Book Reviews

THE NEW MODES
OF CONTRACTARIANISM:
A CRITICAL REVIEW

Morals By Agreement. By David Gauthier. Oxford: The University Press, 1986.

Persons, Rights, and the Moral Community. By Loren Lomasky. Oxford: The University Press, 1987.

Both of these books are seminal contributions to the sub-genre of contractarian social philosophy. Gauthier's represents the culmination and synthesis of over two decades of work scattered throughout many journals and anthologies, while Lomasky's work, not unlike Nozick's *Anarchy, State, and Utopia*, emerges in full form with little prior journal exposure. They also differ radically in style and method, although they share a common destination: a contractarian derivation of neo-liberal rights and institutions.

Before examining each work separately, it behooves us to see what they share in common. Both take radical exception to a prevailing meta-ethical doctrine, which Gauthier labels *universalism* and Lomasky terms *impartiality*. Gauthier contrasts the *universalistic* conception of rationality with what he calls the *maximizing* conception, the latter of which he endorses.

> On the maximizing conception it is not interests in the self, that take oneself as object, but interests of the self, held by oneself as subject, that provide the basis of rational choice and action. On the universalistic conception it is not interests in anyone, that take any person as object, but interests of anyone, held by some person as subject, that provide the basis for rational choice and action. If I have a direct interest in your welfare, then on either conception I have reason to promote your welfare. But your interest in your welfare affords me such reason only given the universalistic conception. (Gauthier, p. 7)

Similarly, Lomasky erects as what he terms 'the foil' a picture of impartiality as the hallmark of moral rationality:

> On this account, morality involves treating all persons alike, though of course not in the simple-minded sense of acting in precisely the same way toward

everyone. Rather, impartiality involves weighing all ends by the same standard
of value and striving to maximize the sum of this value irrespective of whose
particular interests turn out to be favored. Impartiality is not indifferent *to*
persons; it is indifferent *among* persons. (Lomasky, p. 23)

Both Gauthier and Lomasky, to their credit, attempt to offer an alternative
to this standard 'impartial universalism'. For both, it is the considered
preferences, or interests, of individuals that form the basis for (moral)
rationality. Both attempt to arrive at a neo-liberal social philosophy from
the starting point of quasi-Hobbesian individuals via contractarian methods.
The radically different styles and methods of the authors, however, makes
for difficult joint-review. It will be best if we look at each work separately,
in summary fashion, and then step back and compare them.

Gauthier is attempting, in his own words, "the rational reconstruction of
morality." By this he means an answer to the question, "What would rational,
reflective individuals in a 'state of nature' agree to, as they bargain their
way to basic rights, rules, and institutions?" He claims to demonstrate that
morality "can be generated as a rational constraint from the non-moral
premises of rational choice." (p. 4) Gauthier is out to derive a basically Lockean
world from Hobbesian actors and methods, refurbished and updated with
the formidable arsenal of contemporary game—and decision theory.

Gauthier begins with a fairly standard account of economic (practical)
rationality as the maximization of (subjective) utility. This can easily be
paraphrased into 'pursuit of (perceived) self-interest', but Gauthier resists
this, stressing instead merely that agents act to maximize the fulfillment of
considered preferences, whatever the content of these preferences. By
'considered' Gauthier means both informed *and* reflective; thus differing from
a strict doctrine of *revealed* preference, i.e., actual choice. He synthesizes
aspects of various philosophers: Moral principles are derivable and bind
(if at all) through *reason* rather than sentiment. [Kant] Reason is, however,
strictly instrumental [Hume] and preferences are both subjective and relative.
[Hobbes and Hume] Rationality *consists in* satisfying standard consistency
criteria (e.g. transitivity). But utility functions, being subjective and relative,
are essentially non-comparable. Rational agents act to maximize their own
expected utilities, regardless (often merely oblivous) of how this affects others'
utilities (any desire regarding *other* persons is already accounted for in the
agent's utility function).

The entire background problem of Social Contract, i.e., multiple interaction
situations among numerous individuals, is introduced in Chapter 3. In a
multi-person world, the perennial problems of social philosophy arise in
cases of strategic interaction where outcomes depend jointly on the choices
(actions, strategies) or *other* rational agents, i.e., where *each* agent must take
into account the actions (strategies) of others in making his own choices.
Allowing mixed strategies, there will always be *at least one* equilibrium, a
set of strategies from which no agent gains by defecting. Problems arise,
however, in Prisoner's Dilemma-type situations; all equilibria *may* be sub-
optimal in the sense of yielding all agents less utility than they would receive
from some other outcome. (The now-familiar example is an arms race in
which both sides would prefer mutual disarmament to mutual armament,
but each prefers to remain armed whether its opponent disarms or not.
Mutual armament is the only stable equilibrium, but it is sub-optimal.)

Before going on to (re)construct a rationale for cooperation, Gauthier includes a fascinating Chapter 4 on "The Market: Freedom from Morality." Gauthier argues that in perfectly competitive markets Prisoner's Dilemma situations cannot arise, and thus they constitute a "morally free zone" where most moral constraints are superfluous since each person acting in his own interest works to the mutual advantage of all. The only constraints implied in such contexts are those *presupposed* by market processes, i.e., no force or fraud. Problems arise, however, even in market contexts, in two ways: (1) *which* optimal outcome is arrived at crucially depends on the set of *initial endorsements* (possessions, capacities, etc.) [as has been pointed out ceaselessly by e.g., Hillel Steiner]; and (2) markets 'fail", i.e., real markets consist of networks of transactions which generate significant *externalities* (uncompensated effects, on third parties).

Prisoner's Dilemma situations and market failure together give rise to problems of how and on what terms purely rational agents will achieve cooperation. For Gauthier, the core of a rational social morality consists of two essentially Lockean notions: no parasites and no free-riders. The former is any displacement of the costs of one's activities onto others, as when Upstream dumps toxic waste into the river, thus polluting Downstream's drinking water. The latter is any enjoyments of the benefits of cooperation that provides public goods that are worth their cost without being willing to pay one's fair share of those costs. The common principle behind these notions is: so far as possible, moral rules should be such that if the rules are obeyed, the acts (and ensuing benefits or harms) rebound only on the agent himself. Gauthier identifies, and extensively discusses, three principles which flesh out this Lockean morality: (1) constrained maximization; (2) minimax relative concession; and (2) the Lockean Proviso. This latter, as might be expected, is simply the pre-bargaining stipulation that no one can better his position through interactions that *worsen* another. This stipulation is obviously normative, yet pre-contractual (pre-bargaining), which if not a contradiction is at least a paradox, considering that Gauthier's whole project is to generate *moral* principles from *non-moral* principles. What is the rationale for the Proviso? "Without limitations that exclude the taking of advantage, a rational individual would not dispose himself" to comply with cooperative agreements. (p. 255) This seems circular at best, since Gauthier's program is the *generation*, not postulation, of normative principles. To say, in effect, "but my program won't get off the ground without it (the Proviso)" is lame, if not question-begging.

Granting the Proviso, we can now consider Gauthier's solutions to the bargaining problem, i.e., how to 'divvy up' the 'co-operative surplus.' Gauthier sees the problem of just social principles as essentially how to allocate shares of the cooperative surplus generated by interaction among individuals, given the existence of Prisoner's Dilemma situations and externalities. He argues that rational agents would reach the following solution to this problem. For each agent there is a no-agreement utility level corresponding to the initial bargaining position—this is what the agent could expect to obtain *without* cooperative interaction (if the Proviso is satisfied). An agent's *maximal claim* is the highest utility it is possible for the agent to receive while all others receive at least their 'no-agreement' utilities, i.e., the outcome that channels all benefits from cooperation to her (the agent) such that it is just marginally the case that no other cooperator or coalition would do better either withdrawing or excluding her from cooperation. An agent's *concession* at a

given outcome is the difference between her utility at the outcome and her maximal claim. Her *relative concession* is the ratio of her concession to the utility difference between her maximal claim and her no-agreement level. Rational bargainers will agree on that outcome which *minimizes* the largest relative concession that has to be made by someone—this is Gauthier's principle of minimax relative concession, or equivalently, maximum relative benefit.

There are several problems with each of these three principles, but space being at a premium I will adduce just a few. Regarding the Proviso, Gauthier's ultimate rationale for it seems to be this: the prospect of the emergence of society (and hence of an ever-growing cooperative surplus) makes adoption of the Proviso rational for all person. While this effectively rebuts my earlier criticism, in and of itself *this* rationale is questionable, and nowhere in the book does Gauthier argue for it. Regarding the constrained maximization, why should we suppose that Gauthier's Hobbesian egoists will adopt it? Gauthier does argue extensively for this, but, as usual, there are some questions.

Chapter 6 introduces the notions of constrained maximization, along with the notions of straightforward maximization and broad and narrow compliance. A constrained maximizer is one who acts exactly as would a straightforward expected utility maximizer except that the former is ready to cooperate if (1) the utility she can expect if others also cooperate is not less than she could expect if everyone acted as a straightforward maximizer, and (2) the utility she can expect if others also cooperate approaches the outcome determined by minimax relative concession.

Gauthier's argument for constrained maximization is essentially in the form of posing a decision problem: Does the *disposition* toward constrained or straightforward maximization (as a standard of practical reason) yield a higher expected utility pay-off for the agent (within a relevant social context)? Gauthier argues that the choice of constrained maximization actually does better in utility terms. Thus, the devil is out-foxed. Or is he?

Assume an environment containing both constrained maximizers and straightforward maximizers. Assume also that all persons are 'translucent', i.e., all persons in this environment can be judged by any other agent to be either a constrained maximizer or a straightforward maximizer with better-than-random probability of being correct. A disposition toward constrained maximization is rational only in an environment *already* dominated by constrained maximizers. What this demonstrates is the rather weak conclusion that a group of constrained maximizers will fare better than a group of straightforward maximizers. But the question is, How, in an environment of iterated Prisoner's Dilemma situations, do straightforward maximizers ever 'switch' to becoming constrained maximizers? An evolutionary learning model (such as Axelrod's) might be invoked to show that a tendency exists for the constrained maximization disposition to arise, and thus vindicate Gauthier's position that constrained maximization as a rule or policy is rational. But it hardly follows, as Gauthier nonetheless seems to believe, that *particular acts* of constrained maximization are rational.

In sum, Gauthier's book is the most ambitious attempt to date to ground social morality in something more rigorous than a set of coherentists' equilibrating reflections. As such, it should be studied, as should the surrounding literature it will no doubt generate. Gauthier wields game and decision theory with aplomb, and provides excellent expositions of more technical points. He also applies his theories and ideas to such questions

as inter-generational justice and the nature of a liberal individual. It is, without doubt, one of the most important contributions to contractarian theory since Rawls' *Theory of Justice.*

Radically different in style and method is Loren Lomasky's *Persons, Rights, and the Moral Community.* As its title indicates, Lomasky ranges over quite a lot of territory which is of contemporary importance, including many topics thought intractable. Lomasky begins by criticizing the contemporary 'rights explosion,' the 'heady proliferation' of rights to various and sundry things. He correctly points out that the term "right' is a very special and powerful term. "Rights stake out chunks of moral turf that others are forewarned not to trespass; they issue demands with which other *must* (the 'must' is moral, not causal or logical) comply." Hence, we should be at least cautious, if not skeptical, about many contemporary rights-claims.

Skepticism regarding specific rights is not to carry over to the idea of rights as such, however. Unless one is prepared to embrace ethical skepticism generally, rights have their proper (and fundamental) place in the scheme of things ethico-political. "Even if a full-blown theory of rights that is massively supported by our best moral reasoning were conveniently at hand, it does not follow that we could thenceforth eschew talk of rights in favor of reference to the underlying theory." (Lomasky, p. 12) Indeed, why should we abandon the shorthand but powerful terminology of rights? "To insist that this cumbersome machinery [of a full-fledged ethics] be hauled out in its entirety each time one ventures into normative analysis is as perverse as the insistence that all talk of 'electrical charge' or 'cold front' be accompanied by the respective physical or meteorological theories within which they function." (pp. 12- 13)

Having established (1) rights terminology cannot be dispensed with, but (2) the concept or rights is such that its domain is *limited,* Lomasky goes on to tackle the tough questions, some plausible answers to which must (logically) be offered to substantiate any theory of basic rights. Indeed, Lomasky understandably (though merely passingly) criticizes Nozick, Dworkin, and even Rawls for not offering such a foundation. What then is Lomasky's foundation; what is his starting point? He builds implicitly on the work of Gewirth, Williams, Norton, and others, but his foundation for rights is nonetheless original and plausible. It is this: Individuals are *project pursuers;* they have unique values and commitments, and they have reason to value those ends that are distinctively theirs in a way no one with different projects does.

While this idea (a modified Bernard Williams notion) has much in common with the groundwork for basic rights formulated by Machan, Mack, Pilon, and Pollock (among others), Lomasky only occasionally alludes to these authors. There are elements in Lomasky's thought which have obvious affinities to these and other philosophers, and it is worth noting a few of these influences, as well as how Lomasky modifies or diverges from them. Lomasky is sympathetic for example to Gewirth's attempt to ground basic ethics (and, fortiori, basic rights) in the fact that human beings are *agents,* i.e., self-directing and conative beings. The "bare fact of agency," however, is insufficiently robust to sustain a coherent set of rights, according to Lomasky. Thus, he borrows Bernard Williams' instructive emphasis on *projects,* i.e., regulative ends which comprise or extend throughout individuals' lives, play a central role in the person's on-going activities, and provide structural stability to the individual's life. Project pursuit implies that persons require some degree

of non-interference in regard to their on-going activities which constitute their lives. Recognition of basic rights in the form of side constraints safeguards individualism as utilitarian morality cannot.

Lomasky, as noticed above, constructs an alternative to what he calls "the foil," i.e., the (meta) ethical view, modeled on *prudential* decision-making, whereby the ethical decision-procedure involves being indifferent or impartial among persons (impartial regarding values and the projects associated with them), and is based on an *impersonal* standard of value. Interestingly, Gauthier, as noted above, provides a very good characterization of this meta-ethical model. "It is not interests in anyone, that take any person as an object, but interests of anyone, held by some person(s) as subject, that provide the basis for rational choice and action." This 'foil' is at odds with project pursuit. Whereas the key element in the foil's conception of moral reasoning is (a specific and not uncontroversial notion of *impartiality*, project pursuit ineluctably implies a kind of partiality and personalism. This is *not* to say that individuals totally lack the ability to empathize, or that

> project pursuit insulates one from all reason to consider the well-being of others and to take another person's good as providing reasons to bear on one's choice of conduct.... It would be entirely fallacious to conclude that persons enjoy carte-blanche to engage in completely selfish behavior, mindless of the well-being of others. Concern for the personal dimension of morality is not equivalent to the endorsement of egoistic rapacity. (Lomasky, pp. 30, 35)

In contrast to "the Foil," then, Lomasky is concerned to defend a sort of (meta) ethical individualism in which the human being, as the subject of conation and conative activities, is given its due.

How, then, do we derive rights—or, in general, *any* other furniture of social morality—from such a context? While Lomasky's persons are not Hobbesian straw men ("rapacious egoists"), they are not merely partially differentiated integers, "partners in the human enterprise to which all...efforts must be devoted. There is no such thing; there are only the various personal enterprises in which individuals enroll themselves and which provide them with irreducibly personal ends that they strive to realize." Rather there is middle course between the Charybdis of "the Foil" and the Scylla of Hobbesian atomism: moral community is possible because project pursuit almost inescapably involves essential reference to the well-being of others. Hence, a theory of basic rights as side constraints is *prima facie* plausible.

After a Chapter 3 defending project pursuit as a basis for (social) ethics against various possible criticism, in which he discusses the idea that some (though by no means all) value is posterior to choice and therefore personal, Lomasky tackles the rough core question: "How can one go beyond the bare recognition of others as project pursuers to a rational motivation to respect them as project pursuers?" (p. 62) Granted that individuals, as project pursuers, have reason to value their *ability-to-pursue-projects*, what is the motivation to universalize this in the inter-personal realm, and thus establish rights as inviolable constraints? Lomasky suggests a combination of three reasons.

First, there is what could be called the sociobiological or species-solidarity argument. Human beings are *not* organisms who totally lack empathy for their fellow human beings. "Rather, human beings are social animals whose survival is predicated upon their being the beneficiaries of altruistic concern

of limited yet crucial scope." This quasi-Aristotelian, sociobiological explanation of intra-species empathy and altruism does not go far (enough) in establishing interpersonal, inviolable side-constraints.

The second argument is one suggested by Thomas Nagel, and is complementary to the first. Nagel argues that the ability to recognize oneself as one person in a world containing other persons is logically sufficient to provide the basis for transmission of rational motivation. A's having end E' obviously provides B *at least some minimal reason* to act to advance E'. "Some reason" because value is *not* completely impersonal, and A's reason for promoting E' is different and stronger than B's reason. In other words, the recognition by B of a reason R as a reason for E' implies that B is not totally indifferent to E' obtaining. R is why E' should obtain. B's acknowledging that A has reason (understood personally) to bring about E' implies that there is (impersonal) reason to bring about E'; thus, B has some reason to advance A's pursuing E', only if B's commitment to his own project entails pursuing E', which is incompatible with E', does B not have reason on balance to promote E'. Thus, a bridge is provided between someone's *having a reason* and there *being a reason*, while not conflating the two.

This second argument, by itself, is also insufficient to generate a robust theory of rights. Lomasky offers a third argument. He asks us to imagine a world (not unlike the actual world) in which each person has reason to undertake activity to eliminate interference by others. "Because each project pursuer values his own ability to be a project pursuer, each has reason to act to bring about circumstances in which he will be able to lead a coherent life responsive to his own conception of the good." (p. 65)

But what strategy merits adoption? Lomasky adduces three possible strategies and illustrates them by means of a game-matrix. (pp. 65-75) Essentially, the possible strategies are: (1) "Active Aggression," the deliberate and systematic attempt to *remove* any and all other's ability to interfere with and thus impede one's own designs; (2) "General Neglect," wherein each person goes about his business, oblivious of others except when their activities clash, at which time they compete for success; and (3) "Active Deference," which is the deliberate and systematic attempt to avoid interference with and thus impeding the other's designs. Lomasky demonstrates that, for pure Hobbesian egoists, while mutual Active Deference has the highest joint-value payoff, with mutual General Neglect having the second highest joint-value payoff, both are unstable and will tend to degenerate to mutual Active Aggression, i.e., Hobbes' "war of all against all."

This situation is, of course, the now classic Prisoner's Dilemma. Lomasky, however, argues that we need not start with what I term reductive egoists (Hobbesian rapacious brutes). The first two arguments above suggest that we can reject this model of the human agents in favor of a more complex (but for all that more realistic) model, wherein human beings, while nonetheless self-interested expected utility maximizers, are also capable of empathy and even limited altruism. From such a revised starting point, Lomasky provides a cogent evolutionary-learning argument (based on what amounts to iterated extended Prisoner's Dilemma-type situations) to the effect that the persons will, by "invisible hand" processes, adopt a stable regime of mutual Active Deference. It is (at least partially) an *invisible hand* emergence because, while both A and B prefer deference from the other and have some initial willingness to defer, the resultant stable mutual Active Deference

equilibrium is not specifically aimed at by either person. If the solution is extended to a multi-person context, the invisible hand may or may not lead to a multilateral equilibrium of Active Deference, but Lomasky argues, such an equilibrium is the only *moral* equilibrium.

Lomasky considers next the amount and character of deference which should be the case in a moral community. One can imagine a 'fanatic' who prefers a state of mutual Active Deference. While being able to do little regarding the fanatic, Lomasky suggests the following principle which minimizes the number of persons who find the deference they receive from others insufficient compensation for the deference they must supply. "Generally: a stable regime of equal rights for all requires that the amount of required deference be close to the level of deference ideal for the least deferential members of the community."

Chapter 5 discusses the "two concepts of liberalism," classical and welfarist. The former (of which contemporary libertarianism is a radical sub-species) maintains that all (or very nearly all) the rights individuals possess are negative in character, i.e., entail non-interference or forbearance. The latter maintains that individuals have, in addition to negative rights, positive rights, i.e., rights entailing the provision by some individual or institution of a valued item(s). Lomasky reviews several arguments in favor of a welfarist interpretation of liberalism, including the argument from need, and concludes that while liberty-rights must be accorded overall primacy, there is still room, in extreme contexts, for the recognition of welfare-rights.

Chapter 6 provides a discussion of property rights within a liberal order that, while one of the most important and well-argued chapters in the book, is difficult to summarily discuss. I will only hit the highlights. Lomasky adduces a variation of the Lockean Proviso which is part of the basis for his theory of justice in property holdings: "Each person has reason to value a liberty to acquire and use goods equal to the liberty to acquire and use that which every other project pursuer enjoys. Not equality of holdings but equal liberty to acquire holdings is entailed by the normative theory of basic rights." (p. 123) Lomasky also argues that those in exigent straits may demand requisite welfare goods as a matter of right. (p. 126) His reasoning against the strict libertarian disconnects the theory of basic rights from its foundation project pursuer's practical reason. The entire rationale and system of rights is put in jeopardy should it be the case that in certain dire contexts one can either continue to respect others' rights or be able to pursue projects, *but not both.*

The most innovative aspect of Lomasky's book is his application of his theory to children. In the case of children, rights ascriptions are based on the fact that they are potential (and would-be) project-pursuers, and on the fact that they are related to actual project pursuers in bonds of recognition that make their good a concern for others. While the biological family holds the *primary* responsibility regarding the nurture of a child, this does not entail that exclusive responsibility lies therein. The rights of a child, both positive and negative, impose obligations on everyone. After discussing state education, and persuasively arguing for *laissez-faire* in educational matters, Lomasky turns to what I call the Indoctrination Problem, i.e., the potential for abuse such as to turn out not independently thinking and evaluating persons, but "evaluational clones' of the parent/educator, Beings who lack individuality or self-determination. Lomasky cautiously concludes that the burden of proof for interference with familial attempts to inculcate specific ideas or loyalties in children lies with those who propose interference.

In his penultimate chapter "At the Margin," Lomasky further puts his theories to the test by considering, in turn, fetuses and abortion, the mentally defective, the dead, and animals. I will briefly deal with his treatment of the first and last of these. Abortion is a controversial, and almost intractable, issue because various intuitions and established conventions collide here. The fetus *is* living, it *is* human, but is it a *person?* Does it have *rights?* Lomasky argues that even the issue of parenthood is not so important as the question of recognition and individuation. Infants are vastly more individuatedly recognizable and stand in vastly more individuating social relationships than do fetuses. Thus "it is reasonable *as a general policy* to recognize in infants, but not in fetuses, full standing as rights holders." The weak conclusion to be drawn is that abortion is *permissible,* though not necessarily *right.*

Lomasky rejects animal rights for the very straightforward reason that no animal qualifies as a project-pursuer. But to deny animal rights is *not* to deny animals moral standing. Most, if not all, of what can be said on the moral behalf of animals ca be said without resort to rights. That animals are sentient, minimally conative, etc., implies moral consideration. Animals can feel pain, for example, and therefore it is morally wrong to inflict pain on them. And it is precisely *here* that I come to a slight disagreement with Lomasky, for while I am not of the same persuasion as, say, Tom Regan or Peter Singer, I nonetheless hold that it is coherent and even accurate to speak of certain contextual rights for animals, especially higher-order animals: an animal has a (near-) absolute right not to be tortured, for instance, by virtue of the fact that they would suffer (avoidably). On the whole, however, I am much closer (and sympathetic) to Lomasky's view than to Regan's or Singer's.

The two books here reviewed are both solid contributions to social philosophy. It is almost a cliche, yet nonetheless true, that a reviewer can rarely do justice to the subtlety, breadth, and depth of a good work in philosophy. But what a review *can* do is evaluate and recommend; I have tried to do the former, and I shall now do the latter. Both of these books should be read and absorbed by anyone working in moral and/or social philosophy. Both are rich in innovative argument, even sprinkled on occasion with wit and humor. There is much that I have not been able to touch upon, such as Lomasky's treatment of individualist anarchism or Gauthier's concept of the Liberal Individual; but what I have tried to do is provide both a summary and bit of criticism along the way. I will end with what I intend as a high complement indeed: Both works will undoubtedly generate a surrounding literature (an 'industry', if you will), most of it constructively critical, *all* of it in these authors' debts.

MARK C. PHILLIPS

Ft. Worth, Texas