

# Non Serviam

by Stanisław Lem

*This extract is from the 1971 collection of fictitious reviews entitled [A Perfect Vacuum](#) by the SF author Stanisław Lem. ([A Perfect Vacuum](#), Stanisław Lem, 1971, Harcourt Publishers Ltd, ISBN: 978-0156716864)*

Professor Dobb's book is devoted to personetics, which the Finnish philosopher Eino Kaikki has called 'the cruelest science man ever created'. Dobb, one of the most distinguished personeticists today, shares this view. One cannot escape the conclusion, he says, that personetics is, in its application, immoral; we are dealing, however, with a type of pursuit that is, though counter to the principles of ethics, also of practical necessity to us. There is no way, in the research, to avoid its special ruthlessness, to avoid doing violence to one's natural instincts, and if nowhere else it is here that the myth of the perfect innocence of the scientist as a seeker of facts is exploded.

We are speaking of a discipline, after all, which, with only a small amount of exaggeration, for emphasis, has been called 'experimental theogony'. Even so, this reviewer is struck by the fact that when the press played up the thing, nine years ago, public opinion was stunned by the personetic disclosures. One would have thought that in this day and age nothing could surprise us. The centuries rang with the echo of the feat of Columbus, whereas the conquering of the Moon in the space of a week was received by the collective consciousness as a thing practically humdrum. And yet the birth of personetics proved to be a shock.

The name combines Latin and Greek derivatives: 'persona' and 'genetic'—'genetic' in the sense of formation, or creation. The field is a recent offshoot of the cybernetics and psychonics of the eighties, crossbred with applied intellectronics.

Today everyone knows of personetics; the man in the street would say, if asked, that it is the artificial production of intelligent beings—an answer not wide of the mark, to be sure, but not quite getting to the heart of the matter. To date we have nearly a hundred personetic programs. Nine years ago identity schemata were being developed—primitive cores of the 'linear' type—but even that generation of computers, today of historical value

only, could not yet provide a field for the true creation of personoids.

The theoretical possibility of creating sentience was divined some time ago, by Norbert Wiener, as certain passages of his last book, *God and Golem*, bear witness. Granted, he alluded to it in that half-facetious manner typical of him, but underlying the facetiousness were fairly grim premonitions. Wiener, however, could not have foreseen the turn that things would take twenty years later. The worst came about—in the words of Sir Donald Acker—when at MIT "the inputs were shorted to the outputs".

At present a 'world' for personoid 'inhabitants' can be prepared in a matter of a couple of hours. This is the time it takes to feed into the machine one of the full-fledged programs (such as BAAL 66, CREAN IV, or JAHVE 09).

Dobb gives a rather cursory sketch of the beginnings of personetics, referring the reader to the historical sources; a confirmed practitioner-experimenter himself, he speaks mainly of his own work—which is much to the point, since between the English school, which Dobb represents, and the American group, at MIT, the differences are considerable, both in the area of methodology and as regards experimental goals.

Dobb describes the procedure of "6 days in 120 minutes" as follows. First, one supplies the machine's memory with a minimal set of givens; that is—to keep within a language comprehensible to laymen—one loads its memory with substance that is 'mathematical'. This substance is the protoplasm of a universum to be 'habitated' by personoids. We are now able to supply the beings that will come into this mechanical, digital world—that will be carrying on an existence in it, and in it only—with an environment of nonfinite characteristics. These beings, therefore, cannot feel imprisoned in the physical sense, because the environment does not have, from their standpoint, any bounds.

The medium possesses only one dimension that resembles a dimension given us also—namely, that of the passage of time (duration). Their time is not directly analogous to ours, however, because the rate of its flow is subject to discretionary control on the part of the experimenter. As a rule, the rate is maximized in the preliminary phase (the so-called creational warm-up), so that our minutes correspond to whole eons in the computer, during which there takes place a series of successive reorganizations and crystallizations—of a synthetic cosmos. It is a cosmos completely spaceless, though possessing dimensions,

but these dimensions have a purely mathematical, hence what one might call an 'imaginary' character. They are, very simply, the consequence of certain axiomatic decisions of the programmer, and their number depends on him. If, for example, he chooses a ten-dimensionality, it will have for the structure of the world created altogether different consequences from those where only six dimensions are established.

It should be emphasized that these dimensions bear no relation to those of physical space but only to the abstract, logically valid constructs made use of in systems creation. This point, all but inaccessible to the nonmathematician, Dobb attempts to explain by adducing simple facts, the sort generally learned in school. It is possible, as we know, to construct a geometrically regular three-dimensional solid—say, a cube—which in the real world possesses a counterpart in the form of a die; and it is equally possible to create geometrical solids of four, five,  $n$  dimensions (the four-dimensional is a tesseract). These no longer possess real counterparts, and we can see this, since in the absence of any physical dimension No. 4 there is no way to fashion genuine four-dimensional dice. Now, this distinction (between what is physically constructible and what may be made only mathematically) is, for personoids, in general nonexistent, because their world is of a purely mathematical consistency. It is built of mathematics, though the building blocks of that mathematics are ordinary, perfectly physical objects (relays, transistors, logic circuits—in a word, the whole huge network of the digital machine).

As we know from modern physics, space is not something independent of the objects and masses that are situated within it. Space is, in its existence, determined by those bodies; where they are not, where nothing is—in the material sense—there, too, space ceases, collapsing to zero. Now, the role of material bodies, which extend their 'influence,' so to speak, and thereby 'generate' space, is carried out in the personoid world by systems of a mathematics called into being for that very purpose. Out of all the possible 'maths' that in general might be made (for example, in an axiomatic manner), the programmer, having decided upon a specific experiment, selects a particular group, which will serve as the underpinning, the 'existential substrate,' the 'ontological foundation' of the created universum. There is in this, Dobb believes, a striking similarity to the human world. This world of ours, after all, has 'decided' upon certain forms and upon certain types of geometry that best suit it—best, since most simply (three-dimensionality, in order to remain with what one began with). This notwithstanding, we are able to picture 'other worlds' with 'other properties'—in the geometrical and not only in the geometrical realm. It

is the same with the personoids: that aspect of mathematics which the researcher has chosen as the 'habitat' is for them exactly what for us is the 'real-world base' in which we live, and live perforce. And, like us, the personoids are able to 'picture' worlds of different fundamental properties.

Dobb presents his subject using the method of successive approximations and recapitulations; that which we have outlined above, and which corresponds roughly to the first two chapters of his book, in the subsequent chapters undergoes partial revocation—through complication. It is not really the case, the author advises us, that the personoids simply come upon a ready-made, fixed, frozen sort of world in its irrevocably final form; what the world will be like in its specificities depends on them, and this to a growing degree as their own activeness increases, as their 'exploratory initiative' develops. Nor does the likening of the universum of the personoids to a world in which phenomena exist only to the extent that its inhabitants observe them provide an accurate image of the conditions. Such a comparison, which is to be found in the works of Sainter and Hughes, Dobb considers an 'idealist deviation'—a homage that personetics has rendered to the doctrine, so curiously and so suddenly resurrected, of Bishop Berkeley. Sainter maintained that the personoids would know their world after the fashion of a Berkeleyan being, which is not in a position to distinguish *esse* from *percipi*—to wit, it will never discover the difference between the thing perceived and that which occasions the perception in a way objective and independent of the one perceiving.

Dobb attacks this interpretation of the matter with a passion. *We*, the creators of their world, know perfectly well that what is perceived by them indeed exists; it exists inside the computer, independent of them—though, granted, solely in the manner of mathematical objects. And there are further clarifications. The personoids arise germinally by virtue of the program; they increase at a rate imposed by the experimenter—a rate only such as the latest technology of information processing, operating at near-light speeds, permits. The mathematics that is to be the 'existential residence' of the personoids does not await them in full readiness but is still 'in wraps,' so to speak—unarticulated, suspended, latent—because it represents only a set of certain prospective chances, of certain pathways contained in appropriately programmed subunits of the machine. These subunits, or generators, in and of themselves contribute nothing; rather, a specific type of personoid activity serves as a triggering mechanism, setting in motion a production process that will gradually augment and define itself; in other words, the world surrounding these beings

takes on an unequivocalness only in accordance with their own behaviour.

Dobb tries to illustrate this concept with recourse to the following analogy. A man may interpret the real world in a variety of ways. He may devote particular attention—intense scientific investigation—to certain facets of that world, and the knowledge he acquires then casts its own special light on the remaining portions of the world, those not considered in his priority-setting research. If first he diligently takes up *mechanics*, he will fashion for himself a *mechanical* model of the world and will see the Universe as a gigantic and perfect clock that in its inexorable movement proceeds from the past to a precisely determined future. This model is not an accurate representation of reality, and yet one can make use of it for a period of time historically long, and with it can even achieve many practical successes—the building of machines, implements, etc. Similarly, should the personoids 'incline themselves,' by choice, by an act of will, to a certain type of relation to their universum, and to that type of relation give precedence—if it is in this and only in this that they find the 'essence' of their cosmos—they will enter upon a definite path of endeavours and discoveries, a path that is neither illusory nor futile. Their inclination 'draws out' of the environment what best corresponds to it. What they first perceive is what they first master. For the world that surrounds them is only partially determined, only partially established in advance by the researcher-creator; in it, the personoids preserve a certain and by no means insignificant margin of freedom of action— action both 'mental' (in the province of what they think of their own world, of how they understand it) and 'real' (in the context of their 'deeds'—which are not, to be sure, literally real, as we understand the term, but are not merely imagined, either). This is, in truth, the most difficult part of the exposition, and Dobb, we daresay, is not altogether successful in explaining those special qualities of personoid existence—qualities that can be rendered only by the language of the mathematics of programs and creational interventions. We must, then, take it somewhat on faith that the activity of the personoids is neither entirely free—as the space of our actions is not entirely free, being limited by the physical laws of nature—nor entirely determined—just as we are not train cars set on rigidly fixed tracks.

A personoid is similar to a man in this respect, too, that man's 'secondary qualities'— colors, melodious sounds, the beauty of things—can manifest themselves only when he has ears to hear and eyes to see, but what makes possible hearing and sight has been, after all, previously given. Personoids, perceiving their environment, give it from out of themselves those experiential qualities which exactly correspond to what for us are the charms of a

beheld landscape— except, of course, that they have been provided with purely mathematical scenery. As to 'how they see it,' one can make no pronouncement, for the only way of learning the 'subjective quality of their sensation' would be for one to shed his human skin and become a personoid. Personoids, one must remember, have no eyes or ears, therefore they neither see nor hear, as we understand it; in their cosmos there is no light, no darkness, no spatial proximity, no distance, no up or down; there are dimensions there, not tangible to us but to them primary, elemental; they perceive, for example—as equivalents of the components of human sensory awareness—certain changes in electrical potential. But these changes in potential are, for them, not something in the nature of, let us say, pressures of current but, rather, the sort of thing that, for a man, is the most rudimentary phenomenon, optical or aural—the seeing of a red blotch, the hearing of a sound, the touching of an object hard or soft.

From here on, Dobb stresses, one can speak only in analogies, evocations. To declare that the personoids are 'handicapped' with respect to us, inasmuch as they do not see or hear as we do, is totally absurd, because with equal justice one could assert that it is we who are deprived with respect to them—unable to feel with immediacy the phenomenalism of mathematics, which, after all, we know only in a cerebral, inferential fashion. It is only through reasoning that we are in touch with mathematics, only through abstract thought that we 'experience' it. Whereas the personoids *live* in it; it is their air, their earth, clouds, water, and even bread—yes, even food, because in a certain sense they take nourishment from it. And so they are 'imprisoned,' hermetically locked inside the machine, solely from our point of view; just as they cannot work their way out to us, to the human world, so, conversely—and symmetrically—a man can in no wise enter the interior of their world, so as to exist in it and know it directly. Mathematics has become, then, in certain of its embodiments, the life-space of an intelligence so spiritualized as to be totally incorporeal, the niche and cradle of its existence, its element.

The personoids are in many respects similar to man. They are able to imagine a particular contradiction (that *A* is and that *not-A* is) but cannot bring about its realization, just as we cannot. The physics of our world, the logic of theirs, does not allow it, since logic is for the personoids' universum the very same action-confining frame that physics is for our world. In any case—emphasizes Dobb—it is quite out of the question that we could ever fully, introspectively grasp what the personoids 'feel' and what they 'experience' as they go about their intensive tasks in their nonfinite universum. Its utter spacelessness is no prison—that

is a piece of nonsense the journalists latched onto—but is, on the contrary, the guarantee of their freedom, because the mathematics that is spun by the computer generators when 'excited' into activity (and what excites them thus is precisely the activity of the personoids) — that mathematics is, as it were, a self-realizing infinite field for optional actions, architectural and other labours, for exploration, heroic excursions, daring incursions, surmises. In a word: we have done the personoids no injustice by putting them in possession of precisely such and not a different cosmos. It is not in this that one finds the cruelty, the immorality of personetics.

In the seventh chapter of *Non Serviam* Dobb presents to the reader the inhabitants of the digital universum. The personoids have at their disposal a fluency of thought as well as of language, and they also have emotions. Each of them is an individual entity; their differentiation is not the mere consequence of the decisions of the creator-programmer but results from the extraordinary complexity of their internal structure. They can be very like, one to another, but never are they identical. Coming into the world, each is endowed with a 'core,' a 'personal nucleus,' and already possesses the faculty of speech and thought, albeit in a rudimentary state. They have a vocabulary, but it is quite spare, and they have the ability to construct sentences in accordance with the rules of the syntax imposed upon them. It appears that in the future it will be possible for us not to impose upon them even these determinants, but to sit back and wait until, like a primeval human group in the course of socialization, they develop their own speech. But this direction of personetics confronts two cardinal obstacles. In the first place, the time required to await the creation of speech would have to be very long. At present, it would take twelve years, even with the maximization of the rate of intracomputer transformations (speaking figuratively and very roughly, one second of machine time corresponds to one year of human life). Secondly, and this is the greater problem, a language arising spontaneously in the 'group evolution of the personoids' would be incomprehensible to us, and its fathoming would be bound to resemble the arduous task of breaking an enigmatic code— a task made all the more difficult by the fact that such a code would not have been created by people for other people in a world shared by the decoders. The world of the personoids is vastly different in qualities from ours, and therefore a language suited to it would have to be far removed from any ethnic language. So, for the time being, linguistic evolution *ex nihilo* is only a dream of the personeticists.

The personoids, when they have 'taken root developmentally,' come up against an enigma

that is fundamental, and for them paramount—that of their own origin. To wit, they set themselves questions—questions known to us from the history of man, from the history of his religious beliefs, philosophical inquiries, and mythic creations: Where did we come from? Why are we made thus and not otherwise? Why is it that the world we perceive has these and not other, wholly different properties? What meaning do we have for the world? What meaning does it have for us? The train of such speculations leads them ultimately, unavoidably, to the elemental questions of ontology, to the problem of whether existence came about 'in and of itself,' or whether it was the product, instead, of a particular creative act—that is, whether there might not be, hidden behind it, invested with will and consciousness, purposively active, master of the situation, a Creator. It is here that the whole cruelty, the immorality of personetics manifests itself. But before Dobb takes up, in the second half of his work, the account of these intellectual strivings—these struggles of a mentality made prey to the torment of such questions—he presents in a series of successive chapters a portrait of the 'typical personoid,' its 'anatomy, physiology, and psychology'.

A solitary personoid is unable to go beyond the stage of rudimentary thinking, since, solitary, it cannot exercise itself in speech, and without speech discursive thought cannot develop. As hundreds of experiments have shown, groups numbering from four to seven personoids are optimal, at least for the development of speech and typical exploratory activity, and also for 'culturization'. On the other hand, phenomena corresponding to social processes on a larger scale require larger groups. At present it is possible to 'accommodate' up to one thousand personoids, roughly speaking, in a computer universum of fair capacity; but studies of this type, belonging to a separate and independent discipline—sociodynamics—lie outside the area of Dobb's primary concerns, and for this reason his book makes only passing mention of them.

As was said, a personoid does not have a body, but it does have a 'soul'. This soul—to an outside observer who has a view into the machine world (by means of a special installation, an auxiliary module that is a type of probe, built into the computer)—appears as a 'coherent cloud of processes,' as a functional aggregate with a kind of 'centte' that can be isolated fairly precisely, ie, delimited within the machine network. (This, *nota bene*, is not easy, and in more than one way resembles the search by neurophysiologists for the localized centers of many functions in the human brain.)

Crucial to an understanding of what makes possible the creation of the personoids is

Chapter 11 of *Non Serviam*, which in fairly simple terms explains the fundamentals of the theory of consciousness. Consciousness—all consciousness, not merely the personoid—is in its physical aspect an 'informational standing wave,' a certain dynamic invariant in a stream of incessant transformations, peculiar in that it represents a 'compromise' and at the same time is a 'resultant' that, as far as we can tell, was not at all planned for by natural evolution. Quite the contrary; evolution from the first placed tremendous problems and difficulties in the way of the harmonizing of the work of brains above a certain magnitude—ie, above a certain level of complication—and it trespassed on the territory of these dilemmas clearly without design, for evolution is not a deliberate artificer.

It happened, simply, that certain very old evolutionary solutions to problems of control and regulation, common to the nervous system, were 'carried along' up to the level at which anthropogenesis began. These solutions ought to have been, from a purely rational, efficiency-engineering standpoint, canceled or abandoned, and something entirely new designed— namely, the brain of an intelligent being. But, obviously, evolution could not proceed in this way, because disencumbering itself of the inheritance of old solutions— solutions often as much as hundreds of millions of years old—did not lie within its power. Since it advances always in very minute increments of adaptation, since it 'crawls' and cannot 'leap,' evolution is a dragnet 'that lugs after it innumerable archaisms, all sorts of refuse,' as was bluntly put by Tammer and Bovine. (Tammer and Bovine are two of the creators of the computer simulation of the human psyche, a simulation that laid the groundwork for the birth of personetics.)

The consciousness of man is the result of a special kind of compromise. It is a 'patchwork,' or, as was observed, eg, by Gebhardt, a perfect exemplification of the well-known German saying: *Aus einer Not eine Tugend machen* (in effect: 'To turn a certain defect, a certain difficulty, into a virtue'). A digital machine cannot of itself ever acquire consciousness, for the simple reason that in it there do not arise hierarchical conflicts of operation. Such a machine can, at most, fall into a type of 'logical palsy' or 'logical stupor' when the antinomies in it multiply. The contradictions with which the brain of man positively teems were, however, in the course of hundreds of thousands of years, gradually subjected to arbitrational procedures. There came to be levels higher and lower, levels of reflex and of reflection, impulse and control, the modeling of the elemental environment by zoological means and of the conceptual by linguistic means. All of these levels cannot, do not 'want' to tally perfectly or merge to form a whole. What, then, is consciousness? An expedient, a

dodge, a way out of the trap, a pretended last resort, a court allegedly (but only allegedly!) of highest appeal. And, in the language of physics and information theory, it is a function that, once begun, will not admit of any closure—ie, any definitive completion. It is, then, only a *plan* for such a closure, for a total 'reconciliation' of the stubborn contradictions of the brain. It is, one might say, a mirror whose task it is to reflect other mirrors, which in turn reflect still others, and so on to infinity. This, physically, is simply not possible, and so the *regressus ad infinitum* represents a kind of pit over which soars and flutters the phenomenon of human consciousness. 'Beneath the conscious' there goes on a continuous battle for full representation—in it—of that which cannot reach it in fullness, and cannot for simple lack of space; for, in order to give full and equal rights to all those tendencies that clamour for attention at the centres of awareness, what would be necessary is infinite capacity and volume. There reigns, then, around the conscious a never-ending crush, a pushing and shoving, and the conscious is not—not at all—the highest, serene, sovereign helmsman of all mental phenomena but more nearly a cork upon the fretful waves, a cork whose uppermost position does not mean the mastery of those waves....

The modern theory of consciousness, interpreted informationally and dynamically, unfortunately cannot be set forth simply or clearly, so that we are constantly—at least here, in this more accessible presentation of the subject—thrown back on a series of visual models and metaphors. We know, in any case, that consciousness is a kind of dodge, a shift to which evolution has resorted, and resorted in keeping with its characteristic and indispensable *modus operandi*, opportunism—ie, finding a quick, extempore way out of a tight corner. If, then, one were indeed to build an intelligent being and proceed according to the canons of completely rational engineering and logic, applying the criteria of technological efficiency, such a being would not, in general, receive the gift of consciousness. It would behave in a manner perfectly logical, always consistent, lucid, and well ordered, and it might even seem, to a human observer, a genius in creative action and decision-making. But it could in no way be a man, for it would be bereft of his mysterious depth, his internal intricacies, his labyrinthine nature....

We will not here go further into the modern theory of the conscious psyche, just as Professor Dobb does not. But these few words were in order, for they provide a necessary introduction to the structure of the personoids. In their creation is at last realized one of the oldest myths, that of the *homunculus*. In order to fashion a likeness of man, of his psyche, one must deliberately introduce into the informational substrate specific

contradictions; one must impart to it an asymmetry, acentric tendencies; one must, in a word, both *unify* and *make discordant*. Is this rational? Yes, and well-nigh unavoidable if we desire not merely to construct some sort of synthetic intelligence but to imitate the thought and, with it, the personality of man. Hence, the emotions of the personoids must to some extent be at odds with their reason; they must possess self-destructive tendencies, at least to a certain degree; they must feel internal tensions—that entire centrifugality which we experience now as the magnificent infinity of spiritual states and now as their unendurably painful disjointedness. The creational prescription for this, meanwhile, is not at all so hopelessly complicated as it might appear. It is simply that the logic of the creation (the personoid) must be disturbed, must contain certain antinomies.

Consciousness is not only a way out of the evolutionary impasse, says Hilbrandt, but also an escape from the snares of Godelization, for by means of paralogistic contradictions this solution has sidestepped the contradictions to which every system that is perfect with respect to logic is subject. So, then, the universum of the personoids is fully rational, but they are not fully rational inhabitants of it. Let that suffice us—Professor Dobb himself does not pursue further this exceedingly difficult topic.

As we know already, the personoids have souls but no bodies and, therefore, also no sensation of their corporeality. "It is difficult to imagine", has been said of that which is experienced in certain special states of mind, in total darkness, with the greatest possible reduction in the inflow of external stimuli—but, Dobb maintains, this is a misleading image. For with sensory deprivation the function of the human brain soon begins to disintegrate; without a stream of impulses from the outside world the psyche manifests a tendency to lysis. But personoids, who have no physical senses, hardly disintegrate, because what gives them cohesion is their mathematical milieu, which they do experience. But how? They experience it, let us say, according to those changes in their own states which are induced and imposed upon them by the universum's 'external-ness'. They are able to discriminate between the changes proceeding from outside themselves and the changes that surface from the depths of their own psyche. How do they discriminate? To this question only the theory of the dynamic structure of personoids can supply a direct answer. And yet they are like us, for all the awesome differences.

We know already that a digital machine can never spark with consciousness; regardless of the task to which we harness it, or of the physical processes we simulate in it, it will remain

forever apsychic. Since, to simulate man, it is necessary that we reproduce certain of his fundamental contradictions, only a system of mutually gravitating antagonisms—a personoid—will resemble, in the words of Canyon, whom Dobb cites, a "star contracted by the forces of gravity and at the same time expanded by the pressure of radiation". The gravitational center is, very simply, the personal 'I,' but by no means does it constitute a unity in either the logical or the physical sense. That is only our subjective illusion! We find ourselves, at this stage of the exposition, amid a multitude of astounding surprises.

One can, to be sure, program a digital machine in such a way as to be able to carry on a conversation with it, as if with an intelligent partner. The machine will employ, as the need arises, the pronoun 'I' and all its grammatical inflections. This, however, is a hoax! The machine will still be closer to a billion chattering parrots—howsoever brilliantly trained the parrots be—than to the simplest, most stupid man. It mimics the behavior of a man on the purely linguistic plane and nothing more. Nothing will amuse such a machine, or surprise it, or confuse it, or alarm it, or distress it, because it is psychologically and individually No One. It is a Voice giving utterance to matters, supplying answers to questions; it is a Logic capable of defeating the best chess player; it is—or, rather, it can become—a consummate imitator of everything, an actor, if you will, brought to the pinnacle of perfection, performing any programmed role—but an actor and an imitator that is, within, completely empty. One cannot count on its sympathy, or on its antipathy. It works toward no self-set goal; to a degree eternally beyond the conception of any man it 'doesn't care,' for as a person it simply does not exist.... It is a wondrously efficient combinatorial mechanism, nothing more.

Now, we are faced with a most remarkable phenomenon. The thought is staggering that from the raw material of so utterly vacant and so perfectly impersonal a machine it is possible, through the feeding into it of a special program—a personetic program—to create authentic sentient beings, and even a great many of them at a time! The latest IBM models have a top capacity of one thousand personoids. (The number is mathematically precise, since the elements and linkages needed to carry one personoid can be expressed in units of centimeters-grams-seconds.) Personoids are separated one from another within the machine. They do not ordinarily 'overlap,' though it can happen. Upon contact, there occurs what is equivalent to repulsion, which impedes mutual 'osmosis'. Nevertheless, they are able to interpenetrate if such is their aim. The processes making up their mental substrates then commence to superimpose upon each other, producing 'noise' and

interference. When the area of permeation is thin, a certain amount of information becomes the common property of both partially coincident personoids—a phenomenon that is for them peculiar, as for a man it would be peculiar, if not indeed alarming, to hear 'strange voices' and 'foreign thoughts' in his own head (which does, of course, occur in certain mental illnesses or under the influence of hallucinogenic drugs). It is as though two people were to have not merely the same, but *the same* memory; as though there had occurred something more than a telepathic transference of thought—namely, a 'peripheral merging of the egos'.

The phenomenon is ominous in its consequences, however, and ought to be avoided. For, following the transitional state of surface osmosis, the 'advancing' personoid can destroy the other and consume it. The latter, in that case, simply undergoes absorption, annihilation—it ceases to exist (this has already been called murder). The annihilated personoid becomes an assimilated, indistinguishable part of the 'aggressor'. We have succeeded—says Dobb—in simulating not only psychic life but also its imperilment and obliteration. Thus we have succeeded in simulating death as well. Under normal experimental conditions, however, personoids eschew such acts of aggression. 'Psychophagi' (Castler's term) are hardly ever encountered among them. Feeling the beginnings of osmosis, which may come about as the result of purely accidental approaches and fluctuations—feeling this threat in a manner that is of course nonphysical, much as someone might sense another's presence or even hear 'strange voices' in his own mind—the personoids execute active avoidance maneuvers; they withdraw and go their separate ways.

It is on account of this phenomenon that they have come to know the meaning of the concepts of 'good' and 'evil'. To them it is evident that 'evil' lies in the destruction of another, and 'good' in another's deliverance. At the same time, the 'evil' of one may be the 'good' (ie, the gain, now in the nonethical sense) of another, who would become a 'psychophage'. For such expansion—the appropriation of someone else's 'intellectual territory'—increases one's initially given mental 'acreage'. In a way, this is a counterpart of a practice of ours, for as carnivores we kill and feed on our victims. The personoids, though, are not obliged to behave thus; they are merely able to.

Hunger and thirst are unknown to them, since a continuous influx of energy sustains them—an energy whose source they need not concern themselves with (just as we need not go to

any particular lengths to have the sun shine down on us). In the personoid world the terms and principles of thermodynamics, in their application to energetics, cannot arise, because that world is subject to mathematical and not thermodynamic laws.

Before long, the experimenters came to the conclusion that contacts between personoid and man, via the inputs and outputs of the computer, were of little scientific value and, moreover, produced moral dilemmas, which contributed to the labeling of personetics as the cruelest science. There is something unworthy in informing personoids that we have created them in enclosures that only *simulate* infinity, that they are microscopic 'psychocysts,' capsules in our world. To be sure, they have their own infinity; hence Sharker and other psychoneticians (Falk, Wiegand) claim that the situation is fully symmetrical: the personoids do not need our world, our 'living space,' just as we have no use for their 'mathematical earth'.

Dobb considers such reasoning sophistry, because as to who created whom, and who confined whom existentially, there can be no argument. Dobb himself belongs to that group which advocates the principle of absolute nonintervention—'non-contact'—with the personoids. They are the behaviorists of personetics. Their desire is to observe synthetic beings of intelligence, to listen in on their speech and thoughts, to record their actions and their pursuits, but never to interfere with these. This method is already developed and has a technology of its own—a set of instruments whose procurement presented difficulties that seemed all but insurmountable only a few years ago. The idea is to hear, to understand—in short, to be a constantly eavesdropping witness—but at the same time to prevent one's 'monitorings' from disturbing in any way the world of the personoids.

Now in the planning stage at MIT are programs (APHRON II and EROT) that will enable the personoids—who are currently without gender—to have 'erotic contacts,' make possible what corresponds to fertilization, and give them the opportunity to multiply 'sexually'. Dobb makes clear that he is no enthusiast of these American projects. His work, as described in *Non Serviam*, is aimed in an altogether different direction. Not without reason has the English school of personetics been called 'the philosophical Polygon' and 'the theodicy lab'.

With these descriptions we come to what is probably the most significant and, certainly, the most intriguing part of the book under discussion—the last part, which justifies and

explains its peculiar title. Dobb gives an account of his own experiment, in progress now for eight years without interruption. Of the creation itself he makes only brief mention; it was a fairly ordinary duplicating of functions typical of the program JAHVE VI, with slight modifications. He summarizes the results of 'tapping' this world, which he himself created and whose development he continues to follow. He considers this tapping to be unethical, and even, at times, a shameful practice. Nevertheless, he carries on with his work, professing a belief in the necessity, for science, of conducting such experiments also—experiments that can in no way be justified on moral—or, for that matter, on any other non-knowledge-advancing—grounds. The situation, he says, has come to the point where the old evasions of the scientists will not do. One cannot affect a fine neutrality and conjure away an uneasy conscience by using, for example, the rationalization worked out by vivisectionists—that it is not in creatures of full-dimensional consciousness, not in sovereign beings that one is causing suffering or only discomfort.

In the personoid experiments we are accountable twofold, because we create and then enchain the creation in the schema of our laboratory procedures. Whatever we do and however we explain our action, there is no longer an escape from full accountability. Many years of experience on the part of Dobb and his coworkers at Oldport went into the making of their eight-dimensional universum, which became the residence of personoids bearing the names ADAN, ADNA, ANAD, DANA, DAAN, and NAAD. The first personoids developed the rudiment of language implanted in them and had 'progeny' by means of division. Dobb writes, in the Biblical vein, "And ADAN begat ADNA. ADNA in turn begat DAAN, and DAAN brought forth EDAN, who bore EDNA....". And so it went, until the number of succeeding generations had reached three hundred; because the computer possessed a capacity of only one hundred personoid entities, however, there were periodic eliminations of the 'demographic surplus'. In the three-hundredth generation, personoids named ADAN, ADNA, ANAD, DANA, DAAN, and NAAD again make an appearance, endowed with additional numbers designating their order of descent. (For simplicity in our recapitulation, we will omit the numbers.) Dobb tells us that the time that has elapsed inside the computer universum works out to—in a rough conversion to our equivalent units of measurement—from 2 to 2.5 thousand years.

Over this period there has come into being, within the personoid population, a whole series of varying explanations of their lot, as well as the formulation by them of varying, and contending, and mutually excluding models of 'all that exists'. That is, there have arisen

many different philosophies (ontologies and epistemologies), and also 'metaphysical experiments' of a type all their own. We do not know whether it is because the 'culture' of the personoids is too unlike the human or whether the experiment has been of too short duration, but, in the population studied, no faith of a form completely dogmatized has ever crystallized—a faith that would correspond to Buddhism, say, or to Christianity. On the other hand, one notes, as early as the eighth generation, the appearance of the notion of a Creator, envisioned personally and monotheistically.

The experiment consists in alternately raising the rate of computer transformations to the maximum and slowing it down (once a year, more or less) to make direct monitoring possible. These changes in rate are, as Dobb explains, totally imperceptible to the inhabitants of the computer universum, just as similar transformations would be imperceptible to us, because when at a single blow the whole of existence undergoes a change (here, in the dimension of time), those immersed in it cannot be aware of the change, because they have no fixed point, or frame of reference, by which to determine that it is taking place. The utilization of 'two chronological gears' permitted that which Dobb most wanted—the emergence of a personoid history, a history with a depth of tradition and a vista of time.

To summarize all the data of that history recorded by Dobb, often of a sensational nature, is not possible. We will confine ourselves, then, to the passages from which came the idea that is reflected in the book's title. The language employed by the personoids is a recent transformation of the standard English whose lexicon and syntax were programmed into them in the first generation. Dobb translates it into essentially normal English but leaves intact a few expressions coined by the personoid population. Among these are the terms 'godly' and 'ungodly,' used to describe believers in God and atheists.

ADAN discourses with DAAN and ADNA (personoids themselves do not use these names, which are purely a pragmatic contrivance on the part of the observers, to facilitate the recording of the 'dialogues') upon a problem known to us also—a problem that in our history originates with Pascal but in the history of the personoids was the discovery of a certain EDAN 197. Exactly like Pascal, this thinker stated that a belief in God is in any case more profitable than unbelief, because if truth is on the side of the 'ungodlies' the believer loses nothing but his life when he leaves the world, whereas if God exists he gains all eternity (glory everlasting). Therefore, one should believe in God, for this is dictated very

simply by the existential tactic of weighing one's chances in the pursuit of optimal success.

ADAN 300 holds the following view of this directive: EDAN 197, in his line of reasoning, assumes a God that requires reverence, love, and total devotion, and not only and not simply a belief in the fact that He exists and that He created the world. It is not enough to assent to the hypothesis of God the Maker of the World in order to win one's salvation; one must in addition be grateful to that Maker for the act of creation, and divine His will, and do it. In short, one must serve God.

Now, God, if He exists, has the power to prove His own existence in a manner at least as convincing as the manner in which what can be directly perceived testifies to His being. Surely, we cannot doubt that certain objects exist and that our world is composed of them. At the most, one might harbour doubts regarding the question of what it is they do to exist, how they exist, etc. But the fact itself of their existence no one will gainsay. God could with this same force provide evidence of His own existence. Yet He has not done so, condemning us to obtain, on that score, knowledge that is roundabout, indirect, expressed in the form of various conjectures—conjectures sometimes given the name of revelation. If He has acted thus, then He has thereby put the 'godlies' and the 'ungodlies' on an equal footing; He has not compelled His creatures to an absolute belief in His being but has only offered them that possibility. Granted, the motives that moved the Creator may well be hidden from His creations. Be that as it may, the following proposition arises: God either exists or He does not exist. That there might be a third possibility (God did exist but no longer does, or He exists intermittently, in oscillation, or He exists sometimes 'less' and sometimes 'more,' etc.) appears exceedingly improbable. It cannot be ruled out, but the introduction of a multivalent logic into a theodicy serves only to muddle it. So, then, God either is or He is not. If He Himself accepts our situation, in which each member of the alternative in question has arguments to support it—for the 'godlies' prove the existence of the Creator and the 'ungodlies' disprove it—then from the point of view of logic we have a game whose partners are, on one side, the full set of the 'godlies' and 'ungodlies,' and, on the other, God alone.

The game necessarily possesses the logical feature that for unbelief in Him God may not punish anyone. If it is definitely unknown whether or not a thing exists—some merely asserting that it does and others, that it does not—and if in general it is possible to advance the hypothesis that the thing never was at all, then no just tribunal can pass judgment

against anyone for denying the existence of this thing. For in all worlds it is thus: when there is no full certainty, there is no full accountability. This formulation is by pure logic unassailable, because it sets up a symmetrical function of reward in the context of the theory of games; whoever in the face of uncertainty demands full accountability destroys the mathematical symmetry of the game; we then have the so-called game of the non-zero sum. It is therefore thus: either God is perfectly just, in which case He cannot assume the right to punish the 'ungodlies' by virtue of the fact that they are 'ungodlies' (ie, that they do not believe in Him); or else He will punish the unbelievers after all, which means that from the logical point of view He is not perfectly just. What follows from this? What follows is that He can do whatever He pleases, for when in a system of logic a single, solitary contradiction is permitted, then by the principle of *ex falso quodlibet* one can draw from that system whatever conclusion one will. In other words: a just God may not touch a hair on the head of the 'ungodlies,' and if He does, then by that very act He is not the universally perfect and just being that the theodicy posits.

ADNA asks how, in this light, we are to view the problem of the doing of evil unto others. ADAN 300 replies: whatever takes place here is entirely certain; whatever takes place 'there'—ie, beyond the world's pale, in eternity, with God—is uncertain, being but inferred according to the hypotheses. Here, one should not commit evil, despite the fact that the principle of eschewing evil is not logically demonstrable. But by the same token the existence of the world is not logically demonstrable. The world exists, though it could not exist. Evil may be committed, but one should not do so, and should not, I believe, because of our agreement based upon the rule of reciprocity: be to me as I am to thee. It has naught to do with the existence or the nonexistence of God. Were I to refrain from committing evil in the expectation that 'there' I would be punished for committing it, or were I to perform good, counting upon a reward 'there,' I would be predicating my behaviour on uncertain ground. Here, however, there can be no ground more certain than our mutual agreement in this matter. If there be, 'there,' other grounds, I do not have knowledge of them as exact as the knowledge I have, here, of ours. Living, we play the game of life, and in it we are allies, every one. Therewith, the game between us is perfectly symmetrical. In postulating God, we postulate a continuation of the game beyond the world. I believe that one should be allowed to postulate this continuation of the game, so long as it does not in any way influence the course of the game here. Otherwise, for the sake of someone who perhaps does not exist we may well be sacrificing that which exists here, and exists for certain.

NAAD remarks that the attitude of ADAN 300 toward God is not clear to him. ADAN has granted, has he not, the possibility of the existence of the Creator: what follows from it?

ADAN: Not a thing. That is, nothing in the province of obligation. I believe that—again for all worlds—the following principle holds: a temporal ethics is always independent of an ethics that is transcendental. This means that an ethics of the here and now can have outside itself no sanction which would substantiate it. And this means that he who does evil is in every case a scoundrel, just as he who does good is in every case righteous.

If someone is prepared to serve God, judging the arguments in favour of His existence to be sufficient, he does not thereby acquire *here* any additional merit. It is his business. This principle rests on the assumption that if God is not, then He is not one whit, and if He is, then He is almighty. For, being almighty, He could create not only another world but likewise a logic different from the one that is the foundation of my reasoning. Within such another logic the hypothesis of a temporal ethics could be of necessity dependent upon a transcendental ethics. In that case, if not palpable proofs, then logical proofs would have compelling force and constrain one to accept the hypothesis of God under the threat of sinning against reason.

NAAD says that perhaps God does not wish a situation of such compulsion to believe in Him—a situation that would arise in a creation based on that other logic postulated by ADAN 300. To this the latter replies: an almighty God must also be all-knowing; absolute power is not something independent of absolute knowledge, because he who can do all but knows not what consequences will attend the bringing into play of his omnipotence is, *ipso facto*, no longer omnipotent; were God to work miracles now and then, as it is rumoured He does, it would put His perfection in a most dubious light, because a miracle is a violation of the autonomy of His own creation, a violent intervention. Yet he who has regulated the product of his creation and knows its behaviour from beginning to end has no need to violate that autonomy; if he does nevertheless violate it, remaining all-knowing, this means that he is not in the least correcting his handiwork (a correction can only mean, after all, an initial non-omniscience), but instead is providing—with the miracle—a sign of his existence.

Now, this is faulty logic, because the providing of any such sign must produce the impression that the creation is nevertheless improved in its local stumblings. For a logical

analysis of the new model yields the following: the creation undergoes corrections that do not proceed from it but come from without (from the transcendental, from God), and therefore miracle ought really to be made the norm; or, in other words, the creation ought to be so corrected and so perfected that miracles are at last no longer needed. For miracles, as *ad hoc* interventions, cannot be *merely* signs of God's existence: they always, after all, besides revealing their Author, indicate an addressee (being directed to someone *here* in a helpful way). So, then, with respect to logic it must be thus: either the creation is perfect, in which case miracles are unnecessary, or the miracles are necessary, in which case the creation is not perfect. (With miracle or without, one may correct only that which is somehow flawed, for a miracle that meddles with perfection will simply disturb it, more, worsen it.) Therefore, the signaling by miracle of one's own presence amounts to using the worst possible means, logically, of its manifestation.

NAAD asks if God may not actually want there to be a dichotomy between logic and belief in Him: perhaps the act of faith should be precisely a resignation of logic in favour of a total trust.

ADAN: Once we allow the logical reconstruction of something (a being, a theodicy, a theogony, and the like) to have internal self-contradiction, it obviously becomes possible to prove absolutely anything, whatever one pleases. Consider how the matter lies. We are speaking of creating someone and of endowing him with a particular logic, and then demanding that this same logic be offered up in sacrifice to a belief in the Maker of all things. If this model itself is to remain noncontradictory, it calls for the application, in the form of a metalogic, of a totally different type of reasoning from that which is natural to the logic of the one created. If that does not reveal the outright imperfection of the Creator, then it reveals a quality that I would call mathematical inelegance—a *sui generis* unmethodicalness (incoherence) of the creative act.

NAAD persists: perhaps God acts thus, desiring precisely to remain inscrutable to His creation—ie, non-reconstructible by the logic with which He has provided it. He demands, in short, the supremacy of faith over logic.

ADAN answers him: I follow you. This is, of course, possible, but even if such were the case, a faith that proves incompatible with logic presents an exceedingly unpleasant dilemma of a moral nature. For then it is necessary at some point in one's reasonings to

suspend them and give precedence to an unclear supposition—in other words, to set the supposition above logical certainty. This is to be done in the name of unlimited trust; we enter, here, into a *circulus vitiosus*, because the postulated existence of that in which it behooves one now to place one's trust is the product of a line of reasoning that was, in the first place, *logically correct*; thus arises a logical contradiction, which, for some, takes on a positive value and is called the Mystery of God. Now, from the purely constructional point of view such a solution is shoddy, and from the moral point of view questionable, because Mystery may satisfactorily be founded upon infinity (infiniteness, after all, is a characteristic of our world), but the maintaining and the reinforcing of it through internal paradox is, by any architectural criterion, perfidious. The advocates of theodicy are in general not aware that this is so, because to certain parts of their theodicy they continue to apply ordinary logic and to other parts, not. What I wish to say is this, that if one believes in contradiction,\* one should then believe *only* in contradiction, and not at the same time still in some noncontradiction (ie, in logic) in some other area. If, however, such a curious dualism is insisted upon (that the temporal is always subject to logic, the transcendental only fragmentarily), then one thereupon obtains a model of Creation as something that is, with regard to logical correctness, 'patched,' and it is no longer possible for one to postulate its perfection. One comes inescapably to the conclusion that perfection is a thing that must be logically patched.

EDNA asks whether the conjunction of these incoherencies might not be love.

ADAN: And even were this to be so, it can be not any form of love but only one such as is blinding. God, if He is, if He created the world, has permitted it to govern itself as it can and wishes. For the fact that God exists, no gratitude to Him is required; such gratitude assumes the prior determination that God is able not to exist, and that this would be bad—a premise that leads to yet another kind of contradiction. And what of gratitude for the act of creation? This is not due God, either. For it assumes a compulsion to believe that to be is definitely better than not to be; I cannot conceive how that, in turn, could be proven. To one who does not exist surely it is not possible to do either a service or an injury; and if the Creating One, in His omniscience, knows beforehand that the one created will be grateful to Him and love Him or that he will be ungrateful and deny Him, He thereby produces a constraint, albeit one not accessible to the direct comprehension of the one created. For this very reason nothing is due God: neither love nor hate, nor gratitude, nor rebuke, nor the hope of reward, nor the fear of retribution. Nothing is due Him. A God who craves such

feelings must first assure their feeling subject that He exists beyond all question. Love may be forced to rely on speculations as to the reciprocity it inspires; that is understandable. But a love forced to rely on speculations as to whether or not the beloved exists is nonsense. He who is almighty could have provided certainty. Since He did not provide it, if He exists, He must have deemed it unnecessary. Why unnecessary? One begins to suspect that maybe He is not almighty. A God not almighty would be deserving of feelings akin to pity, and indeed to love as well; but this, I think, none of our theodicies allow. And so we say: We serve ourselves and no one else.

We pass over the further deliberations on the topic of whether the God of the theodicy is more of a liberal or an autocrat; it is difficult to condense arguments that take up such a large part of the book. The discussions and deliberations that Dobb has recorded, sometimes in group colloquia of ADAN 300, NAAD, and other personoids, and sometimes in soliloquies (an experimenter is able to take down even a purely mental sequence by means of appropriate devices hooked into the computer network), constitute practically a third of *Non Serviam*. In the text itself we find no commentary on them. In Dobb's Afterword, however, we find this statement: 'ADAN's reasoning seems incontrovertible, at least insofar as it pertains to me: it was I, after all, who created him. In his theodicy I am the Creator. In point of fact, I produced that world (serial No. 47) with the aid of the ADONAI IX program and created the personoid gemmae with a modification of the program JAHVE VI. These initial entities gave rise to three hundred subsequent generations. In point of fact, I have not communicated to them—in the form of an axiom—either these data or my existence beyond the limits of their world. In point of fact, they arrived at the possibility of my existence only by inference, on the basis of conjecture and hypothesis. In point of fact, when I create intelligent beings, I do not feel myself entitled to demand of them any sort of privileges—love, gratitude, or even services of some kind or other. I can enlarge their world or reduce it, speed up its time or slow it down, alter the mode and means of their perception; I can liquidate them, divide them, multiply them, transform the very ontological foundation of their existence. I am thus omnipotent with respect to them, but, indeed, from this it does not follow that they owe me anything. As far as I am concerned, they are in no way beholden to me. It is true that I do not love them. Love does not enter into it at all, though I suppose some other experimenter might possibly entertain that feeling for his personoids. As I see it, this does not in the least change the situation—not in the least.

Imagine for a moment that I attach to my BIX 310 092 an enormous auxiliary unit, which will be a 'hereafter'. One by one I let pass through the connecting channel and into the unit the 'souls' of my personoids, and there I reward those who believed in me, who rendered homage unto me, who showed me gratitude and trust, while all the others, the 'ungodlies,' to use the personoid vocabulary, I punish—eg, by annihilation or else by torture. (Of eternal punishment I dare not even think—that much of a monster I am not!) My deed would undoubtedly be regarded as a piece of fantastically shameless egotism, as a low act of irrational vengeance—in sum, as the final villainy in a situation of total dominion over innocents. And these innocents will have against me the irrefutable evidence of logic, which is the aegis of their conduct.

Everyone has the right, obviously, to draw from the personetic experiments such conclusions as he considers fitting. Dr Ian Combay once said to me, in a private conversation, that I could, after all, assure the society of personoids of my existence. Now, this I most certainly shall not do. For it would have all the appearance to me of soliciting a sequel—that is, a reaction on their part. But what exactly could they do or say to me, that I would not feel the profound embarrassment, the painful sting of my position as their unfortunate Creator?

The bills for the electricity consumed have to be paid quarterly, and the moment is going to come when my university superiors demand the 'wrapping up' of the experiment—that is, the disconnecting of the machine, or, in other words, the end of the world. That moment I intend to put off as long as humanly possible. It is the only thing of which I am capable, but it is not anything I consider praiseworthy. It is, rather, what in common parlance is generally called 'dirty work'. Saying this, I hope that no one will get any ideas. But if he does, well, that is his business'.

---

\* *Credo quia absurdum est* (Prof Dobb's note in the text).