Modal logics are not truth-functional; one cannot, for example, deduce the truth-value of '$\Diamond(A\&B)$' from the truth-values of '$\Diamond A$' and '$\Diamond B$'. And it seems for the same reason that a discursive logic cannot be truth-functional either. Suppose, for example, that we have two propositions A, and B, both of which are assertible from (possibly distinct) standpoints, and hence both true in the Jaina system. What is the truth-value of $A\&B$? It seems that this proposition could be either true, false, or both.

It is possible to offer a defence of the Jaina position here. For simplicity, let us restrict ourselves to the Jaina system with only optimal standpoints and just three truth-values. If my suggested defence works here, its extension to the full Jaina system **J7**, would not be especially problematic. Consider again the three-valued logic, **M3**, whose values were defined as follows:

$$|p| = 1 \text{ iff } \forall\sigma\ \sigma : p.$$
$$|p| = 2 \text{ iff } \forall\sigma\ \sigma : \neg p$$
$$|p| = 3 \text{ iff } \exists\sigma\ \sigma : p\ \&\ \exists\sigma\ \sigma : \neg p.$$

These correspond to the three possible states of a totality of optimal standpoints. When we try to construct the truth-table for conjunction in such a system, we find that it is non-truth-functional. Thus, consider the truth-value of '$p\&q$', when $|p| = |q| = 3$. Here, $|p\&q|$ might itself be 3, but it might also be 2. Thus, the truth-value of the conjunction is not uniquely determined by those of its conjuncts. What is uniquely determined, however, is that the truth-value belongs to the class (2, 3). To proceed, we can appeal to an idea first introduced by N. Rescher in his

paper "Quasi-truth-functional systems of propositional logic".[57] A quasi-truth-functional logic is defined there as one in which "some connectives are governed by many-valued functions of the truth-values of their variables". The entries in the truth-table of such a logic are typically not single truth-values but sets of values. It is clear that the system set up just now is, in this, sense, quasi-truth-functional. Now, as Rescher himself points out, a quasi-truth-functional logic will always be equivalent to a multi-valued strictly truth-functional system. The idea, roughly, is that we can treat a class of truth-values as constituting a new truth-value. Typically, if the quasi-truth-functional system has n truth-values, its strictly truth-functional equivalent will have 2^n - 1 values (Rescher notes that "in the case of a three-valued (T, F, I) quasi-truth-functional system we would need seven truth-values, to represent: T, F, I, (T, F), (T, I), (F, I), (T, F, I)" but argues that there are special reasons entailing that for a two-valued quasi-truth-functional system we need four rather than three values.). The seven-valued system which results in this way from the three-valued logic sketched above has, in fact, been studied notably by Moffat[58]. I will therefore call it **M7**. An initially tempting idea is to identify the Jaina system **J7** with **M7**. This, however, will only work if the fourth value, u, is defined thus:

$$|p| = u \text{ iff } \forall\sigma\ \sigma : p \vee \forall\sigma\ \sigma : \neg p.$$

For then '*tu*' in the Jaina system will be identical with '1' in the Moffat system, etc. This is, however, not an interpretation which receives any textual support.

Instead, let us observe that there is a close connection between **M7** and the restricted Jaina system, **J3**. For note that the value (1, 3) in **M7** is such that

$$\begin{array}{lll} |p| = (1,3) & \text{iff} & |p| = 1 \vee |p| = 3 \\ & \text{iff} & \forall\sigma\ \sigma : p \vee (\exists\sigma\ \sigma : p\ \&\ \exists\sigma\ \sigma : \neg p) \\ & \text{iff} & \exists\sigma\ \sigma : p. \end{array}$$

Thus (1, 3) in **M7** is just the value 'true' in **J3**. Similarly, (1, 2) in **M7** is just the value 'false' in **J3**. Thus, *although **J3** is not strictly truth-functional, its truth-tables are embedded in those of the Moffat logic, **M7**.*

It is presumably possible to find a quasi-truth-functional system whose truth-tables embed those of **J7**, the full Jaina system, in an entirely analogous way. Thus, although the loss of Adjunction means that the Jaina logic **J7**, is not truth-functional, its truth-table is embedded in a suitable

[57] Rescher, N.: 1962, "Quasi-truth-functional systems of propositional logic", *Journal of Symbolic Logic*, **27**: 1 - 10.

[58] Moffat, D. C. and Ritchie, G. D.: 1990, "Modal queries about partially-ordered plans", *J. Expt. Theor. Artif. Intell.*, **2**: 341 - 368. See also Priest, G.: 1984, "Hyper-contradictions," *Logique et Analyse*, **107**: 237–43.

quasi-functional system. The lack of truth-functionality is not, after all, a fatal flaw in the Jaina approach.

3.5 Axiomatisation of the Jaina System

We have shown that it is possible to use many-valued truth-tables to formalise the Jaina system. This was, in effect, the approach of the Jaina logicians themselves. Yet it would surely be much better to proceed by axiomatising the modal standpoint operator, ∇. Once again we look to Rescher[59]. His work on what he calls "assertion logics" is an extension of the work of Jaśkowski. Rescher introduces a system **A1**, with the following axiomatic basis:

(A1)	$(\exists p)\sigma : p$	[Nonvacuousness]
(A2)	$(\sigma : p \,\&\, \sigma : q) \supset \sigma : (p \,\&\, q)$	[Conjunction]
(A3)	$\neg\sigma : (p \,\&\, \neg p)$	[Consistency]
(R)	If $p \vdash q$, then $\sigma : p \vdash \sigma : q$	[Commitment]

Note that one effect of the rule (R) is to ensure that the notion captured is not merely explicit assertion but 'commitment to assert', for (R) states that from a standpoint one may assert anything entailed by another of the assertions. I believe that the Jainas would accept each of the axioms (A1) to (A3). Bharucha and Kamat, it may be noted, would reject (A3), while Matilal, as I have represented him, would reject (A2). I have already argued that these claims are mistaken. In particular, with regard to (A2), although it is true that the Jainas reject Adjunction, what this means is that assertions made from within different standpoints cannot be conjoined, not that assertions made within the same standpoint cannot be conjoined.

We now introduce the modal standpoint operator, ∇ "arguably", via the definition:

$$\nabla p \text{ iff } (\exists\sigma)\sigma : p,$$

and add the axioms of $\mathbf{S3}^0$ or some other suitable modal system.

Rescher defines some further systems by adding further axioms, none of which, I think, the Jainas would accept. For example, he defines **A2** by adding to **A1** the axiom that anything asserted by everyone is true $[(\forall\sigma)\sigma : p \supset p]$. There is no reason to suppose the Jainas commit themselves to this. The system **J3,** however, is distinguished by the new axiom (A4):

(A4) $\neg(\exists\sigma)(\neg\sigma : p \& \neg\sigma : \neg p)$ [Optimality]

Rescher too proposes a "three-valued approach" to assertion logic, via the notion of 'the truth status of the assertion p with respect to an assertor',

[59] Rescher, N.: 1968, *Topics in Philosophical Logic*, Reidel, Dordrecht, chapter xiv.

written '$|p|\sigma$', and the definitions:

$$\begin{aligned} |p|\sigma &= T \text{ iff } \sigma : p, \\ &= F \text{ iff } \sigma : (\neg p), \text{ and} \\ &= I \text{ iff } \neg(\sigma : p)\&\neg(\sigma : \neg p), \end{aligned}$$

and he shows that using the axioms of **A1**, we can derive a quasi-truth-functional logic for this system. These are not quite the Jaina values, as introduced earlier, for they do not quantify over standpoints or assertors. It is clear, however, that the Jaina system is of the same type as a modalised Rescher assertion logic. Their innovation is to introduce three truth-values via the definitions given before ($|p|_\Sigma = t$ iff $(\exists\sigma)(\sigma : p)$; $|p|_\Sigma = f$ iff $(\exists\sigma)(\sigma : \neg p)$; and $|p|_\Sigma = u$ iff $(\exists\sigma)(\neg(\sigma : p)\&\neg(\sigma : \neg p))$, where '$|p|_\Sigma$' stands for 'the status of the assertion p with respect to the total discourse Σ'). It is this attempt to take a many-valued approach to the modalised, rather than the unmodalised, version of assertion logic which generates the extra complexity of the Jaina system. I have already noted that, since the axiom "$p \supset \nabla p$" is lacking, the modal structure of the system will be no stronger than that of $\mathbf{S3}^0$. Yet in principle there seems no reason to think that the Jaina system cannot in this way be given an axiomatic basis.

3.6 Pluralism, Syncretism, and the Many-faceted View of Reality

The Jainas avoid dogmatism and a one-sided view of the world simply by noting that assertions are only justified in the background of certain presuppositions or conditions. It is perfectly possible for an assertion to be justified given one set of presuppositions, and for its negation to be justified given another different set. The Jainas' ingenuity lies in the skill with which they developed a logic of discourse to make more precise this natural idea. However, they also went beyond this, for they added that every standpoint reveals a facet of reality, and that, to get a full description of the world, what we need to do is to synthesise the various standpoints. As Matilal puts it, "The Jainas contend that one should try to understand the particular point of view of each disputing party if one wishes to grasp completely the truth of the situation. The total truth ... may be derived from the integration of all different viewpoints".[60] But is this further step, the step from pluralism to syncretism, a coherent step to take? In particular, how is it possible to integrate inconsistent points of view? The point is made by Priest and Routley, who, commenting on the Jaina theory, state that "...such a theory risks trivialization unless some (cogent) restrictions are imposed on the parties admitted as having obtained partial truth — restrictions of a type

[60] Matilal, B. K.: 1977, *The Central Philosophy of Jainism*, Calcutta University Press, Calcutta.

that might well be applied to block amalgamations leading to violations of Non-Contradiction".[61]

Perhaps we can understand the Jaina position as follows. The so-called 'integration' of two points of view, σ_1 and σ_2, does not mean the creation of some new standpoint, which is the combination of the first two. For this would lead to the formation of inconsistent standpoints unless implausible constraints were placed on what can constitute a standpoint. Instead, what it means is that, if p is assertible from some standpoint σ_1, then this fact, that p is assertible from σ_1, can itself be asserted from σ_2 and every other standpoint. In this way, each disputant can recognise the element of truth in the other standpoints, by making explicit the presuppositions or conditions under which any given assertion is made.

If correct, this idea has an interesting consequence. In moving from pluralism to syncretism, the Jainas commit themselves to the claim that we are led to a *complete* account of reality by integrating of all the different points of view . It follows from this that every true proposition must be asserted within some standpoint, i.e. "$p \supset (\exists\sigma)(\sigma : p)$ or "$p \supset \nabla p$". Hence the move from pluralism to syncretism is a move from a logic of assertibility based on $\mathbf{S3}^0$ or weaker to one based on **S3** or stronger.

To conclude, we have seen how the Jainas developed a plausible and interesting logic of philosophical discourse, how they did not (or need not) commit themselves to the strongly paraconsistent position normally attributed to them, and how, as they strengthened their position from one of pluralism to one of syncretism, they had also to strengthen correspondingly the modal logic underlying the operator "*syāt*".

4 LOGIC IN NAVYA-NYĀYA: THE METAPHYSICAL BASIS OF LOGIC

4.1 The use of graphs in interpreting Vaiśeṣika Ontology

Let us turn now to the Navya-Nyāya school, a school and a set of thinkers predisposed towards the study of the metaphysical structure of the natural world, and to the logical theory that is integral to this ontology. Three revisionary Nyāya thinkers – Bhāsarvajña (c. AD 950), Udayana (c. AD 1050), and Raghunātha (c. AD 1500) – saw in effect that there is a graph-theoretic basis to the classical Vaiśeṣika notion of a category. I will show how the graph-theoretic interpretation of their ideas lends itself to a distinctive treatment of negation, logical consequence and number.

Classical Vaiśeṣika lists six *kinds* of thing: substance, quality, motion, universal, individuator, inherence. Later Vaiśeṣika adds a seventh: absence.

[61] Priest, G., Routley, R. Norman, J. eds.: 1989, *Paraconsistent Logic: Essays on the Inconsistent*, Philosophia Verlag, Munchen, p.17.

The basic stuff of the cosmos in the Vaiśeṣika world-view is atomic. Atoms are uncreatable, indestructible, non-compound substances. Atoms can coalesce into composite substances and can move. Indeed, the only changes in this cosmos are changes in the arrangement, properties and positions of the atoms. Creation is a matter of coalescing, destruction of breaking (and even God does not *create* the cosmos *ab nihilo*, but only 'shapes' it, as a potter shapes clay into a pot). A compound substance is a whole, composed out of, and inhering simultaneously in each of, its parts. These substances are individuated by the type and organisation of their parts. A 'quality' in classical Vaiśeṣika is a property-particular – for example, a particular shade of blue colour or a distinct flavour (what one would now call a 'thin' property). Qualities inhere in substances and in nothing other than substances. A 'motion' is another sort of particular; it too inheres in a substance and in nothing but a substance. Universals inhere in substances, qualities and motions. A universal inheres simultaneously in more than one, but has nothing inhering in it. Lastly, the 'individuator' (*viśeṣa*) is a distinctive and eponymous component in classical Vaiśeṣika ontology. An individuator inheres in and is unique to a particular atom: it is that by which the atomic, partless substances are individuated.[62]

Two principles lie at the heart of the Vaiśeṣika system: a principle of identity and a principle of change. The Vaiśeṣika principle of change is this: a becomes b iff the parts of a rearrange (perhaps with loss or gain) into the parts of b. 'Motions' are that in virtue of which the parts rearrange or stay together. There are basic or partless parts, the atoms, which, precisely because they have no parts, are incapable of becoming anything else. They move about but are eternal and indestructible. The Vaiśeṣika principle of identity is this: $a = b$ iff the parts of a are numerically identical to and in the same arrangement as the parts of b. 'Qualities' are that in virtue of which the parts are numerically identical or different. Atoms, precisely because they are partless, require a different principle of identity: atoms are distinct iff they have distinct individuators. Universals are limits on the degree of possible difference and change. One thing cannot change into another thing of an entirely different *sort* (a mouse into a mustard seed). One thing a can become another thing b iff the same universal resides in both a and b, that is, if a and b are of the same sort (as Udayana puts it, universals *regulate* causality).

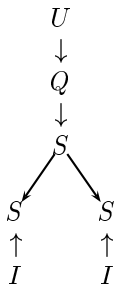
This is the motivation for there being six 'types' of thing (substances, qualities, motions, universals, individuators, inherence). The problem is to find a proper philosophical basis for the notion of a 'type' of thing thus appealed to. In his *Lakṣaṇāvalī*, Udayana reconstructs the categories in a

[62] An excellent review of the details of Vaiśeṣika ontology is Karl Potter ed., "Indian Metaphysics and Epistemology – The Tradition of Nyāya-Vaiśeṣika up to Gaṇgeśa", in *The Encyclopedia of Indian Philosophies*, vol. 2 (Delhi: Motilal Banarsidass, 1977), introduction.

new way, a way which I shall claim explicates the notion of a type *graph-theoretically*. A graph is a simple sort of algebraic structure, consisting of set of nodes or vertices, and a set of edges (an edge being defined as a pair of nodes). A graph is 'directed' if the edges have a direction. Graphs, like many other mathematical structures, are realised in natural phenomena. A striking example is molecular structure: it is because the structure of a molecule is a graph that one can use a graph to depict one:

$$H - O - H$$

The implicit structure of the Vaiśeṣika ontology is that of a directed graph. The inherence relation connects things in the ontology in inheror-inheree pairings. So the substances, qualities, motions, universals and individuators are represented as the nodes of a graph whose set of edges represent the inherence relation. A fragment of the graph might look like this:



This graph represents the following state of affairs: a universal U inheres in a quality Q which inheres in a substance S. That substance is a dyad composed of two atoms in which it inheres, and each of which has inhering in it an individuator I. The structure of the world is a directed graph.

The nodes in a graph can be classified according to the number of edges terminating in them, and the number of edges starting from them: so the *valency* of a node in a directed graph is an ordered pair of integers (n, m). What Udayana saw in the *Lakṣaṇāvalī* is that things of different types in the Vaiśeṣika ontology correspond to nodes of different valencies. His brilliant idea is to use the idea of valency to *define* the categories of substance, quality, motion, universal, and individuator. He begins with a classification of the categories into the four valency-groups $(+,+), (+,0), (0,+)$ and $(0,0)$:[63]

5. Noneternal [= compound] substance, quality, motion, universal, and individuator inhere.

[63] Numbering of the verses in the *Lakṣaṇāvalī* follows Musashi Tachikawa, *The Structure of the World in Udayana's Realism: A Study of the Lakṣaṇāvalī and the Kiraṇāvalī* (Dordrecht: Reidel Publishing Company, 1981).

6. Eternal [i.e. atomic] substance, inherence, and absence lack the property of inhering.

7. Substance, quality, and motion are inhered in.

8. Universal, individuator, inherence, and absence have nothing inhering in them.

In particular then, atoms have valency $(+, 0)$, universals and individuators have valency $(0, +)$, while compound substances, qualities and motions have valency $(+, +)$.

Notice that Udayana says that the inherence relation itself has a valency — $(0, 0)$. We should not take this to mean that the inherence relation is to be represented by a node disconnected from the rest of the graph, but rather that it does not correspond to any *node* in the graph at all. The first and most fundamental graph-theoretic type distinction is the distinction between a node and an edge, and the inherence relation is represented in a graph by the set of edges, not by any node. The set of edges represents the extension of the inherence relation.

If the categories are to be distinguished from one another according to the valency of the nodes in that graph which is isomorphic to the world of things, then further specification is needed. The distinction between universals and individuators is simple: an individuator has valency $(0, 1)$ while a universal has valency $(0, m)$, with $m > 1$:

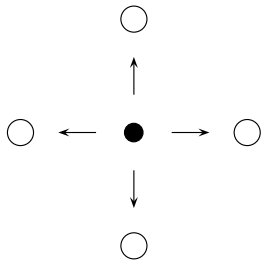
202. A universal has nothing inhering in it, inheres, and is co-located with every difference.

203. Individuators lack the property of being inhered in, inhere, and lack the property of inhering by being co-located with every difference.

Udayana's phrase 'co-located with every difference' is a technical device for expressing the idea that a universal inheres in more than one. For if an inheror inheres in exactly one thing x, then all other things are loci of difference–from–x, and the inheror is not co-located with difference–from–x. However, if the inheror inheres in two things x and y, then difference–from–x is located in y and difference–from–y is located in x, and the inheror is co-located with both differences. So something co-located with every difference–from each of the things in which it inheres is necessarily located in more than one thing. Notice that in classical Vaiśeṣika, individuators are said to have no universals inhering in them precisely because they are fundamental units of individuation, having nothing in common with one another.

Any node with valency $(0, m)$ with $m > 1$ is now to be called a 'universal', and any node with valency $(0, 1)$ is to be called an 'individuator':

UNIVERSAL



INDIVIDUATOR

The valency of atoms is different from that or qualities or motions, but we still need a general definition of substance, covering both atomic and compound substances. For compound substances, like universals but unlike atoms, inhere in other things (their parts). Udayana in fact offers four definitions, of which the first three repeat older definitions. The fourth definition, however, is completely original:

9. A substance is not a substratum of absence of quality.

10. Or, it belongs to such a kind as inheres in what is incorporeal, inheres in what is not incorporeal and does not inhere in what inheres in what is not corporeal.

11. Or, it belongs to such a kind as inheres in space and in a lotus but not in smell.

12. Or, it is that in which inheres that in which inheres that which inheres.

The first of these definitions is the classical one in Vaiśeṣika[64] — a substance is that which possesses qualities. Udayana returns to this definition in his famous but conservative commentary, the *Kiraṇ āvalī*. He thinks of replacing it in the more experimental *Lakṣaṇāvalī* with a definition that makes no reference to any other category and indeed is phrased entirely in terms of the notion of inherence: a substance is 'that in which inheres that in which inheres that which inheres'. In other words, a substance is to be represented by a node like this:

SUBSTANCE

The point of the definition is that a substance possesses qualities, and qualities possess universals, and nothing else in the ontology possesses something which possesses something. For universals and individuators possess nothing, while qualities and motions possess universals and nothing else.

Let us define a 'path' between one node and another in the obvious way: there is a path from node **x** to node **w** if there is a sequence of nodes {**x**, **y**, ..., **w**} such that there is an edge from **x** to **y**, an edge from **y** to **z**, ... , an edge from **v** to **w**.[65] Define the 'length' of a path as the number of edges between the first and the last node. Udayana's definition of a substance is now: a node is a substance iff there is a path at least of length 2 leading to it. Substances inhere in their parts; so the definition entails that every part of a substance is a substance.

The classical conception of qualities and motions makes them almost identical: they both inhere only in substances, and they both are inhered in only by universals.[66] Praśastapāda's remark[67] that the qualities other than contact, breaking, number and separateness 'inhere in one thing at a time' should not be construed as implying that they inhere in only one thing, but only that this group of qualities are *monadic* (non-relational) properties. These features are enough to distinguish qualities and motions

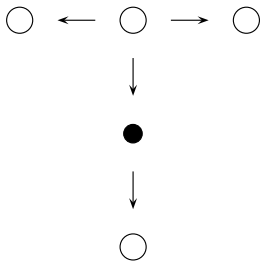
[64] *Vaiśeṣikasūtra* 1.1.14: "The characteristic of a substance is to possess actions, qualities and to be [their] inherence cause."

[65] In what follows, bold roman letters denote nodes in the graph, and italic letters denote the entities those nodes represent.

[66] *Vaiśeṣikasūtra* 1.1.15–6. *Padārthadharmasaṃgraha* 18. Section numbering in the *Padārthadharmasaṃgraha* follows Karl Potter ed., "Indian Metaphysics and Epistemology – The Tradition of Nyāya-Vaiśeṣika up to Gaṇgeśa", in *The Encyclopedia of Indian Philosophies*, vol. 2 (Delhi: Motilal Banarsidass, 1977), pp. 282–303.

[67] *Padārthadharmasaṃgraha* 50–51.

from all else: from universals and individuators (which do not have anything inhering in them), and from substances (which are inhered in by things that are themselves inhered in). It explains too why qualities cannot inhere in qualities — if they did then they would be equivalent graph-theoretically to substances.



What is difficult is to find any principled way to distinguish between qualities and motions. There was indeed a persistent revisionary pressure to assimilate these two categories. Bhāsarvajña[68] heads the revisionary move, stating unequivocally that motions should be treated as qualities because, like qualities, they reside in substances and possess universals. From a graph-theoretic perspective, this revision is well motivated: qualities and motions are represented by nodes of the same valency, and so are things of the same *type*. Udayana chooses the harder way, and tries to formulate definitions that will accommodate the distinction. The classical Vaiśeṣika idea[69] that motions are what cause substances to come into contact with one another is reflected in his definitions:

126. A quality belongs to such a kind as inheres in both contact and non-contact, and does not inhere in the non-inherent cause of that sort of contact which does not result from contact.

190. A motion belongs to such a kind as inheres in the non-inherent cause of contact and does not inhere in contact.

These definitions introduce two new relations, contact and causation, neither of which are explicable in terms of inherence nor belongs to the graph-theoretic interpretation of the categories. The very success of that interpretation gives a rationale to the revisionary pressure. Finding a pattern into which all but a few items of some phenomenon fit grounds a presumption that those items are in some way discrepant. This is a general principle of scientific and rational inquiry, and we can see it been used by Bhāsarvajña

[68] *Nyāyabhūṣaṇa*, p. 158.
[69] *Vaiśeṣikasūtra* 1.1.16.

to motivate revisions in the classical Vaiśeṣika theory. Rationality appears here in the form of *principled* revision.[70]

Let us define a *Vaiśeṣika graph* as a connected directed graph each of whose nodes is a substance, quality, motion, universal or individuator, where:

> A *substance* is a node terminating a directed path of length 2.
>
> A *quality* or *motion* is a node v with valency $(+,+)$, such that the initial node of any edge terminating in v has zero invalency [i.e. such that qualities are not substances].
>
> A *universal* is a node with valency $(0,n)$ with $n > 1$.
>
> An *individuator* is a node with valency $(0,1)$.

Let us say further that node x inheres in a node y iff xy is an edge in a Vaiśeṣika graph. Then we can easily prove some results well-known to the Nyāya-Vaiśeṣika logicians:

LEMMA 1. *No quality inheres in a quality.*

Proof. The invalency of a quality x is nonzero, so any node in which x inheres terminates a path of length 2. ■

LEMMA 2. *Substances inhere only in substances.*

Proof. Any node x in which a substance inheres terminates a path of length 2. ■

THEOREM 3. *A Vaiseṣika graph has no directed cycles.*

Proof. The elements of a directed cycle must have valency $(+,+)$. So no universal or individuator can be a member of such a cycle, because neither has nonzero invalency. No quality or motion can be a member of a cycle, because only universals and individuators inhere in qualities and motions, and there are no universals or individuators in a cycle. No atomic substance can be a member of a cycle, because atoms have zero outvalency. That leaves only cycles of compound substances. But there can be cycles of substances only if a substance can have as a part something of which it is a part and so (if the part-of relation is transitive) be a part of itself. ■

DEFINITION 4. The *level* of a node in a Vaiśeṣika graph is the length of the longest directed path leading to it. (Note — this is well-defined because there are no directed cycles.)

THEOREM 5.

[70] For later comment on Bhāsarvajña's revision: Karl Potter and Sibajiban Bhattacharyya eds. "Indian Philosophical Analysis — Nyāya-Vaiśeṣika from Gaṅgeśa to Raghunātha Śiromaṇi", in *The Encyclopedia of Indian Philosophies*, vol. 6 (Delhi: Motilal Banarsidass, 1993), pp. 323, 525–528.

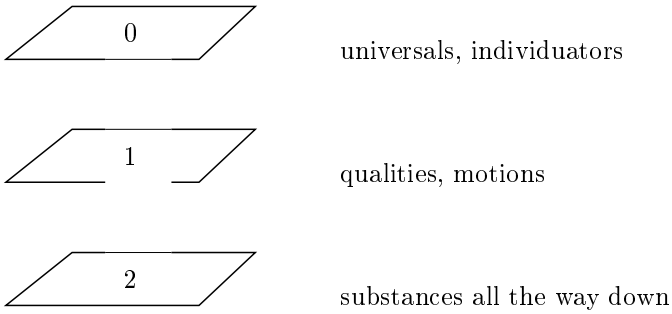
(i) All and only universals and individuators belong to level 0;

(ii) All and only qualities and motions belong to level 1;

(iii) All and only substances belong to levels 2 and below.

Proof. (i) Only universals and individuators have zero invalency. (ii) By Lemmas 1 and 2, qualities and motions are inhered in only by universals and individuators, so belong to level 1. Substances do not belong to level 1 by definition. (iii). By definition, any node in level 2 is a substance, and by Lemma 2, any node in level $n > 2$ is a substance. ■

So the structure of the Vaiśeṣika graph is like this:



To what extent are we justified in adopting the graph-theoretical interpretation of Navya-Nyāya? I propose the following *Methodological Test*: The graph-theoretic interpretation is confirmed to the extent that it explains or predicts revisions made to the classical Vaiśeṣika system. Revisions include the introduction of a seventh category absence, the assimilation of qualities and motions, the elimination of individuators, the identification of co-extensive universals, the new account of number. Let us say that a node x is *redundant* in G if its deletion, together with the deletion of any edge incident to it, preserves all paths in G not containing x. The resulting graph $G*$ is a *conservative contraction* of G. Then we have, in effect, the following revisions being proposed by Nyāya authors — (I) All individuators are redundant in a Vaiśeṣika graph (Raghunātha); (II) Two universals are co-extensive only if at least one is redundant (Udayana); and (III) Qualities and motions are entities of the same type (Bhāsarvajña).

4.2 Negation as absence

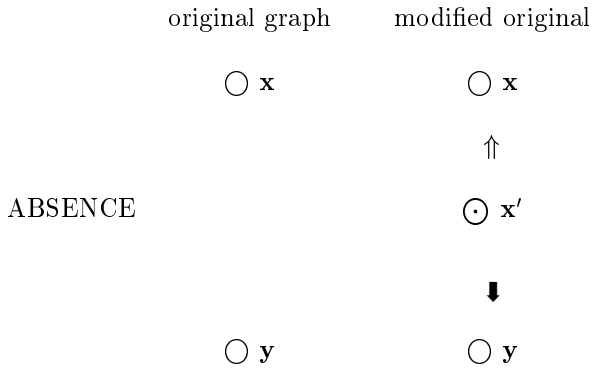
'Absence' in Navya-Nyāya is not the same as nonexistence. Fictional characters, dream-objects and hallucinations are nonexistent: they do not exist as it were *by nature*. It would be an absurdity to go in search of Hamlet in order to find out whether he really exists or not — his nonexistence is not a merely contingent lack in the world of things. The absence of water on the moon, on the other hand, is a contingent and concrete fact; so too is the absence of colour in my cheeks. Notice the role of the phrases 'of water', 'of colour' here: an absence has an *absentee* — that which the absence is an absence of. It also has a location (e.g. the moon, my cheeks), and a time. So the proposal is to reparse the sentence 'x does not occur in y at time t' as 'an absence-of-x occurs in y at t.' For it is often the case that the absence of something somewhere is more salient than any fact about what is present there.

There is one relatively straightforward way to interpret the idea of absence graph-theoretically. If x does not inhere in y, then there is no edge ($\mathbf{x}$, $\mathbf{y}$) in the graph. Now for every graph, there is a dual. The dual has the same nodes as the original graph, but has an edge between two nodes just in case the original does not. So the dual graph does have an edge ($\mathbf{x}$, $\mathbf{y}$). Following this idea, one would be led to say that absences are things of a different type to any presence because they are edges in the dual graph, rather than edges or nodes in the original.

For various reasons, the Vaiśeṣika do not consider this to be an adequate explanation of the category. One problem is that it makes absences more like relations than 'things', and this does not keep to the spirit of the Vaiśeṣika idea that absences are entities. In fact, absences do display much relation-like behaviour — after all, absence is always the absence of x in y. Another objection, however, is if absence is a new category, its introduction should result in an extension of the original graph, and not in the introduction of a new graph, let alone a graph completely disconnected from the original. For the connected world of things ought not be represented by a pair of disconnected graphs. A third problem arises if we admit something called 'unpervaded' occurrence, as we will see.

The Vaiśeṣika idea is represent absences as *nodes*, related in new ways to the nodes of the original graph. Here is how to do it. For each unconnected pair of nodes ($\mathbf{x}$, $\mathbf{y}$), create a new node $\mathbf{x}'$ in the original graph. This new node will have edges to $\mathbf{x}$ and to $\mathbf{y}$, but they will be edges of two new types. The edge ($\mathbf{x}'$, $\mathbf{x}$) is an edge belonging to the extension of the absentee-absence (*pratiyogitā*) relation, which I shall signify as '$\Rightarrow$.' This represents the relation between an absence and what the absence is of. The edge ($\mathbf{x}', \mathbf{y}$) is an edge belonging to the extension of the 'absential special relation' (*abhāvīya-svarūpa-sambandha*), signified here by '$\blacktriangleright$'. This represents the occurrence relation between an absence and its location. The

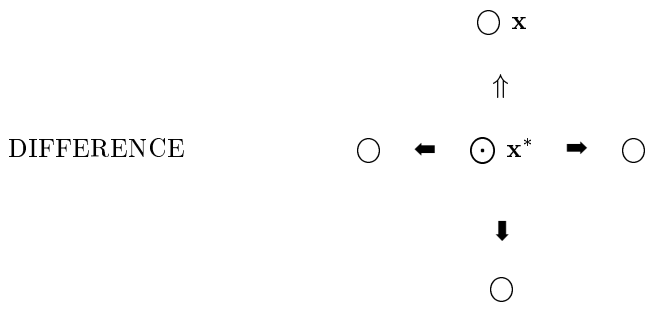
relation between an absence and its location is clearly not the same as the relation between a presence and its location (inherence, contact), for it is clear that when a person is is absent from a room, their absence is not in the room in the same sense that the other things in the room are.



These new nodes belong in a domain outside the system of levels, for they inhere in nothing and nothing inheres in them (inherence, and the whole system of levels, is a structure on presences). The modified graph is instead a concatenation of the original graph of nodes and edges with a new structure of 'absential nodes' and 'absential edges.'

Navya-Nyāya theory of absence draws a type distinction between *simple absence* (*atyantābhāva*) and *difference* (*anyonyābhāva*). Difference is the absence of a relation of identity between two things. Here '$x \neq y$' is paraphrased as 'a difference-from y occurs in x'. Graph-theoretically, the distinction between absence and difference is a distinction between a negation on edges and a negation on nodes in the original graph. For, trivially, every node is such that it is *different* from every other node. One way to represent this would be to introduce a new kind of 'nonidentity' edge into the graph, an edge which connects every node with every other node. The Naiyāyika, however, wants to the category of absence to correspond to a domain of things rather than relations; so in the graph-theoretic representation, differences have to be represented as nodes rather than edges. So let us say that for every node **x** in the original graph, there is a new node **x***. Call it an 'antinode'. **x*** is connected to *every* node in the graph. It is connected to **x** by an edge of the absentee-absence type, and to every node other than **x** by an absential location edge. There is a one-one correspondence between the new domain of antinodes and the domain of original nodes.

The leading idea behind the graph-theoretic interpretation of the categories is that a type of thing is a type of node, and node-types are determined by patterns of possible valencies in the graph. It was for this reason that



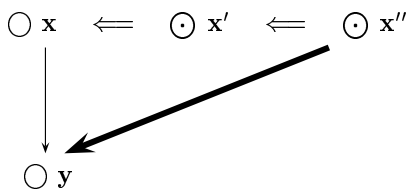
we did not need earlier to the *label* the nodes. With the introduction of the category of absence, we have two higher-order type distinctions: the distinction between positive and negative nodes, and the distinction among the negative nodes between absential nodes and antinodes. Do these distinctions have a graph-theoretic explanation, or must we allow ineliminable node-labels to demarcate presence nodes, absential nodes and antinodes? What we do have now are three different types of edge — corresponding to the relations of inherence, absence-absentee, and absential location. So we might hope to distinguish between positive and negative nodes as those which are not and those which are at the end of an absential edge. That is, we make it a requirement that no positive node absentially qualify any other node. Clearly, the suggestion will work only if the absence of an absence is not identical to a presence. We will see in the next section that the graph-theoretically oriented Raghunātha indeed denies that this is so. So as not to beg the question at this point, and for the sake of pictorial clarity, I will continue to *mark* positive nodes ◯ and negative nodes ⊙ differently.

What about the distinction between absential and antinodes? The traditional way of making the distinction is to say that simple absence is the denial of inherence (or some other nonidentity relation) and difference is the denial of identity. Graph-theoretically, the distinctive feature of an antinode **x*** is that it absentially qualifies every node other than **x**, while an absential node **x**′ does not. Does this difference fail when x is something which inheres in nothing (an atom, an individuator)? No, because such things do not inhere in themselves — so **x**′ unlike **x*** absentially qualifies **x**. Indeed, this second contrast is itself sufficient to discriminate absential nodes and antinodes.

The above treatment of absence is in effect a procedure for introducing new nodes into the original graph. One set of new nodes fills the 'gaps' in that graph: whenever there is no edge between two nodes, an absential node is introduced between, and linked to, them. Another set of new nodes exactly mirrors the original graph: for each node in the original, there is one and only one antinode, linked to everything the original node is not. But

now, having supplemented the original graph with two sets of new nodes, nothing is to stop us from repeating the procedure again — generating new sets of *second-order* absence nodes — and to do this again and again. It seems that we have introduced a procedure for the indefinite recursive expansion of the graph. Fortunately this does not in fact happen. As we will now see, no subsequent recursion of the procedure after the first produces any new nodes.

Prima facie, it seems plausible to reason as follows (as we will shortly see, this reasoning turns out to be subject to an important caveat). If $\mathbf{x}$ is in $\mathbf{y}$, then $\mathbf{x}'$, the absence of $\mathbf{x}$, is not in $\mathbf{y}$, and so $\mathbf{x}''$, the absence of $\mathbf{x}'$, is in $\mathbf{y}$. Conversely: if $\mathbf{x}''$ is in $\mathbf{y}$, then $\mathbf{x}'$ is not in $\mathbf{y}$, so $\mathbf{x}$ is in $\mathbf{y}$. Graph-theoretically, we represent this as follows:

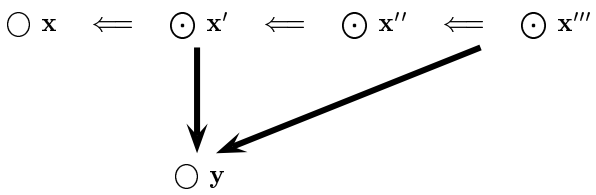


If this is right, then it follows that an entity and the absence of its absence 'occur' in exactly the same set of loci: for all $\mathbf{y}$, there is an inherence edge $(\mathbf{x}, \mathbf{y})$ just in case there is an absential location edge $(\mathbf{x}'', \mathbf{y})$. Can we appeal now to a uniqueness condition for absences, and infer that the absence of an absence of an entity is identical to the entity? The point is controversial, with the majority favouring identification. It is Raghunātha[71] who argues that the identification is unsound, on the ground that nothing can turn an absence into a presence. Here again Raghunātha's intuition agrees with the graph-theoretic reconstruction: the nodes $\mathbf{x}$ and $\mathbf{x}''$ are connected to other nodes by means of different types of edge. So they cannot both represent

[71] Raghunātha, *Padārthatattvanirūpaṇa*, p. 55. Daniel Ingalls, *Materials for the Study of Navya-Nyāya Logic* (Cambridge, Mass.: Harvard University Press, 1951), p. 68. Bimal. K. Matilal, *Logic, Language and Reality* (Delhi: Motilal Banarsidass, 1985), p. 149. Roy W. Perrett, "Is Whatever Exists Knowable and Nameable?" *Philosophy East & West* 49.4 (1999), pp. 410–414, esp. 408–9. I disagree here with the idea of Matilal and Perrett that there is only an intensional difference between an object and the absence of its absence. For me, a type difference in the graph means a type difference in categories of thing.

entities of the same type. Moreover, as we shall see in more detail below, the Naiyāyikas do not accept that it is generally valid to infer from the occurrence of $\mathbf{x}$ in $\mathbf{y}$ to the occurrence of $\mathbf{x}''$ there, although they do allow the converse. This is the caveat in the line of reasoning with which I began this paragraph. The implication is that $\mathbf{x}$ and $\mathbf{x}''$ need not, after all, share the same set of loci.

Let us repeat the procedure once more. If $\mathbf{x}'$, the absence of $\mathbf{x}$, is in $\mathbf{y}$, then $\mathbf{x}''$ is not in $\mathbf{y}$, and so $\mathbf{x}'''$, the absence of $\mathbf{x}''$, is in $\mathbf{y}$. Conversely: if $\mathbf{x}'''$ is in $\mathbf{y}$, then $\mathbf{x}''$ is not in $\mathbf{y}$, so $\mathbf{x}'$ is in $\mathbf{y}$. Graph-theoretically:



It follows that a first-order absence and the absence of *its* absence reside in exactly the same set of loci. But here we *can* appeal to the uniqueness condition, because the edges are all of the same type. So x''' is identical to x', as Raghunātha himself allows.[72] Similarly, x'''' is identical to x'', and so on. There are no absential nodes of order higher than two. The argument is summed up by Annambhaṭṭa in the *Tarkasaṃgraha* [§89]:

> The view of the early thinkers is that the absence of an absence is nothing but a presence; it is not admitted as a new absence for there would then be an infinite regress. According to the new school, however, the absence of an absence is a distinct absence, and there is no regress as the third absence is identical to the first.

Recall that we defined the absence $\mathbf{x}'$ as a node such that $\mathbf{x}'$ is absentially located in $\mathbf{y}$ if there is no edge between $\mathbf{x}$ and $\mathbf{y}$. That definition was adequate for the introduction of first-order absences, because there is only one kind of edge in the original graph, namely the inherence edge. The

[72] *Padārthatattvanirūpaṇa*, pp. 67–69. Daniel Ingalls, *Materials*, pp. 68–69; Bimal. K. Matilal, *Logic, Language and Reality*, pp. 149–150.

expanded graph has another sort of edge, however: the absential edge. So the notion of a second-order absence is underdetermined by our original definition. The new definition we need is:

Rule for Absence:

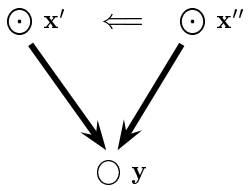
> An absence $\mathbf{x}'$ is absentially located in $\mathbf{y}$ if $\mathbf{x}$ does not inhere in $\mathbf{y}$.

Rule for Higher Order Absence:

> For $i > 1$, an absence $\mathbf{x}^i$ is absentially located in $\mathbf{y}$ iff $\mathbf{x}^{i-1}$ is not absentially located in $\mathbf{y}$.

The second clause implies that absence is a classical negation for $i > 1$, and so, in particular, that an absence of an absence of an absence is identical to an absence. A double negation, however, is a mixture — a negation defined on inherence edges followed by a negation defined on absential qualifier edges – and for that reason behaves non-classically. What I will show in the next section is that Navya-Nyāya logic rejects the classical rule of Double Negation Introduction — the rule that licenses one to infer from p to $\neg\neg p$. What replaces it is a weakened rule — infer from $\neg p$ to $\neg\neg\neg p$. This is because negation is a procedure for filling 'gaps' in the graph: whenever there is no edge between two nodes, the rule for negation licenses us to insert an absential node between them. The classical rule for Double Negation Elimination – the rule that licenses one to infer from $\neg\neg p$ to p — remains valid in Nayva-Nyāya logic (i.e. if $\mathbf{x}'$ is not in $\mathbf{y}$, then $\mathbf{x}$ is in $\mathbf{y}$).[73] The effect of this weakening in the rule for Double Negation Introduction is that one is no longer entitled to infer that if $\mathbf{x}$ is in $\mathbf{y}$, then $\mathbf{x}'$ is not in $\mathbf{y}$. One effect of this is to block the equivalence of a positive entity with the absence of its absence. We can say that $\mathbf{x}'$ is the absence of $\mathbf{x}''$, but we cannot say that $\mathbf{x}$ is the absence of $\mathbf{x}'$. Graph-theoretically, connections of the form $\bigcirc\!\!\!\!\bullet\,\mathbf{x} \Rightarrow \odot\, \mathbf{x}'$ are prohibited, since a positive entity cannot be the absence of *anything*. Also prohibited are triangles of the form below, because negation behaves classically within the domain of absences. What is stranger, however, is the effect the weakened rule has of *permitting* a positive entity to be co-located with its absence. For we are no longer in a position to assert that the presence of an entity is inconsistent with its absence. Let us see how the Nyāya philosophers arrive at the conclusion that one must allow for such an unusual possibility.

[73] Daniel Ingalls draws a comparison between Navya-Nyāya and intuitionist logic (*Materials*, p. 68, n. 135), claiming that it is the elimination rule for double negation that is rejected. However we are able, in Navya-Nyāya logic, to infer from the absence of the absence of an entity to the presence of that entity; conversely, we are not able to infer from the presence of an entity to the absence of its absence — the non-pervasive node is a counter-example.



Whenever something inheres in a compound substance, the question arises: does it also inhere in the parts? An entity is said to be of 'locus-pervading' occurrence just in case it inheres in all the parts of its locus (as well as in the locus itself).[74] It *saturates* its locus. A sapphire is red through-and-through, and sesame oil pervades every part of the seed; but a painted vase is blue only on the outside. Let us say then that x is locus-pervading with respect to y just in case x inheres in y and if z is a part of y then x inheres in z.[75] The only things that have parts are substances, and substances inhere in their parts and in nothing else. So x is locus-pervading with respect to y just in case x inheres in y and if y inheres in z then x inheres in z. Certain types of quality pervade their loci, according to the classical Vaiśeṣika authors.[76] Examples include weight, viscosity, and fluidity. A thing is heavy just in case every part of it is heavy. Colours, tastes, smells can pervade their loci but need not do so.[77] And a compound substance is locus-pervading with respect to each of its parts, if 'part of' is a transitive relation.

The notion of a locus-pervading entity has a distinctive graph-theoretic correlate. An edge $(\mathbf{n}_1, \mathbf{n}_2)$ is locus-pervading just in case there is an edge from $\mathbf{n}_1$ to any node in any path from $\mathbf{n}_2$.

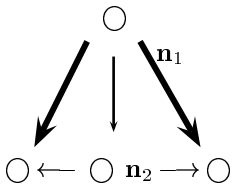
While locus-pervading nodes are straightforwardly definable in the system as so far developed, the concept of 'unpervaded occurrence' (*avyāpya-vṛttitva*) marks a theoretical innovation. The classic Buddhist refutation of realism about wholes is that wholes must be the bearers of contradictory properties. For if some parts of a vase are red and other parts are not red,

[74] Ingalls (1951: 73–74); Bimal Matilal, *The Navya-Nyāya Doctrine of Negation* (Cambridge, Mass.: Harvard University Press, 1968), p. 53, 72, 85; Matilal (1985: 119–122).

[75] Frege's notion of 'divisibility' is formally rather analogous. Gottlob Frege, *The Foundations of Arithmetic*, translated by J.L. Austin (Oxford: Basil Blackwell, 1950), p. 66: "The syllables "letters in the word three" pick out the word as a whole, and as indivisible in the sense that no part of it falls any longer under the same concept. Not all concepts possess this quality. We can, for example, divide up something falling under the concept 'red' into parts in a variety of ways, without the parts thereby ceasing to fall under the same concept 'red."'

[76] Karl Potter ed., "Indian Metaphysics and Epistemology – The Tradition of Nyāya-Vaiśeṣika up to Gaṇgeśa", in *The Encyclopedia of Indian Philosophies*, vol. 2 (Delhi: Motilal Banarsidass, 1977), pp. 114–119.

[77] Raghunātha, *Padārthatattvanirūpaṇa*, pp. 44–6.



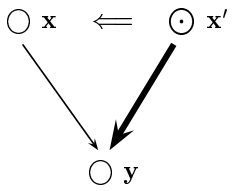
and if the vase as a whole has a colour in virtue of its parts having colour, then one seems forced to admit either that the whole is both red and not red, or that it has no colour at all.[78] The traditional Nyāya-Vaiśeṣika solution is less than satisfactory – it is to say that the whole has a new shade of colour called 'variegated'! Recognising the *ad hoc* nature of such a response, later Naiyāyikas try instead to make sense of the idea that a property can be co-located with its absence.[79] The idea is to capture the sense in which one says that the vase is red, because its surface is red, allowing at the same time that it is not red, because its insides are some other colour. A favourite Nyāya example involves the relation of contact: the tree enjoys both monkey-contact (there is a monkey on one of its branches) and also the absence of monkey-contact (its roots and other branches are in contact with no monkey). This defence of realism is what motivates later writers to allow there to be such a thing as unpervaded occurrence, defined to be an occurrence that is co-located with its absence. That is, an unpervading node is a node **x** such that there is an edge (**x**, **y**) and an edge (**x′**, **y**). Triangles such as the following are now deemed to be permissible in the graph:

The strangeness of such a possibility is ameliorated if one says, as some Naiyāyikas do, that x occurs in y as 'delimited' by one part, and its absence occurs in y as 'delimited' by another part.[80] Gaṅgeśa nevertheless goes to considerable lengths to reformulate logic and the theory of inference in Navya-Nyāya in a way that permits the co-location of an entity with its

[78] Dharmakīrti, *Pramāṇavārttika* II, 85–86; Kamalaśīla, *Pañjikā under Tattvasaṃgraha* 592–598.

[79] Udayana, *Ātmatattvaviveka*, pp. 586–617. Prabal Kumar Sen, "The Nyāya-Vaiśeṣika Theory of Variegated Colour (*citrarūpa*): Some Vexed Problems", *Studies in Humanities and Social Sciences* 3.2 (1996), pp. 151–172.

[80] Ingalls (1951: 73–4); Bimal Matilal, *The Navya-Nyāya Doctrine of Negation* (Cambridge, Mass.: Harvard University Press, 1968), pp. 71–73.



absence. The phenomenon of unpervaded occurrence is not regarded as a minor curiosity in Nyāya, but as the occasion for serious revision in their analysis.[81]

4.3 Definitions of logical consequence

With the introduction of absence, the graph-theoretic ontologies can serve as semantic models for a propositional language. A sentence 'p' is assigned, let us stipulate, an ordered pair of nodes (**x**, **y**). The sentence is true if that pair is an edge in the graph, false if it is not.[82] The negation of that sentence, '$\neg p$', is true if (**x**′, **y**) is an edge, false if it is not. Again, '$\neg\neg p$' is true if (**x**″, **y**) is an edge, false if it is not. If triangles like the one above are possible, then the truth of 'p' does not imply the truth of '$\neg\neg p$', since (**x**, **y**) is an edge but not (**x**″, **y**). So the propositional logic being modelled is, as we have already observed, one in which Double Negation Introduction does not hold. In this theory, we still have these correspondences between truth-value and negation:

(R1) if $\neg T\alpha$ then $T\neg\alpha$. from Rule for Absence

(R2) $T\neg\neg\alpha$ iff $\neg T\neg\alpha$ from Rule for Higher Order Absence

What we no longer have is:

(R3) if $T\neg\alpha$ then $\neg T\alpha$

The reason, as I said before, is that negation is an operation that fills 'gaps' in the graph — it tells us nothing when there is already an edge between two nodes. So the truth of a proposition is consistent, in Navya-Nyāya logic, with the truth of its negation. This element of dialetheism in the theory does not, however, mean that anything is provable or that anything follows from anything else — the correspondences R1– R2 are enough to prevent the system collapsing. Let us see why.

[81] Matilal's property-location language, in which properties have both a 'presence range' and an 'absence range' and the two ranges are permitted to overlap, is a different way to capture the same idea; Matilal (1985: 112–127).

[82] Gaṇgeśa, *Tattvacintāmaṇi*, I, *pramālakṣaṇa*, p. 401.

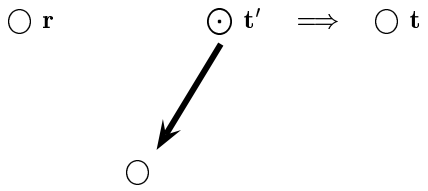
In the modern analysis of valid inference, an inference is valid just in case it is impossible for the premises to be true without the conclusion also being true. In the logic of classical India, validity is a matter of property-substitution, and the problem is to determine the conditions under which the occurrence of a reason property at a location warrants the inference that a target property occurs there too ("*Ta* because *Ra*"). The leading idea is that such property substitutions are valid just in case the reason does not 'wander' or 'deviate' from the target (*avyabhicāra*). In a famous passage called the *vyāptipañcaka*, Gaṅgeśa suggests five ways to make sense of this idea:[83]

> Now, in that knowledge of a pervasion which is the cause of an inference, what is pervasion? It is not simply non-wandering. For that is not
>
> 1. nonoccurrence in loci of the absence of the target, nor
> 2. nonoccurrence in loci of the absence of the target which are different from loci of the target, nor
> 3. non-colocation with difference from a locus of the target, nor
> 4. being the absentee of an absence which resides in all loci of absence of the target, nor
> 5. nonoccurrence in what is other than a locus of the target,
>
> since it is none of these where the target is maximal.

A 'maximal' property is a property resident in everything (*kevalānvayin*). Gaṅgeśa dismisses the five provisional analyses on the grounds that all are formulated in terms of 'absence of the target', and that that phrase is undefined when the target is maximal (the absence of a maximal property — assumed here not to be of unpervaded occurrence — would occur in nothing and so be 'unexampled', contradicting a basic condition of connectedness). In his preferred definition, Gaṅgeśa exploits a trick to overcome this problem.[84] He says that any property whose absence is colocated with the reason is not identical to the target. This implies that the target is not a property whose absence is colocated with the reason, but the contraposed formulation avoids the use of the troublesome phrase 'absence of the target'.

Consider now the difference between the first and second analyses in the list of five. Graph-theoretically, what the first analysis states is that, if **r** is the node representing the reason, and **t** is the node representing the target, then **r** is present in no node where **t** is absent —

[83] Gaṇgeśa, *Tattvacintāmaṇi*, II, *vyāptipañcaka*, pp. 27–31.
[84] Gaṇgeśa, *Tattvacintāmaṇi*, II, *siddhānta-lakṣaṇa*, p. 100.



But what happens if the target has nonpervaded occurrence? Then the first analysis is too strong.[85] For it is not a necessary condition on valid inference that the reason not be present wherever the target is absent, if there are nodes where the target is present as well as absent. What validity precludes is the presence of the reason without the presence of the target. So the proper definition is that the reason is not present wherever the target is not present (and so also absent). This is exactly what the second analysis states. We can make the point in terms of our earlier definitions of truth and negation. The premise in an inference is the statement that the reason occurs in a certain location, the conclusion the statement that the target occurs in that location. What our first analysis asserts is that the premise is not true if *the negation of the conclusion is true* (= absence of target in the location). The second analysis states instead that the premise is not true if the conclusion is *false* (= denial of presence of target in the location). Ironically, then, it is the very element of dialetheism of the Navya-Nyāya system which forces Gaṅgeśa to disambiguate the definition of validity, and to distinguish the correct definition from the one that had been preferred before.

Let α = "the reason r inheres in x", β = "the target t inheres in x". Then $\alpha \vDash \beta$ iff t pervades r. The problem is to solve for 'pervades'. The first solution in Gaṅgeśa is:

1. whatever the value of x, α is *not true* if $\neg\beta$ is true, i.e.
   $\alpha \vDash \beta$ iff under any assignment of value to x, $T\neg\beta \rightarrow T\neg\alpha$

His second solution is:

2. whatever the value of x, α is not true if β is not true, i.e.
   $\alpha \vDash \beta$ iff under any assignment of value to x, $\neg T\beta \rightarrow \neg T\alpha$.

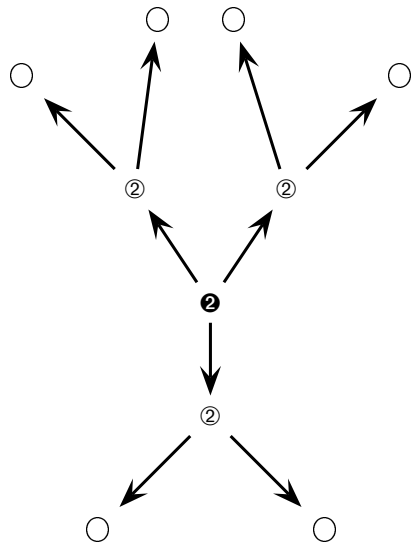
What we have seen is that (2) and not (1) is the correct analysis of logical consequence if R3 is rejected.[86]

[85] I follow here the explanation of Raghunātha. *Vyāptipañcakadīdhiti* text 3–4 (Ingalls (1951: 154)).

[86] R3 is what Graham Priest calls the 'exclusion principle.' For a semantic theory without this principle, see his *In Contradiction: A Study of the Transconsistent* (Dordrecht:

4.4 Number

The classical Vaiśeṣika theory of number is that numbers are qualities of substances.[87] A quality 'two' inheres in both members of a pair of substances, another quality 'two' inheres in another such pair, and all the qualities 'two' have inhering in them a single universal 'twohood' (see graph on page 86.[88]



❷ the universal "twohood"
② the quality-particular "two"

Bhāsarvajña and Raghunātha, as usual, lead the reforming move. Bhāsarvajña's theory[89] is that numbers are not qualities at all, but relations of identity and difference. Thus the sentence 'a and b are one' means simply that $a = b$, while 'a and b are two' means that $a \neq b$. Bhāsarvajña's analysis is echoed, very much later, in Gadādhara's (c. AD 1650) comments on the meaning

Martin Nijhoff Publishers, 1987), chapter 5.

[87] *Vaiśeṣikasūtra* 1.1.9, 7.2.1–8.

[88] For a more detailed description of the classical account: Jonardon Ganeri, "Objectivity and Proof in a Classical Indian Theory of Number," in *Synthese*, 129 (2001), pp. 413–437.

[89] *Nyāyabhūṣaṇa*, p. 159.

of the word 'one'.[90] Gadādhara states that the meaning of 'one F' is: an F as qualified by being-alone, where 'being alone' means 'not being the absentee of a difference resident in something of the same kind.' In other words, 'one F' is to be analysed as saying of something which is F that no F is different to it. If this is paraphrased in a first-order language as '$Fx \ \& \ \neg(\exists y)(Fy \ \& \ y \neq x)$' then it is formally equivalent to a Russellian uniqueness clause '$Fx \ \& \ (\forall y)(Fy \rightarrow y = x)$'. The idea that 'one' expresses uniqueness is in the spirit of Bhāsarvajña's idea that it denotes the identity of a thing. In any case, it is clear that, for Gadādhara, 'one' has a logical role similar to that of the definite article.

Raghunātha is more radical still.[91] The central problem is that things in *any* category in the Vaiśeṣika ontology can be numbered, and Raghunātha concludes that numbers must belong in a new category of their own:

> Number is a separate category, not a kind of quality, for we do judge that there is possession of that [number] in qualities and so on. And this [judgement we make that qualities have number is] not an erroneous one, for there is no [other] judgement which contradicts it.

Raghunātha puts pressure at exactly the right place. The 'is-the-number-of' relation is not reducible to the relation of inherence or any relation constructed out of it, for it is a relation between numbers and *any* type of thing. What is this new relation? Raghunātha points out that while inherence is a distributive relation (*avyāsajya-vṛtti*), the number-thing relation has to be collective (*vyāsajya-vṛtti*). The distinction occurs in the context of sentences with plural subjects. An attributive relation is distributive if it relates the attribute to every subject — if the *trees* are old, then each individual tree is old. A relation is collective if it relates the attribute to the subjects collectively but not individually — 'the trees form a forest' does not imply that each tree forms a forest. Number attributions are collective; if one says that there are two pots here, one does not imply that each pot is two. Inherence, however, is a distributive relation, and so cannot be the relation of attribution for numbers. This new relation is called the 'collecting' (*paryāpti*) relation by Raghunātha:[92]

> The collecting relation, whose existence is indicated by constructions such as "This is one pot" and "These are two", is a special kind of self-linking relation.

[90] *Śaktivāda with Kṛṣṇa Bhaṭṭa's Mañjūṣa, Mādhava Bhattācārya's Vivṛtti and Sāhitya Darśanācārya's Vinodini*, edited by G. D. Sastri (Benares: Kashi Sanskrit Series no. 57, 1927). p. 189.

[91] *Padārthatattvanirūpaṇa*, pp. 86–87.

[92] *Avacchedakatvanirukti with Jagadīśa's Jāgadīśī*, edited by Dharmananda Mahabhaga (Varanasi: Kashi Sanskrit Series 203), p. 38.

His commentator Jagadīśa explains:

> It might be thought that the collecting relation is [in fact] nothing but inherence...So Raghunātha states that collecting [is a special kind of self-linking relation]. ... In a sentence like "This is one pot", collecting relates the property pot-hood by delimiting it as a property which resides in only one pot, but in a sentence like "These are two pots", collecting relates the property twoness by delimiting it as a property which resides in both pots. Otherwise, it would follow that there is no difference between saying "These are two" and "Each one possesses twoness".

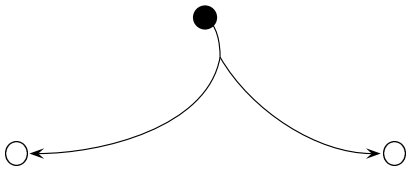
Thus the number two is related by the collecting relation to the two pots jointly, but not to either individually. Raghunātha's idea is clear in the graph-theoretic context. The introduction of numbers requires one final expansion of the graph. We introduce another new domain of nodes (**1**, **2**, **3**,..) and another new type of edge from these nodes. Like ordinary edges, this new type of edge is an ordered pair whose first member is a node, but now the second member is set of nodes. The new edge connects the node **2** with every pair of nodes (**x**, **y**). Likewise, it connects the node **3** with every triple of nodes (**x**, **y**, **z**), and so on. The node **2**, then, is that node from which all edges to pairs begin, the node **3** the node from which all edges to triples begin, and so forth. This is enough to individuate number-nodes graph-theoretically (see graph on 89:

The nodes to which the new edge can connect a number-node can be of any type. In particular, they can themselves be number-nodes. Indeed, the new edge connects **2** with pairs of nodes one of whose members is **2** itself (see graph on page 90:

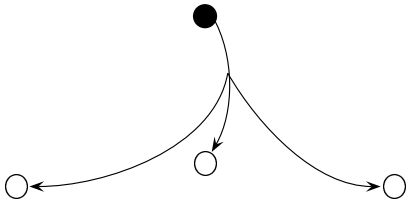
This solves the cross-categorial problem. Number-nodes are related by the new kind of bifurcating edges to nodes of any and every type in the graph, including number-nodes themselves.

The graph-theoretic approach is, I think, full of potential. It offers a new way to read and interpret Navya-Nyāya logic. One might proceed by looking for further situational constraints on what constitutes a permissible graph and applying graph theory to analyse the structure of those graphs. One might also try to establish the relationship between such graphs and classical or nonclassical logics. The treatment of negation suggests a comparison with dialetheic logic, and the idea of self-linking nodes perhaps with non-wellfounded set theory. My aim here has been to expose the logical basis of Vaiśeṣika theory, and to draw a conclusion about the nature of logical thinking in India. The conclusion is simply this. The idea that nature instantiates mathematical structure is not remote from the Indian understanding of natural philosophy, contrary to what has generally been believed, but is in fact a fundamental aspect of it.

NUMBER 2



NUMBER 3



BIBLIOGRAPHY

Selected Indian Logical Texts

[Mogalliputta Tissa, c. 3rd BC] Mogalliputta Tissa (c. 3^{rd} BC). Kathāvatthu. Translation – Aung (1915). Discussion – Bochenski (1956), Ganeri (2001), Matilal (1998: 33–7), Schayer (1932–33).

[Milinda-paṅhā, c. 1st AD] *Milinda-paṅhā* (c. 1st AD). Translation – T. W. Rhys Davids (1890).

[Agniveśa, c. 100 AD] Agniveśa (c. 100 AD). *Carakasaṃhitā*. Translation – Sharma (1981-94). Discussion – Gokhale (1992), Matilal (1998: 38–43), Prets (2000), Solomon (1976, chapter 2).

[Kaṇāda, c. 200 AD] Kaṇāda (c. 100 AD). *Vaiśeṣikasūtra*. Translation – Sinha (1911). Discussion – Nenninger (1994), Nozawa (1991), Schuster (1972)

[Nāgārjuna, c. 200 AD] Nāgārjuna (c. 200 AD). *Vaidalyaprakaraṇa*, Translation – Tola & Dragonetti (1995).

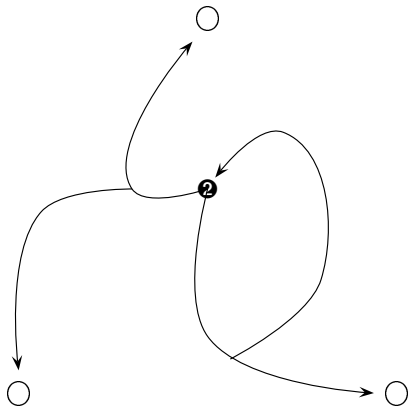
[Nāgārjuna, c. 200 AD] Nāgārjuna (c. 200 AD). *Upāyahṛdaya*. Discussion – Tucci (1929b).

[Gautama, c. 150 AD] Gautama Akṣapāda (c. 150 AD – 250 AD). *Nyāyasūtra*. Translation – Gangopadhyay (1982). Discussion – Bochenski (1956), Chakrabarti (1977), Ganeri (2000), Ganeri (2001), Gokhale (1992), Matilal (1985), Matilal (1998), Prets (2001), Schayer (1933), Randle (1924),

[Vasubandhu, c. 400 AD] *Vasubandhu* c. 400 AD – 480 AD). *Vādavidhi, Vādavidhāna, Tarkaśāstra*. Discussion – Tucci (1929a), Tucci (1929b).

[Vātsyāyana, c. 350 AD] Vātsyāyana (c. 350 AD – 425 AD). *Nyāyabhāṣya*. Translation – Gangopadhyay (1982). Discussion – Bochenski (1956), Matilal (1998).

[Diṅnāga, c. 480 AD] Diṅnāga (c. 480 AD – 540 AD). *Pramāṇasamuccaya*. Translation – Hayes (1988). Discussion – Bochenski (1956), Ganeri (2001), Hayes (1980), Hayes



(1988), Herzberger (1982), Katsura (1983), Katsura (1986a), Matilal (1998), Matilal & Evans eds. (1986), Oetke (1994).

[Diṅnāga, c. 480 AD] Diṅnāga (c. 480 AD – 540 AD). *Hetucakranirṇaya*. Translation – Chatterji (1933), Chi (1969). Discussion – Bharadwaja (1990), Bochenski (1956), Chi (1969), Randle (1924)

[Śankarasvāmin, c. 500 AD] Śankarasvāmin (c. 500 AD – 560 AD). *Nyāyapraveśa*. Translation – Tachikawa (1971). Discussion – Chi (1969), Gillon & Love (1980), Oetke (1996).

[Uddyotakara, c. 500 AD] Uddyotakara (c. 550 AD – 625 AD). *Nyāyavārttika*. Translation – Jha (1984). Discussion – Gokhale (1992).

[Dharmakīrti, c. 600 AD] Dharmakīrti (c. 600 AD – 660 AD). *Pramāṇavārttika*. Discussion – Gokhale (1992), Hayes (1987), Katsura ed. (1999), Matilal (1998), Matilal & Evans eds. (1986), Steinkellner (1973), Steinkellner ed. (1991).

[Dharmakīrti, c. 600 AD] Dharmakīrti (c. 600 AD – 660 AD). *Nyāyabindu*. Translation – Gangopadhyay (1971), Stcherbatsky (1930, volume 2). Discussion – Gokhale (1992).

[Dharmakīrti, c. 600 AD] Dharmakīrti (c. 600 AD – 660 AD). *Vādanyāya*. Translation – Gokhale (1993). Discussion – Chinchore (1988).

[Siddhasena, c. 700 AD] Siddhasena (c. 700 AD). *Nyāyāvatāra*. Translation and discussion – Balcerowicz (2001).

[Udayana, c. 1050 AD] Udayana (c. 1050 AD). *Nyāyavārttikatātparyapariśuddhi, Nyāyapariśiṣta, Lakṣaṇāvalī*.

[Gaṇgeśa, c. 1325 AD] Gaṇgeśa (c. 1325 AD). *Tattvacintāmaṇi*. Discussion – Bhattacharyya (1987), Bochenski (1956), Gangopadhyay (1975), Goekoop (1967), Ingalls (1951), Matilal (1968), Matilal (1985), Matilal (1998), Staal (1988), Vattanky (2001), Wada (1990), Wada (forthcoming).

[Aung, 1915] S. Z. Aung. *Points of Controversy, or, Subjects of Discourse: Being a translation of the Kathāvatthu from the Abhidhammapītaka*, eds. S.Z. Aung and C.A.F. Rhys Davids. Pali Text Society. Routledge and Kegan Paul, London, 1915.

[Bagchi, 1953] S. Bagchi. *Inductive Reasoning: A Study of Tarka and its Role in Indian Logic*. Munishchandra Sinha, Calcutta, 1953.

[Balcerowicz, 2001] P. Balcerowicz. *Epistemology in Historical and Comparative Perspective: Critical Edition and English Translation of Logical-Epistemological Treatises: Nyāyāvatāra, Nyāyāvatāra-vṛti and Nyāyāvatāra-t ippana with Introduction and Notes*, Franz Steiner Verlag, Hamburg, 2001.

[Balcerowicz and Mejor, 2000] P. Balcerowicz and M. Mejor, eds. *On the Understanding of other cultures: Proceedings of the International Conference on Sanskrit and Related Studies to Commemorate the Centenary of the Birth of Stanisław Schayer*, Warsaw 1999. Oriental Institute, Warsaw University. 2000.

[Bharadwaja, 1990] V. Bharadwaja. *Form and Validity in Indian Logic*. Indian Institute of Advanced Study, Shimla, 1990.

[Bhattacharyya, 1987] S. Bhattacharyya. Some Aspects of the Navya-Nyāya Theory of Inference, In *Doubt, Belief and Knowledge*, S. Bhattacharyya, pp. 245–267. Indian Council of Philosophical Research, Delhi, 1987.

[Bochenski, 1956] J. M. Bochenski. The Indian Variety of Logic. In *A History of Formal Logic*, 2nd edn, J. M. Bochenski, pp. 416–447. Trans. I. Thomas, Chelsea Publ. Co., New York, 1961.

[Chakrabarti, 1977] K. K. Chakrabarti. *The Logic of Gotama*. University of Hawaii Society for Asian and Comparative Philosophy Monograph, no. 5. University Press of Hawaii, 1977.

[Chatterji, 1933] D. Chatterji. Hetucakranirṇaya - A Translation. *Indian Historical Quarterly*, **9**, 266–272, 511–514, 1933.

[Chi, 1969] R. S. Y. Chi. *Buddhist Formal Logic: A Study of Dignāga's Hetucakra and K'uei-chi's Great Commentary on the Nyāyapraveśa*. The Royal Asiatic Society of Great Britain. London, 1969.

[Chinchore, 1988] M. Chinchore. *Vādanyāya – A Glimpse of a Nyāya-Buddhist Controversy*. Sri Satguru Publications, Delhi, 1988.

[Davids, 1890] T. W. R. Davids. *The Questions of King Milinda* (2 volumes). Clarendon Press, Oxford, 1890.

[Galloway, 1989] B. Galloway. Some Logical Issues in Madhyamaka Thought. *Journal of Indian Philosophy*, **17**, 1–35, 1989.

[Ganeri, 1999] J. Ganeri. Dharmakīrti's Semantics for the Particle eva ("only"). In Katsura ed., pp. 101–116, 1999.

[Ganeri, 2000] J. Ganeri. Rationality as a method of research into the Nyāya system. In Balcerowicz & Mejor. 147–156, 2000.

[Ganeri, 2001] J. Ganeri. *Philosophy in Classical India: The Proper Work of Reason*. Routledge, London, 2001.

[Ganeri, 2001a] J. Ganeri, ed. *Indian Logic: A Reader*. Curzon, London, 2001.

[Ganeri, 2001] J. Ganeri. Argumentation, Dialogue and the Kathāvatthu. *Jouranl of Indian Philosophy*, 29, 485–493, 2001.

[Gangopadhyaya, 1971] M. Gangopadhyaya. *Vinītadeva's Nyāyabinduṭīkā*. Indian Studies Past & Present, Calcutta, 1971.

[Gangopadhyay, 1975] M. Gangopadhyay. Gaṅgeśa on Vyāptigraha: The Means For The Ascertainment of Invariable Concomitance. *Journal of Indian Philosophy*, **3**, 167–208, 1975.

[Gangopadhyay, 1982] M. Gangopadhyay. *Gautama's Nyāya-Sūtra with Vātsyāyana's Bhāṣya*. Indian Studies.Past & Present, Calcutta, 1982.

[Gangopadhyay, 1984] M. Gangopadhyay. *Indian Logic In Its Sources*. Munshiram Manoharlal, Delhi, 1984.

[Gillon and Love, 1980] B. Gillon and M. L. Love. Indian Logic Revisited: Nyāyapraveśa Reviewed. *Journal of Indian Philosophy*, **8**, 349–384, 1980.

[Gillon and Hayes, 1982] B. Gillon and R. Hayes. The Role of the Particle eva in (Logical) Quantification in Sanskrit. *Wiener Zeitschrift für die Kunde Süd-und Ostasiens*, **26**, 195–203, 1982.

[Gillon, 1999] B. Gillon. Another Look at the Sanskrit Particle eva. In Katsura ed., pp. 117–130, 1999.

[Gillon, 2001] B. Gillon, ed. Proceedings of the Panel on Logic in Classical India, ICANAS Montreal 2000. *Journal of Indian Philosophy Special Issue*. vol. 29, 2001.

[Goekoop, 1967] C. Goekoop. *The Logic of Invariable Concomitance in the Tattvacintāmaṇi*. Reidel, Dordrecht, 1967.

[Gokhale, 1991] P. P. Gokhale. The Logical Structure of Syādvāda. *Journal of Indian Council of Philosophical Research*, **8**, 73–81, 1991.

[Gokhale, 1992] P. P. Gokhale. *Inference And Fallacies Discussed In Ancient Indian Logic*. Sri Satguru Publications, Delhi, 1992.

[Gokhale, 1993] P. P. Gokhale. *Vādanyāya of Dharmakīrti: The Logic of Debate*. Sri Satguru Publications. Delhi, 1993.

[Gupta, 1895] S. N. Gupta. The Nature of Inference in Indian Logic. *Mind*, **4**, 159–175, 1895.

[Hayes, 1980] R. P. Hayes. Diṅnāga's Views on Reasoning. *Journal of Indian Philosophy*, **8**, 219 - 277, 1980.

[Hayes, 1987] R. P. Hayes. On The Reinterpretation Of Dharmakīrti's svabhāvahetu. *Journal of Indian Philosophy*, **15**, 319–332, 1987.

[Hayes, 1988] R. P. Hayes. Diṅnāga on the Interpretation of Signs. *Studies of Classical India*, vol 9. Kluwer, Dordrecht, 1988.

[Herzberger, 1982] H. H. Herzberger. Three Systems of Buddhist Logic. In B. K. Matilal and R. D. Evans eds., pp. 59–76, 1982.

[Hoffman, 1982] F. J. Hoffman. Rationality in early Buddhist four-fold logic. *Journal of Indian Philosophy*, **10**, 309–337, 1982.

[Ingalls, 1951] D. H. H. Ingalls. *Materials for the Study of Navya-Nyāya Logic*, Harvard University Press, Harvard, 1951.

[Jha, 1984] G. Jha. *The Nyāya-Sūtras of Gautama with the Bhāṣya of Vātsyāyana and the Vārttika of Uddyotakara*. Motilal Banarsidass (reprint), Delhi, 1984.

[Katsura, 1983] S. Katsura. Dignāga on trairūpya. *Journal of Indian and Buddhist Studies*, **32**, 15–21, 1983.

[Katsura, 1986a] S. Katsura. On trairūpya Formulae. In *Buddhism and Its Relation To Other Religions: Essays in Honour of Dr. Shozen Kumoi on His Seventieth Birthday*, pp. 161–172, 1986.

[Katsura, 1986b] S. Katsura. On the Origin and Development of the Concept of Vyāpti. *Tetsugaku*, **38**, 1–16, 1986.

[Katsura, 1999] S. Katsura, ed. *Dharmakīrti's Thought and its Impact on Indian and Tibetan Philosophy: Proceedings of the Second International Dharmakīrti Conference*, Hiroshima, Verlag Der Österreichischen Akademie Der Wissenschaften, Wien, 1999.

[Katsura, 2001] S. Katsura. Indian logic: induction, deduction or abduction? In Gillon (2001).

[Matilal, 1968] B. K. Matilal. *The Navya-Nyāya Doctrine of Negation*. Harvard University Press, Harvard, 1968.

[Matilal, 1971] B. K. Matilal. *Epistemology, Logic and Grammar in Indian Philosophical Analysis*. Mouton, The Hague, 1971.

[Matilal, 1985] B. K. Matilal. *Logic, Language and Reality: An introduction to Indian Philosophical Studies*. Delhi Motilal Banarsidass, 1985. Second edn. under new subtitle, *Indian Philosophy and Contemporary Issues*, 1990.

[Matilal, 1990] B. K. Matilal. *The Word and the World*, Appendix 2. Oxford University Press, Delhi, 1990.

[Matilal, 1998] B. K. Matilal. *The Character of Logic in India*, edited by Jonardon Ganeri and Heeraman Tiwari. State University of New York Press, Albany, 1998.

[Matilal and Evans, 1986] B. K. Matilal and R. D. Evans, eds. *Buddhist Logic and Epistemology: Studies in the Buddhist Analysis of Inference and Language. Studies of Classical India*, vol. 7. Kluwer, Dordrecht, 1986.

[Mullatti, 1977] L. C. Mullatti. *The Navya-Nyāya Theory of Inference*. Karnatak University Press, Dharwad, 1977.

[Müller, 1853] M. Müller. Indian Logic. Printed as an Appendix to Thomson, W. *An Outline of the Necessary Laws of Thought.* 3rd edition. Longmans, Green, and Co, London, 1853.

[Nenninger, 1994] C. Nenninger. Analogical Reasoning in Early Nyāya-Vaiśeṣika. *Asiatische Studien*, **48**, 819–832, 1994.

[Nozawa, 1991] M. Nozawa. Inferential Marks in the Vaiśeṣikasūtras. *Saṃbhāṣā: Nagoya Studies in Indian Culture and Buddhism*, **12**, 25–38, 1991.

[Oetke, 1994a] C. Oetke. *Studies On The Doctrine Of Trairūpya*, : Wiener Studien zur Tibetologie und Buddhismuskunde, Wien, 1994.

[Oetke, 1994b] C. Oetke. *Vier Studien zum Altindischen Syllogismus.* Reinbek, 1994.

[Oetke, 1996] C. Oetke. Ancient Indian Logic as a Theory of Non-Monotonic Reasoning. *Journal of Indian Philosophy*, **24**, 447–539, 1996.

[Ono, 1999] M. Ono. Dharmakīrti on asāsāraṇānaikāntika. In Katsura ed., pp. 301–316, 1999.

[Peckaus, 2001] V. Peckaus. Dignāga's Lgoic of Invention. Lecture delivered at the First International Conference of the New Millenium on History of Mathematical Sciences, Indian National Science Academy. University of Delhi.

[Prets, 2000] E. Prets. Theories of Debate, Proof and Counter-Proof in the Early Indian Dialectical Tradition. In Balcerowicz & Mejor eds. pp 369–382, 2000.

[Prets, 2001] E. Prets. Proof and counterproof in early indian dialectic and logic. In Gillon ed., (2001).

[Randle, 1924] H. N. Randle. A Note on the Indian Syllogism. *Mind*, **33**, 398–414, 1924.

[Randle, 1930] H. N. Randle. *Indian Logic in the Early Schools.* Oxford University Press, Oxford, 1930.

[Robinson, 1957] R. H. Robinson. Some Logical Aspects of Nāgārjuna's System. *Philosophy East and West*, **6**, 291–308, 1957.

[Schayer, 1932] St. Schayer. Studien zur Indischen Logik. 1. Der Indische und der Aristotelische Syllogismus. 2: Altindische Antizipationen der Aussangenlogik. *Bulletin International de l'Academie Polonaise des Sciences et des Lettres, Classe de Philologie*, Krakow, nr. 4–6, pp. 98–102 (1932) and nr. 1–6, pp. 90–96 (1933). Krakow.

[Schayer, 1933] St. Schayer. Über die Methode der Nyāya-Forschung. In O. Stein and W. Gambert eds., *Festschrift für Moritz Winternitz*, pp. 247–257. Leipzig, 1933.

[Schuster, 1972] N. Schuster. Inference in the Vaiśeṣikasūtras, *Journal of Indian Philosophy*, **1**, 341–395, 1970.

[Sharma, 1981] R. K. Sharma. *Caraka-saṃhitā: Agniveśa's Treatise Refined and Annotated by Caraka.* Text with English Translation. Varanasi: Chaukhambha Orientalia, 1981–84.

[Sinha, 1911] N. Sinha. *The Vaiśeṣikasūtras of Kaṇāda, with the commentary of Śaṅkara Miśra.* The Panini Office, Bhuvaneswari Asrama, Allahabad, 1911. ,

[Solomon, 1976] E. Solomon. *Indian Dialectics*, 2 volumes. B. J. Institute of Learning and Research, Ahmedabad, 1976.

[Staal, 1988] J. F. Staal. *Universals: Studies in Indian Logic and Linguistics.* University of Chicago Press, Chicago and London, 1988.

[Stcherbatsky, 1930] Th. Stcherbatsky. *Buddhist Logic.* Vols 1 and 2, Bibliotheca Buddhica, 26. Leningrad.

[Steinkellner, 1973] E. Steinkellner. On the Interpretation of the svabhāvahetu? *Wiener Zeitschrift Für Die Kunde Süd-Und Ostasiens*, **18**, 117–129, 1973.

[Steinkellner, 1991] E. Steinkellner. The Logic of the svabhāvahetu in Dharmakīrti's Vādanyāya. In E. Steinkellner, ed., 1991a.

[Steinkellner, 1991a] E. Steinkellner, ed. *Studies in the Buddhist Epistemological Tradition. Proceedings of the Second International Dharmakīrti Conference*, Vienna, l989. Verlag Der Österreichischen Akademie Der Wissenschaften, Wien, 1991.

[Tachikawa, 1971] M. Tachikawa. A Sixth-Century Manual of Indian Logic. (A Translation of the Nyāyapraveśa), *Journal of Indian Philosophy*, **1**, 11–145, 1971.

[Tillemans, 1990] T. F. Tillemans. On sapakṣa, *Journal of Indian Philosophy*, **18**, 53–80, 1990.

[Tola and Dragonetti, 1995] F. Tola and C. Dragonetti. *Nāgārjuna's Refutation of Logic (Nyāya): Vaidalyaprakaraṇa.* Motilal Banarsidass, Delhi, 1995.

[Tucci, 1929a] G. Tucci. Buddhist Logic before Diṅnāga (Asaṅga, Vasubandhu, Tarkaśāstras). *Journal of the Royal Asiatic Society*, 451–88, 1929. Corrections: ibid. 870–1, 1929.

[Tucci, 1929b] G. Tucci. *Pre-Diṅnāga Texts on Logic from Chinese Sources*. Gaekwad Oriental Series, no. 49, Baroda, 1929.

[Tucci, 1930] G. Tucci. *The Nyāyamukha of Dignāga: The Oldest Buddhist Text on Logic*. Materialen zur Kunde des Buddhismus, no. 15. Heidelberg: Otto Harrasowitch, 1930.

[Tuske, 1998] J. Tuske. Dinṅnāga and the Raven Paradox. *Journal of Indian Philosophy*, **26**, 387–403, 1998.

[Vidyabhusana, 1921] S. C. Vidyabhusana. *A History of Indian Logic: Ancient, Mediaeval and Modern Schools*. Calcutta University, 1921.

[Warder, 1963] A. K. Warder. The earliest Indian logic, Trudi Dvadtsat Pyatogo Mejdunarodnogo Kongressa Vostokovedov, Moscow, Izdatelstvo Vostochnoi Lieraturi, vol. IV, 1963

[Uno, 1993] A. Uno. Vyāpti in Jainism. In N. K. Wagle and F. Watanabe eds., *Studies on Buddhism in Honour of Professor A. K. Warder*, pp. 160–167. University Of Toronto, Toronto, 1993.

[Vattanky, 2001] J. Vattanky. *A System of Indian Logic: The Nyāya Theory of Inference*. Routledge, London, 2001.

[Wada, 1990] T. Wada. *Invariable Concomitance in Navya-Nyāya*. Sri Satguru. Delhi, 1990.

[Wada, forthcoming] T. Wada. The Origin of Navya-Nyāya and its Place within the History of Indian Logic. In the *Felicitation Volume for M. Tachiwaka* (forthcoming).