UNIVERSITÀ CATTOLICA DEL SACRO CUORE

DIPARTIMENTO DI ECONOMIA INTERNAZIONALE
DELLE ISTITUZIONI E DELLO SVILUPPO

Carlo Beretta

Can common knowledge of rationality make information incomplete? The case of the centipede

N. 0603

V&P

UNIVERSITÀ CATTOLICA DEL SACRO CUORE

DIPARTIMENTO DI ECONOMIA INTERNAZIONALE
DELLE ISTITUZIONI E DELLO SVILUPPO

Carlo Beretta

Can common knowledge of rationality make information incomplete? The case of the centipede

N. 0603

V&P

I Quaderni del Dipartimento di Economia internazionale delle istituzioni e dello sviluppo possono essere richiesti alla Segreteria: (Tel. 02/7234.3788 - Fax 02/7234.3789 - E-mail: segreteria.diseis@unicatt.it).
www.unicatt.it/dipartimenti/diseis

Università Cattolica del Sacro Cuore, Via Necchi 5 - 20123 Milano

www.vitaepensiero.it

ISBN 978-88-343-1687-0

Carlo Beretta

Can common knowledge of rationality make information incomplete? The case of the centipede[1]

[1] Trattandosi di una prima stesura di appunti destinati a studenti, correzioni, suggerimenti e commenti sono particolarmente desiderati.

Abstract

In a centipede game, call reasonable a strategy that, coupled with its best response, leads to the highest payoff among those dominating the Nash equilibrium for the player that adopts it. Reasonableness requires playing pass till the end of the game, the best reply to reasonableness is called an accommodating strategy. Backward induction (bi) makes reasonableness irrational but, given common knowledge of rationality, bi can be applied only if reasonableness is rational. The rationality of reasonableness can be argued as the choice one would make in an underlying game in which the strategies are playing the overt game according to reasonableness or to bi. The underlying game has approximately the form of a stag hunt game, in particular one in which the player who should move last plays reasonable and the other accommodating. This equilibrium dominates that in which both use backward induction. Reasonable players would select such an equilibrium but anyway, the plurality of equilibria injects uncertainty about the reasoning it is rational to use in the overt game.

Common knowledge – rationality – backward induction – reasonableness – incomplete information – centipede game

JEL: C72

Introduction

Substantive rationality, choosing a maximal element given a set of alternatives and an ordering, can be taken as a fact of life. Simple casual observation militates against this. If it is not a fact of life, its use must be justified. When the use is intentional, it is generally justified through its instrumentality in the pursuit of one's aims.

In games, the alternatives are strategies, ordered by dominance. In this context, it is known that straightforward application of substantive rationality can lead to unpalatable conclusions. Examples range from Newcomb's paradox to the centipede game and the finitely repeated prisoner's dilemma. One assumes that a rational person cannot but be substantively rational; one claims that substantive rationality implies deletion of dominated strategies and thereby the use of backward induction; then, substantive rationality with enough knowledge necessitates the conclusions (Dekel – Gul (1997)).

A well known critique, advanced, for example, in Bicchieri (1989) and (1993), is that this requires the use of counterfactuals: one has to justify one's decision by behaviour at a stage one knows that will never be reached if the assumption of common knowledge of rationality is satisfied. In the case of the centipede, the use of such a counterfactual is not simply justifying one's decisions through irrelevant eventualities. To use repeatedly backward induction, one must assume that a certain stage will, or at least might, be reached; this implies that, up to that stage, one of the commonly known to be rational players has behaved "irrationally" and one discounts the fact that after that stage that same player will start to behave "rationally". If one cannot give a reason for this inconsistency in the way in which he decides, this adds irrationality in the behaviour of the "rational" player in question.

To avoid the bleak predicament to which common knowledge of rationality appears to lead, one suggests destroying it by playing "irrationally" as soon as one can. On the one hand, if the move is successful, one is left with no hunches on what to do of the "irrationality" one has injected (Reny (1992) and (1995)). On the other, one can find reasons, perhaps even rationality, in the decision to play "irrational", so whether common knowledge of rationality has actually been destroyed is quite unclear.

If one takes Bicchieri's point seriously, however, one does not need to destroy common knowledge of rationality in order to prevent the use of backward induction; the latter simply cannot be used in such circumstances. But, without backward induction, rationality by itself is unable to say anything about how to play.

The usual way to reconcile actual, admittedly sensible, behaviour in games of this kind is to inject some doubts about the rationality of the other player (Kreps, et alt. (198)). Following Bicchieri, here it is argued that it is precisely the common knowledge of rationality that makes uncertain how a rational person would interpret and play the game in question.

After introducing the notion of "reasonableness", the kind of behaviour it leads to is discussed. Reasonableness[2] is necessary to have ground for applying backward induction but backward induction beats reasonableness, and makes the latter apparently "irrational" in the overt game. To reconcile reasonableness with substantive rationality, it is argued that the choice on whether to follow reasonableness or backward induction is made in an underlying game which has the characteristics of a stag hunt, and therefore has more than one equilibrium.

In this paper, only overt games of the centipede type are considered, but the reasoning applies to many multistage games that have just one Pareto inefficient Nash equilibrium in dominating strategies, reached by deletion of dominated strategies.[3]

---

[2] Or something leading to the same behaviour.

[3] A companion paper deals with the case of the finite prisoner's dilemma.

Centipede games

Call 1 the first player to move and 2 the other. Fig. 1 gives the usual representation of the centipede game for N even.
It is easy to check that, if stage n is reached and down is chosen at it, then player 1 gets n/2 +1/2, and player 2 gets n/2 – 1/2, if n is odd, while player 1 gets (n/2) – 1 and player 2 gets 1 + (n/2), if n is even.

Common knowledge of rationality[4] is taken to entail that both players use backward induction and this implies that the first player to move has to stop the game as soon as he has the chance to do it, whatever gains might be obtained if the game went on till the end. I will refer to this as the backward induction strategy or view, bi for short.

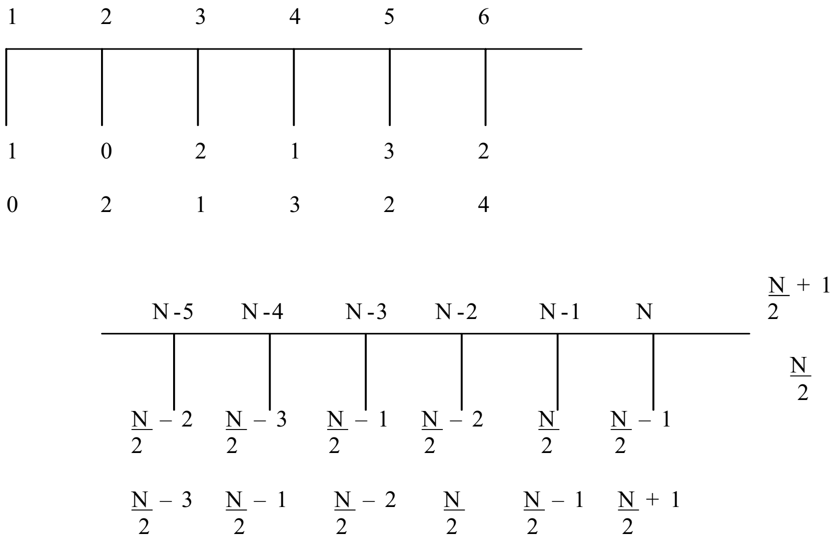

Fig. 1

[4] Or better, common knowledge that both will adopt what appears to be the substantively rational strategy in the overt game, since whether this is the substantively rational thing to do is what is questioned.

A definition of reasonableness

In this game, for most of the stages each player has strategies which admit a best response by the other that would lead to a result which is better for both than that obtained by following bi. In the case at hand, there are very many strategies of this kind, but they can be easily ordered. Call *reasonable* any undominated strategy with this property. "Always pass" comes out on top.

The problem with this strategy is that it seems one cannot believe both that a player is rational and that he will actually implement it. There is however a strategy, discussed in Appendix 1, that, if he knows to be taken as rational, a player has reasons to implement[5] up to a certain stage, perhaps is uncertain on whether to go on to implement it or shift to bi at some stage(s),[6] knowing that from some stage onward both players, if they are rational, must follow bi. As it will be seen, this strategy is consistent with playing pass "almost to the end of the game".[7] In particular, in the case of fig. 1, it is consistent with player 1 playing pass till the stage N – 7 included, makes him doubtful on whether to play pass or down at stage N – 5, definitely discourages playing pass at N – 3.

The decisions to play pass, to behave as if randomising between pass and down, or to play down, are taken consistently, by using always the same reasoning framework, without the use of conterfactuals, and always with the aim of maximizing one's payoff. What is claimed is that its adoption is consistent with the knowledge of rationality.

A *reasonable player* is one that in an overt game always chooses a strategy which, coupled with its best response, leads to a maximal pay-off for the player in question and dominates the Nash equilibrium. A reasonable player, then, always chooses a reasonable strategy which can be instrumentally justified, when such a strategy exists; otherwise, he plays a Nash equilibrium strategy. A best response to the strategy followed by a reasonable player is called an

---

[5] If he can, i. e., if the second does not play down before.

[6] And which, therefore, generates uncertainty also in the other player.

[7] This point will be made precise in a short while.

accommodating strategy.

Almost by definition, when it has any scope, a reasonable strategy is not a best response to the best response to itself and, in this sense, reasonableness forbids any use of backward induction by the player that adopts it. To make the choice of reasonableness consistent with rationality, player 1, for example, must have reason to believe player 2 will adopt a best response to 1's strategy, so 2 must be aware of the possibility of reasonableness, have at least a rough[8] idea about which sequence of moves in the overt game its adoption implies for 1, but, above all, be convinced of the rationality of 1 in using reasonableness. 2 does not need to give up using backward induction, but the belief that 1 plays reasonable limits[9] its use by 2 to that necessary for discovering his own best response. 1, then, cannot exclude 2 will distrust 1's reasonableness and play down as soon as he is given the chance[10] but knows that, if 2 accepts 1's reasonableness, 2 will limit the use he will make of backward induction.

There are two steps in the reasoning that argues that reasonableness is consistent with rationality.

The first step relies on the fact that to support the rationality of playing down as soon as one has the chance, either one needs to use a counterfactual inconsistent with the assumption one wants to prove[11] or to show that reasonableness, besides dictating a strategy which is consistent with the assumption that one will reach almost the end of the game, without introducing counterfactuals is aalso consistent with substantive rationality.

Appendix 1 shows that assuming that any stage up to N – 2 might be reached is not inconsistent with reasonableness while it is with backward induction and defines what to do on grounds that are not logically weaker than those of backward induction.[12] This means a player has at least two alternatives between which to choose in order

[8] But, as it will be shown, not very precise.

[9] If 2 wants to exploit the reasonableness of 1.

[10] Possibly because he pushes backward induction to its limit.

[11] Or, equivalently, to assume that, at some stage, one takes a decision based on a reasoning inconsistent with that which must have been followed in order to reach that stage.

[12] And is independent from it.

to decide how to play the game: play down as soon as one can or follow reasonableness. If both are consistent with rationality, substantive rationality is to be used not in discriminating which of them satisfies its requirements, but in choosing which of them must be adopted.

Appendix 2 shows that, for most of the game,[13] playing tricks, feigning to be reasonable and therefore playing pass at the current stage in order to lure the other to play pass and then play down as soon as one's turn comes, is inconsistent with reasonableness and is dominated by it, at least till almost the end of the game.

Consistency of reasonableness with rationality

The second step of the argument goes as follows. Suppose to be at stage 1. At that stage, actually even before the first move, both have to decide what to do in all possible future situations and, at that stage, 1 knows that, 2 is in a very similar position to his own with respect to deciding whether to play reasonable (or accommodating) or use bi and play down as soon as he has a chance. Assuming 1 did have a choice between adopting backward induction, and therefore following the bi strategy in the overt game from the start, and playing in accordance with reasonableness, at least for some time, one must discuss which is the substantively rational choice.

If 1 knows 2 is bound[14] to follow backward induction and has no choice about it, or that one's deviation from the strategy implied by backward induction will not[15] affect the belief of 2 about the fact that 1 is bound to use backward induction from the immediately following stage onward, and so on, 1 must use backward induction. However, to assume that 2 uses backward induction and is rational, implies that either he must doubt the rationality of 1 or must assume that there is a strategy a rational 1 can follow which will allow to reach that stage. The only possibility is for 1 always to play pass, but this is not a rational strategy so this possibility must be dis-

[13] Actually till stage N – 5.

[14] But could this be justified as a consequence of substantive rationality?

[15] But can it not?

carded. Given common knowledge of rationality, 1 knows that 2 knows stage N – 1 cannot be reached, but, as argued in the Appendices, common knowledge of rationality is consistent with both believing that stage N – 7, and possibly even stage N – 5, can be reached if they find a way to make reasonableness consistent with rationality.[16]

The rationality of reasonableness can be argued by moving from the choice of the strategy to employ in the overt game to the choice made in an underlying game in which the alternatives are those of following reasonableness and following backward induction. In this game, 1 knows that, if 2 knows about the possibility of reasonableness, they are playing a stag hunt game. As Appendix 3 shows, they both can determine the payoffs they would get if the player who should have the last move, if that stage is reached, played reasonable, while the other played accommodating. They know that playing accommodating to a reasonable opponent is better than playing consistently reasonable. But they also know that, if both are consistently rational, nobody will consider playing accommodating to an accommodating player.[17] But playing accommodating to a reasonable opponent leads to the same strategy in the overt game as playing reasonable till the very final stages.

There are then three equilibria, two in pure strategies, in the first of which one player plays either reasonable and the other accommodating, in the second both use backward induction, besides an equilibrium in mixed strategies. Which equilibrium will be selected is indeterminate.[18] In this game, the adoption of reasonableness cannot be discarded on rationality grounds and is instrumentally superior to immediate defection.

Assume 1 opts for reasonableness and plays pass at stage 1. It is the turn of 2 to make his move. He can choose to think irrelevant the

[16] Notice that if this is impossible, they cannot use backward induction, so that they would be in the situation discussed in Reny's papers.

[17] To consider accommodation to accommodation would in fact declare reasonableness irrational, but if reasonableness is irrational no rational player can choose it.

[18] though reasonableness would point to the Pareto dominating one.

move by 1 and simply adopt the strategy chosen before any move was made. Observing a behaviour at variance with that dictated by backward induction, one can claim that the only thing 2 can say is that that strategy has not been followed.[19] But if mistakes and irrationality are ruled out and 2 knows 1 is rational, knows about the possibility of being reasonable, and can see that reasonableness[20] can be an instrumentally justified choice for 1, deviation from bi by 1 must induce 2 to consider the possibility of 1 playing reasonable or accommodating to a reasonable 2.[21] In other words, if 2 had decided to use backward induction and play down, he has reasons to revise his choice. If 2 had chosen to play reasonable or accommodating to a reasonable 1, he would be strengthened in his conviction.
Till any agent plays down, stage after stage they know they are called to play a centipede game that differs slightly from the original one. At stage n, the agent that plays pass shifts to the other a gain of 1 and asks him to decide whether to play reasonably or according to backward induction a centipede game of length $N - n - 1$, knowing that when and if stage $N - 1$ is reached, one must follow backward induction. Stage after stage, what one has to give up remains constant, while the additional maximal potentially reachable payoffs diminish. In the mixed strategy equilibrium, the probability of backward induction must then increase. But the reasoning given above in favour of reasonableness remains unaltered.
How long one will go on playing reasonable or accommodating to a reasonable opponent and therefore the strategy actually implemented

[19] 1 may simply have made a mistake, if mistakes are allowed. Or 1 may be an "irrational" player, meaning a 1 that decides what to do without considering properly the characteristics of the game,[19] but then, to decide what is best for him to do, 2 must form ideas about the process which generates mistakes or the kind of irrationality 1 is affected from. All these cases, however, require considering the behaviour observed as not, or at least not necessarily, intendedly chosen as instrumentally efficient for the pursuit of the aims of the agent that implements it, as not consistent with the substantive rationality of 1.
[20] Playing pass "almost" to the end of the game.
[21] As shown in Appendix 2, the possibility that 1 is playing a trick can be discarded till the final stages of the game.

can be left undefined.[22] Some fuzziness on this point is not important.[23] What is relevant is that, if 1 needs to consider what to do if stage n > 1 were reached and believes that rationality requires consistency in the motivations of one's choices, following reasonableness seems a likely, if not the only, possibility. 2 may not believe that reasonableness will be followed till the very end of the game but this does not mean that the assumption of rationality forbids him to believe 1 could use it at least for some time. [24]

Closing remarks

Actually, one way of looking at the game is that what players must decide is just the moment in which it is optimal to distrust the other will go on being reasonable or accommodating. It is optimal to start distrusting just before the other does, but to distrust before this moment leads to losses of potential gains. The payoffs associated to reasonableness give a measure of the potential gains that can still be reaped, which must be traded off with the sure gains that playing down can give. In a world of backward induction, this game is not available;[25] in the world of reasonableness, it is and it is not worth to consider distrusting till almost the end of the game.[26]

In fact, what reasonableness does in this game is to point out that, in the situation at hand, there is an element of conflict embedded in an environment in which cooperation can be very rewarding. One can play out the conflict immediately or delay it in time. Playing it out immediately gives one of the players the upper hand, at the cost of giving up the potential benefit of delaying it. Reasonableness is simply the choice of delaying it as much as possible, consistently

---

[22] A reasonable 1 can play pass till N – 3 included or stop at N – 5.

[23] Actually it may even be useful making uncertain when it is time for 2 to start using backward induction.

[24] Or, if one prefers, has a trivial solution.

[25] Or, if one prefers, has a trivial solution.

[26] Given the substantial similarity of the games in this sequence, till that stage is reached, the way in which the preceding trust games have been played can reinforce one's decision to trust the other is trustful and trustworthy.

with preserving rationality.

The main advantage of reasonableness is that it is a natural assumption to make when the instrumental rationality of backward induction, the "substantive rationality" in the choice among strategies in the overt game, is in doubt, or outright violated, but the conditions of knowledge usually stipulated about the rationality of the other player still hold. Furthermore, in the particular case in exam, the reasonable strategy to adopt is easy to figure out at least broadly.[27] Admitting reasonableness associates to the overt game an underlying sequence of games in the choice of rationality that has the structure of a trust game, actually, it transforms the choice of strategy in the overt game in choices on how to play a sequence of trust games: one knows the sequence will end, actually that the trust game will degenerate into a game in which giving up reasonableness in the underlying game and choosing the substantively rational strategy in the overt game become dominant strategies, but there is uncertainty on when exactly this will happen.

Furthermore notice that, to determine whether playing pass is justified, one needs to consider only the end of the game. This allows reasonable players to have preferences among centipedes of different length. A rational player that sees no alternative for a rational agent to using bi cannot do this.[28] A reasonable player always prefers[29] a game with N + 2 stages to one with N, so he strictly prefers a centipede game of length 8 to one of length 6. Therefore, at stage N – 7, to play reasonable for 1 is strictly preferred to playing tricks, and 2 can be sure a reasonable 1 will never play down till that stage is reached, while 1 will be uncertain on whether to play down or to play tricks at stage N – 5, whether he thinks 2 is an accommodating or a reasonable player.

Finally observe that, to induce cooperation, asymmetry of information is not necessary; actually, it is common knowledge of the essen-

---

[27] And, obviously, it is easy to check whether the behaviour kept tallies with reasonableness or not.

[28] Provided he has always the same role, i. e., he is always the first or the second mover.

[29] Of course, if reasonableness is given a chance by the other player.

tial similarity of their situation that seems to help in this regard.
With respect to previous literature, this paper uses Bicchieri's criticism of the use of counterfactuals to argue that, in order to avoid it, one has to introduce a strategy that can be consistently justified till the end of the game: playing a reasonable strategy is an answer to this question. It avoids at least in part the indeterminacy that playing pass has in Reny's construction, allows to preserve in the other player confidence in the rationality of the agent that played pass and therefore allows him a rational reply. Finally, it allows to generate uncertainty endogenously, instead of having it exogenously introduced as in Kreps et alt.(1982).

References

Bicchieri C. (1989) Self-refuting theories of strategic interaction: a paradox of common knowledge, *Erkenntniss*, vol. 30, pp. 69-85

Bicchieri C. (1993) *Rationality and coordination*, Cambridge University Press, Cambridge

Dekel E. - Gul F. (1997) "Rationality and knowledge in game theory" in Kreps D. M. - Wallis K. F. (eds.) *Advances in economics and econometrics: theory and applications*, Cambridge University Press, Cambridge, vol. I, pp. 87-172

Kreps D. M. - Milgrom P. - Roberts J. - Wilson R. (1982) "Rational cooperation in the finitely repeated prisoners' dilemma", *Journal of Economic Theory*, vol. 27, pp. 245-252

Reny P. J. (1992) Rationality in extensive form games, *Journal of Economic Perspectives*, vol. 6, n. 4, pp. 103-18

Reny P. J. (1995) Rational behaviour in extensive form games, *Canadian Journal of Economics*, vol. XXVIII, n. 1, pp. 1-16

Appendix 1: Playing reasonable

If N is even, 2 has the last move. At N, a rational 2 has no choice but playing down and whether 1 is rational or not is irrelevant for the choice of 2 at this stage.
However, if one is at the beginning of the game, to consider what to do at stage N requires having reason to assume that it will ever be reached. In planning to play down at stage N, 2 is not using backward induction, so he is not inconsistent with not having used it before getting there. But 2 must assume that neither he nor 1 have used backward induction and that they had reason not to use it before. Anyway, reaching this stage is inconsistent not only with backward induction but also with the reasonableness of 1 at stage N – 1 and with 2 being reasonable at stage N – 2.[30] Reaching this stage is then inconsistent with the assumption that each of them is instrumentally rational, whatever content is given to rationality. A consistently justified strategy can tell what to do if stage N is reached but should not use the fact that N will be reached in its justification.
To make it sensible to consider the possibility to have to decide at stage N – 1, 1 too must assume that it can be reached when players use rational strategies. Also in this case, a consistently justified strategy can tell what to do if stage N – 1 is reached but should not depend on N – 1 being reached for its justification.
Both if he uses backward induction or follows reasonableness, he must play down. If he uses, actually is the first to use, backward induction, he must play down because this is the last time he has a move and knows what is optimal for 2 to do if he plays pass. But playing down is also the reasonable choice, since if he plays pass, the best response of 2 would make 1 worse off. In these circumstances, ignoring backward induction and playing pass would simply declare him irrational, but that would not change what is optimal for 2 to do if stage N were reached. So an instrumentally rational 1

---

[30] If 2 plays pass at stage N – 2, he knows that even a reasonable 1 must play down at stage N – 1, and all the more so a 1 which follows backward induction.

has no alternative and must play down.
If backward induction is the only reasoning justifying his decision, using it at this stage is however inconsistent with not having used it before, and inconsistency is at odds with the fact that 1 is rational, knows it, and knows that his rationality is common knowledge. If he wants to go on believing to be rational, he must find a rational explanation for why he did not use backward induction before. If he is reasonable, he has such an explanation for himself; it can be shown that a reasonable 1 can play pass at N – 3, and therefore that changing move[31] does not entail inconsistency in the justification of this decision on his part. However, given common knowledge of the rationality of both players, he must find such an explanation not only with respect to himself, but also as concerns 2, and playing pass at N – 2, requires inconsistency both with the use of bi and of reasonableness by 2 at that stage.[32]
Reaching stage N – 2 is inconsistent with rationality, if backward induction is the only way of reasoning admitted, since it entails that till that stage nobody used backward induction. However, it could be reached if both players are reasonable, or if 1 is reasonable and 2 plays accommodating. At this stage, 2 must play down, both if he is a backward induction type or if he is reasonable but a backward induction type shows inconsistency, while a reasonable one does not. So, considering what to do at stage N – 2 involves a counterfactual in a bi world, while it does not in a world in which reasonableness is given a chance, i. e. if both are reasonable, or 1 is reasonable and 2 plays accommodating.
Reaching stage N – 2 assumes a pass at stage N – 3. Reaching this stage is consistent with reasonableness of both players, or with player 1 being reasonable and 2 playing accommodating; it is not consistent with rationality if this implies that any of the players must be a bi type. At stage N – 3, 1 has "always play pass" as a reasonable strategy;[33] he knows that 2 knows it and cannot discard its

[31] Going from pass at N – 3 to down at N – 1.
[32] Furthermore, to believe that 2 is rational is necessary in order to know what 2 must and will do at stage N.
[33] If 2 plays the best response to this strategy, 1 would be no worse off than

adoption without attributing to 1 the use of backward induction, i. e. without disbelieving the reasonableness of 1.
If 1 is consistently reasonable, he will play pass up to stage N – 3 included; and for 2, to play pass till stage N – 4 is consistent both with him playing reasonable or playing accommodating against a reasonable opponent. So, if reasonableness is consistent with rationality, considering what to do at stage N – 3, actually even at stage N – 2, does not involve the use of a counterfactual: it is just one of the situations which could arise, consistently with all the assumptions being satisfied.
To be consistently reasonable, 1 must play pass till N – 3 included. The best accommodating strategy for 2 is then to play pass till stage N – 2 is reached, when he will play down.
To be consistently reasonable, 2 must play pass till N – 2 is reached, when he will play down. The best accommodating strategy for 1 is to play pass till stage N – 3 is reached, when he will play down.
Notice that for 2, consistent reasonableness and accommodation to a consistently reasonable 1 lead to the same strategy; for 1, reasonableness implies a pass at N – 3, while accommodation requires a down at that stage. For 2, consistent reasonableness is no better than backward induction from stage N – 4 onwards, while for 1 this happens at N – 3.

---

playing down at stage N – 3, while 2 would be strictly better off.

Appendix 2: Playing tricks

One can ask whether it is rational, or at least instrumentally justified, to be consistently reasonable and play pass till stage N – 3 included. Playing reasonable means persisting in not using backward induction, but that is what both must have been doing all along. If one is seriously considering to have to decide at that stage, what has to be justified is not the giving up of backward induction, since it has never been used, but whether 1 will stick to the reasonable strategy also for the rest of the game or has now reasons to abandon it he did not have before.

1 can play pass not with the intention of sticking to the reasonable strategy but simply to trump and play a trick on the other, i. e. to play pass at the current stage intending to play down as soon as he is called to move again. Playing a trick can be justified without using backward induction, actually, the sheer possibility of playing it assumes backward induction is kept at bay. What keeps 1 from playing pass at N – 3 is the fact that he knows 2 knows that, for 1, playing a trick[34] is better than playing reasonable till the end of the game. This makes optimal for 2 to play down at N – 2, if there were a pass at N – 3. So, at this stage, 1 has a reasonable strategy but he knows its adoption and implementation are not credible.

At N – 4, 2 has both the possibility of playing reasonable and of playing a trick. Playing a trick gives him a higher payoff than playing reasonable, but what is interesting is that playing reasonable, i. e., playing pass also at N – 2, gives him the same payoff than playing down immediately. In a sense, playing reasonable, if the other plays accommodating or reasonable, leaves him no worse off than using backward induction at stage N – 4.

At N – 5, 1 knows that, for 2, playing a trick at N – 4 is better than playing consistently pass; both these choices would however require 2 to play pass at N – 4 and justify a pass at N – 5.[35] But 1 also knows that, for 2, playing consistently pass is no better than playing

[34] I. e., playing pass at N – 3 and then playing down at N – 1, if that stage is reached.

[35] Giving 1 the possibility to play down at N – 3.

down at N – 4,[36] and this is the reason that can decide 1 to play down immediately. In a sense, 1 is not forced by the knowledge of the rationality of 2 to assume that 2 will use backward induction at stage N – 4, though he cannot discard this possibility.

At N – 5, 1 can play pass either because he chooses to be reasonable or because he tries to play a trick. What is new is that playing a trick successfully does not give a higher payoff than sticking to reasonableness also at stage N – 3, if he can assume he will be taken as reasonable at that stage. In any case, both strategies imply a pass at N – 5, and a pass is irrational for 1 only if 2 is sure to use backward induction at stage N – 4, but, by the reasoning above, there is no certainty on this point: a bi type knows he would never reach that stage and if that stage is reached, backward induction cannot have been used.

At stage N – 6, again 2 has both the possibility of reasonableness and of playing a trick, besides that of using backward induction. Both reasonableness and tricks imply a pass, both promise higher payoffs than immediate use of backward induction and now playing a trick is strictly dominated by playing consistently pass. Backward induction is necessitated only if it is sure that 1 will not play reasonable or tricks at N – 5, but of this there is no certainty.

The reasoning can then be repeated stage after stage till the first. It is obvious that playing a trick too early in the game is inconsistent with rationality. Furthermore, if there are still enough stages to go, i. e., if N – n is sufficiently large, provided it can be argued that reasonableness is not inconsistent with rationality, no player is constrained by his own rationality and common knowledge of rationality to use backward induction and knows the same will be true also for the other player at the stage immediately following the one considered, especially if there is a pass at the current stage.

Reaching stage N – 7, then, comes to depend only on the fact that 2 can believe that 1 is reasonable, at least up to that stage. What has to be shown is that if 2 believes that 1 is substantively rational in the pursuit of his aims and observes that 1 has played pass, he has no choice but to believe that 1 is reasonable: if 1 is rational, only if he

---

[36] In a sense, no better than using backward induction at that stage.

is willing to use a reasonable strategy can play pass till this stage.

All the reasoning comes down to claiming that, if a player has always used pass in all previous stages, one needs a reason to justify his shifting to the use of down. One cannot use backward induction alone, since this would make the player inconsistent and therefore irrational. The reason is found in the fact that, at some stage, playing a trick comes to dominate playing consistently pass in the future. There is a stage at which the two strategies are indifferent: if at this stage the player in question plays pass, this indifference makes uncertain for the other player whether it is better to play pass or down when it comes his time to move. What motivates the fact that at some stage one will play down is not backward induction but the ordering of playing a trick with respect to playing consistently reasonable till the end.[37]

Appendix 3: In support of reasonableness

If reasonableness cannot be discarded through rationality when playing the underlying game, actually it is required also at that level to be able to justify using bi, then reasonableness can be considered in deciding how to play in the overt game. In this game, what 2 has to decide is whether to accommodate and use the best response to reasonableness, which will be denoted as strategy a, to shift he too to reasonableness (rea), when this is possible, or, finally, to play immediately the apparently substantively rational strategy, bi.[38]

For player 2, choosing reasonableness comes however in two variants. He can opt for consistent reasonableness,[39] or he can choose to be reasonable till the stage in which persisting in reasonableness does not promise more than starting to use backward induction, at which point he shifts to bi or randomises in the choice between reasonableness and bi.[40] As it has been seen, in this case, choosing ac-

[37] Uncertainty about the stage one will shift from pass to down would anyway make difficult the use backward induction.

[38] Play down as soon as one has the choice.

[39] And play pass till N – 2 is reached and then play down.

[40] And if bi is finally chosen, he will play down at N – 4.

commodation to a persistently reasonable opponent leads to the same choice as being consistently reasonable.
Similarly, 1 can choose to be reasonable till reasonableness becomes indifferent to using backward induction and at that stage start to randomise or outright to use backward induction,[41] or he can choose to be persistently reasonable.[42] For 1, the accommodating strategy to a consistently reasonable opponent is identical to that associated to the use of backward induction at N – 3.[43]
Different combinations of choices will lead to different payoffs, but, for simplicity I will group both a and rea under the banner R. Of course, any choice within R implies a pass at least till stage N – 4 is reached, so they will be indistinguishable for most of the game.[44]
Can 2 have reason to use R at the following stage? To answer this question is important also for determining whether it can ever be rational for 1 to choose R and, in turn, the reasons 1 can have for his choice are important for the choice of 2.
Consider the choice to be made at any stage $1 \leq n \leq N - 3$.

---

[41] In which case he can play down at N – 3.
[42] And then play pass also at N – 3.
[43] One could introduce an accommodating strategy for the possibility the other starts to use bi from the stage at which this is not dominated by consistent reasonableness, but this would unnecessarily complicate the reasoning without introducing new possibilities. I exclude the possibility for both agents to consider adopting the best response to a best response by the other because it would require inconsistency in the choice of actions at different stages on the part of a player assumed to adopt reasonableness, and this would contradict the knowledge of rationality. Besides, if a player admits this possibility for himself, he must admit it for the other player; then iterated delation of dominated strategies would start a process of backward induction that would lead to the use of bi from the start, a possibility already considered.
[44] They can pass also in order to play a trick, but this would be irrational for most of the game.

n odd

1 \ 2	d_{n+1}	d_{n+3}	d_{N-4}	d_{N-2}
d_n	$\frac{n-1}{2}+1;\ \frac{n-1}{2}$	$\frac{n-1}{2}+1;\ \frac{n-1}{2}$	$\frac{n-1}{2}+1;\ \frac{n-1}{2}$	$\frac{n-1}{2}+1;\ \frac{n-1}{2}$
d_{n+2}	$\frac{n+1}{2}-1;\ \frac{n+1}{2}+1$	$\frac{n+1}{2}+1;\ \frac{n+1}{2}$	$\frac{n+1}{2}+1;\ \frac{n+1}{2}$	$\frac{n+1}{2}+1;\ \frac{n+1}{2}$
d_{N-3}	$\frac{n+1}{2}-1;\ \frac{n+1}{2}+1$	$\frac{n+1}{2};\ \frac{n+1}{2}+2$	$\frac{N}{2}-3;\ \frac{N}{2}-1$	$\frac{N}{2}-1;\ \frac{N}{2}-2$
d_{N-1}	$\frac{n+1}{2}-1;\ \frac{n+1}{2}+1$	$\frac{n+1}{2};\ \frac{n+1}{2}+2$	$\frac{N}{2}-3;\ \frac{N}{2}-1$	$\frac{N}{2}-2;\ \frac{N}{2}$

Fig. 2

Figg. 2 and 3 give the payoff matrix, distinguishing the case in which n is odd or even. Each d has an index that says in which stage the player in question plans to play down. For example, in the case of fig. 2, at n it is 1 to move: d_n is to play down immediately,[45] d_{n+2} is to play down at n + 2,[46] and finally d_{N-3} (or d_{N-1}) is the decision to play down at stage N – 3[47] (or N – 1).[48]

[45] The strategy associated to immediate use of backward induction.
[46] The strategy associated to deciding at stage n to play a trick.
[47] Play the accommodating strategy or use backward induction from that stage.
[48] The strategy implied by consistent reasonableness.

n even

2 \ 1	d_n	d_{n+2}	d_{N-4}	d_{N-2}
d_{n+1}	$\frac{n}{2} - 1$; $\frac{n}{2} + 1$	$\frac{n}{2} - 1$; $\frac{n}{2} + 1$	$\frac{n}{2} - 1$; $\frac{n}{2} + 1$	$\frac{n}{2} - 1$; $\frac{n}{2} + 1$
d_{n+3}	$\frac{n}{2} - 1$; $\frac{n}{2} + 1$	$\frac{n}{2}$; $\frac{n}{2} + 2$	$\frac{n}{2} + 2$; $\frac{n}{2} + 1$	$\frac{n}{2} + 2$; $\frac{n}{2} + 1$
d_{N-3}	$\frac{n}{2} - 1$; $\frac{n}{2} + 1$	$\frac{n}{2}$; $\frac{n}{2} + 2$	$\frac{N}{2} - 3$; $\frac{N}{2} - 1$	$\frac{N}{2} - 1$; $\frac{N}{2} - 2$
d_{N-1}	$\frac{n}{2} - 1$; $\frac{n}{2} + 1$	$\frac{n}{2}$; $\frac{n}{2} + 2$	$\frac{N}{2} - 3$; $\frac{N}{2} - 1$	$\frac{N}{2} - 2$; $\frac{N}{2}$

Fig. 3

Looking at payoffs, d_{N-3} dominates d_{N-1}, for player 1, while d_{N-4} dominates d_{N-2}, for player 2.[49] This reflects the fact that both concern stages in which persisting in reasonableness does not promise higher gains than starting to use backward induction, unless the other too persists to be reasonable, i. e. in which it is optimal to start to use backward induction if the other does not persist in reasonableness. Deletion of the dominated strategy requires use of backward induction, and a reasonable person has alternatives to this choice which do not violate rationality and, if both persist in reasonableness, they get higher payoffs than if the first to reach indifference adopts bi. However, for most of the game, one can reason as if only the first and third row and column were relevant.

At each stage, what one observes of a player is just if he plays pass or down. This does not reveal much about the strategy chosen by the

[49] Once d_{N-1} has been deleted.

player unless he adopts bi and ends the game; otherwise, he can play pass because he has chosen one of the alternatives in the last two colums or row, lumped together in R.[50]

If player 1 plays pass at stage n, he is declaring he accepts to pay 1 for the possibility of playing a centipede game of length N – n – 1, with the other player having the first[51] move. At each stage, the loss to playing R with respect to choosing bi when the other chooses bi for the stage immediately following is 1, while the gain to going bi just before the other is also limited to 2, both small compared with prospective gains at least when there are still enough stages to go. Furthermore, the dominated equilibrium in pure strategies is not even strict.

The move chosen at the each stage is just the result of the decision on how to play such a finite sequence of trust games, in which players alternate in the role of row and column player, all with a similar payoff matrix in which what changes is just n, the stage reached, and with it the level of payoffs associated to what remains of the game. Also this sequence of games can be solved backwards, and so one can conclude that each player must play bi as soon as he has a chance, but only at the cost of using a counterfactual which implies inconsistent behaviour from a person, oneself included, known to be rational. If one insists on consistency, one is bound to assume that also this sequence of games can be played reasonably, and then that R can be adopted in the overt game, almost till the end.[52]

---

[50] One could also play a trick, but a rational player would not do this till n almost reaches N.

[51] And the last.

[52] In a centipede, it is natural to look at choices as made in succession, but remember that actually, at each stage, both players simultaneously are a-sked to make a choice between being an R or a bi type, with choosing R implying to play pass at the current step and bi to play down, also if only one of the players reveals the move he has chosen. Assume, for example that it is player 2 that does not reveal the move he has chosen because he has the move only in the following stage. 2 can revise his choice once he observes the choice of 1. If 1 plays bi in the current stage, revision by 2 is irrelevant; if 1 plays R, 2 may have reason to shift from bi to R, but he is not constrained, not even by rationality, to do so, and is unlikely to have

In this view, in the initial steps of the sequence, at each stage they are facing an assurance game with the usual three equilibria, two in pure strategies, with (R, R) somewhat undefined but clearly Pareto dominating (bi; bi),[53] and one in mixed strategies where both choose R with positive probability.[54] Stage after stage, the (bi, bi) equilibrium is associated to higher and higher payoffs, converging to those associated to the (R, R) equilibrium. In the final steps, the game degenerates into a prisoners' dilemma with just one equilibrium in dominating strategies.

Again, one can solve the sequence backwards, and conclude for the optimality of playing bi as soon as one has a chance, but at the cost of inconsistency. If inconsistency is to be avoided, one must look at the game forward. If one does so, in the assurance phase, rationality does not allow to single out which equilibrium will be played or even give a probability to the event a particular one is reached. One consequence of this fact is that one knows there is an upper bound to the number of stages the game can go on, N, but the actual number of stages that will be played is uncertain. This, by itself, blocks backward induction and makes the adoption of R in the assurance section consistent with rationality in the choice of the kind of rationality one intends to follow. Consistence with reasonableness implies however to single out coordination on (R, R) in this phase.

---

reason for shifting from R to bi. In this sense, this way of looking at the game tilts choices towards R.

[53] At least if N is sufficiently large.

[54] As near to 1 as one likes, provided N is chosen sufficiently high.

Quaderni dell'Istituto di economia internazionale, delle istituzioni e dello sviluppo dell'Università Cattolica del Sacro Cuore

(DAL 2002 QUADERNI DEL DIPARTIMENTO)

9401 Beretta C. *"Is economic theory up to the needs of ethics?"* (Part I) (trad. it. "Le scelte individuali nella teoria economica" pubblicata in M. Magrin (a cura di) (1996) "La coda di Minosse", Franco Angeli, Milano)

9402 Beretta C. *"Alcune radici del problema dell'autonomia individuale"*

9403 Beretta C. *"Asimmetrie informative ed autonomia: le strutture contrattuali e la formazione dei mercati"* (Parte I)

9404 Merzoni G. *"Delega strategica e credibilità delle minacce nella contrattazione tra sindacato e impresa"*

9405 Beretta C. *"Alcune funzioni e caratteristiche delle regole"* (pubblicato in Rivista Internazionale di Scienze Sociali, a. CII, n. 3, luglio-settembre, pagg. 339-55)

9501 Beretta C. *"Having alternatives, being free and being responsible"* (pubblicato in Cozzi T. - Nicola P.C. - Pasinetti L.L. - Quadrio Curzio A. (a cura di) "Benessere, equilibrio e sviluppo. Saggi in onore di Siro Lombardini", Vita e Pensiero, Milano)

9502 Beretta C. - Beretta S. *"Il mercato nella teoria economica"* (pubblicato in Persone & Imprese, n. 2, 1995)

9503 Beretta S. - Fortis M. - Draetta U. *"Economic Regionalism and Globalism"* (Europe-Iran Roundtable, Third Session, may 26, 1995)

9504 Beretta S. *"World Trade Organization: Italia ed Europa nel nuovo assetto globale"* (pubblicato su Rivista Internazionale di Scienze Sociali, a. CIII, n. 3, luglio-settembre 1995, p. 415-456)

9505 Colangelo G. - Galmarini U. *"Ad Valorem Taxation and Intermediate Goods in Oligopoly"*

9601 Beretta S. *"Disavanzi correnti e movimenti finanziari. Una survey molto selettiva e qualche (ragionevole) dubbio"*

9602 Beretta C. *"Strumenti per l'analisi economica - 1"*
9603 Beretta C. *"Dottrina sociale della Chiesa e teoria economica"*
9604 Venturini L., *"Endogenous sunk costs and structural changes in the Italian food industry"*
9701 Natale P., *"Posted Vs. Negotiated Prices under Incomplete Informaion"*
9702 Venturini L. - Boccaletti S. - Galizzi G., *"Vertical Relationships and Dual Branding Strategies in the Italian Food Industry"*
9703 Pieri R., Rama D., Venturini L., *"Intra-Industry Trade in the European Dairy Industry"*
9704 Beretta C., *"Equilibrio economico generale e teoria dei contratti"* (pubblicato in Istituto Lombardo - Accademia di Scienze e Lettere, Incontro di studio n. 14, Disequilibrio ed equilibrio economico generale, Milano, 1998)
9705 Merzoni G., *"Returns to Process Innovation and Industry Evolution"*
9801 Beretta C., Beretta S., *"Footpaths in trade theory: Standard tools of analysis and results from general equilibrium theory"*
9802 Beretta C., *"Alcuni problemi di giustizia, dal punto di vista dell'economista"*
9803 Beretta C., *"La scelta in economia"*
9901 Merzoni G., *"Observability and Co-operation in Delegation Games: the case of Cournot Oligopoly"*
9902 Beretta C., "*Note sul mercantilismo e i suoi antecedenti"*
9903 Beretta C., *"A Ricardian model with a market for land"*
0001 Beretta S., *"Disavanzi nei pagamenti e commercio intertemporale: alcuni spunti di analisi 'reale'"*
0002 Beretta S., "*Strumenti finanziari derivati, movimenti di capitale e crisi valutarie degli anni Novanta: alcuni elementi per farsi un'idea"*
0003 Merzoni G., *Stategic Delegation in Firms and the Trade Union*
0101 Colombo F. – Merzoni G., "*Reputation, flexibility and the optimal length of contracts"*
0102 Beretta C., *Generalità sulla scelta in condizioni di certezza*

QUADERNI EDITI DA VITA E PENSIERO*

---

* Nuova linea di Quaderni DISEIS stampata grazie ad un accordo con l'Editrice Vita e Pensiero dell'Università Cattolica.
(*) Testo disponibile presso il DISEIS

0503 Beretta C., *"Mercato, società e stato in un'economia aperta - Parte II"*
0601 Beretta C., *"L'ipotesi di razionalità - Parte I"*
0602 Beretta C., *"L'ipotesi di razionalità - Parte II"*

Finito di stampare
nel mese di dicembre 2007
da Gi&Gi srl - Triuggio (MI)

ISBN 978-88-343-1687-0