Pragmatic Rationalizability

Gerhard Jäger Christian Ebert
Dept. of Linguistics and Literature
University of Bielefeld

{gerhard.jaeger,christian.ebert}@uni-bielefeld.de

Abstract
We present a formal game-theoretic model towards the explanation of implicatures based on the computation of *iterated best responses*: literal meaning of signals constitutes their default interpretation, and rational communicators decide about their communicative strategies by iteratively calculating the best response to this default strategy. We demonstrate by means of several examples how the resulting *pragmatically rationalizable* strategies account for different types of implicatures.

1 Signaling Games

In order to introduce the basic concepts of the underlying game-theoretic model (see e.g. Osborne and Rubinstein, 1994, for an introductory textbook on game theory), let us look at a simple scenario where communication makes a decisive difference. Suppose Robin invited Sally for dinner where he wants to serve some thai curry. While slicing the chili he realizes that he is unsure about whether Sally likes her curry hot or not. Robin obviously wants to offer his guest the curry in her preferred way. In other words, they both prefer the outcome where Sally receives her favorite type of curry over the other outcome, where she finds it inedible because of the lack or the abundance of chili.

We may formalize this scenario as follows. There are two possible worlds, w_1 and w_2. In w_1 Sally prefers mild curry; in w_2 she likes it hot. Robin has a choice between two actions: preparing a mild curry would be action a_{mild}, and preparing a hot curry action a_{hot}. Sally knows how she likes her curry, i.e. she knows which world they are in, but poor Robin does not. But although he does not know for sure, he may have some *a priori* belief about Sally's liking, i.e. about the probabilities of each world. Maybe Robin has seen her eat jalapeño spiced taco burgers on another occasion such that he assumes w_2 to be more likely than w_1. In our concrete example, let us assume that he is totally clueless – he assigns both worlds an *a priori* probability of 50%.

This scenario can be represented formally as follows. In game theory, the preferences are usually encoded by assigning numerical values, called *utilities* or *payoffs*, to each outcome for each player. For instance, as Robin prefers the outcome of performing action a_{mild} in world w_1 over the outcome of performing the same action in w_2, we may assign 1 to the first and 0 to the second outcome and use the $\geq$-order on natural numbers to reflect the preference order. Continuing like that we arrive at the following *utility matrix*. Rows represent possible worlds and columns represent Robin's actions. The first number in each cell gives Sally's payoff for this configuration, and the second number Robin's payoff.

	a_{mild}	a_{hot}
w_1	$1,1$	$0,0$
w_2	$0,0$	$1,1$

(1)

Without any further coordination between the players, Robin will remain clueless and he will have to prepare one type of curry hoping to guess the right one. His *expected utility/payoff* for performing action a is as follows (given his prior belief p^*, the set of possible worlds W and his utilities $u_R(w,a)$ for the outcome (w,a)):

$$\text{EU}(a) = \sum_{w \in W} p^*(w) \cdot u_R(w,a) \tag{2}$$

In the case at hand he receives an expected payoff of 0.5 for either action, and (because of their identical preferences/utilities) Sally will also receive an expected payoff of 0.5. They can do better though if they communicate. Sally might simply tell Robin her favourite type of curry. Suppose Robin expects that Sally says "*mild*" in w_1 and "*hot*" in w_2. Then the rational course of action for Robin is to perform a_{mild} if he hears "*mild*", and to perform a_{hot} upon hearing "*hot*". In other words, Robin learns the actual world from Sally's utterance, i.e. revises his belief, and acts accordingly. On the other hand, if Sally beliefs that Robin will react to these signals in this way, it is rational for her to say "*mild*" in w_1 and "*hot*" in w_2. If they follow this rational course of behaviour they both will obtain an overall payoff of 1.

So adding the option for communication may improve the payoff of both players. Technically, the original scenario (which is not really a game but a decision problem because Sally has no choice between actions) is transformed into a *signaling game*. Here the sender (Sally in the example) can send signals, and she can make her choice of signal dependent on the actual world. Formally, the sender's behaviour is given by a *sender strategy*, which is a function from possible worlds to signals. The receiver (Robin in the example) can condition his action on the signal received. So a *receiver strategy* is a function from signals to actions. We will represent strategies graphically as tables indicating the corresponding functions:

$$\begin{array}{cc} \text{Sally's strategy } s: & \text{Robin's strategy } r: \\[1ex] \begin{bmatrix} w_1 & \rightarrow & \text{“}\textit{mild}\text{”} \\ w_2 & \rightarrow & \text{“}\textit{hot}\text{”} \end{bmatrix} & \begin{bmatrix} \text{“}\textit{mild}\text{”} & \rightarrow & a_{\text{mild}} \\ \text{“}\textit{hot}\text{”} & \rightarrow & a_{\text{hot}} \end{bmatrix} \end{array} \tag{3}$$

The above example suggests that rational players will benefit from the option of communication. Things are not that simple though. Consider the following pair of strategies:

$$\begin{array}{cc} \text{Sally's strategy } s': & \text{Robin's strategy } r': \\[1ex] \begin{bmatrix} w_1 & \rightarrow & \text{“}\textit{hot}\text{”} \\ w_2 & \rightarrow & \text{“}\textit{mild}\text{”} \end{bmatrix} & \begin{bmatrix} \text{“}\textit{hot}\text{”} & \rightarrow & a_{\text{mild}} \\ \text{“}\textit{mild}\text{”} & \rightarrow & a_{\text{hot}} \end{bmatrix} \end{array} \tag{4}$$

If Sally and Robin play these strategies they will also end up with the maximal payoff of 1. Pure reason does not provide a clue to decide between these two ways to coordinate. It is thus consistent with rationality that Sally assumes Robin to use r' and thus to signal according to s', while Robin assumes Sally to use s, and thus will interpret her signals according to r. In this situation, Robin will perform a_{hot} in w_1 and a_{mild} in w_2. Both players would receive the worst possible expected payoff of 0 here.

These considerations ignore the fact that the two signals have a conventional meaning which is known to both players. In our example, we would say that the conventional meaning of "*mild*" is the proposition $[\![mild]\!] = \{w_1\}$ whereas the meaning of "*hot*" is $[\![hot]\!] = \{w_2\}$. Then (s,r) is a priori more plausible than (s',r') because in (s,r) Sally always says the truth and Robin always believes the literal meaning of Sally's message.

However, rational players cannot always rely on the honesty/credulity of the other player. Consider the following scenario. Rasmus also invites Sally for dinner but he cannot stand her. He wants to annoy her by preparing the curry the way she does not like. So while Sally will still prefer to receive her favoured type of curry, Rasmus will be happy only if he manages to prepare her disfavoured type.

	a_{mild}	a_{hot}
w_1	$1,-1$	$-1,1$
w_2	$-1,1$	$1,-1$

(5)

Here the interests of Sally and Rasmus are strictly opposed; everybody can only win as much as the other one looses. Again we assume that Sally can send two signals "*mild*" and "*hot*" with the conventional meaning as above. If Rasmus is credulous, he will react to "*mild*" with a_{hot} and to "*hot*" with a_{mild}. But if Sally believes this and is rational, she will be *dishonest* and send "*hot*" in w_1 (where she actually likes mild curries) and "*mild*" in w_2 (where she actually likes hot curries). But Rasmus might anticipate this. If he is not quite so credulous, he may switch his strategy accordingly, and react to "*mild*" with a_{mild} etc. This again might be anticipated by Sally and she might revert to the lying strategy, which again might be anticipated by Rasmus, etc. In fact, it turns out that *any* strategy is rationalizable in this game.[1] In other words, no real communication ensues. The lesson here is that communication might help in situations where the interests of the players are aligned, but it does not make a difference if these interests are completely opposed.

Gricean Reasoning

The kind of reasoning that was informally employed in the last section is reminiscent to pragmatic reasoning in the tradition of Grice (1975). For instance, information can only be exchanged between rational agents if it is in the good interest of both agents that this information transfer takes place. This intuition is captured by Grice's *Cooperative Principle*. Furthermore, we mentioned the default assumption that messages are used

[1] A strategy s is *rationalizable* if there is a consistent set of beliefs such that s maximizes the expected payoff of the player, given these beliefs and the assumption that rationality of all players is common knowledge.

according to their conventional meaning, unless overarching rationality considerations dictate otherwise. This corresponds to Grice's Maxim of Quality.

To illustrate how game theoretic reasoning can account for pragmatic reasoning let us consider the prime example of a scalar implicature, namely the strengthening of the conventional meaning of "*some*" to "*some but not all*". You can imagine that Robin wants to know who was at the party last night, and Sally knows the answer. In $w_\forall$ all girls were at the party and in $w_{\exists\neg\forall}$ some but not all girls were there. Again, Robin is completely unsure, i.e. he considers each world to be equally likely. Considering Robin's actions let us assume that there are three of them: two actions $a_\forall$ and $a_{\exists\neg\forall}$ that are appropriate in and only in worlds $w_\forall$ and $w_{\exists\neg\forall}$, respectively, and a kind of default action $a_?$. For each world, both Sally and Robin prefer Robin to perform the appropriate action to Robin performing the default action, which they in turn prefer to Robin performing the inappropriate action. The following payoff structure reflects this preference order.

	$a_\forall$	$a_{\exists\neg\forall}$	$a_?$
$w_\forall$	$10,10$	$0,0$	$9,9$
$w_{\exists\neg\forall}$	$0,0$	$10,10$	$9,9$

(6)

Furthermore we have three different messages with their corresponding conventional meaning.

$$
\begin{array}{rcll}
f_\forall & = & \text{“\textit{All girls were at the party.}”} & [\![f_\forall]\!] = \{w_\forall\} \\
f_{\exists\neg\forall} & = & \text{“\textit{Some but not all girls were at the party.}”} & [\![f_{\exists\neg\forall}]\!] = \{w_{\exists\neg\forall}\} \\
f_\exists & = & \text{“\textit{Some girls were at the party.}”} & [\![f_\exists]\!] = \{w_\forall, w_{\exists\neg\forall}\}
\end{array}
$$

Obviously $f_{\exists\neg\forall}$ is more complex than the other two messages, which are approximately equally complex. This is covered by the assignment of *costs* to signals which the sender has to pay. Formally this is implemented by a cost function c that assigns some numerical value to every signal. Let us say that in this example we have $c(f_\forall) = c(f_\exists) = 0$ and $c(f_{\exists\neg\forall}) = 2$. So the sender's utility is now a three-place function u_S that depends on the actual world, the message sent, and the action that the receiver takes. If $v_S(w,a)$ is the distribution of sender payoffs that is given in the payoff table (6) above, the sender's overall utility is now

$$u_S(w,f,a) \ = \ v_S(w,a) - c(f) \tag{7}$$

According to Gricean pragmatics, Sally would reason about her strategy roughly as follows:

If I am in $w_\forall$ I want Robin to perform $a_\forall$ because this gives me a utility of 10. $a_\forall$ is what he would do if he believed that he is in $w_\forall$. I can try to convince him of this fact by saying $f_\forall$. It is not advisable to say $f_{\exists\neg\forall}$, because if Robin believed it, he would perform $a_{\exists\neg\forall}$, which gives me a utility of a mere -2. Also saying $f_\exists$ is not optimal. If Robin believes it, this will not settle the issue which world we are in for him and thus he will perform $a_?$, because his expected utility in this case is 9 while his expected utility for the other two actions is only 5. This would give me also a utility of 9. So it seems reasonable to send $f_\forall$ in $w_\forall$.

If I am in $w_{\exists\neg\forall}$, it might seem reasonable to say $f_{\exists\neg\forall}$ because if Robin believes it, he will perform $a_{\exists\neg\forall}$, which is my favorite outcome. However, I will have to pay the costs of 2, so my net utility is only 8. If I say $f_\exists$ and Robin believes it, he will perform $a_?$. Since $f_\exists$ is costless for me, my net utility is 9 in this case, which is better than 8. So in $w_{\exists\neg\forall}$ I will send $f_\exists$. After this reasoning, Sally will hence settle on the following strategy:

$$\begin{bmatrix} w_\forall & \rightarrow & f_\forall \text{ (``\textit{All girls were at the party.}'')} \\ w_{\exists\neg\forall} & \rightarrow & f_\exists \text{ (``\textit{Some girls were at the party.}'')} \end{bmatrix} \tag{8}$$

Robin in turn will anticipate that Sally will reason this way: If I am confronted with the message $f_\forall$, I know that the world is $w_\forall$, hence I will perform $a_\forall$. If I hear $f_\exists$, I know that the world is $w_{\exists\neg\forall}$, hence I will perform $a_{\exists\neg\forall}$ after all. Therefore his strategy will look as follows:

$$\begin{bmatrix} f_\forall \text{ (``\textit{All girls were at the party.}'')} & \rightarrow & a_\forall \\ f_\exists \text{ (``\textit{Some girls were at the party.}'')} & \rightarrow & a_{\exists\neg\forall} \end{bmatrix} \tag{9}$$

Sally, being aware of this fact, will reason: Taking into consideration Robin's reasoning and his eventual strategy, it is even more beneficial for me to send $f_\exists$ if I am in $w_\exists$ because this will give me the maximal payoff of 10. So I have no reason to change the plan of sending $f_\forall$ in $w_\forall$ and $f_\exists$ in $w_{\exists\neg\forall}$.

Hence she will stick to her strategy in (8). In a further round of deliberation Robin will realize this and thus also stick to his strategy (9). Any further deliberation of Sally and Robin will not change anything.

This iterated reasoning procedure explains the emergence of the scalar implicature. It leads to a sender strategy where $f_\exists$ is sent if and only if $\{w_{\exists\neg\forall}\}$ is true. In other words, the literal meaning of $f_\exists$, which is $\{w_\forall, w_{\exists\neg\forall}\}$, has been pragmatically strengthened to a proper subset $\{w_{\exists\neg\forall}\}$. The information that $w_\forall$ is not the case is a scalar implicature — "*some*" is pragmatically interpreted as "*some but not all*".

As in the examples discussed in the previous section, the inferences that are used here start with a default assumption that messages are used according to their literal interpretation, but this is only a provisional assumption that is adopted if this is not in contradiction with rationality.

2 Iterated Best Response

The reasoning pattern that is used here makes implicit use of the notion of the *best response* of a player to a certain probabilistic belief. A best response (that need not be unique) to such a belief is a strategy that maximizes the expected payoff of the player as compared to all other strategies at his disposal, given this belief state. For a player to be *rational* means then to always play some best response, given his belief.

Let us now assume the position of an external observer who wants to formally model the notion of a best response, say Sally's best response to her belief about Robin's strategies. If we denote the set of strategies available to Robin at some point with R, we know that 1. Sally believes that Robin will play some strategy from R, 2. Sally holds all strategies in R possible, i.e. she cannot exclude any strategy for sure. Despite that we do not have any further information about Sally's belief – maybe she holds all strategies in R equally possible, or maybe she considers some strategies more likely than others. The best we can do as external observer is to take all possible beliefs for Sally into account.

Formally we can do this by modeling a belief of Sally as a probability distribution over the set of strategies R such that it does not assign zero probability to any element of R (i.e. Sally cannot exclude any strategy for sure). Let us therefore define the following sets of probability distributions over X for a non-empty and finite set X:

$$\Delta(X) \doteq \left\{ p \in X \to [0,1] \,\middle|\, \sum_{x \in X} p(x) = 1 \right\} \tag{10}$$

$$\Delta^{+}(X) \doteq \left\{ p \in X \to (0,1] \,\middle|\, \sum_{x \in X} p(x) = 1 \right\} \tag{11}$$

The difference is subtle but important. Both $\Delta(\cdot)$ and $\Delta^{+}(\cdot)$ can be used to model probabilistic beliefs. If we say that a player holds a belief from $\Delta(X)$, say, this means that he may exclude some elements from X with absolute certainty. On the other hand, if he holds a belief from $\Delta^{+}(X)$, then he may have certain guesses, but he is not able to exclude any element from X with certainty. In the case discussed above, Sally's believe about Robin's strategies R is modeled some $\rho \in \Delta^{+}(R)$. Hence any best response of Sally's to any such belief is a *potential best response* for Sally against R. All that we as an external observer can predict with certainty if we assume Sally to be rational, is that she will play some potential best response against R.

The iterative inference process that was used in the computation of the implicature above can be informally described as follows. At start, Sally provisionally assumes that Robin is entirely credulous, and that he conditions his actions only on the literal interpretation of the message received. Let us call the set[2] of credulous strategies R_0.

Sally's turn. Sally might ponder any strategy that is a potential best response against R_0. Let us call this set of strategies S_0.

Robin's turn. Robin might ponder all strategies that are potential best responses against S_0. The set of these strategies is R_1.

Sally's turn. Sally might ponder any strategy that is a potential best response against R_1. Let us call this set of strategies S_1.

Robin's turn ...

In general, S_n and R_{n+1} are the set of strategies that are potential best responses against R_n and S_n, respectively. If a certain strategy cannot be excluded by this kind of reasoning, i.e. if there are infinitely many indices i such that it occurs in S_i or R_i, then we call it a *pragmatically rationalizable strategy*.

Contexts

In the scalar implicature example, the described reasoning of Sally and Robin went in circles at some point. Therefore, all strategies they considered possible at this point were pragmatically rationalizable. These were exactly the strategies in (8) and (9), which described the scalar implicature.

[2] There might be more than one credulous strategy because several actions may yield the same maximal payoff for Robin in certain situations.

Taking another close look at the payoff structure in (6), we see that the scalar implicature arises because the difference between $v_S(w_{\exists\neg\forall}, a_{\exists\neg\forall})$ and $v_S(w_{\exists\neg\forall}, a_?)$ is smaller than the costs of sending $f_{\exists\neg\forall}$. Suppose the utilities would be as in (12), rather than as in (6). Then the pragmatically rationalizable outcome would be that Sally uses $f_{\exists\neg\forall}$ in $w_{\exists\neg\forall}$, while $f_\exists$ would never be used.

	$a_\forall$	$a_{\exists\neg\forall}$	$a_?$
$w_\forall$	$10,10$	$0,0$	$6,6$
$w_{\exists\neg\forall}$	$0,0$	$10,10$	$6,6$

(12)

At this point, we introduce another level of uncertainty concerning the payoff structure (in addition to the uncertainty of the player about the actual strategy of the other player). Robin might actually not know for sure what Sally's precise preferences are. If we call the utility matrix (6) *context* c_1, and the utilities in (12) context c_2, Robin might hold some probabilistic belief about whether Sally is in c_1 or in c_2. Likewise, Sally need not know for sure which context Robin is in. Now in each round of the iterative reasoning process, the players will ponder each strategy that is a potential best response not only to any probability distribution over strategies of the previous round as before, but also to any probability distribution over contexts. Sally will compute her first set of best responses S_0 by assuming a credulous Robin as follows: In $w_\forall$ I will definitely send $f_\forall$, no matter which context is the actual one. Now for $w_{\exists\neg\forall}$: If the actual context is c_1 it is better to send $f_\exists$ because the costs of sending the more explicit message $f_{\exists\neg\forall}$ exceed the potential benefits. But if it is c_2 and it is advisable to use $f_{\exists\neg\forall}$ nevertheless.

Robin, in turn, will reason as follows to compute his best responses R_1: If I hear $f_\forall$, we are definitely in $w_\forall$, and the best thing I can do is to perform $a_\forall$, no matter which context we are in. If I hear $f_{\exists\neg\forall}$ we are in $c_2/w_{\exists\neg\forall}$ and I will perform $a_{\exists\neg\forall}$. If I hear $f_\exists$ we are in $c_1/w_{\exists\neg\forall}$ and I will also play $a_{\exists\neg\forall}$.

So in S_1 Sally will infer: $f_\forall$ will induce $a_\forall$, and both $f_{\exists\neg\forall}$ and $f_\exists$ will induce $a_{\exists\neg\forall}$, no matter which context Robin is in. Since $f_\exists$ is less costly than $f_{\exists\neg\forall}$, I will hence always use $f_\forall$ in $w_\forall$ and $f_\exists$ in $w_{\exists\neg\forall}$, regardless of the context I am in.

Robin, in R_1, will thus conclude that his best response to $f_\forall$ is always $a_\forall$, and his best response to $f_\exists$ is $a_{\exists\neg\forall}$. Nothing will change in later iterations. So here, the scalar implicature from "*some*" to "*some but not all*" will arise in all contexts, even though context c_2 by itself would not license it.

The Formal Model

In this section we will present a formal model that captures the intuitive reasoning from the last section. A *semantic game* is a game between two players, the *sender* and the *receiver*. It is characterized by a finite set of contexts C, a finite set of worlds W, a finite set of signals (or forms) F, a finite set of actions A,

- a probability distribution $p^* \in \Delta^+(W)$ specifying the receiver's *a priori* probability for each world,
- an interpretation function $[\![\cdot]\!] : F \to \mathrm{Pow}(W)$,

- and utility functions

  $u_S\colon C \times W \times F \times A \to \mathbb{R}$ for the sender and

  $u_R\colon C \times W \times A \to \mathbb{R}$ for the receiver.

As in (7), we will give the sender's utility function by separating the context/outcome utilities v_S from the signalling costs $c\colon F \to \mathbb{R}$ in the following. The structure of the game is common knowledge between the players.

Definition 1. The space of *pure sender strategies* $\mathcal{S} = C \times W \to F$ is the set of functions from context/world pairs to signals. The space of *pure receiver strategies* $\mathcal{R} = C \times F \to A$ is the set of functions from context/signals pairs to actions. A *sender belief* is a pair of probability distributions $(\rho, p) \in \Delta(\mathcal{R}) \times \Delta(C)$ and a *receiver belief* is a pair of probability distributions $(\sigma, p) \in \Delta(\mathcal{S}) \times \Delta(C)$.

The central step in the iterative process described above is the computation of the set of strategies that maximize the expected payoff of a player against his belief about the strategies of the other player and the context. The notion of a *best response* captures this.

Definition 2. Let (σ, p) be a receiver belief and (ρ, p) a sender belief. The set $\mathrm{BR}_R(\sigma, p)$ of *best responses of the recevier to* (σ, p) and the set $\mathrm{BR}_S(\rho, p)$ of *best responses of the sender to* (ρ, p) are defined as follows:

$$\mathrm{BR}_R(\sigma, p) \doteq \left\{ r \in \mathcal{R} \;\middle|\; \forall c \in C: \; r \in \operatorname*{argmax}_{r \in \mathcal{R}} \sum_{s \in \mathcal{S}} \sigma(s) \sum_{c' \in C} p(c') \sum_{w \in W} p^*(w)\, u_R(c, w, r(c, s(c', w))) \right\}$$

$$\mathrm{BR}_S(\rho, p) \doteq \left\{ s \in \mathcal{S} \;\middle|\; \forall c \in C\, \forall w \in W: \; s \in \operatorname*{argmax}_{s \in \mathcal{S}} \sum_{r \in \mathcal{R}} \rho(r) \sum_{c' \in C} p(c')\, u_S(c, w, s(c, w), r(c', s(c, w))) \right\}$$

Based on this definition of best responses to a certain belief we may define the set of *potential best responses* against some set P of strategies of the opposing player as the set of strategies that are best responses to some belief state that assigns positive probability exactly to the elements of P.

Definition 3. Let $S \subseteq \mathcal{S}$ and $R \subseteq \mathcal{R}$ be a set of sender and receiver strategies, respectively. The set $\mathrm{PBR}(S)$ of *potential best responses of the receiver to S* and the set $\mathrm{PBR}(R)$ of *potential best responses (of the sender) to R* are defined as follows:

$$\mathrm{PBR}(S) \doteq \{\, r \in \mathrm{BR_R}(\sigma, p) \mid (\sigma, p) \text{ a receiver belief with } \sigma \in \Delta^+(S) \,\}$$

$$\mathrm{PBR}(R) \doteq \{\, s \in \mathrm{BR_S}(\rho, p) \mid (\rho, p) \text{ a sender belief with } \rho \in \Delta^+(R) \,\}$$

Suppose we know that Sally, being the sender, knows which context and world she is in, she believes for sure that Robin will play a strategy from R, and there is no more specific information that she believes to know for sure. We do not know which strategy from R Sally expects Robin to play with which likelihood, and which context Sally believes to be in. Under these conditions, all we can predict for sure is that Sally will play some strategy from $\mathrm{PBR}(R)$ if she is rational.

The same seems to hold if we only know that Robin, the receiver, expects Sally to play some strategy from S. Then we can infer that Robin, if he is rational, will certainly play a strategy from $\mathrm{PBR}(S)$. However, we may restrict his space of reasonable strategies even further. Suppose none of the strategies in S ever make use of the signal f (formally put, $f \in F - \bigcup_{s \in S} \mathrm{range}(s)$). We call such a signal *unexpected*. Then it does not make a difference how Robin would react to f, but he has to decide about how to react to f nevertheless because receiver strategies are *total* functions from context/form pairs to actions. It seems reasonable to demand (and it leads to reasonable predictions, as we will see below) that Robin should, in the absence of evidence to the contrary, still assume that f is true. For instance, if Sally speaks English to Robin, and she suddenly throws in a sentence in Latin that Robin happens to understand, Robin will probably assume that the Latin sentence is true, even if he did not expect her to use Latin.

If Robin encounters such an unexpected signal, he will have to revise his beliefs. Robin will have to figure out an explanation why Sally used f despite his expectations to the contrary, and this explanation can bias his prior beliefs in any conceivable way. We have to assume though that the result of this believe revision is a consistent belief state, and that Robin will act rationally according to his new beliefs. Formally speaking, he should only consider strategies that react to an unexpected signal f in a way that maximizes his expected utility, given that f is interpreted literally for some belief about W.

We can now proceed to define the iterative reasoning procedure that was informally described in the previous section, taking into account the treatment of unexpected signals detailed above (recall that p^* is the receiver's *a priori* probability distribution).

Definition 4.

$$
\begin{aligned}
R_0 &\doteq \left\{ r \in \mathcal{R} \;\middle|\; \forall c \in C \forall f \in F : r(c,f) \in \operatorname*{argmax}_{a \in A} \sum_{w \in [\![f]\!]} p^*(w) u_R(c,w,a) \right\} \\
S_n &\doteq \mathrm{PBR}(R_n) \\
R_{n+1} &\doteq \left\{ r \in \mathrm{PBR}(S_n) \;\middle|\; \forall f \in F - \bigcup_{s \in S_n} \mathrm{range}(s) \forall c \in C \right. \\
&\qquad \left. \exists p \in \Delta^+(W) : r(c,f) \in \operatorname*{argmax}_{a \in A} \sum_{w \in [\![f]\!]} p(w) u_R(c,w,a) \right\}
\end{aligned}
$$

R_0 is the set of credulous strategies of the receiver. It consists of those strategies r that yield, in each context c and for each signal f, some action $a \in A$ that is optimal for the receiver (i.e. that maximize his expected utility, cf. (2)), given that his *a priori* belief p^* is updated with the information that f is used literally. S_n is the set of potential best responses of the sender against R_n. Likewise, R_{n+1} is the set of potential best responses of the receiver if he assumes that the sender plays a strategy from S_n in which he always tries to make sense of unexpected messages under the assumption that they are literally true.

The sets of *pragmatically rationalizable strategies* are the set of sender strategies and receiver strategies that cannot be excluded for sure by the iterative reasoning process, no matter how deeply the reasoning goes.

Definition 5. $(\mathbf{S},\mathbf{R}) \in \mathrm{Pow}(\mathcal{S}) \times \mathrm{Pow}(\mathcal{R})$, the sets of *pragmatically rationalizable strategies*, are defined as follows:

$$\begin{aligned}\mathbf{S} &\doteq \{s \in \mathcal{S} \mid \forall n \in \mathbb{N} \exists m > n : s \in S_m\}\\ \mathbf{R} &\doteq \{r \in \mathcal{R} \mid \forall n \in \mathbb{N} \exists m > n : s \in R_m\}\end{aligned}$$

Note that there are only finitely many strategies in $\mathcal{S}$ and $\mathcal{R}$ (because we are only considering pure strategies over finite sets). Therefore there are only finitely many subsets thereof. The step from (S_n, R_n) to (S_{n+1}, R_{n+1}) is always deterministic. It follows that the iterative procedure will enter a cycle at some point. This ensures that $(\mathbf{S},\mathbf{R})$ is always defined.

3 Applying the IBR Model

In light of this formal definition, let us consider some of the previous examples again. For the ease of exposition we will specify signals as $f_{x_1 \dots x_n}$ with the convention that $[\![f_{x_1 \dots x_n}]\!] = \{w_{x_1}, \dots, w_{x_n}\}$. Furthermore, if the utilities of the players are identical for each outcome, we will show only one number in the utility matrix. If the *a priori* probability p^* is not explicitly stated, we assume that it is the uniform distribution on W that assigns all worlds equal probability.

Completely aligned interests. We assume that all signals f are costless, i.e. $c(f) = 0$. There is only one context and v_S and u_R are given in table (1). Here is the sequence of iterated computation of potential best responses, starting with the set R_0 of credulous strategies.

$$\begin{aligned}\mathbf{R} = R_0 &= \left\{ \begin{bmatrix} f_1 & \to & a_{\mathrm{mild}} \\ f_2 & \to & a_{\mathrm{hot}} \\ f_{12} & \to & a_{\mathrm{mild}} \end{bmatrix}, \begin{bmatrix} f_1 & \to & a_{\mathrm{mild}} \\ f_2 & \to & a_{\mathrm{hot}} \\ f_{12} & \to & a_{\mathrm{hot}} \end{bmatrix} \right\}\\ \mathbf{S} = S_0 &= \left\{ \begin{bmatrix} w_1 & \to & f_1 \\ w_2 & \to & f_2 \end{bmatrix} \right\}\end{aligned}$$

In the following we will abbreviate the specifications of the strategy sets by dropping the set brackets and by conflating the strategies to one representation. The original set can be recovered by combination of all possible argument/value pairs.

Completely Opposing Interests. Again all messages are costless and there is only one context. The utilities are as in (5). Here the iterative procedure enters a never-ending cycle:

$$R_0 = \begin{bmatrix} f_1 & \to & a_{\text{hot}} \\ f_2 & \to & a_{\text{mild}} \\ f_{12} & \to & a_{\text{mild}}/a_{\text{hot}} \end{bmatrix} \qquad S_0 = \begin{bmatrix} w_1 & \to & f_2 \\ w_2 & \to & f_1 \end{bmatrix}$$

$$R_1 = \begin{bmatrix} f_1 & \to & a_{\text{mild}} \\ f_2 & \to & a_{\text{hot}} \\ f_{12} & \to & a_{\text{mild}}/a_{\text{hot}} \end{bmatrix} \qquad S_1 = \begin{bmatrix} w_1 & \to & f_1 \\ w_2 & \to & f_2 \end{bmatrix}$$

$$R_2 = R_0 \qquad S_2 = S_0$$

$$\mathbf{R} = \begin{bmatrix} f_1/f_2/f_{12} & \to & a_{\text{mild}}/a_{\text{hot}} \end{bmatrix} \qquad \mathbf{S} = \begin{bmatrix} w_1/w_2 & \to & f_1/f_2/f_{12} \end{bmatrix}$$

So if the interests of the players are completely opposed, *any* strategy is pragmatically rationalizable and no communication will ensue.

Scalar Implicatures and the Q-Heuristics. Next we will reconsider the example of the scalar implicature discussed above. There are two contexts c_1 and c_2 with utilities as in (6) and (12), respectively. The signals and their costs are also as above: $c(f_\forall) = c(f_\exists) = 0$ and $c(f_{\exists\neg\forall}) = 2$.

$$R_0 = \begin{bmatrix} (c_1, f_\forall)/(c_2, f_\forall) & \to & a_\forall \\ (c_1, f_{\exists\neg\forall})/(c_2, f_{\exists\neg\forall}) & \to & a_{\exists\neg\forall} \\ (c_1, f_\exists)/(c_2, f_\exists) & \to & a_? \end{bmatrix} \qquad S_0 = \begin{bmatrix} (c_1, w_\forall)/(c_2, w_\forall) & \to & f_\forall \\ (c_1, w_{\exists\neg\forall}) & \to & f_{\exists\neg\forall} \\ (c_2, w_{\exists\neg\forall}) & \to & f_\exists \end{bmatrix}$$

$$\mathbf{R} = R_1 = \begin{bmatrix} (c_1, f_\forall)/(c_2, f_\forall) & \to & a_\forall \\ (c_1, f_{\exists\neg\forall})/(c_2, f_{\exists\neg\forall})/(c_1, f_\exists)/(c_2, f_\exists) & \to & a_{\exists\neg\forall} \end{bmatrix}$$

$$\mathbf{S} = S_1 = \begin{bmatrix} (c_1, w_\forall)/(c_2, w_\forall) & \to & f_\forall \\ (c_1, w_{\exists\neg\forall})/(c_2, w_{\exists\neg\forall}) & \to & f_\exists \end{bmatrix}$$

The previous example illustrated how pragmatic rationalizability formalizes the intuition behind Levinson's (2000) *Q-heuristics* "What isn't said, isn't." This heuristics accounts, *inter alia* for scalar and clausal implicatures like the following:

(1) a. *Some boys came in.* $\leadsto$ *Not all boys came in.*

b. *Three boys came in.* $\leadsto$ *Exactly three boys came in.*

(2) a. *If John comes, I will leave.* $\leadsto$ *It is open whether John comes.*

b. *John tried to reach the summit.* $\leadsto$ *John did not reach the summit.*

The essential pattern here is as in the example above: There are two expressions A and B of comparable complexity such that the literal meaning of A entails the literal meaning of B. There is no simple expression for the concept "B but not A". In this scenario, a usage of "B" will implicate that A is false.

The I-Heuristics. Levinson assumes two further pragmatic principles that, together with the Q-heuristics, are supposed to replace Grice's maxims in the derivation of generalized conversational implicatures. The second heuristics, called *I-heuristics*, says: "What is simply described is stereotypically exemplified." It accounts for phenomena of pragmatic strengthening, as illustrated in the following examples:

(3) a. *John's book is good.* $\rightsquigarrow$ *The book that John is reading or that he has written is good.*

b. *a secretary* $\rightsquigarrow$ *a female secretary*

c. *road* $\rightsquigarrow$ *hard-surfaced road*

The notion of "stereotypically exemplification" is somewhat vague and difficult to translate into the language of game theory. We will assume that stereotypical propositions are those with a high prior probability and that simplicity of descriptions can be translated into low signaling costs. So the principle amounts to "Likely propositions are expressed by cheap forms".

Let us construct a schematic example of such a scenario. Suppose there are two possible worlds (which may also stand for objects, like a hard surfaced vs. soft-surfaced road) w_1 and w_2, such that w_1 is *a priori* much more likely than w_2, say $p^*(w_1) = 3/4$ and $p^*(w_2) = 1/4$. There are three possible actions for Robin: he may choose a_1 if he expects w_1 to be correct, a_2 if he expects w_2, and a_3 if he finds it too risky to choose.

There are again three signals, f_1, f_2 and f_{12}. This time the more general expression f_{12} (corresponding for instance to "*road*") is cheap, while the two specific expressions f_1 ("*hard-surfaced road*") and f_2 ("*soft-surfaced road*") are more expensive: $c(f_1) = c(f_2) = 5$, and $c(f_{12}) = 0$.

The interests of Sally and Robin are completely aligned, except for the signaling costs which only matter for Sally. There are three contexts (13). In c_1 and c_2, it is safest for Robin to choose a_3 if he decides on the basis of the prior probability. In c_3 it makes sense to choose a_1 if he only knows the prior probabilities because the payoff of a_3 is rather low (but still higher than making the wrong choice between a_1 and a_2). In c_1, but not in c_2 it would be rational for Sally to use a costly message if this is the only way to make Robin perform a_1 rather than a_3.

c_1:

	a_1	a_2	a_3
w_1	28	0	22
w_2	0	28	22

c_2:

	a_1	a_2	a_3
w_1	28	0	25
w_2	0	28	25

c_3:

	a_1	a_2	a_3
w_1	28	0	10
w_2	0	28	10

(13)

$$R_0 = \begin{bmatrix} (c_1,f_1)/(c_2,f_1)/(c_3,f_1)/(c_3,f_{12}) & \to & a_1 \\ (c_1,f_2)/(c_2,f_2)/(c_3,f_2) & \to & a_2 \\ (c_1,f_{12})/(c_2,f_{12}) & \to & a_3 \end{bmatrix}$$

$$\mathbf{S} = S_0 = \begin{bmatrix} (c_1,w_1)/(c_3,w_1) & \to & f_1/f_{12} \\ (c_1,w_2)/(c_3,w_2) & \to & f_2 \\ (c_2,w_1) & \to & f_{12} \\ (c_2,w_2) & \to & f_2/f_{12} \end{bmatrix}$$

$$\mathbf{R} = R_1 = \begin{bmatrix} (c_1,f_1)/(c_2,f_1)/(c_3,f_1)/(c_3,f_{12}) & \to & a_1 \\ (c_1,f_2)/(c_2,f_2)/(c_3,f_2) & \to & a_2 \\ (c_1,f_{12})/(c_2,f_{12}) & \to & a_1/a_3 \end{bmatrix}$$

Here both f_1 and f_2 retain their literal meaning under pragmatic rationalizability. The unspecific f_{12} also retains its literal meaning in c_2. In c_1 and c_3, though, its meaning is pragmatically strengthened to $\{w_1\}$. Another way of putting is to say that f_{12} is *pragmatically ambiguous* here. Even though it has an unambiguous semantic meaning, its pragmatic interpretation varies between contexts. It is noteworthy here that f_{12} can never be strengthened to mean $\{w_2\}$. Applying it to the example, this means that a simple non-specific expression like "*road*" can either retain its unspecific meaning, or it can be pragmatically strengthened to its stereotypical instantiation (like "*hard-surfaced road*" here). It can never be strengthened to a non-stereotypical meaning though.

M-Heuristics. Levinson's third heuristics is the *M-heuristics*: "What is said in an abnormal way isn't normal." It is also known, after Horn (1984), as *division of pragmatic labor*. A typical example is the following:

(4) a. *John stopped the car.*

b. *John made the car stop.*

The two sentences are arguably semantically synonymous. Nevertheless they carry different pragmatic meanings if uttered in a neutral context. (4a) is preferably interpreted as *John stopped the car in a regular way, like using the foot brake.* This would be another example for the I-heuristics. (4b), however, is also pragmatically strengthened. It means something like *John stopped the car in an abnormal way, like driving it against a wall, making a sharp u-turn, driving up a steep road, etc.*

This can be modeled quite straightforwardly. Suppose there are again two worlds, w_1 and w_2, such that w_1 is likely and w_2 is unlikely (like using the foot brake versus driving against a wall). Let us say that $p^*(w_1) = 3/4$ and $p^*(w_2) = 1/4$ again. There are two actions, a_1 and a_2, which are best responses in w_1 and w_2 respectively. There is only one context. The utilities are given as follows:

	a_1	a_2
w_1	5	0
w_2	0	5

(14)

Unlike in the previous example, we assume that there are only two expressions, f and f', which are both unspecific: $[\![f]\!] = [\![f']\!] = \{w_1, w_2\}$. f' is slightly more expensive than f, say $c(f) = 0$ and $c(f') = 1$.

$$\begin{array}{rcl} R_0 &=& \left[\begin{array}{ccc} f/f' & \rightarrow & a_1 \end{array}\right] \\ R_1 &=& \left[\begin{array}{ccc} f & \rightarrow & a_1 \\ f' & \rightarrow & a_1/a_2 \end{array}\right] \\ \mathbf{R} \;=\; R_2 &=& \left[\begin{array}{ccc} f & \rightarrow & a_1 \\ f' & \rightarrow & a_2 \end{array}\right] \end{array} \qquad \begin{array}{rcl} S_0 &=& \left[\begin{array}{ccc} w_1/w_2 & \rightarrow & f \end{array}\right] \\ S_1 &=& \left[\begin{array}{ccc} w_1 & \rightarrow & f \\ w_2 & \rightarrow & f/f' \end{array}\right] \\ \mathbf{S} \;=\; S_2 &=& \left[\begin{array}{ccc} w_1 & \rightarrow & f \\ w_2 & \rightarrow & f' \end{array}\right] \end{array} \tag{15}$$

The crucial point here is that in S_0, the signal f' remains unused. Therefore any rationalizable interpretation of f' which is compatible with its literal meaning is licit in R_1, including the one where f' is associated with w_2 (which triggers the reaction a_2). Robin's reasoning at this stage can be paraphrased as: If Sally uses f, this could mean either w_1 or w_2. Since w_1 is *a priori* more likely, I will choose a_1. There is apparently no good reason for Sally to use f'. If she uses it nevertheless, she must have something in mind which I hadn't thought of. Perhaps she wants to convey that she is actually in w_2.

Sally in turn reasons: If I say f, Robin will take action a_1. If I use f', he may take either action. In w_1 I will thus use f. In w_2 I can play it safe and use f, but I can also take my chances and try f'.

Robin in turn will calculate in R_2: If I hear f, we are in w_1 with a confidence between 75% and 100%. In any event, I should use a_1. The only world where Sally would even consider using f' is w_2. So if I hear f' we are surely in w_2 and I can safely choose a_2. If Robin reasons this way, it is absolutely safe for Sally to use f' in w_2.

4 Conclusion

We proposed a game theoretic formalization of Gricean reasoning that both captures the intuitive reasoning patterns that are traditionally assumed in the computation of implicatures. The essential intuition behind the proposal is that the literal meaning of signals constitutes their default interpretation, and that rational communicators decide about their communicative strategies by iteratively calculating the best response to this default strategy.

Concerning related work, Franke (2008) proposes to calculate the pragmatically licit communication strategies by starting with a strategy based on the literal interpretation of signals and iteratively computing the best response strategy until a fixed point is reached. So this approach is very similar in spirit to the present one. The main differences are that Franke uses a particular honest sender strategy — rather than the set of all credulous receiver strategies — as the starting point of the iteration process, and that he uses deterministic best response calculation, rather than potential best responses, as update rule.

Acknowledgements

We would like to thank Michael Franke and Robert van Rooij for very illuminating discussions. This work has been supported by the German Science Foundation *DFG* as part of the SFB 673 *Alignment in Communication*, project A2 *Processing of Implicit Common Ground*.

References

Michael Franke. Meaning & inference in case of conflict. In Kata Balogh, editor, *Proceedings of the 13th ESSLLI Student Session*, 65–74. European Summer School in Logic, Language and Information, Hamburg, 2008.

Herbert Paul Grice. Logic and conversation. In P. Cole and J. Morgan, editors, *Syntax and Semantics* **3:** *Speech Acts*, 41–58. Academic Press, New York, 1975.

Laurence Horn. Towards a new taxonomy for pragmatic inference: Q-based and R-based implicatures. In Deborah Schiffrin, editor, *Meaning, Form, and Use in Context*, 11–42. Georgetown University Press, Washington, 1984.

Stephen C. Levinson. *Presumptive Meanings*. MIT Press, 2000.

Martin J. Osborne and Ariel Rubinstein. *A Course in Game Theory*. MIT Press, 1994.