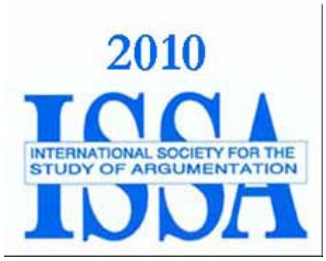
ISSA Proceedings 2010 – Assuring Cooperation: From Prisoner's Dilemmas To Assurance Games To Mutual Cooperation



1. Introduction

How humans should collectively provide for public (and near public) goods – such as, national defense, environmental protection, infectious disease control, and shared moral values – and common pool resources is a topic to which argumentation theorists have paid little attention. Game theorists have usually modeled the problems of providing such goods as a multi-person prisoner's dilemma. Here I will argue that argumentation theorists need to contribute to the understanding of how to deal with both apparent prisoner's dilemmas and with assurance games. I will use classic hypothetical accounts of Thomas Hobbes and Jean Jacques Rousseau to illustrate the problems and the areas to which argumentation theorists should contribute.

2. Prisoner's dilemmas and assurance games

The prisoner's dilemma derives its name from the following story. Row and Column have been accused of some crime. They have agreed with each other not to confess to the crime. But the prosecuting attorney tells Row that if she confesses to the crime and Column remains silent, Row will not be punished. If both confess, both will go to jail for a medium length of time. If both remain silent, both will go to jail for a short time. Of course, since the prosecutor is offering the same deal to Column as she is offering to Row, if Row remains silent and Column confesses, then Row will go to jail for a long time and Column will not be punished. Row must decide whether she should cooperate with Column and remain silent, or defect and confess to the prosecutor. Column also faces this choice.

It would seem that it is most rational for Row to defect from her arrangement with Column and confess to the prosecutor, for if Row defects, she is better off no matter what Column does. That is, if Column defects, Row is better off defecting

(she'll get a medium-length sentence) than she is cooperating (she'll get a long sentence). And if Column cooperates, Row is still better off defecting (she'll receive no time in jail) than she is cooperating (she'll get a short time in jail). The same is true for Column. So if each wants to minimize her jail time, both should defect. But if both defect, both will get a medium-length sentence in jail. If, instead, both had cooperated, both would have had to spend only a short time in jail. The dilemma is simply that by doing what appears to be the rational thing for each to do, both will spend more time in jail than if both had acted irrationally.

		Column	
		cooperate	defect
Row	cooperate	***1***,1	***3***,0
	defect	***0***,3	***2***,2

The prisoner's dilemma in terms of years in jail

If Row wants to stay out of jail, she will defect. If Column wants to avoid jail, she will defect. But if both defect, each will spend two years in jail and collectively they will spend four years. If they both cooperate, they will each spend only one year in jail and collectively only two years. So, if each does that which would appear to keep her out of jail, they (collectively) will actually end up in jail for the longest period of time. (Call such prisoner's dilemmas productive prisoner's dilemmas. The contrast is with destructive prisoner's dilemmas where either cooperate/defect or defect/cooperate outcome is the collectively worst.)

The prisoner's dilemma in terms of the players' preferences:

		Column	
		cooperate	defect
Row	cooperate	***2***,2	***4***,1
	defect	***1***,4	***3***,3

A prisoner's dilemma is any situation in which defect/cooperate, cooperate/cooperate, defect/defect, and cooperate/defect are, in descending order, each player's preference ranking of the outcomes.

The collective action problem of providing for many public goods takes the form of a prisoner's dilemma. Thus peace, either within a society or between societies, refraining from polluting the environment, and having one's children get vaccinated against a potential epidemic all take the form of prisoner's dilemmas. (My refraining from polluting will not, by itself, save the environment and will only cost me extra effort. And my polluting if most others make the extra effort to avoid polluting will not ruin the environment. But that is true for you and for everyone else. So we all pollute and are worse off than if none of us had polluted.)

Game theorists have offered a variety of solutions to prisoner's dilemmas. Hobbes held (in effect) that, without fear of punishment to ensure the existence of devices for creating social cooperation, life for humans would be intolerable. Accordingly, he advanced an authority solution; we should collectively hire someone to institute a system of rules and measures (punishments, primarily) to change the payoffs so that we avoid the undesirable outcome of mutual defection. David Gauthier, the most eloquent and sophisticated of contemporary neo-Hobbesians, has argued that rational individuals seeing that instrumental rationality will lead them to sub-optimal outcomes whenever they face a prisoner's dilemma should change their conception of rationality and become constrained maximizers. Others have offered alternative solutions to the problems posed by prisoner's dilemmas; see, for example, the works of Cave, Danielson, MacIntosh, and Mintoff. But in the real world all the standard solutions to prisoner's dilemmas lead to assurance games. And, in the real world (as opposed to decision theory textbooks), coordinating in assurance games is difficult.

Assurance games are games in which both parties' best outcome is mutual cooperation (cooperate/cooperate). The second-best outcome is lone defection (defect/cooperate). Mutual defection (defect/defect) is ranked third, and lone cooperation (cooperate/defect) is the least-preferred outcome. Thus we get the following matrix.

An assurance game in terms of the players' preferences:

		Column	
		cooperate	defect
Row	cooperate	***1***,1	***4***,2
	defect	***2***,4	***3***,3

Again, the numbers represent the preferences for Row and Column. If we think of

mutual cooperation (cooperate/cooperate) as representing going along with the proposed solution to the prisoner's dilemma being faced by our group, universal cooperation (or as near universal cooperation as is practicable for human beings) is the best outcome for each. But being the only person to go along with the proposed solution is the worst outcome for each. While, in prisoner's dilemmas, individual instrumental rationality argues for defection, in an assurance game it argues both for cooperation (that way may yield the best outcome) and against it (that way may yield the worst outcome).

3. Hobbes's account of the foundation of civil society

In Leviathan, Hobbes tells us of the interaction of a group of individuals, roughly equal in their powers and degrees of vulnerability, who find themselves in circumstances where there is neither law nor morality, circumstances which have come to be called the state of nature. Being thrown together, the individuals are forced to interact, although they are by nature not inclined to cooperate. Hobbes argues that in such circumstances each person will be concerned primarily with his or her own survival. He further argues that, given this concern and the nature of the circumstances and certain general facts about human vulnerability to being harmed by others, each person will find it prudent to attack others before being attacked by them. The unhappy result is that their interaction leads to a condition Hobbes called war, and consequently life for each of them is "solitary, poor, nasty, brutish, and short" (*Leviathan*, Book I, Chapter 13). Hobbes then argues that it would be most rational for each to contract with every other to give his or her allegiance to an authoritarian sovereign in order to end the warfare of the state of nature and improve his or her life.

The circumstances in which Hobbes's contractors find themselves is a prisoner's dilemma, and Hobbes advocates that they adopt an authority solution to that dilemma. In this case it is best for each person individually to defect from paying the costs of a joint project to construct a peaceful civil society and to let others cooperate in paying for that project. This is because the defecting individual gains the benefits of a peaceful civil society without bearing any of the costs. Her second-best outcome is one where she and the others cooperate. In such a case, each person gets the benefits of social cooperation but has to pay some of the costs. The third-best outcome is for each and every one to defect from the project of social cooperation. In this case the defector gets no social benefits, since social cooperation does not occur, but at least she does not pay any costs. Finally, her

worst outcome is to be the lone contributor in trying to produce the benefits which social cooperation can bring and to do so while everyone else defects. In this case, no social cooperation comes into existence because only our lone co-operator has contributed in the attempt to bring it about and one person's cooperation is insufficient to create a cooperative civil society. So she bears the costs of this failed venture and gets no benefits. Since it is the case for every individual that she will be better off not contributing whether the others contribute or not, everyone rationally will choose not to contribute (or, in game-theoretic terms, to defect), and consequently no social cooperation will occur.

An authority solution to a prisoner's dilemma changes the payoff structure so that it becomes more rational to cooperate than to defect. As we have seen in a prisoner's dilemma, each agent realizes that she will be better off defecting than cooperating, no matter what the others do, and this fact leads to universal defection and the state of nature. To achieve the benefits of social cooperation, Hobbes proposes a Sovereign who has the nearly absolute power to alter the circumstances of each member of society so that it is in each person's interest to cooperate with the Hobbesian state. Hobbes's Sovereign, through threats of severe punishment for any defection from the cooperative project to build and maintain a peaceful civil society, changes the payoff structure so that it becomes most rational to cooperate in doing one's part to bring about and maintain civil society.

4. David Gauthier's account

The best neo-Hobbesian account of the rational foundations for morality and civil society is the one provided by David Gauthier in *Morals by Agreement*. Gauthier revises Hobbes's account in two ways. First he holds, in effect, that Hobbes was mistaken in characterizing his contractors as having an overriding concern with ensuring their own survival. Obviously, in order for the contractarian justification for our political arrangements to apply to all rational agents, it must take people as they are, regardless of their preferences. Hobbes's contractors, being primarily concerned with their own survival, are inordinately risk-averse. One can put a smaller premium on personal survival than Hobbes did and still be perfectly rational. Second, Gauthier holds that no external solution to the prisoner's dilemma is adequate. For the contractarian theorist to show that it is rational to accept the constraints of morality, it must be shown not just that it would be rational, in effect, to appoint or hire someone to make the world such that it

would be in our interest to cooperate; rather, the contractarian must show that it actually *is* in our rational self-interest to be moral. That is to say, Gauthier holds that any legitimate solution to the problem posed by the prisoner's dilemma-like structure of human interaction in the state of nature must be an *internal* solution, one that shows that it is rational to be, or to become, moral. Hobbes's solution is *external*, showing only that it is rational to create circumstances where, out of fear of the Sovereign, it is rational to behave *as though* one were a moral person.

Gauthier begins by arguing that instrumentally rational individuals will always defect in prisoner's dilemma situations. He calls such individuals straightforward maximizers. He notes that if individuals could jointly cooperate in prisoner's dilemmas, it would be in the individual interest of each to do so, but that this course of action is not going to be chosen because, for each actor, defecting when others cooperate is still better. Gauthier then argues – and this is his most important contribution to decision theory – that fully rational individuals who foresee that they will be in prisoner's dilemmas with others will change their conception of rationality. Seeing that they are frequently going to be in prisoner's dilemmas and seeing that they will continually get the third-best (second-worst) outcome if they remain straightforward maximizers, they rationally ought to change their conception of rationality and adopt the principle of constrained maximization. A constrained maximizer, as Gauthier calls those who adopt this conception of rationality, is one who maximizes expected utility when in individual choice situations and who, when in prisoner's dilemma games, defects unless she is playing with another constrained maximizer, in which case she cooperates. Thus, a group of constrained maximizers will cooperate to produce socially beneficial outcomes for themselves and they will do so entirely because of considerations internal to instrumental rationality. Consequently, the need for a Hobbesian Sovereign is removed.

From the point of view of game theory, perhaps the most important aspect of Gauthier's argument is that it reveals that the instrumental conception of rationality is far richer than had initially been thought. It may be that the conception of rationality which, on the surface, only tells one how to get what one wants also tells one what the limits of what one can rationally want actually are. This is a Hobbesian result which Hobbes himself never realized.

5. *Hobbesian contractarianism*

We can sum up neo-Hobbesian contractarianism as follows.

(1) We should not presume that morality exists prior to human interaction.
(2) The function of morality is to constrain human interaction to make that interaction more likely to further the interests of those involved.
(3) Individuals in a state of nature are in a prisoner's dilemma.
(4) Such individuals take no interest in the interests of others but seek only to further their own interests (they measure their well-being solely in terms of their own utility).
(5) Such individuals are able to follow long and complex arguments about what to do in the state of nature. In Hobbes's case, the arguments show them that they should pre-emptively attack others and, realizing that this is true for everyone, that they should appoint an authority to impose law and morality upon them. In Gauthier's case, the arguments lead them to change their conception of rationality to make themselves into more cooperative individuals. (6) The chosen social arrangements favour bourgeois stability. (For a more developed statement of these characteristics, see Wein 1986.)

6. Rousseau's critique

In Part II of his *Discourse on Inequality*, Rousseau mounts an insightful critique of bourgeois society. He tries to show that bourgeois social arrangements are attractive, stable, and nevertheless the principal sources of our misery. In the midst of this critique, Rousseau tells what has come to be known as the stag hunt story, a story of a group of hunters who go out into the forest to hunt for game. If each hunts on his own, he will be able to catch a few rabbits and survive. Alternatively, the hunters can cooperate and together hunt for a stag, surround it, kill it, and then eat very well. But if even one hunter abandons the cooperative stag hunt to catch rabbits, the stag will escape through the "hole" that the hunter who has gone after a rabbit has left in the "fence". It is rational for each to continue to cooperate in the stag hunt rather than to defect to hunt for rabbits if, and only if, each hunter has adequate assurance that all others will also continue to cooperate. If any hunter lacks the assurance that all the others will continue to cooperate in the stag hunt, then she should abandon the stage hunt and go chase rabbits. This assurance that the other hunters will hunt the stag rather than chasing a rabbit is something every hunter needs and something that every hunter knows every other hunter needs.

The hunters are in an assurance game. The best outcome for each is for joint cooperation resulting in lots of venison for everyone. The next-best outcome is to

hunt rabbits on one's own. The worst outcome is to continue the stag hunt when even one other hunter has abandoned it to chase rabbits.

So far as the circumstances of the state of nature, and the character of individuals in it, go, Rousseau is actually more hard-nosed than either Hobbes or Gauthier is. About the individual hunter who goes after a passing rabbit, Rousseau says, "there can be no doubt that he pursued it without scruple, and that having obtained his prey, he cared very little about having caused his Companions to miss theirs". So Rousseau's noble savages are completely free of scruples and of guilt or remorse for knowingly doing things that harm others. (The others are harmed in one of two ways. Those who continue the now-futile stag hunt miss their chance to eat. Those who go rabbit hunting are also harmed in that a successful stag hunt is not a real option for them, so their negative liberty is decreased.) Hobbes utilizes emotions (especially fear) to motivate his contractors. Rousseau avoids reliance on this crutch.

In addition, Rousseau thinks that, by nature, humans in such a situation will not cooperate. This is because, unlike Hobbes's and Gauthier's contractors, they are unwilling to follow long trains of reasoning about what is in their individual best interest and thus are such utter strangers to foresight that "far from being concerned about a distant future, they did not even think of the next day". By contrast, Hobbes's natural humans do so much thinking about the future that they work themselves through difficult chains of reasoning to conclude that each should launch a pre-emptive strike against others, a conclusion which leads them collectively into a "war of all against all" in which each of them lives a life that is "solitary, poor, nasty, brutish, and short".

Furthermore, unlike Hobbes's contractors, Rousseau's hunters have no strong emotions to motivate them: "having obtained his prey, he cared very little about having caused his fellows to miss their opportunity". Thus, even though Rousseau's hunters are in a situation in which the cooperative outcome would seem to be easier to attain than it is for either Hobbes's or Gauthier's rational maximizers, Rousseau's hunters do not cooperate. Given this, the common portrayal of Hobbes as tough-minded and Rousseau as soft-minded simply does not wash. We cannot dismiss Rousseau as not being realistic enough – or as being overly optimistic – about the nature of pre-social humans.

From the hard-nosed perspective of contemporary neo-Hobbesian contractarian

theory, there is much to admire in Rousseau's argument. If it is correct, it shows that accounts like Hobbes's and Gauthier's (which are frequently criticized for portraying human nature in an unkind light) are, if anything, overly optimistic. They succeed in showing that cooperation is rational *only if* they imbue their contractors either with strong emotions (as Hobbes does with fear of death) or with a level of prudence which is far beyond our natural capacities. Scholars who have studied the arguments are still divided over what Hobbes's and Gauthier's arguments actually are and whether they succeed. Yet Hobbes's and Gauthier's contractors have to have the ability to follow long trains of reasoning and see that somewhere – perhaps far down the road – it is in their interest to cooperate with each other (whether by appointing a Sovereign to make them afraid not to cooperate, as Hobbes suggests, or by changing their conception of rationality to come to develop commitments to cooperation, as Gauthier suggests). Rousseau shows that, given how humans *actually are*, rationality conceived of as maximization of one's self-interest will not lead to mutually beneficial cooperation even in simple assurance games, let alone in prisoner's dilemmas. So, Rousseau's simple stag hunt story provides the basis for a devastating critique of the entire Hobbesian contractarian project.

7. *Rousseau's assurance*

There is at least one respect in which Rousseau's way of looking at the problem of how to characterize our collective-action problems is deeper than the Hobbesian approach is. Of course, both thinkers set up the state of nature in such a way that there is good reason both for us all to cooperate and for each of us not to cooperate with others. Thus, both capture the core issue confronting those who would offer a rigorous account of human sociability. But Rousseau's account goes deeper in just this respect: every solution to a prisoner's dilemma really just moves one from a prisoner's dilemma into an assurance game. (Of course, no theoretical *solution* turns a prisoner's dilemma into an assurance game, since the term "solution" is a success term, and something that moves people from one game where they will reach a sub-optimal outcome to another where they will also reach a sub-optimal outcome is not a success and hence not a "solution".) But, in practical terms, all the real-life practices that would most closely mimic the various theoretical solutions to the problem of ending up with a sub-optimal outcome in a prisoner's dilemma do lead to problems that are, in effect, best modeled as assurance games. Thus, in practical terms, one always needs to know whether, as a matter of fact (rather than of rational decision theory) enough other

people (or nations, religious groups, organizations, *et cetera*) are swayed by the alleged solution to the problem of the prisoner's dilemma to actually act on that solution and avoid the sub-optimal outcome which occurs when people are ignorant of the solution.

We can see this if we consider the sort of solution offered by Hobbes – namely, an authority solution. If I find myself in a collection of people who are in a multi-person prisoner's dilemma and the possibility of an authority solution arises, I need to ask myself whether enough other people are going to take the authority seriously enough for it to really be an authority. Dealing with this question puts one in an assurance game. I should fear the "authority" if and only if I think enough others will fear it. Otherwise it will not be an authority and, hence, I would be foolish to be the only one to obey it in the current circumstances. Of course, everyone else faces the same question, and so we are collectively in an assurance game.

Similarly, on H.L.A. Hart's account of what it is to be a legal system, there has to be a sufficient number of officials who accept the (potential) rule of recognition as binding before it will actually become the legal system's rule of recognition. But each (potential) official needs sufficient assurance that other (potential) officials will take the (potential) rule of recognition to be binding on them in order for it (the potential rule of recognition) to, in fact, be binding and, hence, for there to be a legal system. Each official is in an assurance game with the other potential legal officials. Of course, Hart's legal theory does not claim that "authority solutions" (in the sense used in game theory) are the only solutions to assurance games. Indeed, Hart devotes much of *The Concept of Law* to showing that authority solutions such as that offered by John Austin are not the only solutions – and, indeed are not adequate solutions – to the problem of the true nature of legal authority. Hart's theory is almost universally understood (by both legal positivists and its critics) as a great legal positivist theory about the concept of law. It is better read as an account of the nature (or concept) of the rule of law. We have an authority if, and only if, enough of us take it to be the case that we have an authority. I should cooperate with others if, but only if, I think enough others will cooperate also. If there are enough others cooperating, cooperating becomes my best outcome. But if an insufficient number of others take as an authority what I think to be an authority, I will be worse off obeying the (supposed) authority and better off to simply ignore it (defect). That is, I am in an assurance game, and so

is everyone else. Similar, though, more complex considerations apply to the splendid planning-based theory of law advanced by Scott Shapiro in *Legality*. A society has a legal system if, but only if, enough members of the society engage in the shared cooperative activity needed to instantiate the complex plan that creates, sustains, and is its legal system.

Roughly the same considerations apply to Gauthier's solution to prisoner's dilemmas. Assume that I find myself in a community of straightforward maximizers who have discovered both the wisdom and the capacity to become constrained maximizers. I need to know that enough others really are constrained maximizers (or are about to become such) before it is rational for me to change my conception of rationality from straightforward maximization to constrained maximization, and I need to be sufficiently confident of being able to correctly sort constrained maximizers from straightforward maximizers. Because everyone else is in the same situation, we collectively face an assurance game.

In practice, communities which find themselves in prisoner's dilemmas where there is a game-theoretic solution to their problem, are always moved into an assurance game. Thus, if in real-life, we are going to solve prisoner's dilemmas, we need to solve the assurance problems that (partial game-theoretic) solutions to them always involve. If a group of us finds ourselves in a prisoner's dilemma where some internal solution is open to us – say, we all come to feel there is a moral duty to cooperate whenever such circumstances arise – then in the real world, where there inevitably are going to be some defectors, each reflective person who finds herself in such a situation must ask herself whether she has sufficient assurance that the number of non-defectors – the number of people who are, as a matter of fact, going to do their duty – is great enough to achieve the benefits of collective cooperation. When she lacks such assurance, she benefits both herself and her society if, like Rousseau's hunter, she refuses without scruple to waste her efforts on what she judges to be a futile collectivist project. Since this is true for all reflective persons in the wake of any internal solution to a productive prisoner's dilemma, we all face an assurance problem whenever we develop a would-be solution to a prisoner's dilemma.

As David Lewis shows in *Convention*, some assurance problems can be overcome through the natural development of appropriate conventions, usually those based on focal point solutions. The connections between Lewis's work and argumentation theory have been usefully explored in Eemeren and Grootendorst

(1984). But, as Joseph Heath suggests in *Following the Rules*, "the theory convention provides, at best, only a solution to the problem of coordination. Focal point solutions, at least of the type . . . Lewis consider[s], have absolutely no bite when it comes to resolving cooperation problems" (page, 58). While Heath's criticism is too strong, as can be seen by examining the work done by Andrei Marmor in *Social Conventions: From Language to Law*, the sorts of conventions Lewis discusses can only do limited work in helping humans avoid or overcome cooperation problems. Furthermore, whatever the role of conventions in helping us overcome some *repeated* situations where sub-optimal outcomes threaten to undermine attempts at cooperation, they play at most a secondary role in dealing with one-shot dilemma games. And, as Hobbes and Gauthier both recognize, rational individuals face quite different problems when confronting one-time prisoner's dilemma games than they do in *iterated* prisoner's dilemmas.

Rousseau both sees the problem of how to explain and justify cooperative interaction among humans more clearly than does Hobbes and starts us on the process of offering a deeper, more satisfying account of how to both explain and justify civil society. This is because, unlike Hobbes and Gauthier, whose thoughts on these matters were always put in terms of individual utility maximization, Rousseau thought about things in terms of basic goods. He pondered such issues in terms of what constraints each person would be willing to impose on herself and would want imposed on those with whom she was interacting, realizing both that it is only through self-imposed constraints that we attain full freedom and that society and civilization depend on the reciprocal acceptance of such basic goods.

8. Basic goods

Rousseau, like Kant, held that true freedom consists not simply in the liberty to do what one wants but in the power to act according to rules or principles one has given oneself. Rousseau thinks of social cooperation not (as Hobbes, Gauthier, and most decision theorists do) simply in terms of how to best further the pre-interaction interests of rational individuals but in terms of what constraints it would be rational to impose on oneself (given that others were going to impose the same constraints on themselves) in order for us all to live in civil society. Basic goods, being those characteristics one would be willing to have in oneself and would want in those one expects to be interacting with, are not discussed explicitly by Rousseau. Yet it is clear that Rousseau's approach most closely

mirrors the basic goods approach. He approaches issues about the value of civil society not by asking whether the proposed social arrangements provide more of what is valued in a state of nature but by asking what arrangements can best serve those of us destined to live among others. In so doing, he sees that civilization needs to be viewed from many angles and that its virtues and vices will not be adequately understood if we simply consider – as Hobbes and Gauthier do – whether joining such a society would be a good deal. In this, he anticipates the idea that a developed society is not simply a wealthy society; rather it is a society where each person has the best opportunity to become as fully civilized as is possible, given the resources available to that society. He wants us to evaluate civil society not by a simplistic metric but by having each of us reflect upon how it can best serve to enrich our very existence. While Hobbesians evaluate society by asking if, when living in society, one has more of what one wanted outside society, Rousseau wants us to reflect on how to arrange our mutual interaction so that it enables us to become fully rational and fully civilized.

Put in terms of the stag hunt story, Rousseau envisages a civil society which not only provides us with more meat but which also ensures that our coming to acquire that meat is – and is understood by all as being – the result of a cooperative endeavour among true natural equals. For the noble stag hunters, rumours that the rabbits are but skin and bones would act as an assurance amplifier, giving each person more reason to continue with the stag hunt than to go off hunting rabbits. By contrast, discovery that the other hunters were buying copies of *501 Ways to Stew a Rabbit* would act as an assurance damper, giving each hunter less assurance that others would continue to hunt deer rather than go off to chase rabbits. While Hobbes employs an external Sovereign to introduce a system of punishment to ensure our cooperation (employing fear *within* civil society much as he employed it in his argument that we should *form* civil society), Rousseau sees development of the capacity to cooperate as constitutive of being a fully civilized person. He also helps us see that we need to design our social arrangements so that they are themselves assurance amplifiers, structures which make us willing co-operators not because we fear what will happen if we fail to cooperate but because cooperating with others best expresses what it is to be a civilized person who is truly free.

9. Progress in social theory

Ideas from both Hobbes and Rousseau can be conjoined to help us see the way to

solving many of our increasingly more pressing global collective-action problems. The old attitude that Hobbesians are so conservative and authoritarian that they have little to contribute to contemporary problems, or that Rousseau's insights are too collectivist for contemporary problems, is both simplistic and untenable. While Hobbes held many very conservative political positions, and while many conservatives have been attracted to Hobbes's approach to grounding political obligation, there is nothing inherently conservative about economic contractarianism. (Not all Hobbes's views were conservative. He was one of the first to hold that the state has an obligation to provide welfare payments to the poor; see *Leviathan*, Part 2, Chapter 30, the section titled *Publique Charity*. For an argument for welfare-state liberalism based on neo-Hobbesian ideas, see Wein 1994.)

Furthermore, with so much of the world's economic activity now embodying the neo-liberal ideology of Hobbesian possessive individualism, those who seek to ensure that civil society retains realms where cooperative, caring enterprises are sustained and nurtured need to look to Rousseau's insights for guidance on how best to amplify the assurance each of us may properly have regarding the cooperative capacities and inclinations of her fellows. It is to our detriment that we neglect either Hobbes or Rousseau. (For an argument that, with the demise of deconstructionism and the plunge in popularity of postmodernism, those who seek to develop a rigorous feminist theory of justice should turn to a combination of the insights of Hobbes and Rousseau, see Wein 1997.)

Of course what actually is an assurance damper or an assurance amplifier is an ultimately an empirical question. But to know what data we need to answer that question, a great deal of very careful conceptual analysis – on matters like the distinction between destructive and productive prisoner's dilemmas, the forms and nature of various basic goods, and the limits of human cooperation – is needed. (How else can we in find the relevant data and ascertain how to read those data?) Just as game theorists need to do more work to figure out how best to model the various collective action problems we now face, it is incumbent upon argumentation theorists to develop the conceptual tools for dealing with that information before can we ascertain which things really are assurance dampers (and how to prevent them from arising) and which things actually serve to act as assurance amplifiers (and how we can best go about nurturing them). It is only after we better understand the nature of the parametric choices that confront us

that will we be in a position to go about dealing with these complex empirical issues**[i]**.

NOTE

[i] I benefited from extremely helpful discussion when I presented these ideas at the International Society for the Study of Argumentation conference in Amsterdam on July 1st 2010. I am grateful for helpful comments from two anonymous referees, and for discussion with Wm. Barthelemy, Duncan, MacIntosh, Malcolm Murray, and especially with Thea E. Smith.

REFERENCES

Cave, E. M. (1998). *Preferring Justice: Rationality, Self-Transformation and the Sense of Justice*. Boulder: Westview Press.

Danielson, P. (1992). *Artificial Morality: Virtuous Robots for Virtual Games.* New York: Routledge.

Eemeren, F.H. van and Grootendorst R. (1984). *Speech Acts in Argumentative Discussions*, Dordrecht: Foris Publications.

Gauthier, D. (1986). *Morals by Agreement*. New York: Oxford University Press.

Hart, H.L.A. (1994). *The Concept of Law*. ed. Joseph Raz and Penelope Bullock, 2nd ed. Oxford: Oxford University Press.

Heath, J. (2008). *Following the Rules*. New York: Oxford University Press.

Lewis, D. K. (1969). *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press.

MacIntosh, P.D. (1992). Preference Revision and the Paradoxes of Instrumental Rationality. *Canadian Journal of Philosophy Volume 22*, No. 4, December.

Marmor, A. (2009). *Social Conventions: From Language to Law*, Princeton: Princeton University Press.

Mintoff, J. (2000). *Is Rational and Voluntary Constraint Possible?*. Dialogue Volume 39.

Shapiro, S. J. (2011). *Legality*. Cambridge, MA: Harvard University Press.

Wein, S. (1986). *Problems with Contractarianism Journal of Social Philosophy*, Volume XVI, #3.

Wein, S. (1994). A Hobbesian Foundation for Welfare Rights. In: *The Liberalism-Communitarianism Debate C.F. Delaney* (ed.) Lanham, MD: Rowman & Littlefield.

Wein, S. (1997). Feminist Consciousness and Community Development. *The International Journal of Social Economics*, Volume 24, Issue 12.

Wein, S. (2007). From MADness to SANity. *Peace Review: A Journal of Social*

Justice, 19 (2), 191 – 196.