**“The Convergence of National Rational Self-Interest and Justice in Space Policy; Extending the Contractarian Reduction of Morality to Rationality Into The Final Frontier”**

**Duncan MacIntosh**

**Dalhousie University**[[1]](#endnote-1)

**Abstract**: How may nations protect their interests in space if its fragility makes military operations there self-defeating? This essay claims nations are in Prisoners Dilemmas on the matter, and applies David Gauthier’s theories about how it is rational to behave morally—cooperatively—in such dilemmas. Currently space-faring nations should i) enter into co-operative space sharing arrangements with other rational nations, ii) exclude—militarily, but with only terrestrial force—nations irrational or existentially opposed to other nations being in space, and iii) incentivize all nations into co-operation by redistributing some space-generated wealth to nations that might otherwise ruin space. Further, since all nations are rationally incentivized to ever greater dependence on space, and since each, by threatening space, can hold hostage its benefits to all, each can demand resolution of terrestrial discontents with other nations. So each has a rational interest in treating others morally both in space and on terrestrial issues.

**Key Words: Prisoners Dilemma, Deterrence Dilemma, Cooperative Surplus, Reduction of Morality to Rationality**

**1. Introduction: The Conundrum of Space**

Space presents special challenges to the defense posture of the United States. On the one hand, space is an enormously important site of current and future commerce, current and future support for terrestrial military operations, and possible future military operations on objects in space. So it would be natural to think of it as a domain in which the U.S. should have a strong capacity for military action on space objects for the purpose of defending its various space-related interests.[[2]](#endnote-2)

On the other hand, military actions on objects in space are almost certain to be self-defeating. Use of kinetic weapons in the space environment, for example, will only create cascading sequences of shrapnel multiplication in valuable orbits (the so-called Kessler effect), making space unusable for everyone, the U.S. included; and any sort of build-up of capacity for military action in space is likely to provoke a respondent buildup from adversary nations. This holds not only for buildup of arms whose use could have catastrophic effects, but also for technologies merely capable of relatively subtle effects – e.g., satellites able to maneuver on orbit, and so possibly able to alter the aspects of other satellites[[3]](#endnote-3); or technology, earth- or space-based, able to blind satellites. For those who develop and use such technologies are likely to reap what they sew: other nations will develop similar capacities, and will respond in kind, or worse, to actual uses of these capacities. Moreover, even if not used with malevolent intent, buildups of technology capable of malevolence are likely over time to result in misinterpretations, in accidental conflicts, and so, again, in destructive events likely to ruin space for everyone, the U.S. included.[[4]](#endnote-4)

 Some have argued that these facts make it impossible for the U.S. to achieve military dominance in space, disastrous for the U.S. even to try for it; that, more generally, the world, including the U.S., has no choice but to try for regulations for space any hoped for compliance with which will have to be voluntary rather than able to be militarily enforced; and that in any case, this is all to the good and is consistent with the noble, peaceful conception of conduct in space originally envisaged in the first treaties for space.[[5]](#endnote-5)

 Yet this will seem to many to be a slim basis for hope of good conduct by nations in space. If there is no possibility of enforcement of regulations for space by actions on objects in space, surely the temptation to cheat would be enormous. The best we could hope for is that nations would make and keep promises out of a kind of morality-informed good will – morality requires that we keep our promises and not act in ways disrespectful or exploitative of others. This may seem like a fragile thing on which to rest faith of compliance, especially for nations like the U.S. that are accustomed to a force-forward practice of self-defence, i.e., nations who think good defense is identical with the capacity for good offense and overwhelming retaliation. Such nations will be tempted to develop capacities for military action on space objects as a kind of insurance, something that, as we’ve just seen, is likely to be self-defeating due to entraining correspondent behaviour from other nations, this, again, fraught with the possibility of misunderstandings, accidents and escalations ultimately damaging for everyone.

In this essay, I shall first briefly moot doubts about the facts premising arguments that military action on space objects would be so disastrous as to make acquiring the capacity for such action imprudent. I will also suggest that there are ways for the U.S. to militarily dominate the space environment, should it so choose, without having to maneuver militarily directly on objects in space. But I shall argue that regardless of this fact, the state of geo-politics would make any such ambition unnecessary and not in the interest of the U.S. And I will assuage the worry that nations left to voluntary compliance with mutually beneficial arrangements for space cannot be trusted to comply. I will argue that the worry is based in a mistaken conception of the nature of rationality, and therefore a misunderstanding of what sorts of agreements one could perfectly well expect rationally self-interested agents to make and keep. I will suggest too that a correct understanding of this offers us a clear strategy for the management of space. In fact there are two possible points at which force can figure in compliance with agreements. The one everyone already understands and thinks is the only hope is the threat or use of force to ensure compliance – we use force to manhandle others into compliant behaviour, or we use the threat of force to create the deterrent of possible punishment for failure of compliance. The point of possible use of force that is novel for most people, and the one more appropriate for the space environment, is the threat or use of force to prevent agents who are unlikely to comply with agreements from ever being in a position to fail to comply with them. The idea is that we do not enter into deals with those who cannot be expected to keep to the deals, and we deny these people access to the circumstance in which they would have the opportunity of non-compliance. In this way we avoid making ourselves vulnerable to them in the first place. We may think of this as the difference between, on the one hand, enforcing compliance with the rules of a club by using force or its threat on people in the club, and, on the other hand, enforcing membership criteria on those who seek admission to the club. If we deploy threat or use of force to bar entrance to those likely to be unreasonably violent, we don’t have to worry about how to defend patrons from each other once they are in the club.

The sorts of policies for space that would be in the interest of the U.S. would also be in the interest of all reasonable nations, specifically, policies encouraging co-operation among nations in well-regulated space commerce. Such policies can be expected to be voluntarily complied with by all space-faring nations, provided that the extantly space-faring nations deny access to space by any nations and sub-national groups whose ideologies make them existential threats to any of the extantly space faring nations, and so that make them unlikely to comply with any policies otherwise to the advantage of both them and the nations they seek to destroy.

I will suggest too that the perfect model for the relation between justice and national rational self-interest in space is provided by a variant on David Gauthier’s contractarian moral rationalism. On this view, what is just is whatever contract for mutual advantage would be rationally agreed to by all parties. The rationally agreeable contract would give each signatory the maximum benefit from the surpluses of goods yielded by co-operative arrangements compatible with all signatories getting benefit. Moreover, rational agents, having so contracted, would find it rational to fulfil the contract with others whom they expect will likewise fulfill it. They would find this rational either because Gauthier is right that if it is rational to commit to an action then it is rational to fulfill the commitment, or because (as I have elsewhere argued) if it is rational to commit to an action, then it is thereby rational to have adopted values that would justify fulfilling it.

Whether rational in fulfilling commitments of this sort or not, the only agents to whom we would rationally extend the status of parties to such agreements would be ones pre-selected for their having some antecedent inclination to fulfil these commitments, i.e., an inclination to trustworthiness. Meanwhile, the only agents who could rationally expect to be extended such invitations would be agents who already had, or who were prepared and able to acquire, such inclinations, agents who would be motivated to do this by the advantages that would be conferred by such an invitation.

Finally, I will suggest that space is the great unifier of morality and rationality. Morality is comprised of arrangements to the advantage of everyone. Rational self-interest is comprised of arrangements to the advantage of the self. Space is such a wonderful environment that it is to everyone’s advantage to have the use of objects in space, whether this be use by dint of ownership, rent or gift of those objects. The objects located in space can confer great advantage on everyone. But these objects, once located in space, are so fragile, that they could be destroyed by anyone upon the provocation of anyone else’s malevolence, whether malevolence against objects in space, or in any other venue whose quality is affected by the status of objects in space (which could amount to any venue whatever, since for any venue not obviously connected to space, someone could demand that someone behave a certain way in that venue on penalty of an assault on their assets in space). And this means that it is to no one’s advantage to do anything to the disadvantage of another in these connections, whether to their disadvantage in space or on earth. But this is just to say that, to the degree that everyone’s welfare is tied to activities in space, what is in the interest of everyone is also in the interest of each person. That is, the demands of morality are identical to the demands of self-interest once everyone’s interest is entangled with the integrity of objects in space.[[6]](#endnote-6) This means we all have an incentive to get as many of our enemies as possible enmeshed with the benefits that participation in space confers. For once they are so enmeshed, they will be incentivized to cooperate with us in both space and non-space contexts.

**2. The Defeatist Alleged Facts**

It might be argued that there is room for doubt concerning the defeatism about military options in space whose defeatism is premised on the supposed impossibility of U.S. military dominance of space by military action on space objects, and on this supposedly necessitating peaceful co-existence in space. The alleged facts supporting this defeatism may seem more like facts of convenience – facts believed in only because if they were true they would entail the correctness of certain political visions. Here, such alleged facts would be convenient for those with peaceable, sharing visions of space. And even if the defeatist supposed facts are genuine facts, they will probably be doubted by those for whom the falsehood of these claims would be convenient for their political vision – those who would prefer that the U.S. seek military domination in space, for example, prefer it from such motives as fear and the perceived need for self-defence, from ideological evangelism, from a sense of duty to curate the welfare of oppressed citizens in other space faring nations, from the idea that the first to arrive in space and on astronomical bodies will be able to claim squatters’ rights, or simply from the urge to dominate others. The doubters will believe that non-debris-producing space countermeasures could be developed for neutralizing devices in space; and for clearing, braving or evading such debris fields as may result from kinetic encounters[[7]](#endnote-7); or for compensating for the statistical likelihood of collisions by the multiplication of devices so that there is redundancy against losses to collision with debris. Or the doubters may insist that there are ways of maneuvering militarily in the space environment that don’t involve creating debris fields, e.g., by using a U.S. satellite to alter the aspect of an enemy satellite, or by using lasers to blind the satellite; or by de-orbiting the satellite. Besides, there are some locations in space where risk of debris in Earth orbit is not the issue, e.g., if competing colonies come to exist on other planets or moons in the solar system; or in large structures, like space cities, not particularly proximal to satellite orbits. In fact, the more independent these things became, the less they would be reliant on launches from earth through debris fields for staffing, re-provisioning, and so on. As for independent, competing colonies on other bodies, especially ones with atmospheres, military conflict could proceed apace on land, sea and air on those bodies in the conventional forms already developed terrestrially (although there would remain the problem of the space environment around these bodies).

So however much some of us may favour the courses of action that many think would be dictated were the defeatist claims of fact true, it is important that we not have as our rationale for such courses of action premises that are not firm.

There are replies to these concerns, however. First, the state of technology involved in dealing with debris is not very advanced. So in the short term, debris from kinetic action would be a problem. Secondly, even the less violent methods of interfering in others nations’ satellites are likely to meet with correspondent attacks on one’s own satellites, in an escalation of grey-zone conflict moves.

But the strongest reply comes, again, from David Gauthier, whose model of rationality I said in the introduction I would seek to apply to space scenarios: it is far cheaper to pre-empt the need for conflict than to engage in it; and far cheaper to avoid putting one’s self in the position of having to respond to a betrayal than to respond to a betrayal. (Well, there are expected utility calculations to be done here: the possible harm one might experience from a defector from any deal you make must be balanced against the possible reward if he does not defect, and the odds of each occurring. The deal becomes more rational the greater its possible payoff if it works out, the less the possible harm if it doesn’t, and the greater the odds that it will work out.) That is, as we shall see, instead of preparing to fight battles in the space environment, it would be better to prevent problematic agents from accessing it in the first place. This is where military effort should be applied.

**3. What Follows From the Defeatist Alleged Facts About Military Options**

In any case, it does not follow from the concerns about, for example, debris, that there can be no military maneuvering in regard to space. At most it would follow only that such maneuvering cannot be in the space environment. Nations who mean each other harm don’t need to attack each others’ assets in space to cause harm: they can attack ground, sea and air targets from ground, sea and air; they can attack each others’ cyber assets and informational assets using terrestrial computers; they can freeze each others’ financial assets; by the foregoing means they can attack each others’ civilian populations; and by these means and others, they can deny each others’ abilities to launch assets into space, or to benefit from objects they have launched into space; or they can punish or, by using the aforementioned things as threats, deter hostile actions that might be conducted in space.[[8]](#endnote-8)

Likewise, there can be military maneuvers that defend the space environment, but which do not occur in space or by action on objects in space, namely, maneuvers on land or in sea or air, maneuvers like destroying enemy launch sites, attacking enemy space vehicles on ascent or descent, destroying or jamming enemy ground-based communication with objects in space, espionage against other nations’ space efforts (e.g., by inserting spies into crucial enemy space programs), cyber disruptions of other nations’ space-focused computers, and interfering with other nations’ electoral processes or political deliberation processes to hinder the formation of political will against the space operations of a given nation.

Moreover there are any number of ways for a given nation to have positive and defensive impacts on its prospects in space that aren’t really military at all: e.g., the given nations refraining from trading with nations who aspire to access space but could not develop the wealth needed to do so without trade with the given nation; promoting exchange of scientific research and so building global communities of scientists who would tend to have greater loyalty to the peoples of the world generally than to any particular nation seeking advantage in space; funding blockbuster Hollywood-type movies that in effect propagandize whatever attitudes a country seeks to have the world take to it and to its activities in space; weaponizing social media to similar effect; facilitating exchange programs for students; and, more generally, promoting entanglements between itself and the populations of other nations so that relations of mutual affection form, decreasing the likelihood of any enmity that might rise to the level of military conflict.

It is, therefore, problematically one-dimensional thinking to believe that having effect on space requires maneuvering with objects in space against objects in space. There are ways to have such effects by military maneuvers using objects not in space against other objects not in space; and there are non-military ways to maneuver that can have effects as good or better than military ways of maneuvering.

The foregoing considerations are relevant not only as part of broadening our conception of what sorts of maneuverings should be part of a nation’s policy with regard to space, but also as part of recognizing the viability of actions that would be needed to prevent the very occasion of military action in space on a nation’s space technology, something that will be part of the Gauthier strategy I will be recommending a couple of sections from now.

**4. The Lack of Need For Military Offense and Defense Given the Political State of Play**

But then there is the question whether there is the need for any of the foregoing, or the prospect of it. In this section I shall sketch a vision of the current state of world conflict as merely grey-zone conflict, a vision that would further support the Gauthier conception of rationality applied to space that I’m generally arguing for here. I believe there is an emerging consensus that this is indeed the state of play in the world. But whether it is true or not, its truth would be an essential premise for the larger argument I’m engaged in; so I take some time here to flesh it out.

It is widely acknowledged that such conflicts as there are now between the major powers are, at worst, of the grey-zone sort.[[9]](#endnote-9) Michael Mazaar summarises this idea nicely: Mazaar sees grey zone conflicts as caused by nations whom he classes as incremental revisionists, nations who want to see changes in the world order without destroying that order, and who want to be partners in peace but with more influence in and more profit from that order.[[10]](#endnote-10) On this conception of the state of play in the world, none of the major powers any longer mean to eliminate any of the others; each respects the right to existence of the other. Moreover, none of the major powers propose to export their ideologies into the others by violent means. In fact it is increasingly recognized that this would be impossible: each of the great powers offers its citizens various degrees of negative rights to non-interference in their lives and positive rights to freedom from want; and the tastes of each powers’ populations, shaped as they are by their individual histories, more or less accept what is on offer. Each may sometimes be tempted by what others offer – whether this be more freedom, or more economic security. But those so tempted want more of what they lack only insofar as it can be incorporated into the extant traditions and power dynamics of their own country.

 Each major power does propose to exemplify its preferred ideology to others. Each seeks a sphere of influence in parts of the world adjacent to it, or associated with it by long tradition. And each seeks allies in less adjacent regions, and seeks to have influence in those regions, this in order to create ideological buffer zones between competitor ideologies. That is, each seeks to have client states, defined as states accepting the patronage of larger states and prosecuting shared agendas. And this sometimes results in the different powers having proxy conflicts with each other, conflicts between their respective client states; and in different powers interfering in each others’ internal affairs in various nuisance ways (e.g., the recent disinformation campaigns by Russia in U.S. elections). Here, one power tries to weaken the other powers and prevent them from obstructing the first power’s activities in its aspirational regions of influence. But such conflict as there is, is increasingly less about ideology and more about access to markets. More and more the major nations are political brands whose primary relations with each other are economic in the form of trade, shared commercial or technological ventures, and co-maneuvering against other states, proto states, failed states, or groups of non-state actors who are threats to the major states.

All of this means that there is no real basis for intense military conflict between the major powers anywhere, let alone in space, or about anything, let alone about space. Instead, all nations would do well with arrangements that promote peaceful commerce between them. Any remaining disposition to violent conflict is more the result of inertia from past conflicts than premised in genuine current pretext. Each nation’s military and intelligence agencies, for example, are in the habit of conflict with each other; and each nation’s industrial base has a stake in at least the continued appearance of the need for violent conflict in order to maintain orders for military hardware. Indeed, all the major nations have the need for perceived outside enemies because the resulting postures of defense and possible offense fuel their economies. But this is an entirely different basis for a war footing than the need to defend against an actual threat, or the need to mount a rescue of oppressed populations in other nations. There are, to be sure, other nuisance operations between nations, e.g., copyright violations, industrial espionage, and problematic business practices like goods dumping. But this is mere criminality, not existential threat; and even categorizing this as criminality would be contested, on grounds that criminality is ideology-relative, and nations differ in ideologies. Moreover, the perception in one country of these things being permissible to perpetrate on other countries is, again, a holdover attitude from times of conflict over foundational matters. The current infractions are not the kinds of things themselves worth going to war over.

Further, there is no foreseeable means by which one state could attain such asymmetrical military power over any of the others as to be invulnerable to catastrophic attacks or reprisals. No offensive system will be perfect; nor any defensive system either. So there is no scenario, including no space based scenario, in which any of the major powers will lack the ability to hold the others’ civilization hostage to catastrophic military attack. Perhaps there was a time when we were to each other as people with rifles standing in the corners of a large room, some of us sometimes muttering that we could get the drop on the others; but more and more a better analogy is that we are all standing together in an elevator and everyone’s got a grenade with the pin pulled. We can bring major harm to each other with ever greater speed; and we depend on environments, e.g., the global climate, and the regions of space that comprise useful orbits for satellites, in the peculiar way that, to make these environments worse for our enemy, we’d also have to make them worse for ourselves. So, for example, we can deliver nuclear attacks with a speed three times greater than in the nineteen-fifties[[11]](#endnote-11), but we’d all suffer radioactive contamination and devastating nuclear winter if we ever engaged in such an exchange; and we can destroy each others’ satellites, but in so doing we would make it impossible for our own to function, either immediately because of the debris fields resulting of kinetic attacks, or almost immediately from the inevitable effects of retaliation. I predict we’ll all get along fabulously.

But even if there were a perfect offensive or defensive weapon, no state would want to use it on any of the others. No major state is such that it would be better off if the other state simply ceased to exist. All that would mean is a vast reduction in markets in which to sell, resources, skills and products to trade for, and so on, not to mention the cost in sympathy of watching others suffer and die, nor the crisis of conscience that would follow, the generations of self-recrimination. No major state on earth is such that it would answer in the affirmative to the question, if you had a button that would kill every man, woman, child, and dog of your enemy’s state, would you push it? (For example, Palestinians may want the Jews out of Israel, but they don’t necessarily want them all dead.)

This is not to say there is no point to seeking to militarize space (by which I mean, create and emplace the technology needed to maneuver militarily in space against other nations’ technology in the space environment – space is already militarized, of course, in the sense that nations have assets in space that give them military capacity in non-space environments, e.g., satellites that give them situational awareness and command-and-control in terrestrial military operations). One point to militarizing space would be that many industries in all of these countries would profit enormously from developing and marketing the technologies involved to governments. But there may be more peaceful technologies in whose production these companies could be tasked instead. And the presumptive risks associated with an arms race in space may be decisive on what technologies we should be asking and permitting industries to develop and manufacture.

At any rate, since the nations no longer represent existential challenges to each other, they have all the usual incentives of any agents to enter into co-operative arrangements with each other with a view to producing and sharing the surplus goods that co-operation is so good at generating – the so-called cooperative surplus. Put another way, with the existential issue settled, interactions need no longer be thought of decision-theoretically as zero sum games in which, if one party gains, the other party must lose. Instead, both parties can gain, because they can produce greater gains by co-operating than by fighting. The only resistance to this will come not from ideological objections, but from those who have managed to attain power and wealth in whatever system they live, and who will want to preserve that status. But they can be mollified by the offer of business arrangements to their advantage that again enhance the co-operative surplus, the surplus in which, again, they can expect to share.

**5. The Problem of and Possible Need for Disruptor Technologies**

A further objection to a given nation’s seeking military dominance in space is this: while it is apparently difficult to develop and deploy useful technologies for space, it is apparently much less difficult to develop disruptor technologies, technologies that would ruin other people’s space emplaced objects. In fact let us assume for the sake of argument that every political entity of any significant size anywhere on the earth has or could acquire the means to launch weapons that would ruin space for everyone, cluttering it with prohibitive levels of debris that would make satellite technology unworkable. So: China, Russia, North Korea, Iran, any number of the Western powers, as well as India, Pakistan, probably combinations of radical organizations of whatever religious or political orientation – any and all could ruin everything for everyone.

This fact, if it is a fact, tends to bring into convergence two considerations about space: first, what is it technically possible to take from space by force if necessary against possible competitors? Second, what sort of conduct is ethically required in space? In answer to the first, arguably it is not possible for one agent to militarily dominate another in space, whether that agent be a state, a collection of non-state actors, or a corporation. Any such attempt at domination would immediately cause, or be met with by way of retaliation, ruinous disruption, and so no domination.[[12]](#endnote-12)

In answer to the second question, the question what is ethically required in space, there are competing answers. Joan Johnson-Freese and Kenneth Smith effectively review several of these answers, namely, the approaches of utilitarianism, rights theory and Kantianism, the fairness approach, the common good approach and the virtue ethics approach.[[13]](#endnote-13) Each approach seems to give normative answers either indeterminate as between competing recommendations, or determinate, typically in the direction of peaceful co-operative arrangements rather than aggression and bullying. But either way, these moral recommendations infamously are not bullets. That is, they don’t afford a motivation to comply with their recommendations for those with differing moral concepts, or with no morals – arguments don’t hold you at gunpoint. In light of this fact, I suggest the following. First, the whole object of morality is that it is supposed to make everyone’s life better. But second, whether this is true or not, no one is interested in any concept of their life’s being better that doesn’t involve improvements by the measure of things they actually care about. Third, morality conceived in that way is a set of principles for conduct that can make everyone’s life better in that sense only if some principles are such that if we all followed them we’d all be better off in that sense than if we all violated them. Fourth, therefore, the only supposed moral obligations all parties can be expected to be interested in are ones that prescribe cleaving to principles with the foregoing properties. From all of this follows the conclusion that if anyone has ever proposed, or ever would propose, that morality is anything other than a contract for mutual advantage, it is questionable why on earth we should we listen to them. Why should we do what they say?

But what would the contract for mutual advantage look like? This is a question in bargaining theory. And arguably if everyone has equal power to disrupt, then no one can demand a benefit others would be unwilling to give, since then the others would rebel and take steps to ruin all possibility of benefit. Thus in this situation, any reason for one person to have a benefit would also have to be a reason for all others to agree to confer the benefit. It might seem therefore that everyone should ask for the same, since to ask for more would be special pleading, and to settle for less would be to have accepted someone else’s special pleading. On the other hand, since I do better by any positive arrangement than by none, arguably I should accept any arrangement that gives me any positive benefit, even if it gives you vastly more. What then is to compel mutual movement to the middle ground, or indeed, to any fixed position whatever? How is agreement possible? One hope would be that we all have an inherited tendency as primates to demand roughly equal sharing in such situations; and it is therefore a known fact about us that we will non-rationally ruin deals that don’t approximate equality, a fact which has the effect that, if you know of it, rationally you won’t act so as to trigger it in others. This will give all parties a rational basis for moving to the middle in bargaining in light of a fixed fact itself intractable to rational suasion.

Even supposing the bargaining problem of negotiating to a mutually acceptable deal can be solved, however, what is to motivate people to keep to the deal? What is to prevent them from trying to take more than their negotiated share secretly?

Here again, David Gauthier’s theory is instructive: having formed commitments to comply with contracts compliance with which would roughly equally benefit everyone, rational agents would comply either because Gauthier is right in his view about rationality that if it’s rational to commit to an action then it is rational to fulfill the commitment, or because (as I have elsewhere argued) if it is rational to commit to an action, then it is thereby rational to have adopted values (desires) that would rationally justify one in fulfilling it. (Sincerely committing to a self-limiting contract is the same process as coming to have an inclination towards fulfilling it, at which point, fulfilling it would not longer be an act of self-limitation; instead, it would become an act that expresses the (new) values of the self; it would be what one has come to desire to do.)

Of course an additional reason to comply would be to avoid punishment. But we are here assuming that committing to these deals would give one access to space, and that, again, once in space, large-scale punishing in space will only ruin everything for everyone, and is therefore not a feasible option. So we only want to extend the deal to those who can be expected to comply voluntarily. The only agents who will rationally be offered the status of parties to such agreements will be agents pre-selected for their having some antecedent inclination to fulfil these commitments, i.e., an inclination to trustworthiness. It is rational for us to so limit the offers, rational for other agents to cultivate this inclination so as to earn the offer.

I suppose another method of guaranteed compliance that would work would be to demand that parties allow themselves to be made vulnerable to some consequence for non-compliance the imposition of which would not require kinetic battles in orbit, nor any other action on objects in orbit, e.g., taking a poison that required weekly administration of antidotes, or signing contracts permitting their wages to be garnished should they break the deal, or giving substantial assets to some mutually trusted third party who will release them back to their donors (possibly with interests from the profits of commerce in space) only proportionately and commensurately with their good conduct in space. With such penalties in play, the rationality of compliance would be uncontroversial even on the more standard theory of rational actions as actions that are directly self-interest advancing.

Note that the sort of commitment involved above is exactly not that recommended by Freese and Hitchens, namely, a commitment to a moratorium on use of force in space until, among other things, powers like the U.S. get a better idea of how to mount use of force in space. This is the opposite of a strategy of reassurance to others. [[14]](#endnote-14) Moreover, the related but more deeply logically deceitful strategy known as hedging – pledging peaceful co-operation while developing technologies that can be tasked for offense and defense -- is not a coherent messaging strategy. In fact, it can only cause disruption. Instead what would be required would be the sorts of things others (and indeed, Hitchens herself in another work) have recommended, things like transparency, trust-building, information sharing, and resource sharing.[[15]](#endnote-15) Further steps in this direction would be exchange programs, shared business arrangements, free trade deals, and indeed anything that maximizes the mutual entanglement of our interests. In fact since even if we do not seek to give space technologies dual capacities – civilian and military – any such technologies will tend inevitably to have dual capacities, as we develop and deploy these technologies, we will need simultaneously to be doing other things to reassure others that we will not avail ourselves of the possibilities of their military use in aggressions against other people. We could do this by creating and complying with rules requiring us to impair the military use capacity of space objects fit for civilian use. E.g., we could have our space objects operated by source codes available to all, or build them to a code of fragility making them unable to survive impact with other satellites. There are any number of ways to ensure the non-dual usability of our technologies; and it would be worth developing yet others.

We’ve just seen, then, that the fact that in-space aggression can be met with in-space disruption that will ruin everything for everybody, and that compliance with arrangements cannot, therefore, be enforced by force applied to objects in space, recommends the moral approach known as contractarianism.[[16]](#endnote-16) Indeed, in the circumstances of space, what is recommended by morality and what is recommended by rational self-interest are the same. In fact, given the defeatist view of the facts about the efficacy of in-space aggression we have assumed here, three ideas are pulled together: libertarianism, according to which each person has the right to do whatever she wants provided her doing it is compatible with all others having a similar right; contractarianism, according to which each person has the right and duty to do whatever would be specified in a contract regulating conduct which all people would find it rational to sign given their several interests and powers; and a particular theory of rationality known as constrained maximization, according to which each party would find it rational given her interests and powers and those of everyone else to commit to certain limits on her own behaviour, and each party would find it rational to act on those commitments with those prepared to commit and act likewise. Here, the first idea announces maximum mutual freedom as an ideal, the second idea gives a specification of what that freedom amounts to given that it must be had by all, and the third idea explains a way of reconciling the apparent constraint involved in the second idea with a theory of freedom in choice that presents constrained choice as free rather than restricted.[[17]](#endnote-17)

6**. The Scope of Inclusiveness of Deals Regarding Space**

One standard knock against Gauthier’s theory of the reduction of morality to rationality – as applied in this case, it is a theory of the identification of the national interest with the interest of all nations -- is that it explains how there can be a rational justification for recognizably moral self-limitation of behavior for the benefit of others only between agents who can threaten harm or confer benefits on each other – these possible harms and benefits are the things the contracts between agents are supposed to regulate. And it is thought that there could be people too strong to need such contracts (they can take any good they want by force), or too weak to be able to demand inclusion into them (they have neither the power to threaten, nor anything of value with which to bribe). But interestingly, in the case of space everyone has the power of a disruptor, or has a champion who has that power[[18]](#endnote-18); and so everyone can successfully demand to be included in the circle of those whose welfare others will find it rational to self-limit their behaviors to respect.

Well, almost everyone. You get to participate only if your values don’t put you in logical opposition to the welfare of others. By this I mean that you cannot be brought in if your welfare is defined as requiring the illfare of others, that is, if you are malevolently tuistic towards them.[[19]](#endnote-19) But I conjecture that one party can be malevolently tuistic to others only by dint of artificial and mistaken pretext, e.g., by being party to a religion defined as requiring the extermination of parties to any other religion. Since religions are by definition not demonstrably true, they can have allegiance only epistemically irationally, and cannot rationally found a basis for conflict. This means no one is an epistemically rational party to any such malevolent attitude, and so someone can have been conscripted into it only by ignorance or desperation of circumstance in which some likely reasonable need of theirs will be met only if they adopt otherwise unreasonable attitudes and commitments.

Such persons cannot be allowed access to vulnerable environments, like space; and they cannot be allowed standing in circumstances that would give them power over others, this again barring them from space. No Jihadists allowed.

On the other hand, since necessarily such people have these attitudes only from desperation of circumstance, they can be freed of the attitudes by improvement of their circumstances and by being offered incentives to join the rest of civilization. Accordingly I propose that they be included as beneficiaries of the dividends that the commercialization of space has the prospect of providing. And as this gradually affords improvement in their circumstances, their attitudes will adjust, whence they too can be brought fully into the circle of contributors towards and full beneficiaries of, cooperative actions in space generating a shareable surplus.

**7. The Tail Wags the Dog: Political and Economic Identities Imposed on Nations by the Conditions on Viable Participation in the Economy of Space**

An interesting consequence of the exigencies of well-regulated relations among nations who would participate economically in space is that, whatever the political and economic character of any given nation, all nations who maneuver in space would have to act, in relation to other nations, as if they were liberal democracies with regulated capitalist economies and more or less strong welfare states.

Roughly speaking, both justice and rational self-interest require nations to maneuver in space under the auspices of in effect being participants in a liberal democracy -- but with nations rather than individual citizens as constituents -- since they would have to deliberate to consensus about how to maneuver together. Further, they would in effect be maneuvering as if in a capitalist economy -- regardless of whatever economic system may feature within a given nation -- since each will be operating as an independent business in space. But this would have to be an economy regulated to avert tragedies of the commons in space, something all parties would happily agree to for mutual advantage.[[20]](#endnote-20) The parties would also create a welfare state, one paying a “space dividend” to all persons, this in order to allay any discontent that might lead to some agents seeking to disrupt space in a way ruinous for everyone. Finally, all policing of these arrangements could be done terrestrially and in the course of denying access to mutually beneficial deals by entities unlikely to keep to the arrangements, rather than by interfering with objects in space in an exercise of enforcement of deals, or of punishment for violation of deals.

**8. Conclusion and Speculation**

Reliance on space makes those so reliant vulnerable. But for that very reason it also creates opportunities. For the leverage of it converts all terrestrial relations into prisoners dilemmas, situations where we can all do well by making and keeping agreements to cooperate and to share the fruits of cooperation, but where each of us is tempted to violate such agreements for greater individual advantage; and by the same token, the universal reliance on the space environment means that no one who would violate arrangements about space could be admitted into the circumstances of a space prisoners dilemma and its prospective profits. Therefore, no one would violate the agreement to cooperate. So the solution to the problem of cooperation in space is also the solution to the problem of peace on earth.

There is disagreement about whether we are safer with or without being weaponized against space objects. The argument against our being safer with weaponization is that the actual use of any such weapons would be disastrous, and creating the capacity for their use increases the likelihood of their use. The argument for us being safer with weaponization is that, if everyone is armed to the teeth, no one will be tempted to mess with anyone and peace will prevail. The decisive argument decides against arming, and goes like this: arming is prohibitively costly, it is unnecessary given the compatibility of the interests of most peoples and our ability to exclude from space those whose interests are diametrically opposed to ours, together with the consolation from David Gauthier’s decision theory that forming a self-advantaging commitment to refrain from exploitative action in space entrains the rationality of fulfilling that commitment.[[21]](#endnote-21)

**9 Postscript: The logic of Bidding for Entry into the Space Environment**

It might be thought that those not yet in the space club have the incentive to create weapons that could make things worse for those in the club, this to extort those already in the club to give membership to those not already in. But creating such weapons will only successfully extort membership into the club if it is accompanied by credible assurances that the weapons will not be used once one is accepted into the club. Perhaps the logic of space use is the same as that of the deterrence paradox – mutual assured destruction: each of us is and/or should be such that if anyone ruins space in some degree for us, we will ruin it worse for everyone, including, inevitably, ourselves. This threat keeps the peace, and nations acquire standing in space by having the capacity to credibly threaten, and to follow through on the threat. This is, of course, why some nations seek nuclear capability: they are trying to extort a seat at the international bargaining table.

But what, then, of the Jihadists who I have said should be barred admission into the space environment? What if they acquire the capacity to be disruptors? It depends: if they are doing this to gain entrance into a system of benefits, that’s great, since the possibility of such entrance will limit what they are prepared to do – they won’t want to behave so badly as to induce others to bar them, or to eject them; or so badly that no one will make a concession to them in negotiating a deal. That is, the circumstance of them wanting entry for that reason is the same as the circumstance in which they have ceased to be Jihadists. Now, what remains is just bargaining amongst those whose interests can be co-advanced with the right bargain.

On the other hand, if the Jihadist seek to be disruptors just as part of an all out war against those whom they would disrupt, then they are an existential threat, a categorical enemy to all of space faring civilization, they are a threat to the enormous co-operative surplus that space commerce can afford to every one who can tolerate others as a co-participant in such commerce, and so those prepared to share space are entitled to do whatever it takes to prevent this disruption.

But note that there are other ways to incentivize people into letting one have access to space than threatening to destroy the space environment if one is not given access to it. For example, one might be such that one can confer vast benefits on others if only one is admitted. Perhaps one has proprietary technology that would be a benefit to everyone were one admitted into space. Or perhaps one is prepared to work cheap; and cheap labour in space would be of extra value, just as it is on earth. Or perhaps one has great wealth with which to buy the right of entry into space. Or to bribe one’s way into the right. There are many perfectly legitimate ways to skin this cat.

So one does not need to threaten disruption in order to be granted access to space commerce. The idea that one needs to do this is premised on the assumption that space commerce is a zero sum game, that sharing space means dividing profits into smaller piles, and so that only someone who could threaten disruption would be admitted to a share. But in this domain, as in so many, this assumption is false. The truth is that, in space, first, there is enormous abundance, more than any current players could exploit; and second, specialization of labour means there’s a role for every player. Further, space operations are expensive, and risky, and all parties do better if the risk and expense are shared. So we don’t need to be extorted to want to admit agents into space commerce with us. We just need to be reassured that prospective partners won’t cheat or be disruptors. Therefore the incentive structure of space incentivizes honest cooperators. This means any nation that invests in weapons able to attack space objects is wasting its money.

But what if you are a nation opposed to the welfare of space faring nations, and unlikely to be able to compete technologically in that domain? Then you might find it rational to try to destroy space for everyone, thence to restore equality of advantage to you in terrestrially founded maneuvers, where you have a chance.

1. For useful feedback thanks to my students at Dalhousie University, and to Sheldon Wein, L.W., Max Dysart, and Christian Weisenburger. This essay was originally prepared for the conference, “The Weaponization of Outer Space: Ethical and Legal Implications," organized by the Center for Ethics and the Rule of Law at the University of Pennsylvania Law School, Philadelphia Pennsylvania, April 2018. I’m grateful for comments received from other participants at that conference, especially the co-organizer, Cassandra Steer, and from Kiernan McClelland, both of whom provided extensive written comments. [↑](#endnote-ref-1)
2. Jerry Hendrix and Adam Routh, “A Space Policy for the Trump Administration”, Center for a New American Security, Oct 2017 (ms). [↑](#endnote-ref-2)
3. I.e., the orientation of their antennas, disrupting their ability to send data, or the orientation of their sensors, disrupting their ability to gather data. [↑](#endnote-ref-3)
4. Grego, Laura, “Space and Crisis Stability”, Union for Concerned Scientists (ms.); Cassandra Steer, “Global Commons, Cosmic Commons; Implications of Military and Security Uses of Outer Space”, *Georgetown Journal of International Affairs*, 16.3, Spring 2017; Joan Johnson-Freese and Kenneth Smith, “U.S. Space Dominance: An Ethics Lens," chap. 5 in Cassandra Steer and Matthew Hersch, eds. *War and Peace in Outer Space: Law, Policy, and Ethics* (Oxford University Press, 2021). [↑](#endnote-ref-4)
5. “1967 Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space”, United Nations Office of Outer Space Affairs; P.J. Blount, “Peaceful Purpose for the Benefit of All Mankind: The Ethical Foundations of Space Security," ch. 4 in Steer and Hersch, *War and Peace*; Cassandra Steer, “Overview of the Existing Mechanisms of Global Space Governance”, in: Ram Jakhu and Joe Pelton (eds.), *Global Space Governance: An International Study* (Springer, 2017). [↑](#endnote-ref-5)
6. Is it only atmospheres that make immoral offense possible? And what of the fragility of many other contexts in which we all maneuver, e.g., the climate? The thing about atmospheres is that they in effect attenuate kinetic effects. In space, if you blow up someone’s satellite, you in effect blow up your own too from the Kessler effect – bad Karma comes swiftly in space. But in an atmosphere, if you blow up someone’s valuable thing, the blast forces are cushioned by the atmosphere over the distance between the blast and your wealth and treasure, so blowing up someone else’s thing is not, inevitably, blowing up your thing. It is not only space in which offense is self-destructive, of course. Polluting the climate with waste you do not go to the trouble and expense of disposing of sustainably can be self-destructive too. It’s just that the climate is vastly more resilient than the space environment, and so can absorb a lot more abuse. But in principle there is a level at which one could pollute the commons that would also result in one’s polluting one’s self. Interestingly, the question whether or not a given action is in effect the end game of a paradox of deterrence where, to harm another, one must also harm one’s self, or the end game of defection in a prisoners dilemma, where, in harming another, one is purely benefitting one’s self at the expense of another, is dependent on whether the conditions under which one is maneuvering isolate the aggressor from the effects of aggressing. And this is a contingent and variable factor of environment. It is also, therefore, technology-dependent: if I develop a technology that immunizes me from the global winter effect of using nuclear weapons (assuming it is those effects that would be the harm to me in this instance), and you don’t, then we are no longer in a deterrence paradox, for you have lost the power to deter me by threat of an action that will harm me as well as you. (Actually, there is a disanalogy here: in the deterrence paradox proper, I must threaten to do something to you I intrinsically don’t want to do to prevent you from doing something I intrinsically want even less. While in the case just discussed, I become freed to make a threat against you to do something to you that, by virtue of contingent facts, would, previously also have hurt me. It’s not that I didn’t want to hurt you. I did. I just didn’t want to hurt myself in the process.) [↑](#endnote-ref-6)
7. See for example Eli Meixler, “SpaceX Has Launched an Experimental Space Junk Sweeper Into Orbit”, Time.com, April 3, 2018. (Last accessed April 23, 2018). <http://time.com/5225670/spacex-space-junk-cleaner-launch/>; the University of Surrey, Surrey Space Center, Space Missions brief on the mission, RemoveDEBRIS <https://www.surrey.ac.uk/surrey-space-centre/missions/removedebris>, last accessed April 23, 2018; and Himanshu Goenka, “International Space Station Could Wield Lasers To Prevent Space Debris Collisions”, *International Business Times*, May 16, 2018 (last accessed May 16, 2018). http://www.ibtimes.com/international-space-station-could-wield-lasers-prevent-space-debris-collisions-2681320 [↑](#endnote-ref-7)
8. For yet other options, including public “naming and shaming”, see Jana Robinson ”Deterring Chinese and Russian space hybrid warfare by economic and financial means”, *The Space Review*, Sept. 18, 2017. Last accessed on April 18, 2018. <http://www.thespacereview.com/article/3331/1> It is worth noting, of course (as Casandra Steer pointed out to me), that many of these methods of attack themselves depend on, or would be made more effective by, the use of space assets – e.g., the satellites which relay information around the world for the financial industry, and the GPS satellites which afford navigation information for terrestrial military operations. Still, this can be true compatibly with a military being able to use space assets to attack ground targets, without thereby having to imperil the operability of the space environment. Such activity by a given nation may, however, court attacks on its own assets in the space environment. [↑](#endnote-ref-8)
9. Grey zone conflict is generally represented as conflict that fails to meet the paradigm of an interstate military conflict by virtue of one or more of the following features: it involves aggressive actions insufficiently consequential to merit a military response; it uses the methods of anonymity and misrepresentation and is therefore difficult of attribution; it uses non-kinetic methods or uses kinetic methods only marginally; it is not necessarily conducted by nation states, only by sub-national entities, like companies, hacktivists, and radical sub-groups, but it may be occurring with the indulgence of the nation-states that embed these sub-entities; it is incremental and subtle by deed but can cumulate to large effect; it is aimed at non-military targets, e.g., at businesses, informational infrastructures, the deliberative processes and mechanisms of a polity, or at individuals. (The foregoing paragraph is paraphrased from Duncan MacIntosh, “We Have Met the Grey Zone and He is Us: How Grey Zone Warfare Exploits Our Undecidedness about What Matters To Us,” in Mitt Regan and Aurel Sari, eds., *Hybrid Threats and Grey Zone Conflict: The Challenge to Liberal Democracies* (Oxford University Press, 2024), pp. 62-63.) [↑](#endnote-ref-9)
10. For a survey of the nature of grey zone conflict that supports the conception of world affairs I sketch in the main text, see, Michael Mazaar, *Mastering the Gray Zone: Understanding a Changing Era of Conflict*, United States Army War College Strategic Studies Institute (2015) pp. 126-137. [↑](#endnote-ref-10)
11. For a description of some of the claims made about new military technologies, together with some balancing scepticism, see Tobin Harshaw, “Why Putin’s Nuclear-Powered Superweapon Went Up in Smoke: A Q&A on the deadly Russian missile explosion, China’s growing arsenal and the new nuclear balance”, *Bloomberg Opinion*; Politics & Policy, August 17, 2019 (last accessed August 28, 2019) https://www.bloomberg.com/opinion/articles/2019-08-17/putin-s-nuclear-superweapon-went-up-in-smoke. [↑](#endnote-ref-11)
12. One worry here would be logically possible peoples whose ideology is fully implemented by Paleolithic conditions of existence, and whose best defense against modernity would be the imposition of such conditions on everyone else. Such peoples would have an incentive to destroy all technologies that might endow anyone with a better-than-Paleolithic life; and so these peoples would have an incentive to destroy the space environment, and would be indifferent to reprisals in the space environment, since they don’t care to maneuver there. These peoples would also be indifferent to sabotage of the advantages of space, since they derive none. These peoples, as well as any peoples whose drive is thanotic (death-seeking for themselves, death-imposing on others), would not be moved by the above arguments (except if the thanotic peoples thought they could kill more people using space based technology). But this is just to say that they would fall into the category of those who are our categorical enemies, that is, who cannot co-exist with us in mutual moral respect, and who, therefore, are part of a pair each member of which has an absolute right of self-defence against the other. Each necessarily aims to be an existential threat against the other. Such people would fall into the category of those existentially ideologically opposed to other nations. I have suggested there are no such peoples any more. But if there are, I shall be arguing, they are peoples we have both a pragmatic justification for excluding from having access to space, and a moral right and duty to prevent being able to access it. As for conventional, non-space conflict possibilities with them, we probably have a duty to try to reform them before we resort to military maneuvers against them. For almost certainly they have their attitudes more as a result of, for example, defense from the colonializing aggression of other nations, rather than from some evidence-based conviction that their values are correct. [↑](#endnote-ref-12)
13. Joan Johnson-Freese and Kenneth Smith, “U.S. Space Dominance: An Ethics Lens” (ms.). [↑](#endnote-ref-13)
14. Joan Johnson Freese and Theresa Hitchens, “Toward a New National Security Space Strategy: Time for a Strategic Rebalancing”, Atlantic Council Strategy Paper No. 5 [↑](#endnote-ref-14)
15. See for example the proposals in P.J. Blount, “Peaceful Purpose for the Benefit of All Mankind: The Ethical Foundations of Space Security” (ms.); Theresa Hitchens, "Norm Setting and Transparency and Confidence-Building in Space Governance," ch. 2 in Steer and Hersch, *War and Peace*; Gilles Doucet, "A Proposed Transparencey Measure as a Step Toward Space Arms Control," ch. 10 in Steer and Hersch, *War and Peace*; Icho Kealotswe, "The Rule of Law in Outer Space: A Call for an International Outer Space Authority," ch. 3 in Steer and Hersch, *War and Peace*; Paul Meyer, “Diplomacy: The Missing Ingredient in Space Security," ch. 12 in Steer and Hersch, *War and Peace*; and Yasuhito Fukushima, “Debates over the Military Value of Outer Space in the Past, Present and the Future: Drawing on Space Power Theory in the U.S.” presented to the conference mentioned in note 1, above. [↑](#endnote-ref-15)
16. This would have the additional advantage of recommending conduct compatible with international humanitarian law and the laws of war, both of which are at great risk of being violated by military action in space given the special characteristics of the space environment. (For example, these principles would be violated if anyone blew up a lot of satellites, since the resulting high-velocity shrapnel would make the orbital environment uninhabitable for the foreseeable future, violating the requirement that weapons permissible in war must not have the effect upon use of rendering a large environment uninhabitable.) For a discussion of such possible consequences and their implications for these widely accepted legal strictures on permissible use of force, see, e.g., Dale Stephens and Cassandra Steer, “Conflicts in Space: International Humanitarian Law and its Application to Space Warfare”, *McGill Annals of Air and Space Law*, Vol XXXX (2016). [↑](#endnote-ref-16)
17. For the original version of David Gauthier’s contractarian theory, see David Gauthier, *Morals By Agreement* (Oxford: Oxford University Press, 1986). On the rationality of complying with the above agreements even if there is no threat of punishment for non-complying, see Duncan MacIntosh, "Assuring, Threatening, a Fully Maximizing Theory of Practical Rationality, and the Practical Duties of Agents," *Ethics*, Vol. 123, No. 4 July (2013), pp. 625-656. For a discussion of just how much of our intuitive morality can be captured by this program, see Duncan MacIntosh, "Categorically Rational Preferences and the Structure of Morality," in Peter Danielson, ed., *Modeling Rationality, Morality and Evolution; Vancouver Studies in Cognitive Science, Volume 7* (Oxford University Press, 1998), pp. 282-301. For an application of Gauthier’s ideas to contemporary relations between nations where there is no power compelling the nations to comply with principles they’ve committed to following, see Jens David Ohlin, *The Assault on International Law* (Oxford: Oxford University Press, 2015). [↑](#endnote-ref-17)
18. It is worth pointing out that anyone at the national level who has the power of a disrupter almost certainly also has the power to confer benefits as well or instead – they obviously have an organized society, an infrastructure, a manufacturing base, some level of wealth, and so on, all of which could in principle be recruited to constructive ends and so to the expansion of the shareable co-operative surplus. [↑](#endnote-ref-18)
19. To be tuistic is to be such that one’s welfare is logically affected by the welfare or ilfare of others; and one is malevolently tuistic just in case one’s own welfare rises as that of others falls. What’s crucial to the phenomenon is that your welfare falls not be cause of the circumstance which leads to that of another falling, but because of the very falling of the other’s welfare itself. [↑](#endnote-ref-19)
20. On the idea of tragedies of the commons, see Garrett Hardin, “The Tragedy of the Commons”, *Science*, Vol. 162 (December 1968), 1243-1248. [↑](#endnote-ref-20)
21. I’m arguing that the correct theory of rationality explains how voluntary compliance with arrangements for peace in space can be trusted to be complied with. But suppose people don’t believe this theory: won’t they self-fulfill predictions of their non-trustworthiness? Or is the theory so robustly correct that it will overwhelm this misplaced paranoia? An argument in favor of the latter is that people mostly fulfill the optimistic theory, in spite of their philosophical skepticism about its truth. [↑](#endnote-ref-21)