Conflict, Interest and Strategy: A Risk Limit Approach to Conflict

Matthias Mohr and Kjell Hausken

96/7

Max-Planck-Institut für Gesellschaftsforschung
Lothringer Str. 78
D-50677 Köln
Germany

Telephone +49(0)221/33605-0
Fax +49(0)221/33605-55
E-Mail mpi@mpi-fg-koeln.mpg.de
Home Page http://www.mpi-fg-koeln.mpg.de

MPIFG Discussion Paper 96/7
ISSN 0944–2073
December 1996

Abstract

The paper criticizes Axelrod's (1970) measure of conflict, which is basically a measure of the *conflict of interest* and does not account for the potential for overt conflictful behavior based on how the actors are strategically interpositioned. Differences between the two dimensions of conflict are illustrated, and it is shown how both dimensions are necessary for a theory of conflict. Conflict is conceptualized as a struggle for preferred equilibria in 2x2 games with at least two equilibria. A new measure of conflict for this class of games based on the concept of risk limit is developed in both a static and a repeated game setting. The risk limit approach measures the actors' inclination to *behave* in a conciliatory or challenging way. This is different from Axelrod's approach, which measures the incompatibility – and hence the conflict – of the actors' *preferences*. We show that higher incompatibility of preferences between the actors does not necessarily imply that they are more inclined to behave challengingly. Furthermore, we show that the more actors value future payoffs in terms of present payoffs, the greater their inclination will be to behave challengingly.

Zusammenfassung

In diesem Beitrag wird Axelrods (1970) Konfliktmaß kritisiert, welches als Maß des *Interessenkonflikts* die strategische Entscheidung der Akteure für konflikthaftes Verhalten unberücksichtigt läßt. Es werden Unterschiede zwischen den beiden Dimensionen von Konflikt verdeutlicht, und wir zeigen, daß beide Dimensionen in einer Konflikttheorie wichtig sind. Konflikt wird hier aufgefaßt als Streit um bevorzugte Gleichgewichte in 2x2-Spielen. Wir entwickeln ein neues Konfliktmaß, welches auf dem Konzept der Risikogrenze (risk limit) aufbaut und stellen es in einem statischen und in einem wiederholten Spiel vor. Das Maß mißt die Bereitschaft der Akteure, sich in einer Konfliktsituation konziliant oder herausfordernd zu *verhalten*. Es unterscheidet sich damit von Axelrods Maß, welches auf die Unvereinbarkeiten – und somit den Konflikt – zwischen *Präferenzen* abzielt. Wir zeigen, daß eine größere Unvereinbarkeit der Präferenzen nicht notwendigerweise eine ausgeprägtere Neigung zu konfliktträchtigem Verhalten impliziert. Weiterhin wird gezeigt, daß die Akteure um so stärker zu konflikthaftem Verhalten neigen, je höher sie Auszahlungen in der Zukunft relativ zu Auszahlungen in der Gegenwart bewerten.

Contents

1 Introduction

The argument of this paper is that the nature of conflict is made up of two complementary dimensions, the interest dimension and the strategic dimension. Each by itself would be insufficient to capture what conflict means. The two supplement each other, providing an exhaustive account of conflict from the perspective of rational choice and from that of game theory. The interest dimension has been worked out by Axelrod (1970) in an insightful way. However, as important as it is for understanding conflict, the interest dimension is only one side of the coin. We argue, first, that the other side is represented by the strategic dimension, secondly that the concept of conflict can be reduced to these two complementary dimensions, and thirdly that each dimension is potentially equally important. What relative weight should be assigned to each dimension in an analysis of social conflict depends on what is to be explained and the nature of the social interaction.

The crucial difference between the two dimensions can be illustrated by two questions. The strength of conflict in terms of the interest dimension is determined by answering this question:

1. How large is the payoff one actor gets if the other actor gets her best payoff, how large is the other's payoff if the first actor receives his best payoff, and what are the feasible points which might serve as a compromise?

This first question refers only to the divergence of preferences and therefore to the conflict of interest in a game. However, consider the second question:

2. What might an actor win if she successfully challenges her opponent and what is her chance of winning the challenge?

This second question, which also refers to strategic considerations regarding how an actor can improve her own payoff by exhibiting aggressive, recalcitrant, stubborn or hardheaded – i.e. conflictful – behavior, proceeds beyond the simple first question of how preferences diverge. The crucial difference between the two questions is that the first deals with latent, often hidden conflicts of interest and the second with overt conflictful behavior.

One drawback to limiting attention only to the interest dimension of conflict is that the link between conflict and behavior becomes unclear. Consider, for example, Axelrod's (1970: 80) statement, which we refer to as his proposition concerning conflictful behavior, that "other things being equal, the more conflict of inter-

---

We would like to thank Philipp Genschel, Giandomenico Majone, Renate Mayntz, Thomas Plümper, Andreas Ryll and Fritz W. Scharpf for helpful comments.

est there is, the more probable it is that conflictful behavior will result." Crucial here is how "conflictful behavior" is defined. Axelrod frequently refers to a prisoner's dilemma in which conflictful behavior is well defined. However, he proceeds to argue that his proposition concerning conflictful behavior holds not only in a prisoner's dilemma, but in all cases. Unfortunately, Axelrod does not explicitly define conflictful behavior in general social situations. Rather, he states (Axelrod 1970: 80): "The specific meaning of conflictful behavior will differ from one political process to another, of course, but usually there are types of behavior which are clearly conflictful and hence some hypotheses are easy to specify." We will illustrate these points over the next sections, questioning the validity of Axelrod's hypotheses for conflictful behavior by illustrating that they do not follow from his conflict measure the way he suggests.

We conceptualize conflict as a struggle for preferred equilibria in 2x2 Battle of the Sexes games. A new measure of conflict for this class of games based on the concept of risk limit is developed in both a static and a repeated game setting. It might be argued that a low percentage of games without a mutually best outcome have two Pareto-superior Nash equilibria. However, the significance of such games may be larger. We will not venture into an evaluation of this significance here, because even informed judgments may differ, We would like to point out Knight's (1992) argument that struggle for preferred equilibria is a crucial characteristic, for example, at the start-up of most social institutions. Knight holds that institutions are created to solve these conflicts over preferred equilibria, and that the outcome often tilts in favor of the stronger player. Most institutions are designed to regulate social life in the long term. The players' stakes in the competition for the preferred institutional design may thus be high, depending on the time horizon of the players. This point is elaborated further in Section 4, which deals with the repeated game setting. It is this important role played by Battle of the Sexes situations in real social life which justifies our focus on the Battle of the Sexes.

The risk limit approach measures the actors' inclination to *behave* in a conciliatory or challenging way. We show that higher incompatibility of preferences between the actors does not necessarily imply that they are more inclined to behave challengingly. Furthermore, we show that the more actors value future payoffs in terms of present payoffs, the greater their inclination will be to behave challengingly.

The article is divided into five sections. Section 2 illustrates Axelrod's measure of conflict. Section 3 proposes a new measure of conflict based upon the concept of risk limits (Harsanyi 1977: 280–288), employed in a static and a repeated game setting. Section 4 compares the two concepts more rigorously, analyzing the effects of slight changes of the game structure on the conflict measures, and showing that the results may be contradictory. Section 5 puts the results of the preceding sections into perspective.

2 Evaluation of the Measure of Conflict of Interest

Axelrod's approach to conflict focuses on "the state of incompatibility of the goals of two or more actors" (Axelrod 1970: 5), a state where one actor can get her best payoff only at the expense of the other actor. More specifically, he refers to the proportion of the joint demand which is unfeasible (Axelrod 1970: 57). Consider the battle of the sexes game in Table 2.1, which has two equilibria in pure strategies, (4,3) and (3,4).

		Actor 2: I	Actor 2: II
Actor 1	I	4,3	2,2
	II	1,1	3,4

Table 2.1 Battle of the sexes

Observe that actor 2 only gets her second-best payoff of 3 when actor 1 receives his best payoff of 4, and vice versa. Therefore, their goals are incompatible to a certain extent. Note also that each actor is able to secure a minimax payoff of at least 2 to himself by choosing strategy I and II, respectively. That is, the cell with the payoff distribution (1,1) is irrelevant given the objective in our analysis of determining how incompatible the preferences are. Hence the focus is on the following question: How incompatible are the demands beyond or above the payoffs (the minimax payoffs) the actors are able to secure for themselves? Axelrod's approach can best be illustrated by a graphic representation of the game, Fig. 2.1.

The area of joint demand above the minimax payoffs is the rectangle, in this example a square, spanned out by the minimax point (2,2) and the point determined by the best payoff each actor can possibly obtain under his most favorable circumstances, (4,4), lightly shaded in Fig. 2.1. The area of unfeasible joint de-

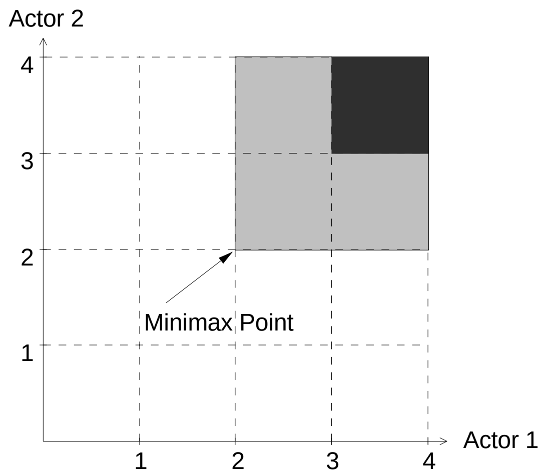

Fig. 2.1 Diagram for the battle of the sexes game in Table 2.1, assuming pure strategies

mand is defined as the polygon spanned out by the best possible payoff each actor can obtain, (4,4), the two points where one of the actors gets his most preferred payoff, (3,4) and (4,3), and the fourth point in Table 2.1, (1,1), if it falls inside the area of joint demand. Since (1,1) falls outside the latter area in Fig. 2.1, the area of unfeasible demand becomes the darkly shaded rectangle. Axelrod defines the degree C_{ax} of conflict as the ratio of the two areas, which in Fig. 2.1 is C_{ax}=1/4. C_{ax} increases when the area of unfeasible joint demand increases and the whole area of joint demand decreases. In Fig. 2.1 we do not allow for mixed strategies. If mixed strategies are accounted for, then all points on the line connecting the points (4,3) and (3,4) are feasible, too. In this case the area of unfeasible joint demand shrinks to the darkly shaded triangle in Fig. 2.2, and C_{ax} reduces to 1/8.

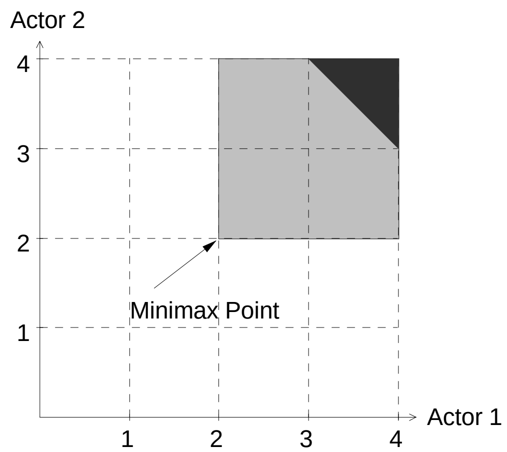

Fig. 2.2 Diagram for the battle of the sexes game in Table 2.1, assuming mixed strategies

Over the following pages, we present an alternative way to define the relevant areas. Since our arguments do not refer merely to the incompatibility of demands, we prefer more general labels than Axelrod's "area of joint demand" and "area of unfeasible joint demand." In Axelrod's measure the area of joint demand appears in the denominator. Hence it serves as a standard of measure, and we refer to it as the "unit area." Axelrod defines the area of the unfeasible joint demand appearing in the numerator as the "outlying area," a term we will also use.

3 A Measure of Conflict Based upon the Concept of Risk Limits

This section presents a measure of conflict based upon the concept of risk limits introduced by Harsanyi (1977: 280–288). This measure clarifies the distinction between conflict of interest and conflictful behavior within a game-theoretical framework. More specifically, the risk limit measure of conflict accounts better for overt strategic conflictful behavior than Axelrod's measure, because the former involves considering the actors' perception of each other's recalcitrance and aggression. In this section, recalcitrance is defined as the threat of switching from an equilibrium strategy to another strategy, or the threat of not following suit if the other actor switches from an equilibrium strategy to another strategy. The approach applies only to games with two or more equilibria since it conceptualizes conflict as a struggle for preferred equilibria.

3.1 The Risk Limit Approach in a One-Shot Game

Consider the battle of the sexes game[1] in Table 2.1, repeated in Table 3.1, which includes variable payoffs used later for generalization. a_i and b_i refer to actor i's payoff in his preferred equilibrium and in the equilibrium he prefers less, respectively. c_i refers to i's payoff in the threat point. Hence $a_i > b_i > c_i$ and $a_i > d_i$, i=1,2.

1 In the literature, the off-diagonal payoffs in the Battle of the Sexes game are considered to be equivalent. We distinguish between the two because our objective is to model situations where the players have two kinds of preferences: 1. They have a strong preference for a certain state, independent of the choice of the other player. 2. They prefer a coordinated to an uncoordinated outcome. If the off-diagonal payoffs were the same, the importance of the second preference would be emphasized rather than the first. Observe, however, that the risk limit measure developed in this section holds also when the off-diagonal payoffs are equivalent.

		Actor 2 I	Actor 2 II
Actor 1	I	4,3 a_1,b_2	2,2 c_1,c_2
Actor 1	II	1,1 d_1,d_2	3,4 b_1,a_2

Table 3.1 Two-person two-strategy game

Each of the actors prefers a different equilibrium: Actor 1 prefers (4,3) while actor 2 prefers (3,4). Consider (4,3) the starting point of our analysis by assuming that this is the status quo point. Actor 2 would prefer the other equilibrium and she could try to force actor 1 to that equilibrium by changing from her first to her second strategy. If actor 1 sticks to his first strategy, the actors receive (2,2). In this case a conflict occurs in the sense that both of the actors adopt conflictful strategies: Actor 2 by challenging actor 1, and actor 1 by withstanding the challenge. Therefore we call (2,2) the conflict point. In the conflict point, both of the actors have an incentive to give in: If he gets a payoff in the equilibrium he prefers less, an actor is still better off than he is with his payoff in the conflict point. On the other hand, since both actors might reasonably expect the other one to give in, both of them also have an incentive to stick to their conflictful strategies in order to receive the payoff in the preferred equilibrium. What then is the highest probability that such a conflict might occur? Define p_1 as the probability that actor 1 sticks to his first strategy when challenged by 2, and p_2 as the probability that actor 2 sticks to her challenge when actor 1 does not give in. Then we can calculate actor 2's expected payoff of a challenge as $2p_1 + 4(1-p_1)$. Define the value of p_1 which makes actor 2 indifferent between her conflictful strategy and her first strategy as $2p_1+4(1-p_1)=3$, which gives $p_1^* = (4-3)/(4-2) = r_2$. If the probability that actor 1 remains hardheaded is higher than p_1^*, then actor 2 would not try to challenge actor 1. Thus, $r_2 = p_1^*$ is the risk limit of actor 2: It is the highest risk of a conflict that she is willing to take by challenging actor 1. In the same way we define the risk limit for actor 1. Given that actor 2 chooses her second strategy, the expected payoff to actor 1 if he does not give in is $2p_2 + 4(1-p_2)$, and his risk limit then is $r_1 = p_2^* = (4-3)/(4-2)$. This is the highest probability of conflict actor 1 is willing to take by withstanding the challenge of actor 2.

Of course, the risk limits are the same for both actors since the game is symmetric. Recapitulating, the highest risk actor 2 is willing to take by challenging actor 1 is

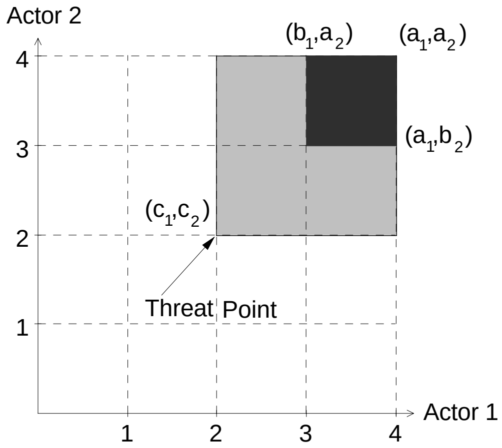

Fig. 3.1 Diagram for the risk limit game in Table 3.1

the same as the highest probability that actor 1 sticks to his first strategy when challenged by actor 2, hence r_2=p_1*. Analogously, r_1=p_2*. Thus p_1* and p_2* are maximum individual probabilities of individual conflictful behavior, and r_1 and r_2 are the maximum individually acceptable risks. Given this, we now propose the product r_1r_2 = p_2*p_1* as our measure of conflict, which is the joint maximum probability that both of the actors will choose their respective conflictful strategies. It is a maximum since each of the constituents p_1* and p_2* are maximum individual probabilities of conflictful behavior. As shown in Fig. 3.1, this conflict measure can be graphically represented as the relation of the darkly shaded area of conflictful behavior to the lightly shaded unit area – similar to Axelrod's approach.

This yields a degree C_{rl} =1/4 of conflict, which is twice as large as Axelrod's C_{ax} =1/8 shown in Fig. 2.2. However, if mixed strategies are excluded from the analysis so that randomization between the two payoffs (3,4) and (4,3) is not possible, then Axelrod's conflict measure would increase to C_{ax} =1/4 (see Fig. 2.1) and thus be the same as the conflict measure C_{rl}=1/4 from the risk limit approach. Finally, observe that the conflict measure according to the risk limit approach does not depend on the equilibrium from which one starts. The areas would have been the same had we started from the equilibrium (3,4) and let actor 1 challenge actor 2. This is always the case in games with two equilibria, in symmetric games as well as in asymmetric games.

The areas might be interpreted as follows: The larger the unit area and the smaller the outlying areas are, the smaller the conflict measure is. The size of the unit area can be interpreted as a measure of the joint demand above or beyond the conflict point, and the darkly shaded area can be interpreted as the jointly unfeasible expectation of an additional gain in case of a conflict. Note however that the respective gains and losses of the actors are not added up but multiplied. For example, if actor 1 expects a large gain by challenging actor 2, while actor 2 expects no benefit from a struggle, then actor 2 will give in and no conflict will occur. On the other hand, if one of the actors has nothing to lose in a conflict, the risk of a conflict increases considerably. Therefore the respective gains (and losses) have to be multiplied rather than added up in order to construct the conflict measure.

Now, referring to Table 3.1, we formalize the risk limit approach, starting with (a_1,b_2). If actor 2 switches to II, she will evaluate the probability p_1 that actor 1 will stick to his old strategy I. A large p_1 corresponds to actor 1 being perceived by actor 2 as being recalcitrant, withstanding the threat or challenge from actor 2. Actor 2 will switch to II if

$$p_1 c_2 + (1 - p_1) a_2 \geq b_2 \tag{3.1}$$

i.e. if

$$p_1 \leq \frac{a_2 - b_2}{a_2 - c_2} = p_1^{*} = r_2 \quad . \tag{3.2}$$

Consider p_2 as the probability that actor 2 sticks to her second strategy after she has challenged actor 1, i.e. after actor 2 has actually challenged the equilibrium (I,I). Actor 1 will switch to II if

$$p_2 c_1 + (1 - p_2) a_1 \geq b_1 \tag{3.3}$$

i.e. if

$$p_2 \leq \frac{a_1 - b_1}{a_1 - c_1} = p_2^{*} = r_1 \quad . \tag{3.4}$$

The product of the risk limits

$$C_{rl} = r_1 r_2 = \frac{(a_1 - b_1)(a_2 - b_2)}{(a_1 - c_1)(a_2 - c_2)} \tag{3.5}$$

can be considered the maximum degree of conflict between the actors, i.e. the product of the maximum degree of recalcitrance the actors assign to each other.

3.2 The Risk Limit Approach in a Supergame

So far we have analyzed one-shot games only. However, an extension to multistage games will show that the degree of conflict is shaped by the players' evaluation of the future. Hence, the analysis of the multistage game will show an important conceptual distinction between the conflict of interest and the risk limit approach to conflict. We devote the remainder of this section to the analysis of the multistage game.

Consider Fig. 3.2. One actor makes a decision whether or not to challenge her opponent, while the opponent at the same time decides whether to resist or to give in. If the opponent gives in, the game is terminated. In most conflictful social situations, however, interaction continues for several time periods.

Fig. 3.2 assumes the equilibrium path where the challenger sticks to her challenge consistently for a certain number of periods (0 periods if she does not challenge at all), after which she (if the opponent does not back down) backs down consistently in all later periods. The reasoning is affected by some of the literature on reputation (such as Kreps / Wilson 1982) criticizing Selten's (1978) analysis of the chain-store paradox, where players early in a repeated game may seek to acquire a reputation for being "tough" or "recalcitrant" or something else. As is well-known from the "folk theorem," many different equilibrium paths can be sustained by many different strategies in an infinitely repeated game. One example would be to alternate the strategies of challenge and forgiveness according to some pattern. Our approach in this paper, however, is to focus on situations in which a reputation argument plays a role, so that the challenger either challenges all the time or not at all. We believe that such situations occur frequently in real

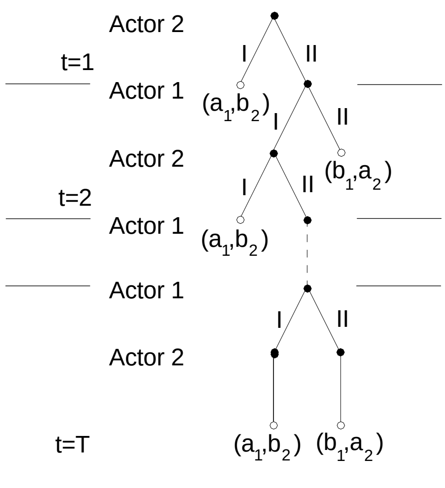

Fig. 3.2 Multistage game with two actors and two strategies

life. Thus, while Fig. 3.2 might not cover all the paths observed in praxis, in our opinion it illustrates plausible equilibrium paths that are observed often enough to justify (or warrant) their analysis.

Given the equilibrium path in Fig. 3.2, actor 2 might try to enforce the payoff a_2 in her preferred equilibrium (b_1,a_2) by changing to her second strategy and sticking to it for several periods rather than giving in and receiving b_2. Actor 1 might withstand the challenge or give in. In every time period in which neither actor gives in, each receives (c_1,c_2). Future payoffs are discounted and the success of actor 2's challenge is uncertain. The crucial question for actor 2 is when to back down from her challenge. Define for actor 2 a time period f after which she ceases to employ her challenging strategy if actor 1 has not conceded before time period f. Define the following parameters and rules of the game:

- δ_2 Discounting factor of actor 2. The present value of a payoff x in period t for actor 2 is $x\delta_2^{t-1}$.
- T Last period of the game. No payoffs are received after time period T.
- f Period of last challenge of actor 2. Actor 2 does not challenge after period f, $0 \le f \le T$.

Observe that if f = 0, then actor 2 does not challenge at all. If actor 2 challenges in period f and actor 1 has not conceded in any of the periods until and including period f, then each gets (c_1,c_2) in the periods 1,2,...,f. Actor 2 then concedes in period f+1 and receives b_2 in the periods f+1, f+2,...,T. (Actor 1 receives a_1 in the periods f+1,f+2,...,T.) If actor 2 challenges in period f and actor 1 concedes in period f, then actor 2 gets c_2 in 1,2,...,f–1, and a_2 in f,f+1,f+2,...,T. (Actor 1 gets c_1 in 1,2,...,f–1, and b_1 in f,f+1,f+2,...,T.) If actor 1 concedes in period e, e < f, then actor 2 gets c_2 in 1,2,...,e–1, and a_2 in e,e+1,...,f,f+1,f+2,...,T. (Actor 1 gets c_1 in 1,2,...,e–1 and b_1 in e, e+1,...,f,f+1,f+2,...,T.)

Actor 2 has some opportunity cost to bear when she behaves conflictfully up to and including period f. If she gives in from the very beginning (i.e., if she does not challenge at all), she would receive $b_2\sum_{t=1}^{T}\delta_2^{t-1}$. Therefore, the expected additional gain $g_2(f)$ she receives if she challenges in every time period until and including period f is

$$g_2(f) = a_2\sum_{t=1}^{f}((1-p)p_1^{t-1}\sum_{s=t}^{T}\delta_2^{s-1}) + c_2\sum_{t=1}^{f}p_1^t\delta_2^{t-1} + b_2p_1^f\sum_{t=f+1}^{T}\delta_2^{t-1} - b_2\sum_{t=1}^{T}\delta_2^{t-1}\,. \quad (3.6)$$

Note that $g_2(f=0) = 0$. Also note that the risk limits derived within the static analysis can be derived by the formula for $g_2(1)$ if T = 1. Then $g_2(1)$ reduces to

$$g_2(1|T=1) = (1-p_1)a_2 + p_1c_2 - b_2 \ . \tag{3.7}$$

$g_2(1\,|\,T=1)$ is the additional gain which actor 2 expects to receive if she challenges actor 1 in period 1 in a game which is played only once (T=1), rather than giving in. We can derive the risk limit r_2 as defined in equation (3.4) if we solve the equation $g_2(1\,|\,T=1) = 0$ for p_1.

Assume now that the interaction continues for a long time, i.e. that T becomes large. If T approaches infinity, (3.6) reduces to

$$\lim_{T\to\infty} g_2(f) = \frac{1-\delta_2^f p_1^f}{(1-\delta_2)(1-\delta_2 p_1)}(a_2(1-p_1)-(b_2(1-\delta_2 p_1)+c_2p_1(1-\delta_2)) \,. \tag{3.8}$$

Observe that according to (3.8), the best player 2 can do is to challenge all the time or not to challenge at all, depending upon whether

$$a_2(1-p_1)-(b_2(1-\delta_2 p_1)+c_2p_1(1-\delta_2)$$

is larger or smaller than zero. Hence setting this term equal to zero allows us to determine the risk limit r_2' for actor 2 when the game in Fig. 3.2 is played an infinite number of times. This leads to

$$p_1^* = r_2' = \frac{a_2 - b_2}{(a_2-c_2)-\delta_2(b_2-c_2)} \,. \tag{3.9}$$

Equation (3.9) defines the value p_1* which makes actor 2 indifferent between challenging actor 1 all the time and conceding in the first time period, given that the game is played an infinite number of times. Observe that Equation (3.9) may be rewritten as

$$r_2' = \frac{r_2}{1-\delta_2(1-r_2)} \,. \tag{3.10}$$

Equations (3.9) and (3.10) are derived analogously for actor 1. As a measure of conflict in a 2x2 game with two equilibria which is played an infinite number of times, we get, using (3.9) and (3.10),

$$C_{rl}' = \frac{(a_1-b_1)(a_2-b_2)}{((a_1-c_1)-\delta_1(b_1-c_1))((a_2-c_2)-\delta_2(b_2-c_2))} \tag{3.11}$$

$$C_{rl}' = \frac{r_1 r_2}{(1-\delta_1(1-r_1))(1-\delta_2(1-r_2))} \,. \tag{3.12}$$

4 Conflict of Interest and the Risk Limit Approach to Conflict: A Comparison

The risk limit approach questions the choice of two crucial points in Axelrod's approach: the minimax payoffs defining the conflict point (also referred to as the zero point or the threat point), and the highest payoffs defining the outmost point. These two points help limit the relevant areas. We show that, according to the risk limit approach, the conflict point is not necessarily the point defined by the minimax payoffs. We also demonstrate that the risk limit approach may lead to another choice of the point defining the outer limits of the relevant area.

Observe that the threat point $(c_1,c_2) = (2,2)$ in Fig. 3.1 is the same as the minimax point in Fig. 2.1. Now change the game in Table 3.1 to the game in Table 4.1, which has the same minimax point as Table 3.1, but where the threat point is changed to $(c_1,c_2) = (1,1)$.

		Actor 2: I	Actor 2: II
Actor 1	I	4,3	1,1
	II	2,2	3,4

Table 4.1 Two-person two-strategy game

Relative to Table 3.1, Table 4.1 reverses the cells (1,1) and (2,2). This does not have any effect on the conflict of interest, which remains the same. That is, the degree $C_{ax}=1/8$ of conflict according to Axelrod is the same for the Tables 3.1 and 4.1. However, the measure of conflict according to the risk limit approach changes. This is due to the fact that (1,1) rather than (2,2) is now the conflict point. More specifically, the graphic representation of Table 4.1 in Fig. 3.2 shows that the degree of conflict according to the risk limit approach is reduced to $C_{rl} = 1/9 < C_{ax}$. This can also be calculated from (3.5) observing that $r_1 = r_2 = (4–3)/(4–1) = 1/3$. Observe that C_{rl} approaches 1 when the parameters in the upper-right corner of Table 4.1, that is $(1,1)=(c_1,c_2)$, approach $(3,3) = (b_1,b_2)$. As (c_1,c_2) approach (b_1,b_2), the probability of the actors' getting stuck at the conflict point (c_1,c_2) *increases*. This can be viewed as an "intermediate stay" from which they may compete or struggle for either of the two equilibria (a_1,b_2) and (b_1,a_2). Increasing (c_1,c_2) toward

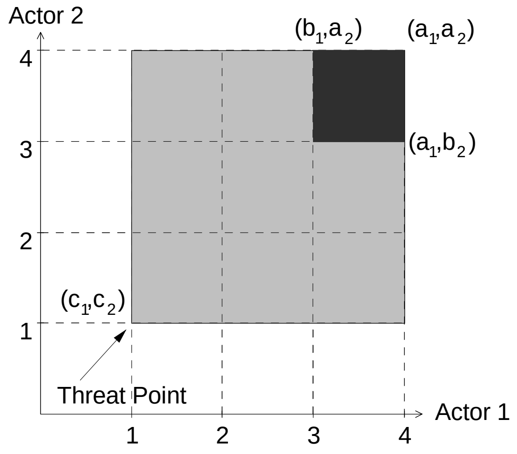

Fig. 4.1 Diagram for the 2x2 game in Table 4.1

(b_1,b_2) thus induces an increasing degree of conflict. Conflict for the game in Table 4.1 is not defined for $c_i>b_i$, i=1,2. Also observe that C_{rl} decreases as c_1 and c_2 approach 0: Decreasing c_1 or c_2 *reduces* the probability of the actors' getting stuck at the conflict point (c_1,c_2), which increases the probability of their remaining permanently in one of the two equilibria (a_1,b_2) and (b_1,a_2). In this case the degree of conflict decreases.

The measure of conflict decreases since the payoffs in the conflict point decrease, which renders conflictful behavior less attractive to both actors. Hence the joint maximum risk of a conflict that the actors are willing to accept decreases, too. Contrary to Axelrod's analysis, the point defined by the minimax payoffs (2,2) does not play any role here. The crucial factors are the payoffs in the two equilibria the actors struggle for, and the payoffs in the conflict point, which is now (1,1) rather than (2,2). Hence the analysis shows how the risk limit approach leads to another definition of the unit area, as demonstrated by Fig. 4.1, than Axelrod's approach does, cf. Fig. 3.1.

We will now show that the definition of the unit area in Axelrod's approach may also be changed by questioning the choice of the outmost point in the upper-right direction. This is the maximum attainable payoff for each actor in Axelrod's analysis. It is defined by the highest payoffs in the game, and it limits both the area of conflicting interest and the unit area. It is obvious that this point is very important in Axelrod's analysis of conflicting interests. However, we will show that strategic reasoning may render this point irrelevant, too. This can be seen in the two-person three-strategy game in Table 4.2, which includes the two-person two-strategy game in Table 2.1, and for expositional convenience preserves the minimax point (2,2).

		Actor 2 I	Actor 2 II	Actor 2 III
	I	4,3	2,2	5,2
Actor 1	II	1,1	3,4	1,1
	III	1,1	2,5	1,1

Table 4.2 Two-person three-strategy game

If actor 1 receives his highest payoff of 5, actor 2 gets her minimax-payoff of 2, and vice versa. Therefore the conflict of interest in the enlarged game is 0.5, which is twice what it was in the two-strategy game illustrated in Table 4.2 by the delimiting bold gray lines. A graphic representation of Table 4.2 is given in Fig. 4.2.

According to Axelrod's measure, the whole relevant area is limited by the points (2,2), (2,5), (5,5), (5,2). The area of conflicting interest is represented by the lightly shaded triangle spanned by (2,5), (5,5), (5,2). Axelrod's measure rises because the additional strategies have sharpened the incompatibility of interest. However, the measure of conflict by risk limits is the same in both games. This is the case because, according to strategic reasoning, only the four cells in the upper-left area of Table 4.2 are relevant. Again we have the two equilibria (4,3) and (3,4). Neither (5,2) nor (2,5) are equilibria. Hence the risk limit approach renders the points (5,2) and (2,5) irrelevant as points from which either actor may challenge the other in order to enforce her most preferred solution. Hence neither of the actors might reasonably hope to get her highest possible payoff. However, as we have shown above, if we start from (4,3), it might be a rational option for actor 2 to challenge actor 1 in order to enforce her preferred equilibrium (3,4). And if we start from (3,4) the same reasoning holds for actor 1 challenging actor 2. Consequently, if we want to calculate the probability of conflictful behavior according to the risk limit approach, the respective upper and right limits are the same in Fig. 3.1 and Fig. 4.2. In Fig. 4.2 the whole relevant area is then bounded by the square spanned by (2,2), (2,4), (4,4), (4,2). The darkly shaded square spanned by (3,3), (3,4), (4,4), (4,3) limits the area of conflictful behavior.

We have already illustrated many important differences between the risk limit approach and Axelrod's approach. What remains is to analyze additional differences when time is explicitly taken into account. Observe that in Axelrod's approach there is no possibility to account for the valuation of future payoffs, i.e. for

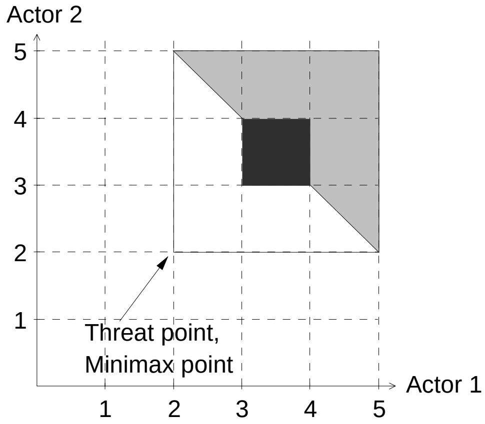

Fig. 4.2 Diagram for the three-strategy game in Table 4.2

the actor's discount rates. In Section 3 we developed the measure C'_{rl} in the supergame in order to demonstrate the effects of discount rates on conflictful behavior. Observe from Equation (3.10) that if future payoffs are of no value to actor 2 ($\delta_2 = 0$), she regards every trial within the supergame as a one-shot game. Then the risk limits for the supergame and the one-shot game coincide. Observe further from (3.10) or (3.9) that r'_2 increases as δ_2 increases. Thus, the less actor 2 discounts future payoffs, the likelier it is that she will behave conflictfully. Generalize this to a hypothesis:

> *Given an enduring structurally given conflict, the more an actor i values the future (δ_i is high), the more he contributes to increased conflict between the actors.*

Also observe from (3.10) that r'_2 reaches its maximum value of 1 if $\delta_2 = 1$, in which case actor 2 does not discount future payoffs at all and will not be content with the payoff b_2. Rather, she has a strong incentive to struggle in order to get a payoff of 1 any time in the future.

Observe that C'_{rl} in (3.11) equals C_{rl} in (3.5) when δ_i=0, since future payoffs are not valued in this case. If future payoffs are valued, that is δ_i>0, observe the respective terms $\delta_i(b_i - c_i)$ which are subtracted in the denominator in (3.11). Since the denominator measures the size of the unit area, it follows that the unit area for the supergame approach to conflict yielding C'_{rl} in (3.11) is smaller than (if δ_i>0) or equal to (if δ_i=0) the unit area for the static game approach to conflict given in (3.5). Hence $C'_{rl} \geq C_{rl}$. This is shown graphically in Fig. 4.3 where, again, the measure of conflict C'_{rl} corresponds to the ratio of the darkly shaded to the lightly shaded areas.

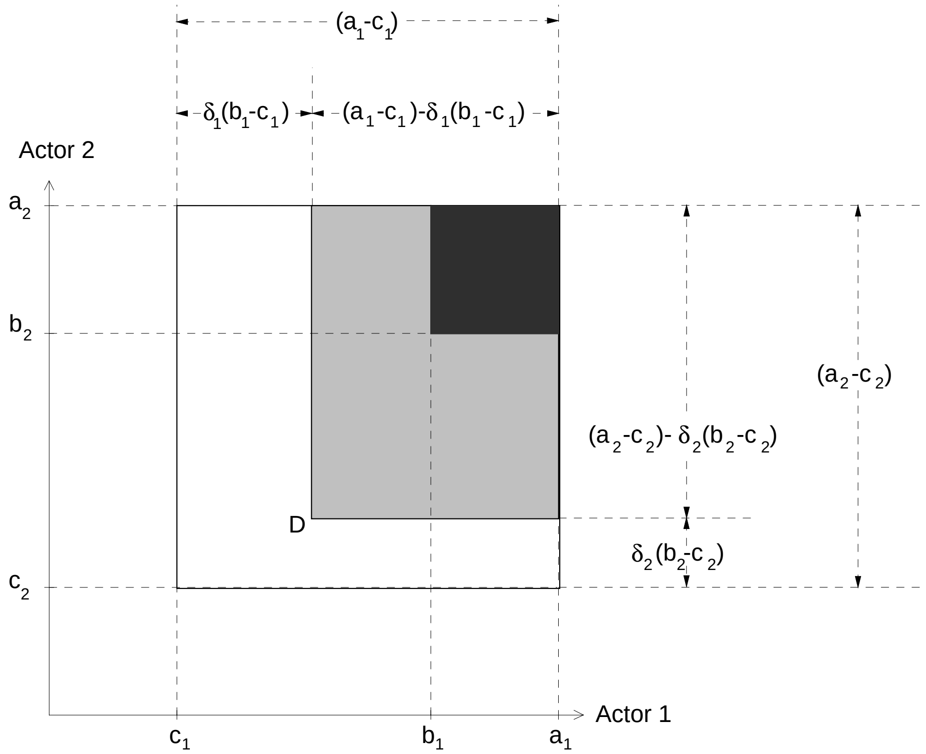

Fig. 4.3 The risk limit approach to conflict in a supergame

More specifically, observe that point D, which refers to the lower left corner of the lightly shaded area in Fig. 4.3, coincides with the threat point (c_1,c_2) in the static game in Fig. 3.1 given that $\delta_i=0$, i=1,2. Point D coincides with point (b_1,c_2) for $\delta_1=1$ and $\delta_2=0$. Point D coincides with point (c_1,b_2) for $\delta_1=0$ and $\delta_2=1$. Finally, point D coincides with point (b_1,b_2) for $\delta_i=1$, i=1,2. In this latter case the lightly and darkly shaded areas in Fig. 4.3 coincide, which yields $C'_{rl}=1$ according to (3.11). This means that if the conflict game is repeated an infinite number of times, and both actors place as much value on each future time period as on the present time period ($\delta_i=1$, i=1,2), then the degree of conflict between the actors according to the risk-limit supergame approach to conflict is $C'_{rl}=1$.

The "shadow of the future" embodied in the discount factors affects the measure by changing the valuation of an expected gain in the future. The higher the actors value present payoffs relative to future payoffs, i.e. the smaller the discount factors δ_i are, the more reluctant they are to risk a conflict. From the perspective of the actors, an engagement into a conflict can be seen as a risky investment. An actor may, through challenging her opponent, accept a loss in the present if she can reasonably hope to make up for it by increasing her gains in the future.

5 Conclusion

Since Axelrod's hypothesis concerning behavior does not follow from his concept of conflict, it is reasonable to search for the implicit assumptions hidden in his approach. One such assumption might be: The greater the difference between the payoff an actor can secure for himself and his best possible payoff, the harder the actor will strive for his best possible payoff. From this additional assumption and his definition of conflict the proposition concerning conflictful behavior can be deduced. However, strategic aspects of social situations are then neglected. The actors try to obtain their best possible payoffs irrespective of the strategic logic of social situations. The risk limit approach however, based on Harasanyi's concept of the risk limit, focuses on the strategic interdependence in social situations and thereby explicates under what conditions the probability of conflictful behavior increases.

The risk limit approach to conflict also makes it possible to consider the time dimension by modeling a supergame and calculating the risk limit for the supergame. We show that the less an actor discounts future payoffs, the higher the probability is that she will behave conflictfully. More generally, the more an actor i values the future, i.e. the greater the discount factor δ_i is, the more she contributes to increased conflict between the actors. Hence, the more an actor values future gains relative to present gains, the more inclined she will be to *invest* in a conflict. While it seems to be common knowledge among social scientists that in situations involving conflict of interest, the actors' interest in the future may help to stabilize cooperative behavior rather than encourage aggressive action, we show that the "shadow of the future" may very well provide an incentive to compete for an equilibrium more favorable to oneself. Of course, since being caught in conflict is costly for both of the actors, they will attempt to develop institutional rules to settle the conflict in the long run. The design of the respective institutions might favor the stronger party (Knight 1992), i.e. the one with the higher risk limit. We do not treat this subject here since it is beyond the scope of this paper.

References

Axelrod, Robert, 1970: *Conflict of Interest.* Chicago: Markham Publishers.

Fudenberg, Drew M. / Eric Maskin, 1986: The Folk Theorem in Repeated Games with Discounting or with Incomplete Information. In: *Econometrica* 54, 533–554.

Harsanyi, John C., 1977: *Rational Behavior and Bargaining Equilibrium in Games and Social Situations.* Cambridge, MA: MIT Press.

Harsanyi, John C. / Reinhard Selten, 1987: *A General Theory of Equilibrium Selection in Games.* Cambridge, MA: MIT Press.

Knight, Jack, 1992: *Institutions and Social Conflict.* Cambridge: Cambridge University Press.

Kreps, David M. / Robert Wilson, 1982: Reputation and Imperfect Information. In: *Journal of Economic Theory* 27, 253–279.

Selten, Reinhard, 1978: The Chain Store Paradox. In: *Theory and Decision* 9, 127–159.