WORKING PAPERS

Uwe Dulleck

The E-Mail Game Revisited – Modeling Rough Inductive Reasoning

November 2002

Working Paper No: 0211

DEPARTMENT OF ECONOMICS

UNIVERSITY OF VIENNA

The E-Mail Game Revisited - Modeling Rough Inductive Reasoning*

Uwe Dulleck [†‡]
University of Vienna

this version November 2002

Abstract

I study the robustness of Rubinstein's (1989) E-Mail Game results towards rough inductive reasoning. Rough induction is a form of boundedly rational reasoning where a player does not carry out every inductive step.

The information structure in the E-Mail game is generalized and the conditions are characterized under which Rubinstein's results hold. Rough induction generates a payoff dominant equilibrium where the expected payoffs change continuously in the probability of "faulty" communication.

The article follows one of Morris' (2001a) reactions to the E-Mail game "that one should try to come up with a model of boundedly rational behavior that delivers predictions that are insensitive to whether there is common knowledge or a large number of levels of knowledge".

JEL-classification: C72

Keywords: bounded rationality, inductive reasoning

---

*Previous versions came without section three and carried different titles. They were titled either "The E-Mail Game Revisited - ..." or "A Note on the E-Mail Game - ...".

†Department of Economics, Hohenstaufengasse 9, A - 1010 Vienna, Austria, E-Mail:uwe.dulleck@univie.ac.at, Ph.: +43-1-4277 37427, Fax: +43-1-4277 9374

‡I am grateful for helpful comments and encouragement by Carlos Alos-Ferrer, Ana Ania Martinez, Ken Binmore, Dirk Engelmann, Ulrich Kamecke, Rudolf Kerschbamer, Georg Kirschsteiger, Jörg Oechssler, Gerhard Orosel, Burkhard C. Schipper, Elmar Wolfstetter, an anonymous referee and seminar participants at University College London (ELSE). Financial support by the Deutsche Forschungsgesellschaft (DFG) through SFB 373 located at Humboldt University in Berlin is gratefully acknowledged.

1 Introduction

Inductive reasoning by economic agents is of great importance in many economic and game theoretic arguments. Maybe the most obvious example is subgame perfection. In this article I am presenting a model of rough inductive reasoning, a form of boundedly rational behavior. Rough inductive reasoning describes the reasoning process where a player does not or cannot carry out every inductive step.

Inductive reasoning seems of special importance to communication processes and the knowledge that is generated by communication. Rubinstein's (1989) Electronic Mail game illustrates this importance and shows with the "famous" E-Mail game paradox the problems created by perfect inductive reasoning. The Electronic Mail game describes differences between common knowledge and "almost common knowledge". In Rubinstein's example a small pertubation of the common knowledge assumption, his "almost common knowledge" changes the equilibrium set and expected payoffs dramatically, players' strategies do not condition on the outcome of the E-Mail communication. There is a major difference between common knowledge and high numbers of levels of knowledge. I generalize the underlying information structure of the E-Mail Game, specify the conditions needed to reach Rubinstein's result and argue that based on the generalized information structure one can describe models of rough inductive reasoning to solve the paradox.

Morris (2001a) argues that Rubinstein's results leads to two competing intuitions how to solve the paradox. Morris (2001a+b) and others[1] follow the notion of approximate common knowledge (ACK). They look for definitions of ACK which imply that rational behavior under ACK is close to rational behavior when there is common knowledge. In this article I advocate the other intuition that situations of common knowledge and high numbers of levels of knowledge are very close in the mind of players. Thus I try to analyze which features of bounded rational behavior are needed such that predictions are insensitive to whether there is common knowledge or a large number of levels of knowledge.

Note, I neither believe that the Electronic Mail game is a particularly realistic model of real world communication nor that the proposed forms of bounded rationality are especially intuitive. I do believe that real life inductive reasoning is rough in the sense specified below. I study the information

[1] Monderer and Samet (1989), Kajii and Morris (1998) and Morris and Shin (1997).

structure of the E-Mail Game to understand the effects of rough inductive reasoning in communication processes. The variants of the Electronic Mail game describe models of rough inductive reasoning for this communication process.

I believe that the hypothesis implied by experimental data that agents do not use induction correctly[2] may be due to the fact that they face limitations on utilizable information which are due to bounded rationality. Agents may actually use some form of induction and reach conclusions that are close to our intuition. The present versions of the E-mail game may be closer to the reasoning processes that agents use in such situations.

The proposed from of bounded rationality was inspired by the analysis of several problems which are also based on inductive reasoning. The absentminded driver (Piccione and Rubinstein (1997))[3], the absentminded centipede (Dulleck and Oechssler (1997)) and the E-Mail game, rest on the structure of the information sets of a decision maker. In many sequential games information sets similar to the one specified below for the E-Mail game arise from limitations of players' mental capacities to perform inductive reasoning.

Besides a characterization of the information structure needed to get Rubinstein's result I apply these idea in variants of Rubinstein's game. The variants of the game present two different assumptions of less than perfectly rational players that capture the idea of rough inductive reasoning. These alter the information structure and the belief structure such that an additional payoff-dominant equilibrium exists. In the equilibria of these games players condition on the E-Mail communication and the expected payoffs to players changes continuously in the probability that a message gets lost.

Dulleck and Oechssler's (1997) model is an example of rough inductive reasoning in the centipede game. They show that predictions under absentmindedness differ from predictions under perfect rationality and common knowledge. The E-mail game is an example where induction under bounded

[2]McKelvey and Palfrey (1992) and Rosenthal (1981) among others present experiments on the centipede game that support this hypothesis. In the literature one finds also competing hypotheses - for example altruistic behavior by subjects - which are supported too.

Colin Camerer (forthcoming) ran experiments on this game. Cabrales, Nagel and Armenter (2001) study a related "global" game.

[3]see also Aumann et al. (1997) and the other articles in the special issue of *Games and Economic Behavior* on the problem of imperfect recall.

rationality yields new (payoff dominant) equilibria. The new equilibrium is close to predictions under perfect rationality and common knowledge whereas in the centipede game a related change in the information structure solved the centipede paradox and the equilibrium was close to intuition and experimental observations.

A different solution to the E-Mail Game paradox is provided by Binmore and Samuelson (2001). In an evolutionary model they explicitly specify costs for observing messages (the cost function is increasing in the maximum number a player can recognize) and for sending messages. By introducing this cost they show that a payoff dominant equilibrium is evolutionarily stable where players choose action B if game G_b prevails and a sufficient number of messages is sent. In one of my setups the restriction of capacities to utilize information can be seen as being due to costs. Compared to their approach, in the present article it is possible that a player may not be able to keep track of messages for some period of time, whereas later on (l messages later) players might pay attention again. Binmore and Samuelson (2001) also do not allow for stochastic restrictions of capacities. Schipper (2001) uses a "simplified model" (related to corollary 1 below) where players for large numbers do not care exactly how many messages have been sent and then studies evolutionary stability between two models players may use to reason in this situation. He shows that both models - Rubinstein's and his simplified model - form evolutionarily stable strategy equilibria.

After introducing Rubinstein's (1989) E-Mail game and presenting his main result, the feasible information structures of the E-Mail Game are studied in section three. The main result of this section characterizes information structures that result in Rubinstein's paradox and those that lead to equilibrium strategies that condition on the E-Mail communication. Section four studies two alternative versions of the E-Mail game applying the result of section three. In these versions equilibrium predictions are close to predictions under perfect information and rationality. Section five concludes.

2 The E-Mail Game

In the Electronic Mail game two players either play a game G_a (with probability $(1-p) > \frac{1}{2}$) or G_b (with probability $p < \frac{1}{2}$). In each game players choose between action A and B. In both games it is mutually beneficial for players to choose the same action. Figure 1 describes the game. In game G_a

(G_b) the Pareto dominant equilibrium is the one where players coordinate on A (B). If players choose different actions the player who played B is punished by $-L$ regardless of the game played. The other player gets 1. It is assumed that the potential loss L is larger than or equal to the gain M and both are positive.

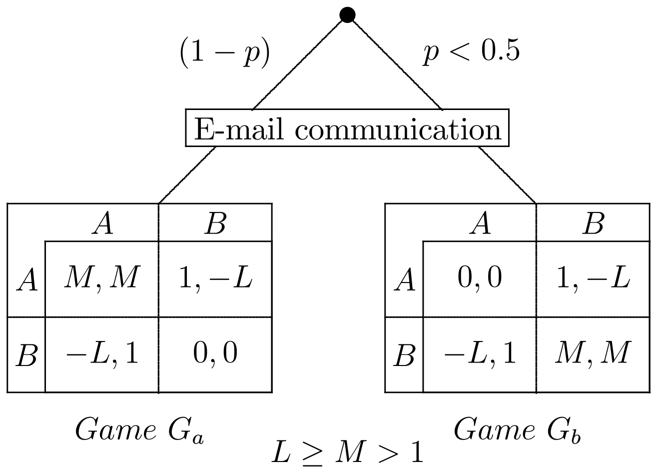

Figure 1: The structure of the E-mail Game (not a game tree!)

Only player one (she) is informed about the game that is actually played. After the game is determined two machines (one for each player) communicate which game is played. Only if game b prevails, player one´s machine sends an E-mail message (a beep) to player two´s (he) machine which is automatically confirmed. This confirmation is confirmed and so on. Let us first assume that messages arrive with certainty.[4] In this case, players have common knowledge about which game is played and a payoff dominant equilibrium exists, where players choose A in game G_a and B in G_b. Two other equilibria exist which have equilibrium strategies that are constant in the communication signals: always to play A and always to play B.

[4] Assume the time needed for the Tth message is equal to $1/2^T$. Under this assumption an infinite number of messages can arrive in a finite amount of time which is equal to 2.

Rubinstein analyzes the situation when the E-Mail communication is almost perfect: With probability ε a message gets lost. Communication stops, when one of the messages (the original message or one of the confirmations) is lost. Players only know how many messages their machine sent to the other player. The number of messages sent is denoted by T_i, $i \in \{1, 2\}$. Once a message is lost they have to make their decision. I assume that messages are likely (ε small) to arrive:

$$\varepsilon < \frac{M-1}{M+L}. \tag{1}$$

The feasible states of the world will be represented by pairs (T_1, T_2) consisting of the numbers of messages sent. I rule out that the machine of player one fails to send a message although we are in game G_b. Therefore a pair is sufficient, because game G_b must prevail if and only if at least one message has been sent by player one´s machine ($T_1 \geq 1$).[5] Player one always observes when her machine sends a message and therefore always knows which game is played. Note however that we do not rule out that this first message gets lost.

The Electronic Mail game represents a slight deviation from common knowledge ("almost common knowledge" in Rubinstein´s terms). Combined with perfect rationality this leads to a discontinuous drop in expected payoffs. Paradoxically in this case the game has an equilibrium, where players never play the payoff dominant equilibrium in one game (b) even if many messages were sent, given they play optimally whenever no message is sent (game G_a). Therefore communication is useless. Rubinstein (1989) proves that there is no Nash equilibrium where players condition on the number of messages sent. I follow the presentation of Osborne and Rubinstein (1994):

Proposition 1 *The electronic mail game has a unique Nash equilibrium, in which both players always choose A.*

The formal proof is provided in Osborne and Rubinstein (1994). The argument of the proof is helpful for the analysis below, therefore I give a short description. Figure 2 helps to understand the proof which is by mathematical

[5] I simplify the notation of Rubinstein (1989) who denotes the feasible states s of the world by triples consisting of the game actually played and the number of messages sent by the machines of player 1 and by player 2, i.e. $s \in \{(a, 0, 0); (b, 1, 0); (b, 1, 1); (b, 2, 1); (b, 2, 2), ...(b, T_1, T_2)...\}$, $T_2 \in \{T_1 - 1; T_1\}$.

induction. The induction starts with the fact that in states $(0,0)$ and $(1,0)$ the dominant strategy for player 1 is to play A thus (A,A) is the only equilibrium given $p < \frac{1}{2}$. The inductive hypothesis is that up to the informational outcome where a player's machine sent $T-1$ messages it is optimal for him to play A. From this hypothesis follows the inductive step that playing A if a player observes that T messages have been sent by his machine is optimal for the player. This is because the consistent belief $z = \frac{\varepsilon}{\varepsilon+\varepsilon(1-\varepsilon)} = \frac{1}{2-\varepsilon}$ to be at the first of two indistinguishable outcomes $(T, T-1)$ and (T,T) for player one [or $(T-1, T-1)$ and $(T, T-1)$ for player two] is greater than $\frac{1}{2}$. Given the other player chooses A at the first of the two nodes in an information set, it is the best reply to choose A because $z > \frac{1}{2}$ and $L \geq M$ which makes the decision independent of the strategy of the other player at the second indistinguishable outcome in the information set. By the principle of mathematical induction it follows that this is true for every observed T.

I will now generalize the structure of the information sets and the resulting beliefs of the E-Mail game and show under which conditions information structures yield Rubinstein´s result.

3 Generalized Information Structures of the E-Mail Game

Figure 2 gives a graphical representation of the communication process, where the automatic moves (by nature) of the machines are represented. The states are the informational "outcomes" of the moves by nature which are observed by the players before they make their decisions.

In Rubinstein´s game, player one cannot distinguish states (T_1, T_1-1) and (T_1, T_1) (and player two cannot distinguish states (T_2, T_2) and (T_2+1, T_2)). In this case the player observes only that T_1 (T_2) messages have been sent by her (his) machine. Player's strategies can condition on the number of messages sent by her (his) machine. The same action is chosen in two states of the world - for at least the two states where player i observes that T_i messages were sent by his machine. In Figure 2 two information sets for each player are marked. These are the sets where players observe that either one or two messages have been sent. A player has to form beliefs about which state in the information set is the actual one. These beliefs are necessary

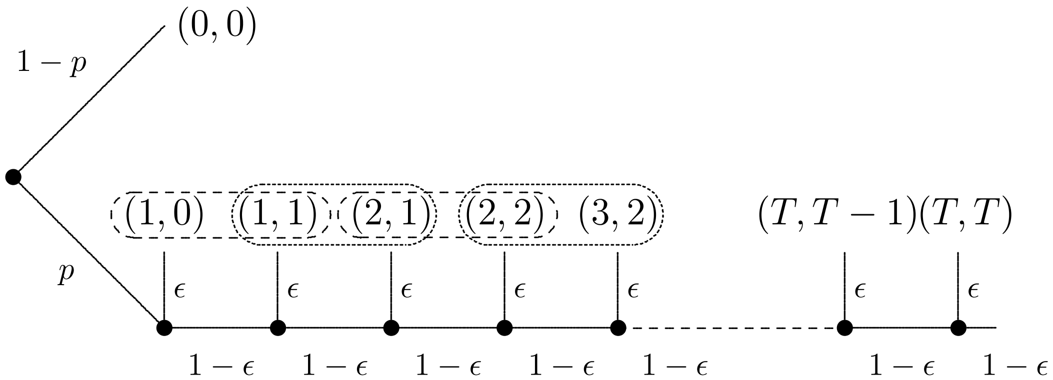

Figure 2: States and information sets in the E-mail Game (numbers are the number of messages sent / information sets marked where either 1 or 2 messages sent by player's machines)

to reach Rubinstein´s results. In Rubinstein´s game, each information set intersects with two information sets of the other player.

To study a generalized version of possible information structures, denote by I_i an information structure of player i. To distinguish the information sets in I_i denote by I_i^l element l in I_i. I define the following two relations to be able to make ordinal statements over information sets. Definition 1 introduces a lexicographic order over states and definition 2 applies this definition to order information sets by comparing the smallest elements/states in two information sets. In the following all statements are based on these relations.

Definition 1 *Let* $<_{Lx}$*be a lexicographic order over pairs of natural numbers such that* $(\tau_1, \tau_2) <_{Lx} (\vartheta_1, \vartheta_2) :\Leftrightarrow$ *(*$\tau_1 < \vartheta_1$*) or (*$\tau_1 = \vartheta_1$ *and* $\tau_2 < \vartheta_2$*).*

Definition 2 *Let* $\precsim_{Lx}$*be a lexicographic order over sets of pairs of natural numbers such that* $I_i^l \precsim_{Lx} I_i^k :\Leftrightarrow \min_{<_{Lx}} I_i^l <_{Lx} \min_{<_{Lx}} I_i^k$ *or* $\min_{<_{Lx}} I_i^l = \min_{<_{Lx}} I_i^k$*. Denote by* $\prec_{Lx}$*the strict relation.*

Assume that each player´s feasible information is represented by a partition over all states of the world with the refinement, that the information sets consist only of consecutive states of the world.[6] Without loss of generality I number the information sets in I_i such that if $l < k$ then $I_i^l \precsim_{Lx} I_i^k$.

[6]To be formally precise, the information sets in the partition must be order intervalls in $<_{Lx}$.

Rubinstein´s E-Mail Game is an example for information sets that fulfill this condition. The information sets in his information structure contain two consecutive states of the world. I allow for any information structure where a player counts the messages in arbitrary units. A different story for this is that players base the decision on the time elapsed in the communication process. If the time needed to send a certain message follows a fixed function, then the time used up before players are asked to make a decision is a signal, at least as informative as the number of messages sent. The information sets above allow players to count the time in different and not necessarily constant units.

To state the result, define the *head* of an information set as the elements in this information set which contain the smallest element and are in the intersection with one (and only one) information set of the other player. The other elements in each information set are called the *tail*. More formally, given an information set I_i^k of player i, let I_j^l be the minimum (with respect to $\precsim_{Lx}$) information set of the other player such that $I_i^k \cap I_j^l \neq \emptyset$. The *head* of I_i^k is defined as $I_i^k \cap I_j^l$ and the *tail* as $I_i^k \backslash I_j^l$.[7]

To refer to Figure 2, (1,1) is the head of the first information set marked for player two and (2,1) is the tail of this information set. Whereas the head is never an empty set, the tail can be empty (if the information set of the other player contains fully the information set of this player) or can contain several information sets of the other player.[8]

For presentational purpose I describe the following from the perspective of player one (she). It holds vice versa for player two (he). A strategy maps the informations sets into the set of feasible actions $\{A, B\}$. Equilibrium strategies of player one are best replies given the strategy of player two. In equilibrium she chooses for each information set the action that is optimal given the actions he chooses in the intersecting information sets (one or more than one). In particular, it is of importance whether actions chosen by the other player differ in the head and the tail of an information set.

[7] For finite sets this is well defined.

[8] To given an example, assume player one recognizes only odd numbers, whereas player two only counts every fourth number, e.g. $1, 5, 9$ etc.. Both recognize if at least one message has been sent. Thus player one's partition of the state space is $\{\{(0,0)\},\ \{(1,0),(1,1),(2,1),(2,2)\},\ \{(3,2),(3,3),(4,3),(4,4)\},\ ...\}$. Whereas the same set of player two is $\{\{(0,0),(1,0)\},\ \{(1,1),(2,1),(2,2),(3,2),(3,3),(4,3),(4,4),(5,4)\},\{(5,5),(6,5),...\}$. The head of the second information set of player two contains the elements $(1,1)$, $(2,1)$ and $(2,2)$.

A strategy that makes use of the communication should be monotone in the number of messages sent. I name a strategy *monotone* if there exists a k such that for all I_i^l with $I_i^l \precsim_{Lx} I_i^k$ action A is chosen whereas otherwise B is chosen.[9] That is, monotone strategies change only once from choosing A to B over the number of messages sent, thus if more than a fixed number of messages is sent then action B will be played, whereas whenever less than this number of messages is sent then A is chosen. If at least one information set contains not too many elements in the head compared to the number of elements in the tail, then there exists a payoff dominant Nash-equilibrium in which players' equilibrium strategies condition on the number of messages sent. The number is determined by the payoffs. For $M = L$ it suffices that the head contains strictly less elements than the tail. Let n be the number of elements in I_i^*- the first information set where player i plays B instead of A - and n^h the number of elements in the head of the information set.

Proposition 2 *If at least for one information set of one player* $\frac{n^h}{n} < \frac{M-1}{L+M-1}$ *then for all* $\varepsilon < \widehat{\varepsilon}$, *a payoff dominant equilibrium exists.*

Proof. Let player i be the player who has the smallest ($\precsim_{Lx}$) information set (of both players) which fulfills the condition of the proposition and denote this information set by I_i^*. Denote by I_j^* the information set of player j that is associated with the head of I_i^*. I prove by construction that an equilibrium exists where players strategies are monotone and player i plays B from I_i^* onwards and player j plays B from the information set following I_j^*.

Denote by $\widetilde{z}$ the belief that player i assigns to being in in the head of I_i^* whenever the state of the world is in I_i^*. This belief is given as

$$\widetilde{z} = \frac{\sum_{i=1}^{n^h} \varepsilon(1-\varepsilon)^{i-1}}{\sum_{i=1}^{n} \varepsilon(1-\varepsilon)^{i-1}} = \frac{1-(1-\varepsilon)^{n^h}}{1-(1-\varepsilon)^n}. \tag{2}$$

For a best reply of player i to play B given player j follows the proposed strategy, the following must hold: $\widetilde{z}(-L) + (1-\widetilde{z})(M) > (1-\widetilde{z})1$. This converges for $\varepsilon \to 0$ to $\frac{n^h}{n} < \frac{M-1}{M-1+L}$.

Given the definition of I_i^*, it is the smallest information set of both players where the condition holds, thus for I_j^* it must be optimal for player j to

[9] I do not consider the case where a player switches once from B to A because this is dominated by playing always B.

play A independent of the action player i in I_i^* . Regarding player j's equilibrium strategy the following holds. For all $I_j^l \precsim_{Lx} I_j^*$ the strategy either the condition is not fulfilled or player i plays A in every state of the world contained in I_j^l. For all $I_j^l \succ_{Lx} I_j^*$ player i plays B in all states of the world contained in I_j^l. ■

To phrase the result in a different way, the proposition says if at least one information set of one of the players has a head that contains less elements than its tail then payoffs $M \leq L$ exist which support a payoff dominant equilibrium with strategies that condition on the E-Mail communication.

4 Rough Induction in the E-Mail Game

The intuition behind Rubinstein´s results is based on the conditional probabilities that whenever a message got lost, it is for both players more likely that the last message her (his) machine sent was lost than that the reply to this message was lost. In the following variations of the E-Mail Game are studied. First the effect of non-distinguishability is analyzed. This represents the case where a player loses track at period T for some periods or a player is due to some cognitive inability or a technical problem of his machine unable to differentiate between some states of the world. This case is also analyzed for forms of stochastic non-distinguishability. In Proposition 3 I state that a critical number of messages sent exists such that whenever more messages than this number are sent players actually use the communication to coordinate on B if G_b prevails. The assumption made is that the probability that (at least) one player loses track for some period of the communication increases over time. Second the effect of counting in different units is studied. This case has a similar intuition as Morris (2001b) studying a "timing" version of the E-Mail game. Both variants lead to the existence of additional payoff dominant equilibria.

4.1 Non-distinguishability

Rubinstein (1989) states that a payoff dominant equilibrium exists if a maximum number of messages can be sent or is recognized (fixed capacity of players). In contrast to this extreme, I do assume that players have an unlimited capacity (there is no maximum number $\overline{T}$ of messages sent they recognize). They may lose track for some numbers of messages sent but are able to pay attention later again (e.g. due to analyzing available extra information). A different way to describe this situation is that the machine of a player may malfunction for some periods but show the correct number of messages sent later again. This weaker form of bounded rationality suffices to generate a payoff dominant Nash equilibrium as a corollary of proposition 2. By reducing the ability to process information the existence of an additional equilibrium is guaranteed. In this payoff dominant equilibrium players condition their strategy on the communication. Let l be the number of elements that cannot be distinguished from a given number T.

Corollary 1 *If exactly one player cannot distinguish the numbers of messages sent if $t \in \{T, T+1, ..., T+l\}$ then a payoff dominant equilibrium exists where players condition their strategies on the communication process if $T > 0$ and $l > l^*$. l^* always exists. In equilibrium both players play B if their machine sent $t \geq T$ and A in all other cases.*

The result of proposition 2 applies because in this case the head only contains one element whereas the tail contains at least three.

One special case of this set-up is that a player loses track at a fixed stage T ($l = \infty$) or his machine stops counting/sending after T messages. If the maximum number of messages the machine sends is restricted then players play B if, and only if, their machine sent this maximum number. This case ($l = \infty$) may be simple to motivate. The machine or the player loses track at a fixed stage T. Justifications for this assumption could be the "overflow" of the machine´s capacities (it can only count up to a certain number) or that real players actually stop counting after they sent a certain number of messages or players have to decide after a certain number of messages has been sent. This is Rubinstein's (1989) solution to the paradox.

I show that the weaker condition that players cannot distinguish between some states is enough to yield a subgame perfect equilibrium with coordination. The machine may not be able to show 18 and therefore it stays on 17 for two turns and then jumps to 19. A player who did not pay attention for a while, may be able to remember later again how many messages where sent (e.g.. he is additionally informed whenever 100 messages have been sent). If the player cannot distinguish/remember whether his machine sent $T, T+1, ...$ or $T+l$ messages, he has to choose one action for all observations in the interval. Assuming $l \to \infty$ and T fixed is an extreme restriction of players' abilities to distinguish numbers. The $l \to \infty$ assumption seems unrealistic because people are known to mess up especially large numbers but in the case numbers differ substantially people do recognize the relation, and whenever prominent numbers are reached, people pay attention again. To give an example 1389 and 1394 may be put in the wrong order whenever the decision maker is under stress. This effect is not present when the numbers in question are 1389 and 10394.[10] It seems to be the case that people do not distinguish sets of numbers, but they realize a substantial difference.

The following remark states an interesting observation about the importance of knowing the underlying fact of the communication process. This results from the case of $T = 0$.

Remark 1 *If players become only aware of the E-Mail communication if a sufficient number of messages has been sent then it matters whether a player knows the underlying fact (that game b is played).*

Given $l > l^*$, corollary 1 also allows for the case where player one (the one that is informed about whether game a or game b is played) cannot distinguish among the set $\{1, 2, ..., l\}$. This models the situation where player one knows which game is played but he only pays attention to (becomes aware of) the communication if a sufficient number of messages has been sent. The result states that if one player knows the underlying fact but only pays attention to the communication if a large enough number of messages has been sent then in almost all cases the payoff dominant equilibrium exists. If player two suffers from non-distinguishability the set of non-distinguishable numbers $\{0, 1, 2, .., l-1\}$ models the situation where player two only pays attention if a high enough number of messages is sent. This case is not covered by proposition 1. If there is no other set of non-distinguishable numbers, the only equilibrium where players play A whenever no message is sent is one where players always choose A, independent of the number of messages sent.

A more plausible assumption is that players are more likely to lose track with higher numbers of messages sent. I analyze this case now. The probability that player one suffers from non-distinguishability increases with the number of messages sent. It gets more and more likely that player one cannot distinguish $l + 1$ consecutive numbers of messages sent. For simplicity of presentation I assume that l is exogenously determined. Results can be generalized in a way similar to corollary 1. For this section the belief $\widetilde{z}$ to be

[10] Tversky(1969, 1977) emphasizes the role of similarities on human reasoning. Theoretical contributions are Rubinstein (1988) and Albers and Albers (1988).

in the head of an information set is given as

$$\widetilde{z} = \frac{\varepsilon}{\sum_{i=0}^{2l+1} \varepsilon(1-\varepsilon)^i} = \frac{\varepsilon}{1-(1-\varepsilon)^{2l+2}}. \tag{3}$$

$\eta(t)$ is the probability that $\{t; t+1; ...; t+l\}$ is a non-distinguishable set of numbers for player one. It is defined to strictly increase in t. If non-distinguishability affects the utilizable information when t messages have been sent by player one´s machine, it will not do so for the following (non-distinguishable) l numbers of messages sent. This implies that if the player cannot distinguish $\{16, 17, .., 20\}$ then $\{17, 18, .., 21\}...\{20, 21, .., 24\}$ cannot be sets of non-distinguishable numbers for this player. The next set of numbers among which player one with probability $\eta(21)$ cannot distinguish is $\{21, 22, .., 25\}$. In the appendix I will define random variables and their distribution to define the process more formally.[11]

$\eta(t)$ is common knowledge among players. The result holds, too, if $\eta(t)$ is a common belief – therefore, it is enough that the player who may suffer from non-distinguishability knows the other player's belief and this is common knowledge. We find the following result:

Proposition 3 *If* $\frac{1-\widetilde{z}}{\widetilde{z}}(M-1) > L$ *and* $\eta(t)$ *is monotone increasing then there exists a critical number of messages sent, such that in a payoff dominant equilibrium players coordinate on action* B *if more than this critical number of messages have been sent by their machines.*

The proof is provided in the appendix.

For an increasing probability that player one cannot distinguish among $l+1$ consecutive numbers, an endogenously determined number T^* of messages sent exists such that players choose always the mutually beneficial actions if more than T^* messages have been sent by both machines. If ε is small this number will be reached almost for sure.

The following strategies form the equilibrium in question: Let T^* be the number where both and especially player two chooses B if he observes $t \geq T^*$ and A otherwise. Player one plays B whenever she observes a $t > T^*$.

[11] I assume a certain form of independence. The result holds too without indepency, in this case one needs to apply *Sylvester's Lemma* (see for example Theorem 19, p. 24 in Mood, Graybill and Boes (1974)) to prove Proposition 3.

If she suffers from non-distinguishability in a way that T^* is in the non-distinguishable set of numbers player one plays B given that the smallest number of messages sent contained in her affected information set is not smaller than $T^* - k^*$. In all other cases she plays A. In the appendix I prove that this is an equilibrium given the stated restrictions on l, M and L. T^* and k^* are also characterized in the appendix. Player one´s strategy relative to player 2's strategy is determined by k^* and depends only on ε given player two plays B if a fixed number of messages is sent. It is independent of player two´s strategy which is characterized by T^*. T^* is determined by k^* and $\eta(t)$.

Actions chosen by players differ in Rubinstein's original setting and the variation presented above - even if players precisely observe the same number of messages has been sent by their machine. Suppose a player in Rubinstein´s original game and a player in the extended game (one player suffers (with positive probability) from non-distinguishability) observe each that exactly τ messages have been sent by their machine. If τ is greater than a critical number of messages sent then best-reply-strategies differ. An informational deficiency which potentially affects the utilizable information at another (smaller / "earlier") informational "outcome" breaks the induction. A local deficiency in information processing abilities changes the optimal strategy even though at the decision making point in time the utilizable information is the same as in the case where no local deficiency exists.

4.2 Counting in different units

To highlight a special feature of Rubinstein's result, I characterize in this section the information structures needed to get Rubinstein's result regarding the potential information structures. Consider a situation where players ($i = 1, 2$) only count or are only informed of every u_ith message sent (they count in units of size $2u_i$ states of the world), e.g. they only count even numbers[12]. Rubinstein´s standard setup is the case $u_1 = u_2 = 1$. In this subsection the situation where players count in different units is analyzed. This may for example be the case, if a player receives the information how long his/her machine sent messages before the break-down of communication. If the time needed follows a fixed function then this information can be as

[12]The reasoning behind the results of this section is similar to the reasoning presented above. Nonetheless, formal proofs are available from the author upon request.

rich as in the original game but it seems more likely that machines count the time in different units.

Consider that both players have $n_i > 1$. Another corollary of proposition 2 shows that Rubinstein's results are only valid for a special case of counting in different units.

Corollary 2 *If both players have the same odd $u_1 = u_2 = u$ and player two (one) starts counting $(u-1)/2$ ($(u+1)/2$) received messages later than player one (two) then Rubinstein's result applies and no equilibrium exists where players condition on the E-Mail communication. In all other cases ($u_1 \neq u_2$, u even or the information sets do not intersect in the prescribed way) then if $L - M$ not too large a payoff dominant equilibrium exists where both players condition on the E-Mail communication. In this equilibrium players choose B if the prevailing state of the world is for both players contained at least in the second information set.*

In Rubinstein's game the information sets of both players intersect with information sets of the other player. In the case that $n_1 = n_2$ each information set intersects with two information sets of the other player, for the case $n_1 = n_2 = 1$ these intersections contain one element only (see Figure 2), thus head and tail of each information set contain an equal number of states of the world. The condition of proposition 2 is fulfilled for sure if the state of the world is contained at least in the second information set.

Let me finally state an interesting aspect about the effect of the size of the information sets on the expected payoffs in the resulting equilibria:

Remark 2 *Rubinstein´s "almost common knowledge" leads to a discontinuous drop with regard to ε in the expected payoff. With non-distinguishability or counting in different units the change in the expected payoff is continuous.*

Consider the expected payoffs of the Electronic Mail Game under the various assumptions on player information. If it is common knowledge ($\varepsilon = 0$) which game is played, the expected payoff of the payoff dominant equilibrium is $\Pi^e = M$. Introducing the probability ε that a message gets lost (Rubinstein´s (1989) article) the expected payoff of the payoff dominant equilibrium drops to $\Pi^e = pM$, because only in game a players choose the mutually beneficial actions. If the messages may get lost but one of the players (potentially) suffers from non-distinguishability, the expected payoff drop changes continuously in ε. For small ε it is almost the same as under common knowledge, i.e. $\lim_{\varepsilon \longrightarrow 0} \Pi^e = M$. Restricting the information to the case where only player one knows about the state of the world, i.e. player two suffers from non-distinguishability with $T = 0$ and $l = \infty$, the expected payoff is the same as in Rubinstein´s game.

5 Conclusions

I introduced a model of rough inductive reasoning and applied it to Rubinstein's (1989) Electronic Mail game. Rubinstein (1989) employs the Electronic Mail game to illustrate that the payoffs vary discontinuously in the assumed information structure, i.e. "almost common knowledge" leads to different optimal behavior compared to the optimal behavior under common knowledge. This paradoxical result has led to a discussion about good definitions of common knowledge in such situation, Morris (2001 a+b) and others advocate approximate common knowledge as an appropriate path to follow. My approach differs. I generalize and study the underlying information structures that yield the results and provide an understanding of rough inductive reasoning, a form of boundedly rational behavior that solves the paradox.

In this article the assumed forms of bounded rationality or rough inductive reasoning - non-distinguishability and counting in different units - lead to the existence of an additional equilibrium in the respective game which is payoff dominant. In the respective equilibrium the expected payoffs change continuously in ε. Given a change in the optimal strategy at one stage, induction leads to an additional payoff dominant equilibrium if a sufficient number of messages has been sent. It is therefore not necessary - as it was in Rubinstein's solution - that bounded rationality affects the decision maker when he has to make his decision.

Instead of introducing an alternative concept for knowledge under faulty communication processes I presented models of bounded rationality that capture the idea of rough inductive reasoning to solve Rubinstein's (1989) paradox.

After generalizing the information structures of the E-Mail Game I showed which structures yield Rubinstein's paradox. Models of "non-distinguishability" and "counting in different units" provided examples for rough inductive reasoning. Both resolved the paradox and resulted in equilibria that predict communication to be used.

6 Appendix - Proof of Proposition 3

We assume that l is large enough, i.e. the following holds:

$$\frac{1-\widetilde{z}}{\widetilde{z}}(M-1) = (M-1)\sum_{i=1}^{2l+1}(1-\varepsilon)^i > L \ . \tag{4}$$

To prove the proposition, I start by proving a lemma that captures the case of a sufficiently large probability η to suffer from non-distinguishability at the states of the world where a player observes $t \in \{T, T+1, ..., T+l\}$. For the ease of presentation, I state the argument for the case where player one may suffer from non-distinguishability.

Lemma 1 *Suppose player one suffers with probability η from non-distinguishability such that he cannot distinguish among the $t \in \{T, T+1, ..., T+l\}$. If $\frac{1-\widetilde{z}}{\widetilde{z}}(M-1) > L$ and $(1-\eta) \leq \frac{(2-\varepsilon)(M-1)}{M+L-1}$ then there exists a payoff dominant Nash equilibrium where both players play B whenever $t \geq T+1$ messages have been sent by their machines. For $M = L$ large, $\eta < \varepsilon$ suffices for the existence of such an equilibrium.*

The following strategies are the equilibrium in question: If player one suffers from non-distinguishability she chooses B if she observes a $t \geq T$ and A otherwise. If player one does not suffer from non-distinguishability, she chooses B whenever she observes a $t > T$ and A otherwise. Player two chooses B if he observes a $t \geq T$ and A otherwise. I will prove that this is an equilibrium given the stated restrictions on M, L and η.

Proof of Lemma 1. I show that the strategies are best reply strategies. Up to "outcome" $(T-1, T-1)$ the equilibrium strategies are proved by the inductive argument of Rubinstein (proposition 1).

Given that player two's strategy is the same as the strategy of player two in corollary 1, the corollary ensures that the described behavior of player one is a best reply whenever she suffers from non-distinguishability (and l is large enough). If player one does not suffer from non-distinguishability her best reply is to play A whenever she observes $t \leq T$ messages have been sent by her machine (see proposition 1). If she observes a $t > T$ then given the stated strategy of player two her best reply is to play B because player two plays B at both of the states in her information set. Therefore given the strategy of player two player one´s strategies are best replies.

Given the strategy of player one, player two plays B whenever he observes that exactly T messages have been sent if his expected payoff is greater than the payoff from playing A, i.e.

$$\eta M + (1-\eta)(z(-L) + (1-z)M) \geq \eta + (1-\eta)(1-z) \tag{5}$$

where $z = \frac{\varepsilon}{\varepsilon+\varepsilon(1-\varepsilon)} = \frac{1}{2-\varepsilon}$ is the consistent belief of the player not suffering from non-distinguishability that the state is (T, T) instead of $(T+1, T)$. Player two´s payoff is 0 if he chooses A. Therefore B is a best reply if (5) holds. This is equivalent to $(1-\eta) \leq \frac{(2-\varepsilon)(M-1)}{M+L-1}$ as stated in the proposition.

Given the observation of any $t > T$ by player two his best reply is to choose B because player one chooses B at both states of the world in the information set in question.

For $M = L$ the condition $(1-\eta) \leq \frac{(2-\varepsilon)(M-1)}{M+L-1}$ simplifies to $\eta \geq \frac{\varepsilon(M-1)+1}{2M-1}$, where $\frac{\varepsilon(M-1)+1}{2M-1} < \varepsilon \Leftrightarrow M > \frac{1}{2\varepsilon}$. ■**(Lemma)**

If this holds and the probability to suffer from non-distinguishability is large compared to the probability that a message gets lost a Nash equilibrium exists where players condition on the number of messages sent.

Inequality (4) is the condition such that the player suffering from non-distinguishability behaves as predicted by the equilibrium beliefs of the other player. It determines a critical $\widehat{l}$ such that for given M, L, ε and $l > \widehat{l}$ the equilibrium strategy of the player suffering from non-distinguishability is a best reply strategy. Whereas the other players strategy is best reply if condition (5) holds. The latter condition determines a critical level of η for given $M, L, \varepsilon M$. Hence, analyzing optimal behavior of the players one needs to look for one player only at the critical level for l and for the other player only at the critical level for η. This form of independence is important for the proof of the proposition.

As an instrument of presentation I define a series of independently distributed random variables $X_t \in \{0; 1\}$, where $P(X_t = 1) := \eta(t)$ is the probability that $X_t = 1$. A player cannot distinguish among a set of numbers $\{t, t+1, ..., t+l\}$ if and only if $X_t = 1$ and t itself is not an element of a non-distinguishable set (more formally $\forall \tau \in \{t-l, ..., t-1\}\ X_\tau = 0$). Given these definitions we prove the result.

First I analyze the behavior of the from non-distinguishability suffering player one. Assume that player two plays B if and only if he observes a

$t_2 \geq T$. Let

$$k^* = \max_k \left\{ k \,\middle|\, \underbrace{\frac{\sum_{i=0}^{2k} \varepsilon(1-\varepsilon)^i}{\sum_{i=0}^{2l+1} \varepsilon(1-\varepsilon)^i}(-L)}_{\text{expected loss } (t<t_2)} + \underbrace{\frac{\sum_{i=2k+1}^{2l+1} \varepsilon(1-\varepsilon)^i}{\sum_{i=0}^{2l+1} \varepsilon(1-\varepsilon)^i}(M-1)}_{\text{expected gain } (t_2 \geq T)} \geq 0 \right\}.$$

k^* is described as follows. The left hand side of the inequality is the expected payoff increase if the player chooses B instead of A given he knows that the number of messages sent is in the set of non-distinguishable numbers and the smallest element in this set is equal to $T-k$ ($X_{T-k} = 1$). $T - k^*$ is the smallest number of messages sent so that player one gets a non-negative expected payoff given player two plays B whenever he observes that more than T messages have been sent by his machine. If the smallest number in a non-distinguishable set is smaller than $T - k^*$ then the expected payoff is negative. Note that k^* depends only on l for a given M, L, ε and is independent of $\eta(t)$ and therefore independent of T.

Next consider the behavior of player two. We know from above that if player one does not suffer from non-distinguishability he will play A if he observes a $t_1 \leq T$. Denote by $\widetilde{\eta}(T)$ the probability that player one plays B because he suffers from non-distinguishability whenever exactly T messages have been sent by player two´s machine,

$$\widetilde{\eta}(T) = 1 - \prod_{i=T-k^*}^{T} (1-\eta(i)).$$

Define $T^* = \min_\tau \left\{\tau \,\middle|\, (1-\widetilde{\eta}(\tau)) \leq \frac{(2-\varepsilon)(M-1)}{M+L-1}\right\}$ as the critical number of messages sent. T^* is the number of messages send such that player two´s expected payoff from playing B is positive.

To prove best reply characteristics of the strategies I show first that given player two´s strategy, player one´s strategy is a best reply. Given the utilizable information is not affected optimality follows from the fact that for $t > T^*$ player two plays B for all states of the world which are element of the information set. If the utilizable information is affected and T^* is among the non-distinguishable numbers, by definition of k^* B is optimal to choose for this information set if the smallest element in the non-distinguishable set is larger than $T^* - k^*$.

Next I prove that player two´s strategy is a best reply. If $(1 - \widetilde{\eta}(T^*)) \leq \frac{(2-\varepsilon)(M-1)}{M+L-1}$ player two´s expected payoff from playing B is positive if he observes that exactly T^* messages have been sent by his machine. The idea is the same as in the proof of lemma 1 equation (5).

This completes the proof of best reply of player one given player two plays as described. Note T^* is defined as the smallest number of messages such that the expected payoff is positive. Therefore player two will never play B when he observes a $t < T^*$.

The existence of T^* is guaranteed by the condition on $\eta(t)$ and the obvious fact that $\widetilde{\eta}(t)$ increases in t. ■

References

[1] Albers, Wulf and Albers, Gertrud (1988), On the Prominence Structure of the Decimal System, in Scholz, R.W. (ed.), *Decision Making Under Uncertainty*, North-Holland, Amsterdam, p. 271-287

[2] Aumann, Robert J., Hart, Sergiu and Perry, Motty (1997), The Absent-Minded Driver, *Games and Economic Behavior*, 20, p. 102-116

[3] Binmore, Ken and Samuelson, Larry (2001), Coordinated Action in the Electronic Mail Game, *Games and Economic Behavior,* 35, pp. 6-30

[4] Cabrales, Antonio, Nagel, Rosemarie and Armenter, Roc (2001), Equilibrium Selection through Incomplete Information in Coordination Games: An Experimental Study, mimeo, Universitat Pompeu Fabra

[5] Camerer, Colin (forthcoming), Iterated Reasoning in Dominance Solvable Games, chapter 5 of Behavioral Game Theory: Experiments on Strategic Interaction, mimeo, California Institute of Technology

[6] Dulleck, Uwe and Oechssler, Jörg (1997), The Absent-Minded Centipede, *Economics Letters*, 55, p. 309-315

[7] Fagin, Ronald, Halpern, Joseph Y., Moses, Yoram and Vardi, Moshe Y. (1995), *Reasoning About Knowledge*, MIT Press, Cambridge

[8] Gray, Jim (1978), Notes on database operating systems, in Bayer, R., Graham, R. M. and Seegmüller (ed.), *Operating Systems: An Advanced Course*, Springer-Verlag, New York

[9] McKelvey, Richard D. and Palfrey, Thomas R. (1992), An experimental study of the centipede game, *Econometrica*, 60, p. 803-836

[10] Kajii, Atsushi and Morris, Stephen (1998), Payoff Continuity in Incomplete Information Games, *Journal of Economic Theory,* 82, p. 267-76

[11] Monderer, Dov and Samet, Dov (1989), Approximating Common Knowledge with Common Beliefs, *Games and Economic Behavior,* 1, pp. 170-190

[12] Mood, Alexander M., Graybill, Franklin A. and Boes, Duane C. (1974), *Introduction to the Theory of Statistics*, 3rd edition, McGraw-Hill, Singapore

[13] Morris, Stephen (2001a), Coordination, Communication and Common Knowledge: A Retrospective on the Electronic Mail Game, draft text prepared for the 2001 European Meetings of the Econometric Society, mimeo

[14] Morris, Stephen (2001b), Faulty Communication: Some Variations on the Electronic Mail Game, *Advances in Theoretical Economics (B.E. Journals in Theoretical Economics)*, Vol. 1, No. 1, Article 5

[15] Morris, Stephen and Shin Hyun S. (1997), Approximate Common Knowledge and Coordination: Recent Lessons from Game Theory, *Journal of Logic, Language and Infomation*, 6, p. 171-190

[16] Osborne, Martin J. and Rubinstein, Ariel (1994), *A course in game theory*, MIT Press, Cambridge

[17] Piccione, Michele and Rubinstein, Ariel (1997), On the Interpretation of Decision Problems with Imperfect Recall, *Games and Economic Behavior*, 20, p.3-24

[18] Rosenthal, Robert (1981), Games of perfect information, predatory pricing and chain-store paradox, *Journal of Economic Theory*, 25, p. 92-100

[19] Rubinstein, Ariel (1988), Similarity and Decision-Making Under Risk, *Journal of Economic Theory*, Vol. 46, p. 145-153

[20] Rubinstein, Ariel (1989), The Electronic Mail Game: Strategic Behavior Under Almost Common Knowledge, *AER*, Vol. 79, No. 3, p. 385-391

[21] Schipper, Burkhard C. (2001), A Note on Rubinstein's Electronic Mail Game, mimeo, University of Bonn

[22] Tversky, Amos, Intransitivities of Preferences, *Psychological Review*, Vol. 76, p. 31-48

[23] Tversky, Amos, Features of Similarity, *Psychological Review*, Vol. 84, p. 327-352