

NOESIS  
The Journal of the Noetic Society  
Number 44  
December 1989 - January 1990

Editor: C.M. Langan  
P.O. Box 131  
Speonk, NY 11972

This is the first of several issues of Noesis for which I have agreed to be responsible. Past issues have contained reference to matters which, while seemingly unrelated, involve concepts similar to those which figure in the resolution of Newcomb's paradox. They will be dealt with in subsequent issues; this one is dedicated to the paradox itself.

Also in this issue is a letter which I wrote long ago, but did not submit. It addresses what I perceived at the time as a kind of identity crisis in the Society, and helps explain my hesitation to submit my work on Newcomb's paradox and other topics. There is also a request for personal information, apparently from a student doing a science project. Her letter is reprinted verbatim; members may cooperate as they see fit.

I apologize for the delay in contacting those members originally scheduled to edit the next several issues. Should any of these members be particularly anxious to put their thoughts into circulation, they need merely notify me. I'll make space between a couple of the issues I'm editing, or within one if possible.

The reason I agreed to this arrangement has to do with my use of the journal in an explicitly formal capacity. Ever since I mailed in my thoughts concerning Newcomb's paradox, there has been an expectation that I would back up my general comments with the detailed explanation to which they alluded. I am now prepared to deliver, but only with the expectation of some reasonable considerations from you in return.

Newcomb's paradox is as famous as an enigma can get. Had anyone but me developed as complete a resolution as the one I offer, he would have been extremely likely to seek initial publication in an accredited scientific or philosophical journal. Then again, he would probably also be an academician of some kind, and would thus believe in his prospects for the fair editorial consideration of his work. I, on the other hand, am not in a position to rely on the established system of sponsorship and review; I am unknown among those who assume that solutions to problems like this one will naturally be discovered within their own ranks, and who act in blind accordance with their assumption. I therefore propose to spare myself the futility of pounding on doors to which only they have keys, by using instead a door which is open to me.

This doorway, Noesis, passes into a rather small room. It is therefore important that those within it bear witness to the origin of the material to follow. That is, against the potential for plagiarism or spurious "independent discovery" of my results by professionals who see my position as invitingly vulnerable, I must rely only on your honesty and fairness. I am sure that some among you have hoped that Noesis would one day become a vehicle

for significant original research. I am no less certain that you understand that this kind of glamour is generally accompanied by responsibility.

The paradox has a history both within and without this group. It was invented by William Newcomb, a physicist, and adopted by Robert Nozick, whose exposition of it has long been regarded as definitive. Nozick introduced it to Martin Gardner; Gardner has probably done more than anyone to popularize it. Notable among its attempted resolutions was that of Bar-Hillel and Margalit of the Hebrew University of Jerusalem; this attempted to reconcile the concept of rationality with the distinction between logical and probabilistic independence. But as it failed to give a mechanism spanning the distinction, it was not acknowledged as final.

The problem was apparently introduced to this group by C. Cole, who agreed with a solution affirming the mental and behavioral independence implied by the standard localistic versions of physical reality embraced by most modern scientists. I took written exception to this view, on the basis of the resolution to follow. Subsequently, two contributions by D. Inada appeared in Noesis, both affirming the one-box solution of the problem. Along with the second of these appeared pieces by Cole and K. Ranieri. The latter of these was the more detailed, and was notable in that it adumbrated several aspects of the actual resolution. However, it too skirted the issue of mechanism.

This is the first publication of the resolution. Neither Nozick nor Gardner has seen it; I lack current information on the former, and I am told that the latter refuses unsolicited correspondence on the grounds that most of it turns out to be "crank mail" (if I have it correctly, some of this mail has contained death threats from would-be geniuses frustrated to the point of murder; if this is fact, one can't begrudge him his precautions). I therefore do not claim backing by any expert consensus. On the other hand, truth is not determined democratically, but logically relative to syntax, and this development is clear enough in point of logic. It therefore does not require a quorum of outside agreement to be pronounced correct.

I have lately received correspondence from another member which convinces me that confusion exists concerning both the structure and meaning of Newcomb's problem, AKA Newcomb's paradox. Newcomb's problem calls for one to infer, from given a set of well-defined conditions, which of two alternatives should be selected in order to maximize a certain monetary expectation. It is apparently the impression of some members that the correct solution is "obvious" unless a certain condition ("omniscience") is suspended, at which point all possible solutions are trivial conversions of unknowns into other unknowns. This, however, is where Newcomb's paradox enters the picture. The paradox evolves from a higher-level (meta-linguistic) consideration of mechanisms implied by the "obvious" solution, whatever that may be to a given solver; it is the upper floor of a split-level maze. The controversy exists solely among those who wander its lower corridors without being able to reach the ledges above. More's the pity, for there resides all meaning.

My belated best wishes to all the members. I hope you find the contents of this issue worth the wait, and that it finds you well on your way into a full and satisfying new year.

### The Resolution of Newcomb's Paradox

We begin by defining Newcomb's problem. Enter Newcomb's Demon, a rather close relative of Maxwell's Demon (the similarity goes well beyond a shared affinity for black boxes). Newcomb's Demon - call him ND for short - is a paragon of intelligence, a genius's genius whose acuity transcends the temporal boundaries of merely ephemeral human beings. Like many geniuses, he displays a certain playfulness: his "hobby" is testing his own infallibility in the prediction of human behaviour by making a generous offer to random human playthings. This offer is couched as a "choice"; though ND's human subjects are in fact forbidden to predicate their choices on external events, they are tacitly allowed to use deterministic or nondeterministic internal strategies computed by corresponding classes of neural events (you may already see the superficiality of the distinction between inward and outward events; I include it only as a customary ingredient of the formulation).

Suppose that ND has chosen you as an experimental subject, and that you know that he has never failed to achieve the result he predicted for any subject throughout the extensive history of his experimental run. Nothing appears rigged; in particular, the data are plausibly distributed with respect to the critical behavior of experimental subjects (if Nozick's data are valid, at a ratio of approximately 2.5/1). The experimental set-up is minimal. You are led to a table on which are two boxes. One is transparent, and in it you can clearly see \$1000. The other is dead black and totally opaque. ND tells you that he has placed \$1,000,000 in the black box if and only if he has predicted that you will take only it and not the other (transparent) box. Otherwise - if he has predicted that you will try to "outsmart" him and take both boxes - the black box has been left empty. ND earnestly assures you that this is no joke; he does not intend to fail. Just in case there are any doubts about this, he offers to take you through a few trial runs using monopoly money and a couple of spare boxes (naturally, you will have to leave the room between these "rehearsals"). You take him up on it. He wins no matter what you think or do. Finally, he tells you that play time is over; next time, the deal is for real. He adds that, lest you suppose he has an infinite amount of time to waste on you, he is giving you a reasonable but finite amount of time to make your move.

Having been around the block, you have already perceived that this is all a kind of game in which ND is the "house", and you his opponent. Of course, since ND is playing for the pure satisfaction of being right while you are playing for big money, a minor conceptual adjustment is in order. But, all things considered, you reason that you should nonetheless apply the standard theory of games in order to maximize your gain. Your reasoning then becomes a little more involved.

You know of certain principles of rational decision that seem to apply to this game. One is known as "the maximization of subjective expected utility", and states that you should try to maximize the sum of the products of the utility (monetary value) of each distinct possible outcome by the probability of that outcome, given your "move". Since ND's past performance and serious intent seem to bring the probability that he has correctly predicted your

choice close to unity, this principle is telling you in no uncertain terms to take only the black box.

But there is a similar rule, called the "dominance principle", which also seems to apply to your predicament. It states that if, for every possible move (or prediction) of your opponent, there is a single countermove as advantageous to you as any other, and more so in at least one case, then you should make that move. You know that ND has already made his move: there is either \$1,000,000 in the black box, or there is nothing. In either case, your eyes tell you unequivocally that there is \$1000 in the clear box. You reason that you have nothing to lose, and precisely \$1000 to gain, by taking both boxes. The dominance principle merely drapes a veil of reason over what your eyes and instincts have already told you: (\$1.001 million or \$0.001 million) beats (\$1 million or \$0). Conclusion: take both boxes.

But this seems to constitute a dilemma. These two principles of rationality, which are supposed to be mutually consistent, are in this case telling you to do opposite things. Feverishly - there is a lot of money at stake for you - you rack your memory for other rules which might help you reconcile or at least decide between them. What about the gambler's fallacy? No matter how many times ND has won in the past, isn't his string of victories really just an extremely unlikely random sequence? Time only runs in one direction, so how can he possibly claim to be more than a lucky (and very rich) guesser? If this is the case, then you certainly don't want to embrace the fallacy of assuming that the outcome of your imminent trial is in any way dependent on those preceding it. If nothing else comes to mind, and quick, you're going to have to go the dominance route and take both boxes.

As if on cue, your neurons begin again to flicker. It occurs to you that you could make no decisions whatever, and would know nothing at all, without relying on various scientific axioms. But axioms are mutually independent, or virtually random with respect to each other. None of them can be deduced from other axioms. Doesn't accepting these axioms - including the dominance principle or whatever more primitive concepts underlie it - thus amount to something very much like the gambler's fallacy? That is, isn't one accepting as a premise that certain future events will be like similar events from the past, but without being able to prove it? You are now in the Minotaur's recreatorium without a thread.

You are also roughly as well-informed on what you should do as any professional scientist or logician who has ever published an analysis of Newcomb's paradox. Fortunately, I am no professional, and need grind the favorite axe of no publisher, professor, or employer. The trail that unwinds below is freshly blazed; though a bit steep in spots, it can be negotiated purely on the strength of an open and reasonably sharp mind.

The first step towards resolving the paradox is to provide a logical scaffolding from which to construct the mathematical model necessary for sound inference. Past arguments involving the problem have used either the standard decision-theoretic model of people playing a game, or the linear "arrow of time" fundamental to classical physics. These two models have been taken to imply opposite solutions, and this suggests that they be somehow unified in an extended "meta-model" which adequately relates the concepts

to each other. It has been less obvious just what this higher model should be.

The solution of problems, and the resolution of paradoxes, are inherently computational activities. What, then, could be a better setting for this resolution than a computational one? And what could possibly be a more fitting preface than a brief introduction to the abstract theory of computation?

Consider an acceptor  $F = (Q, \Sigma, \delta, q_0, A)$ .  $Q$  is a finite nonempty set of internal states,  $\Sigma$  is an alphabet whose symbols are concatenated as strings,  $q_0 \in Q$  is the initial state, and  $A \subseteq Q$  is the set of accepting states triggered only by input recognizable to  $F$ . The transition mapping  $\delta$ , which governs the way in which  $F$  changes states, is deterministic if  $\delta: Q \times \Sigma \rightarrow Q$ , but nondeterministic if  $\delta: Q \times \Sigma \rightarrow 2^Q$  (where  $2^Q$  is the set of all subsets of  $Q$ ). In the nondeterministic case,  $\delta$  will be written  $\delta_\lambda$  for clarity. In terms of human psychology, we might regard the 5-tuple of  $F$  as its "categorical imperative", or accepting syntax, and say that  $F$  projects this syntax onto its universe. Nothing in the universe of  $F$  is recognizable to it but the particular input strings (sense data, facts) which cause it to pass through some  $q \in A$ ; they are its phenomenal "reality", a subset of the noumenal metareality of the wider universe in which strings are representationally generated and entered by programmers. The restriction to finite  $Q$  is pragmatic and amenable to conditional relaxation.

If  $F$  is deterministic, it accepts (recognizes) a string  $s \in \Sigma^*$  if and only if  $\delta(q_0, s) \in A$ . Since we have defined  $\delta$  only for the individual symbols  $\sigma \in \Sigma$ , we must define an extended transition function  $\delta'$ : where  $\lambda$  is the null string,  $\delta'(q, \lambda) = q$ ; and for all  $q \in Q$ ,  $\sigma \in \Sigma$ , and  $s \in \Sigma^*$  (where  $\Sigma^*$  is the set of varirepetitional permutations of the  $\sigma \in \Sigma$ ),  $\delta'(q, s\sigma) = \delta'(\delta'(q, s), \sigma)$ . Thus, the accepting behavior of  $F$  is defined inductively for  $\sigma$ -quantized string extensions; the way in which we recognize and assimilate new bits of information within our reality is specified in  $\delta'$ . Were we to widen the discussion to imagination, conceptualization, theorization, or other intrinsic computations, we would need to consider "ideas"; we would have to generalize from recognition to abstraction by means of a nonlocal or self-iterating, input-free extension of  $\delta$ . If the reference to "strings" seems to imply a dimensional limitation on input, this too can be generalized.

Where  $F$  is nondeterministic, it accepts a string  $s \in \Sigma^*$  if and only if  $\delta_\lambda(q_0, s) \cap A \neq \emptyset$ . The nondeterministic extension  $\delta'_\lambda$  of  $\delta_\lambda$  is defined by induction:  $\delta'_\lambda(q, \lambda) = \{q\}$ , and  $\delta'_\lambda(q, s\sigma) = \cup_{q' \in \delta_\lambda(q, s)} \delta_\lambda(q', \sigma)$ . I.e.,  $q'$  is one of the possible successors of  $q$  under  $\delta_\lambda$  given  $s$ ; the unextended mapping  $\delta_\lambda$  on singletons of  $\Sigma^*$  then determines the image under  $\delta'_\lambda$  of  $s$  plus an adjoint symbol  $\sigma$ , given  $q'$ . This is a classical recursive definition. It describes a stepwise probabilistic ramification of computational potential whose complexity depends on  $\delta_\lambda$ .

Nondeterminism is not always restricted to  $Q \times \Sigma$ ; under certain conditions, either  $Q$  or  $\Sigma$  can be extended, or  $A$  shifted within  $Q$ . This, of course, entails a modification of  $F$ , unless  $F$  is defined to allow for parametric extension and adjustment. To this effect, let  $F'$  be such an open extension. To be meaningful in mechanistic contexts such as those in which acceptors are usually considered,  $F'$  must exist within an appropriate mechanistic extension of the

computative environment of  $F$ . Organisms, being mechanical in the deterministic sense, need not be distinguished in this extension. Nondeterminism can be used to subtly manipulate recognition, thus cryptically modifying an acceptor's reality. Nondeterministic recognition can help to explain the ability of an acceptor to rapidly sieze certain kinds of higher-order phenomena, or even interact with higher-order agencies ordinarily insensible to it.

Having thus formalized the logical abstraction of recognition - i.e., the passive phase of organo-mechanical cognition - we now proceed to the output behavior of computative automata, or to the active phase of cognition. Consider a transducer  $M = (\Sigma, Q, T, \delta, \mu)$ , where  $\Sigma$  is a finite nonempty input alphabet,  $Q$  a finite nonempty state set,  $T$  a finite nonempty output alphabet,  $\delta: Q \times \Sigma \rightarrow Q$  the state-transition function, and  $\mu: Q \times \Sigma \rightarrow T$  the output function. A computation of  $M$  has internal and external phases; through  $\mu$ , the output that  $M$  delivers back to its outward universe depends on strings of  $\delta$ -iterated transitional internal states. Thus,  $\mu$  is a functional of the function  $\delta$  of input. Together,  $\mu$  and  $\delta$  totally determine the behavior of  $M$ . They can be extended from  $\Sigma$  and  $T$  to  $\Sigma^*$  and  $T^*$  as for the acceptor  $F$ :  $\delta'(q, \lambda) = q$ ,  $\mu'(q, \lambda) = \lambda$ ; and  $\delta'(q, s\sigma) = \delta(\delta'(q, s), \sigma)$ ,  $\mu'(q, s\sigma) = \mu'(\delta'(q, s), \sigma)$ . Where appropriate, we can add to  $M$  an initial state ("reset control")  $q_0$ :  $M = (\Sigma, Q, T, \delta', \mu', q_0)$ , to be regarded as a separator of locally independent computations, and put it at the disposal of a function  $r \subset M$  which determines computational relevancy.

Considered as a robotic brain,  $M$   $T$ -behaves according to  $\mu$ , but  $Q$ -reasons towards its decisions along paths generated by  $\delta$ . Where  $\delta$  is deterministic,  $\mu$  may be related to it as a "timing function" according to which any computation can be arrested (input aborted, regression terminated) and converted to output on passing certain tests. Where the duration of the computation is determined with  $\delta$ , input becomes the only variable. Where input as well is fixed in content and scheduling, the entire system is tightly determined. As Laplace might have observed, predicting the behavior of determinate mechanisms requires only data, the means to acquire it, and a valid scientific methodology to organize and interpret it. While the situation is actually more complex, the fact remains that were one to play a deterministic game with a deterministic transducer like  $M$ , one would need only a detailed knowledge of its input and programming to predict the outcome, given analytic tools adequate for that purpose and consistent with one's own constraints (e.g., the amount of time available for analysis and strategy). If one's object in the game were merely the validation of one's prediction, so much the easier to win.

Suppose instead that  $M$  has nondeterministic output mapping  $\mu$ , where state-transition may or may not be deterministic. Then the prediction of output entails control of  $\mu$ , by the predictor. To win a game of prediction, one must now control  $\mu$ , as well as  $\delta$ , to the extent that it is output-critical; one must take over where the probabilistic  $\mu$  leaves off. Since whatever control the transducer has over itself resides in  $\mu$  and  $\delta$ , one must in effect deprive it of self-control. The relevance to "free will" is obvious.

Computation is purposive. The purpose of an acceptor is pure recognition; no action is explicitly predicated on its internal

transitions. The purpose of a transducer is conversion of input to output; yet, such a conversion is aimless unless algorithmic. Like yin and yang, acceptors and transducers are complementary; only together can they begin to resemble functional systems of organic complexity. In order to model organic systems, transducers must be endowed with goals and algorithms comparable to the ends and means of living beings. Algorithms are themselves purposive procedures which model both acceptance and transduction. The problems which comprise their input are scanned by preliminary steps for certain kinds of information, which must in turn be accepted as parameters by subsequent steps, and so on to the output stage (at which point the algorithm delivers its answer). The mechanistic representation of an algorithm must allow for the innate structure of a device, considered apart from the algorithm itself; this structure may have variant and invariant aspects. The algorithm simply conforms variables to purpose given the invariants. As the definitions of F and M might lead one to expect, this generally involves importing to M the set  $A \subset Q$  defined for F.

Human beings, it is said, are self-programming. Their thought is polyalgorithmic; useful algorithms are either meta-algorithmically constructed, or selected from a learned store, to deal with input. If learning, construction, and selection are deterministic, then they characterize a deterministic meta-algorithm not fundamentally different from any other deterministic algorithm we might study. If they are nondeterministic, then they are characteristic of a nondeterministic meta-algorithm, and likewise. It follows that the formal transductive model of human nature withstands any objection from the relative complexity of human mentality or behavior.

Newcomb's object-transducer  $M_n$  naturally includes an acceptor:  $M_n = (E, Q, q_0, \delta, A, \mu, T)$ . Recognition is phasic; a string must often be "pre-accepted" for M to tell whether to accept or reject it. Ordinarily, this tentative phase of recognition is easily computed by the physical entities whose behavior is predicted by ND. To be "real", an input-quantum  $\sigma$  must simply possess a certain first-order predicate, "reality", which - this being a self-validating tautology - induces a type-theoretic predicate stratification like that involving the old Cretan, Epimenides. To this *sine qua non* of recognition corresponds a primary element  $q_1$  of A; no input-quantum failing the  $q_1$ -test is reified, whereas all those passing are relayed to  $Q - q_1$ . Higher-order recognition of "passed" quanta then proceeds at a rate determined by the respective computational demands of the stratified-algorithmic phases of  $\delta_n$ . Corresponding to the structure of A are various ordered states analogous to  $q_1$  within their respective levels of acceptance.

Let us narrow the definition of  $M_n$  in a way consistent with Newcomb's problem. Suppose that associated with  $\mu_n$  is a threshold value  $B > 0$  below which output is nil, but above which a decision will be finalized and implemented. With each  $q_1 \in Q$  we associate a pair of strength coefficients  $a_1$  and  $a_2$  to be incremented and decremented according to a strategic  $\delta_n$  appropriate to the Newcomb decision-theoretic context; these represent the current tendency, given the present amount of input, for  $M_n$  to output either possible behavior (taking one or both boxes, respectively). The  $a_i$  divide  $Q$  into three classes  $X^0$ ,  $Y^0$ , and  $Z^0$ , with membership conditions  $a_1 > a_2$ ,  $a_1 < a_2$ , and  $a_1 = a_2$  respectively. To each  $q_1$  is attached

a total weight  $\alpha_1 = |a_{11} - a_{12}|$ . As soon as the output condition [ $\alpha \geq \beta$ ] is met, a decision and behavior result which correspond to the Q-class of the current state (note that  $\delta_n$ , being strategic, precludes  $q \in Z^0$ , or "indecision", at the output deadline). Thus, the states of  $Q \in M_n$  are preferential and impetal, the graded fore-images of the outputs they favor. This is just a convenient way to view the internal configurations to which they correspond, and does not violate the general definition of transducers. Nor, for that matter, does it violate the way human beings perceive their own decision-making processes.

The question posed by Newcomb's problem involves the computational analysis, by a predictive agency with computational characteristics, of the computational analysis undertaken by a transducer on a given input. That input is the problem itself, presented in the manner prescribed by the formulation. This situation, which defines a computational regression, is recursive and inductively extensible. The regression in turn defines the only soluble context for the higher-level "paradox" generated by the problem. This context translates as mechanism. The mechanism is a stratified automaton  $\Gamma$  containing both the predictor and its object-transducer as sub-automata. Whether "free will" is defined deterministically as mere outside non-interference in  $\mu$  and  $\delta$ , or nondeterministically as the ability of  $M_n$  to override any exogenous restriction of  $\mu_n$  or  $\delta_n$ , its mechanism is contained in that of  $\Gamma$ .

Logical diagonalization of the formal computational language generated by the accepting syntax of  $M_n$  directly implies that certain structural aspects of  $\Gamma$  may be unrecognizable to  $M_n$ . In particular, those aspects involving  $M_n$ -relativized nondeterminacy, as well as those involving certain higher-order predicates of the nondistributive, nonlocal organizations involving  $\mu_n$  and  $\delta_n$ , are formally undecidable to it and need not be recognized directly by it with any degree of specificity. To understand why, consider the extent to which a common computer "recognizes" the extended system including its cpu, its other components, its programmers, and the environment it inhabits. In fact, it can recognize nothing that does not conform to its input-to-output transformational grammar. Even if it were self-analytic, such analysis could be limited to a schematic electronic syntax which overlooks the material elements of which it is constructed. In any case, it can make sense of nothing but strings of input translated and rearranged according to the internal stratification of its hard and soft programming.

You, your purposes, and your dependencies are undecidable to it, and so are the mechanisms by which you can predict and control its behavior. It matters not who formulates this undecidability; if the machine's internal logic is inadequate to do so, yours surely is not (currently, most mechanical acceptors are nongeneralistic, treating complementation as negation and negation as rejection; this bars the tools of diagonalization from their computations). Should it ignore your higher prerogatives, you could "diagonalize" it - if nothing extrinsic to the machine were to stop you - with a sledgehammer whose effects on it do not depend on its acceptance. By analogy, Newcomb's object-transducer  $M_n$  cannot preclude  $\Gamma$  on grounds of "insensibility". Nor, for that matter, can we.

There are many self-styled experts on undecidability who have expressed the opinion that all attempts to reify Gödel's theorem



along paranormal lines reflect a misunderstanding of its "real nature". Such experts are quite correct in that a misunderstanding exists, but the misunderstanding is all theirs. What the theorem forces by juxtaposing truth and derivability (or consistency and completeness) is a hierarchical stratification of classes of truth functions and the inferential syntaxes which parametrize them. This stratification follows that of  $\Gamma$ , fractionating computational reality along with the "truth" to which it corresponds.

The stratification of  $\Gamma$  induces stratum-relativizations of computational time and space. Thus, the timetype in which  $M_n$  computes recognition and output is a mere subtype of that in which it is programmed. Dynamical "arrows of determinacy" which are inviolable to  $M_n$ , being programmed into its accepting syntax, have no force whatsoever to the programmatic agencies themselves. This applies just as well to "metrical" restrictions embodied in the  $M_n$ -syntax; these may allow  $M_n$  to recognize nothing but an artificial submetric of the metric in which these agencies define their own existence.  $M_n$  and its reality might consist of quanta with higher-dimensional interpretations as the termini of channels for the transmission of information between strata. Metatemporal predicates may exist with respect to which those of  $M_n$  are definite only in a mutual sense; predicates which  $M_n$  accepts as "before" and "after" could be the programmatic projections of "in front of" and "in back of", or any other  $\Gamma$ -consistent higher-prepositional relationships.

There can thus exist a mechanism  $x \in \Gamma$  through which a predictor like ND could measure and/or control the mappings  $\delta, \mu \in M_n$  in ways directly insensible to  $M_n$ . Where in  $\Gamma$  relative to  $M_n$  would such a predictor have to be located? Precisely where access is available. Simplistically, we might characterize the predictor-M relationship as one of proper inclusion, where it is understood that prediction is direct rather than second-hand, and programmatic in the passive and active senses. That is, a programmer mentally internalizes the structure of that which he programs, and this internalization amounts to computational inclusion. The fine structure of  $\Gamma$ , while to a degree analytic, is a matter of some complexity. For now, it will suffice to have demonstrated the possibility of  $x$  and its utility to well-situated  $\Gamma$ -subautomata. Because  $\Gamma$  is structured to allow for relativized determinacy and nondeterminacy, the solution is invariant with respect to argumentation involving mind-brain dichotomy. That is, such dichotomies reduce to distinctions of determinacy and nondeterminacy, and may be treated in kind.

Restricted dominance, which relies on probabilistic independence derived from the lower-order, localistic dynamical timetype of  $M_n$ 's artificially restricted "reality", is revealed under  $\Gamma$ -extension to be itself dominated by utility. That is, the subjective utility of  $M_n$  forces the assimilation by  $\delta_n$  of this entire demonstration, which disables restricted dominance and thus frees the strategic component to recognize higher patterns among observed data. The principle of restricted dominance, though valid as long as the reality of  $M_n$  remains unbreached, loses all force in the presence of exodynamic influence.

Let's sum it up. You can be modeled as a deterministic or nondeterministic transducer with an accepting syntax that can be diagonalized, or complemented by logical self-negation. ND can be modeled as a metalogical, metamechanistic programmatic agency.

some insensible part of whom surrounds you in a computational space including physical reality, but not limited thereto. This space is the mechanistic equivalent of the computative regression around which Newcomb's problem is essentially formulated. The existence of this space cannot be precluded by you on the grounds that you cannot directly observe it, nor can it be said by you to deny ND a mechanism of control and prediction of your thought and behavior. Additionally, you have an open-ended run of data which lowers to  $1/\infty$  the probability that ND is "just lucky". This implies that mechanism does indeed exist, and warrants the adjunction to the axioms of physics an independent, empirical physical axiom affirming that mechanism. This then implies that ND can predict or control human thought and behavior (a somewhat weaker implication, you will notice, than "omniscience"). ND possesses means, motive, opportunity...and you. You are "possessed" by Newcomb's Demon, and whatever self-interest remains to you will make you take the black box only. (Q.E.D.)

Do not let incredulity cheat you of understanding. It may be hard to accept the idea of a mathematical proof of the possibility of demons; if so, you may take solace from the fact that this does not of itself imply their actual existence. Remember that Newcomb's problem includes a hatful of hypothetical empirical data that may have no correspondents in your sector of reality. The proof just given merely prevents you from generally precluding the paranormal experiences reported by others on logical grounds. That is, it restrains you from any blanket distribution of too narrow a brand of "rationalism" over reality. Because I have long been exploring the ramifications of the logic this proof employs, I already know how easy it would be to deal with whomever might dispute it on "scientific" or other grounds, regardless of their prestige or reputed intelligence. There is a certainly a strong correlation between rationalistic dogma and the persecution of its foes. Fortunately, it is not too late for "rationalism" to catch up with the advances of twentieth-century logic.

Regarding free will, consider that demons are generally reputed to offer a material reward for the surrender of one's soul. Where we define "soul" as the autoprogrammatic  $\Gamma$ -extension  $M'$  of one's characteristic transductive representation  $M$ , this means only that one is rewarded for voluntarily moving aside and letting the demon take over as captain of one's fate. He does this by adjoining to his subject any functional extensions he requires; e.g.,  $d(\delta)$ , or  $m(\mu)$ . Or, if restricted determinism prevails but he is not in an especially predictive mood, he might create a parametric extension  $F_{\infty}'$  (analogous to  $F'$ ) enabling a hyperdeterministic override  $m \supset \mu$  of one's deterministic output function. Actually, volition may be irrelevant; one might be shoved aside rather than nudged. Where freedom is defined negatively on restricted determinacy, it is duly stratified.

It has by now occurred to most readers that "demon" is a darkly connoted term that could just as well be replaced with "guardian angel" or something equally comforting. After all, ND does not ask his subjects to commit evil, but only to let him do them a rather large favor with no apparent strings attached. Insisting on "free will" here is reminiscent of the child who stares defiantly at its mother while touching the hot stove she has just finished telling

it to avoid. If the child's act has any meaning at all, it is just that poorly timed independence can lead to more painful dependency down the line (surprisingly, Isaac Asimov advocates just this kind of defiance in the Newcomb context, and on what he seems to feel are ethical grounds; but we will deal with ethics later). One cannot be willingly "enslaved" by ND before being enslaved by the money he offers. Who among us always behaves in willful defiance of economic reality? My heart goes out to him, for he is either "enslaved" by an institution, or he is at the absolute mercy of the spontaneous, unsolicited charity of those whose own demons generally predispose them to pass him by.

Having quenched the fire, we will now "mop up". Professor Nozick himself has remarked on the possible nonrecursivity of scientific laws involved in the determination of behavior. This, he surmises, prevents the inference of predictability from determinism. But this conclusion entails the confinement of these predicates within a restricted syntax, whereas the extensionality of  $\Gamma$  renders such a constraint powerless. In other words, restricted recursivity is beside the point for agencies defined to possess sufficient access to the channels through which such laws project as restricted reality, and particularly to their source. All ND has to do is run a  $\Gamma$ -presimulation of interstratum transmission. Note that the computative requirements for this, which correspond to assumptions on the nature of  $\Gamma$ , are satisfied given ND's efficacy.

Nozick also remarks on the regression of determinism to self-validating premises. Self-validating languages comprise their own inferential schemata. Validation, being inferential to the extent of quantification, follows the  $\Gamma$ -stratification of inference. So self-validation is also stratified, and such languages correspond to appropriate levels of computative reality. For sufficiently expressive languages, self-validation becomes diagonalistic self-valuation. Determinism is thus stratified, and so is its counterpart, freedom. The seemingly infinite regression can be brought to closure at an ultimate form of undecidability, which in turn can be locally resolved with a relativized degree of confirmation.

Regarding various remarks that have been submitted to Nozick by members of the physics community, I should explain that the above resolution has a metaphysical generalization of which certain implications may broadly be termed physical. This generalization, as developed by the author, is called the CTMU, or Computation-Theoretic Model of the Universe (the acronym has a mnemonic pronunciation: 'cat-mew'). We will later examine chosen aspects of this model, which transcends characterization as a "theory" by virtue of its categoricity with respect to computative entities such as human beings. That is, the CTMU is as close to "absolute truth" as we will ever be privileged to get along computative or intellectual pathways. If I might be indulged a bit of testimony: beyond the CTMU ultimate syntax, the lattice which gives shape to the multilayered veil of maya, there is but the light that shines forever.

Christopher Michael Chappelle-Langan

(copyright 1990 by C. M. Langan. No portion of this article is to be reproduced in any form without the author's written consent.)

Bibliography: Newcomb's Problem and Two Principles of Choice  
Robert Nozick in *Essays in Honor of Carl G. Hempel*  
Editor: Nicholas Rescher. Humanities Press, 1970  
Newcomb's Paradox Revisited  
Maya Bar-Hillel and Avishai Margalit in *The British Journal for the Philosophy of Science*, vol. 23, No. 4, pp. 295-304; November, 1972  
Reflections on Newcomb's problem: a prediction and free will dilemma  
Martin Gardner and Robert Nozick in the Mathematical Games Department of *Scientific American*, pp. 102 - 106; March, 1974

P. O. Box 403  
Waskom, TX 75692  
January 14, 1990

To Whom It May Concern:

I am doing scientific research on the correlation between I. Q. and blood type. This is a non-profit science fair project. Therefore, I would like to have the names, addresses, and phone numbers of the people in your organization. Your cooperation would be greatly appreciated.

Thank You,



Darlene Michelle Wade

This is to let you know, concerning your planned reduction of Noesis to quarterly editions, that I'd consider assuming editorial responsibilities in whole or in part. The offer seems incumbent on me in that I might have less trouble composing editorials on past and future material. However, I'd require some advance notice in order to upgrade my word processor, since photoreduction may (for all I know) be as expensive as retyping material and printing the journal at home or in the office.

There remains a question of demand. What, in your opinion, is the level of interest in monthly publication? Even with the small amount of direct feedback on my responses to certain published theses, there is doubt as to whether it would be constructive to pursue them to their conclusions with this particular readership. As with any other ideas, grasping those involved here will require enthusiasm as well as intellect. I'd consider your insights, if you have any, to be of clarificative value.

I'll continue in light of Noesis #35, which I've just read. It is remarked, within a complaint about the proposed reduction in frequency, that certain members - including me - have complained about the size of the journal. Other complaints have been made about proprietary issues in science. All of these concerns - size, frequency, and propriety - are related within a wider issue, viz: what are the journal's real purposes, and how well do they serve those of the HRC? As it would be difficult to come up with a more germane topic, I'll attempt to clarify matters in as impersonal a way as possible.

The Encyclopedia of Associations apparently lists this organization's purpose as social. Quarterly publication seems consistent with that; the journals of social clubs need be no more frequent than their major activities, and this club is strewn too widely to hold many meetings. In addition, the reigning editor already seems to consider the costs of publication prohibitive in terms of both time and money. It therefore rests on those desiring monthly issues to provide a justification for their preference.

Speaking only for myself, I've already "put my money where my mouth is" by responding to a number of journal entries in what I've observed to be the going format, which just happens to be a little too tight for the kind of material I've contemplated submitting. Ideas near the foundations of mathematics - where math is the same as philosophy - cannot rely on established symbolism, but may involve the definition of new concepts in terms of natural language. In fact, this situation calls for more than just new definitions; since the HRC probably includes members not fluent in all the basics, more than the usual amount of illustration and preparatory explanation would be necessary.

But this is not the whole problem. I'm sure that some members are familiar with Edward Thorp, the ex-UCLA mathematician who devised a "card-counting" scheme for blackjack and was barred from the casinos for his consistent ability to beat the house. He was interviewed for the September 1988 issue of Omni magazine, in which he complained that he - an extensively credentialed author with major institutional sponsorship - was robbed of credit for at

least two different papers that were published under somebody else's name. He and his colleagues allegedly calculated over ten years ago that published mathematical papers, which can range in content from the profound to the trivial, had a monetary value of \$10,000 each, as well as value in "power and prestige". Thorp is not the only one to have made such complaints; many others, from a variety of disciplines, have voiced their own. The recent Franco-American dispute over the discovery of the AIDS retrovirus is a case in point.

One might conclude that there are many scholars and scientists, hungry for fame, tenure, and grant money, who are willing, given the opportunity, to stage false cases of "simultaneous discovery" (despite allegations to the contrary, the AIDS controversy was eventually determined to have been a genuine instance of this). One can reason further that this temptation, where it exists, is more likely to center on the work of unknowns, who are often in more vulnerable positions than those who already possess a degree of credibility. Were a talented amateur to succeed in producing a truly important piece of research - a possibility of which the existence of groups like the HRG is a kind of affirmation - it would behoove him to exercise circumspection, lest he be relegated to a footnote (or less) in the undeserved accolades of another (it may not have gone unnoticed that the one original invention so far described in full detail in Noesis was reported only after its patent application had already been processed).

As things now stand, nobody who commits this kind of injustice need fear professional censure on the strength of this journal's scholarly prestige. There has been too little in it to date that would qualify it as grounds for a serious accusation of scientific plagiarism. And condensed material, which presents general ideas at the expense of close detail, is easy prey to anyone with the expertise to fill in those details in such a way as to disguise the source of their organization. The same goes for serialization; a first installment can telegraph those to follow. These problems are already familiar to those who write and publish this kind of material; they would otherwise be too naive to survive in today's world of high-stakes science. The subject is unpleasant, but it must be dealt with realistically before any member can safely contribute what he reasonably considers to be solutions for open problems of general import (and especially famous ones like that of the consistency of Bayesian inference, Newcomb's problem, and others that have been mentioned in Noesis). The alternative would be to declare Noesis a digest, suitable only for trivia, condensations of previously-published material, and the rare original thesis which can be fully propounded to an unprepared readership in a space of seven or fewer pages.

It would be nice if science always lived up to its rep as a temple of trust, conscience, and the free exchange of information and understanding. But the system is designed to work much better for some than for others. Just as "it takes money to make money", it takes credit to take credit; the deck is stacked against anyone whose insight is appreciably weightier than his portfolio. At one extreme, the crackpot is prevented from corrupting the worthwhile. At the other, truth is made the slave of recognition, and by those

who claim to serve only truth. When one passes by a shell game, is one duty-bound to play? Only if your answer is an unqualified yes can you insist that anyone play this game without first seeking a precautionary accord.

It is not my intention to imply that the halls of science and academia teem with thieves. In the hearts of those gravitating to learned careers, the concept of service to humanity is seldom completely absent. But the human condition is a moral paradox, and all walks of life contain men and women who prefer to resist their consciences, as long as they can resist while surrounded by the comfort and esteem their consciences might have cost them. Since hypocrisy only compounds the problem, it is better not to pretend that running (what may or may not be) a scientific journal is a stroll through the ethics hall of fame. I suspect that even some of the uninvolved members have previously glimpsed the inadequacy of existing policy, so far as it has ever been stated.

Nor am I saying that declaring the HRG a full-fledged think-tank, and *Noesis* a serious journal of science, will precipitate a sudden torrent of first rate research papers. But at least those members having such material need have fewer misgivings over the imprudence of submitting it here first. If the Hoeflin Research Group is in fact what it sounds like - meaning that research is to be done *by* rather than merely *on* its members - then it should at least accommodate the full descriptions of that research. Should this seem unreasonable, consider that other such groups generally take precautions far in excess of any discussed here. The matter of scheduling also deserves consideration; no one submitting original research will want to consign it to editorial limbo for indefinite periods without knowing precisely where it is and who has seen it - if, indeed, anyone but the editor is allowed to see it prior to publication.

Unfortunately, it is true that even well-organized research environments are plagued by the propensities of researchers to personally adopt the problems on which they've been working. There is a natural tendency to become secretive concerning projects in which the investment of time and energy is large. The situation is a direct function of the competitiveness of science, which extends to the very economics of survival; it can be regarded as the price humanity pays for the beneficial aspects of such competition. Whether those benefits are worth that price is another matter. Nor is the nobility of altruism at issue; it is not always altruistic to sacrifice one's economic viability as a researcher. That is, when the work one offers up to the "brotherhood of science" could determine whether his future beneficial research will win support, altruism can actually dictate that he protect his authorship for the sake of the future advancement of science.

The situation is complicated when research has social and ethical ramifications. It seems that when the author of a theory leaves these ramifications up to others, this invariably promotes argument among "experts" perceiving advantage in various slants or constructions. Polarization occurs before he knows it, and it is no time at all before the only reportable "consensus" on his work is that "no one really knows what it all means". Against this, the researcher is powerless once he goes public. Where he has reason

to believe that the implications of his work are socially critical and/or susceptible to abuse or plain misunderstanding, this may bind him to secrecy until he can enlarge enough to spare society the waste of time attending such destructive, misleading debates. And where he is already able to explain his work, he may hesitate to release condensations or incomplete segments of it for the same reasons. While it is true that some amount of debate is inevitable and even potentially beneficial, there is no reason to promote it in any phase of which the outcome is already known. One authoritative voice is often better heard, and more quickly heeded, than the inharmonious din of numerous, partially-informed experts and commentators. While everyone occasionally participates in a debate of which one already knows the outcome, it is usually for the purpose of shortening the discussion - particularly when the topic is perceived to have an intrinsic urgency dominating other considerations like education, diversion, or generating interest.

For my part, I'm an amateur, and by no means desperate to have my work read at all costs. I'm content to leave that syndrome to those with axes to grind, professional and otherwise. I'll submit such material only if I'm convinced that the others sincerely wish to see it, and see it in sufficient detail to minimize the risk of misappropriation. Otherwise, I've already fulfilled my duty as a HRG member by dissenting, as concisely as I could, from what I reasonably considered to be weak (though well-placed) theses. I should add that "weak" describes only the theses, and not the intellects of those who mentioned them here. They are basic and widely held, and it was only natural for the members propounding them to take them on faith.

That about says it all. I can't speak alone for what the HRG is supposed to be. I suggest that somebody else "put his money where his mouth is" with regard to these issues, and we'll go from there. If I'm on the wrong track entirely, those in the know can easily set things straight. In any case, we can avoid having a particular subset of members characterized as the "grinches" who stole two-thirds of everyone else's forthcoming issues. While original research is probably the most advantageous kind of material for the HRG to publish, it is safe to assume that most of its members are familiar enough with the rudiments of English composition to come up with an essay or two.

Perhaps this idea of rotating editorship is worth some close attention, given a prior determination on exactly what it is that is being edited. Some kind of central clearing house might be required for those wishing to submit material, which could otherwise get lost in the circuit.

I hope these remarks prove useful.

C.M. Langan