Pragmatic Back-and-Forth Reasoning

Michael Franke & Gerhard Jäger

Abstract

We survey a number of game theoretic models that capture speakers' and listeners' pragmatic back-and-forth reasoning about mutual beliefs and linguistic behavior (i.e., utterance choice and interpretation). Two types of models are presented. Firstly, models that rely on rationality of choices and beliefs therein are shown to predict general pragmatic usage and inference patterns. Secondly, we introduce a new probabilistic variant of these reasoning models which parameterizes agents' rationality (and belief therein), thereby enabling fine-grained quantitative predictions of speaker and listener behavior to match data from psycholinguistic experiments.

Language use is often likened to a game that speakers and hearers play. This comparison is helpful for explaining pragmatic inferences and goal-oriented language use. Game theory offers rich tools for representing formal language games and ways of reasoning about them. Game theoretic approaches to pragmatics have been pioneered by Prashant Parikh (e.g. Parikh, 1991, 1992, 2001) but have since been accompanied by several different alternatives with a growing range of applications (see Benz, Jäger, and van Rooij, 2006; Franke, 2013a; Jäger, 2008, for overview). When it comes to tackling pragmatic reasoning along the lines envisaged by Grice (1975), models that spell out pragmatic back-and-forth reasoning are particularly relevant. Pragmatic back-and-forth reasoning is reasoning of speakers and hearers about what the respective other believes, does, believes his interlocutor does and so on. For instance, the general intuitive reasoning scheme behind a scalar inference is pragmatic back-and-forth reasoning of this kind: "I should not interpret 'some' to mean 'all' (although that would not be ruled out by semantic meaning), because, if the speaker had wanted me to do so, he had better said 'all'."

Several concrete formalizations of such reasoning have been proposed (e.g. Benz and van Rooij, 2007; Franke, 2011; Jäger, 2013). We provide an overview of these, and furthermore introduce a novel probabilistic variant. The probabilistic variant is relevant, because, as we argue, one of the main strength of the game theoretic approach to pragmatic reasoning is that it gives us a direct handle to bridge theoretical explanations of pragmatic phenomena, on the one hand, and quantitative data from relevant psycholinguistic experiments, on the other.

Section 1 introduces signaling games as models of the context for pragmatic interpretation. Section 2 gives a general overview of how the models we look at subsequently formalize pragmatic back-and-forth reasoning. Section 3 discusses closely related types of pragmatic reasoning schemes that hinge on the assumption that interlocutors make rational choices. The focus of this section is to show how different design choices lead to different predictions. Section 4 lifts the assumption of (belief in) rationality, making way for a probabilistic variant of pragmatic back-and-forth reasoning. Finally, Section 5 compares the game theoretic approaches sketched here to related approaches, in particular *Gricean* and *Neo-Gricean* theories (e.g. Atlas and Levinson, 1981; Gazdar, 1979; Grice, 1975; Horn, 1984; Levinson, 2000), *bidirectional optimality theory* (Blutner, 1998, 2000), the *intentions first* approach of Geurts (2010), and the Bayesian *rational speech-act model* (Bergen, Levy, and Goodman, 2012; Frank and Goodman, 2012; Goodman and Stuhlmüller, 2013).

1 Signaling Games as Context-Models

Signaling games were invented by David Lewis to counter a regression argument against conventionalist theories of meaning (Lewis, 1969). Since then they have played a major role as a general model of information transmission between agents in ecoomnics and theoretical biology (e.g. Crawford and Sobel, 1982; Grafen, 1990; Nowak and Krakauer, 1999; Spence, 1973). In the present context, we are particularly interested in signaling games where agents use signals that already have a conventionally associated meaning. In this case, we think of these games as representations of the most important contextual features relevant for pragmatic reasoning.

A signaling game involves two players, a sender and a receiver, representing the speaker and the listener. A signaling game captures just one conversational move, namely an utterance of a speaker who has private information that the listener lacks, and the listener's subsequent reaction to this utterance. This reaction can be a concrete physical action or, when we are interested in pragmatic interpretation, an "epistemic action" such as adopting a belief.

A signaling game, as we conceive it here, consists of an ordered and finite set of states T of size n_T. States determine different ways the world could be, as relevant for the conversational exchange (see below). The sender knows the actual state and, conditional on the actual state, selects a message from a given (ordered and finite) set M of size n_M to send to the receiver. The receiver observes the sent message, but doesn't know the actual state. The receiver only has a probabilistic belief about which state is actual given by a vector $\mathbf{p} \in \Delta^+(n_T)$.[1] Each message has a conventionally specified meaning. Conventional meanings of messages are given by a Boolean (n_T, n_M)-matrix B where B_{ij} is the truth value of message j in state i. (We assume that B is a Boolean matrix, i.e., that truth values are classical, but that assumption can be given up easily to allow for pragmatic reasoning about, say, a fuzzy language.) The receiver then selects an action from the (ordered and finite) set A of size n_A. We will assume that players have preferences over outcomes of the game that depend on the actual state and the action chosen by the receiver. Let U_S and U_R be (n_T, n_A)-matrices of utilities that specify the preferences for the sender and receiver. Additionally, sometimes agents may care about differences in *message complexity* (often referred to as *message costs*), captured by a vector $\mathbf{c}$ of length n_M.

For example, a signaling game that represents a generic context of utterance in which Quantity reasoning leading to a scalar inference, such as in example (1), could arise is the following *some-all game.*

(1) a. Joe ate some of the cookies.

b. $\rightsquigarrow$ Joe didn't eat all of the cookies.

There are two states $T = \{t_{\exists\neg\forall}, t_{\forall}\}$. State $t_{\exists\neg\forall}$ is a state in which Joe ate some but not all of the cookies, while $t_{\forall}$ is a state in which he ate all of them. Fixing this set of states as relevant for the conversation works towards implementing the assumption that speaker and listener care about whether Joe ate some or all of the cookies. In other words, in conjunction with the utilities of players (see below), the set of states implements (a generalization of) a question under discussion (see Franke, 2009, 2011, for details on the interpretation of signaling games as pragmatic context models). As usual, we assume that the alternative utterances we compare are "Joe ate some of the cookies" and "Joe ate all of the cookies", which we abstractly represent as $M = \{m_{\text{some}}, m_{\text{all}}\}$. The semantic meaning of the

[1] For $n \in \mathbb{N}$, let $\Delta(n) = \{\langle p_1, \ldots, p_n \rangle \mid \sum p_i = 1 \text{ and } p_i \geq 0\}$ be the set of all probability vectors of size n and $\Delta^+(n) = \{\langle p_1, \ldots, p_n \rangle \mid \sum p_i = 1 \text{ and } p_i > 0\}$ the restriction to all strictly positive probability vectors.

alternatives, relative to the fixed states, is given by matrix:

$$B = \begin{array}{c} \\ t_{\exists\neg\forall} \\ t_{\forall} \end{array} \begin{array}{c} \begin{array}{cc} m_{\text{some}} & m_{\text{all}} \end{array} \\ \left(\begin{array}{cc} 1 & 0 \\ 1 & 1 \end{array} \right) \end{array} .$$

We further assume that the alternatives do not differ with respect to their processing and/or utterance costs, so that $\mathbf{c} = \langle 0, 0 \rangle$. For simplicity, we could assume that the receiver is entirely ignorant about which state is actual, so that his probabilistic beliefs are unbiased: $\mathbf{p} = \langle 1/2, 1/2 \rangle$. Finally, we would assume that speaker and listener cooperatively share the goal of communicating which state is actual. This is implemented by a number of features. Firstly, the set of receiver response actions is identified with the set of states $A = T$. Gricean cooperativity is captured by assuming that $\mathrm{U}_S = \mathrm{U}_R$, and that interlocutors care about communication of the relevant state distinctions is expressed by assumption that the utility matrix U (which is the same for sender and receiver) is the *diagonal matrix* with $\mathrm{U}_{ij} = 1$ if $i = j$ and 0 otherwise.

The some-all game is an example of an *interpretation game* that implements the standard Gricean assumptions of relevance of state distinctions and a cooperatively shared goal of successful communication about these. The example demonstrates that signaling games are a rich means of representing pragmatic contexts that can accommodate the usual Gricean assumptions about relevance and cooperativity. But signaling games are more expressive even and also capture scenarios of, say, language use in uncooperative scenarios (c.f. Franke, de Jager, and van Rooij, 2012; de Jaegher and van Rooij, 2013). Signaling games can also represent markedness differences between messages and between states, as shown in the next example.

The *Horn game* captures reasoning towards Horn's division of pragmatic labor (Horn, 1984). Usually, choosing a *simple* way of expressing a meaning as in (2a) is associated with a *stereotypical* interpretation as in (2b), whereas a *marked* though semantically equivalent expression (3a) is interpreted in a *non-stereotypical* way (3b).

(2) a. Black Bart killed the sheriff.

b. ⇝ Black Bart killed the sheriff in a stereotypical way.

(3) a. Black Bart caused the sheriff to die.

b. ⇝ Black Bart killed the sheriff in a non-stereotypical way.

To represent reasoning towards these pragmatic inferences, we assume that the Horn game has a set $T = \{t, t^*\}$ that distinguishes an unmarked state t and a marked state t^*. There are two messages, one unmarked and one marked: $M = \{m, m^*\}$. The Horn game is an interpretation game, so that $T = A$, $\mathrm{U}_S = \mathrm{U}_R = \mathrm{U}$ and $\mathrm{U}_{ij} = 1$ if $i = j$ and 0 otherwise. The semantic meaning of messages is trivial in this case: $B_{ij} = 1$ for all i, j. Pragmatic reasoning therefore cannot be fueled by Quantity, so to speak, but must rely on manner. A signaling game model of the context of utterance can capture this by assuming, for instance, differences in prior probabilities of states and costs of signals: $\mathbf{p} = \langle 1/2 + \epsilon, 1/2 - \epsilon \rangle$ and $\mathbf{c} = \langle 0, \delta \rangle$ for some small but positive ϵ and δ.

In order to derive more nuanced inferences about the speaker's epistemic state, signaling games can be *lifted*, so that the states of the game represent epistemic states which the speaker could be in (see Franke, 2009, 2011; Jäger, 2013). For reasons of space, we only consider the simple example of epistemic inferences possibly triggered an utterance of a sentence like (1a). If the hearer doesn't pay attention to the epistemic state of a speaker who utters (1a), he might immediately infer what we could call the *base-level implicature* in (1b). But if he does pay attention to the speaker's epistemic states, he might rather draw one of the epistemic scalar inferences in (4).

(1a) Joe ate some of the cookies.

(4) a. Strong epistemic implicature:
The speaker believes that Joe didn't eat all of the cookies.

b. Weak epistemic implicature:
The speaker is uncertain whether Joe ate all of the cookies.

c. General epistemic implicature:
The speaker doesn't believe that Joe ate all of the cookies.

Which epistemic inference the hearer draws depends on his assumptions about the speaker's *competence* (c.f. Sauerland, 2004; Schulz and van Rooij, 2006; Spector, 2006). If the hearer assumes that the speaker is likely competent with respect to the issue of whether Joe ate only some or all of the cookies, then he would best draw the strong inference in (4a). If instead he assumes the speaker to be uninformed about this issue, he would best draw the weak inference in (4b). Finally, if the hearer doesn't know whether the speaker is competent in the relevant respect, he would best only draw the general epistemic implicature in (4c), which subsumes the former two.

To model these different utterance contexts, we construct signaling games with three states $T = \{t_{[\exists\neg\forall]}, t_{[\forall]}, t_{[\exists\neg\forall,\forall]}\}$:

- $t_{[\exists\neg\forall]}$ is a state in which the speaker knows that Joe ate some but not all of the cookies;
- $t_{[\forall]}$ is a state in which the speaker knows that Joe ate all of the cookies;
- $t_{[\exists\neg\forall,\forall]}$ is a state in which the speaker doesn't know whether Joe ate only some or all of the cookies.

Different assumptions of the hearer about the likely competence of the speaker can now be represented as different a priori beliefs. So, if $\mathbf{p} = \langle p_1, p_2, p_3 \rangle$ are the prior probabilities of the above states, we get an a priori belief in speaker competence for $p_1, p_2 > p_3$, an a priori belief in speaker incompetence whenever $p_1, p_2 < p_3$ and uncertainty about it whenever $p_1 = p_2 = p_3$. Almost everything else in this signaling game remains as before. There are two messages m_{some} and m_{all}, which are equally costly. The receiver chooses an interpretation action $T = A$ and interlocutors cooperatively strive for perfect communication of the speaker's epistemic state. The only difference lies in the interpretation of the meaning matrix. Whereas before we interpreted $B_{ij} = 1$ as saying that message j is true in state i. Since we are now dealing with epistemic states, we will interpret this as saying that message j is believed to be true in state i. So, with this, we get:

$$B = \begin{array}{c} \\ t_{[\exists\neg\forall]} \\ t_{[\forall]} \\ t_{[\exists\neg\forall,\forall]} \end{array} \begin{array}{c} \begin{array}{cc} m_{\text{some}} & m_{\text{all}} \end{array} \\ \left(\begin{array}{cc} 1 & 0 \\ 1 & 1 \\ 1 & 0 \end{array}\right) \end{array}.$$

2 Basic Idea of Iterated X-Response Reasoning Schemes

Signaling games represent contexts. Pragmatic reasoning about a given context is formalized by suitable solution concepts. Nash equilibrium and its variations are by-far the most well-known game theoretic solution concepts. Early applications of game theory to formal pragmatics indeed relied on equilibrium notions (Parikh, 1991, 1992, 2001). But here we would like to take an explicitly *epistemic approach* to solving language games, in which we spell out interlocutors' back-and-forth reasoning about each others' possible beliefs and choices (see Franke, 2013a, for concise arguments in favor of this approach). More concretely, we discuss a number of variations of a family of pragmatic reasoning schemes, which we will call IxR-schemes (iterated X-response), where $x \in \{b, c, q\}$ is a variable for the variants "best", "cautious" and "quantal" that we discuss in detail below.

The easiest way of looking at models from the IxR-family is to think about variably sophisticated language users, more abstractly cast as *strategic types* of agents. Strategic types are hierarchically organized. Level-0 agents are unstrategic, in the sense that they do no take the full game situation into account. Their behavior is characterized mainly by the semantic meaning of signals. Intuitively, a level-0 speaker only cares about saying something true, while a level-0 listener interprets every message literally.

The behavior of higher level types is more involved, and it is here that most of the variability between different types of IxR-schemes shows. Generally speaking, a level-$k+1$ agent has some kind of belief about the behavior of his opponent and adapts his behavior in some way or other to that belief. The variability comes from different ways of deriving the beliefs of these agents, and the way that they react to it. Generally speaking, the *behavioral belief* of a level-$k+1$ agent, i.e., his belief about the behavior of his interlocutor, is derived in some fashion from the behavior of strategic types of level $l \leq k$. More concretely, each level-$k+1$ agent "looks down" the type hierarchy, so to speak, and conjectures how likely the opponent's type is k, $k-1$, ... From that, and the given behavior of k, $k-1$, ..., a level-$k+1$ agent obtains an expectation about his interlocutor's behavior. In the simplest case, behavioral beliefs are *myopic*. If so, a level-$k+1$ agent believes that his opponent is exactly of level k. Assuming myopia often makes definitions and computations easier, but it is also not necessarily an unrealistic assumption about resource-bounded human reasoning. Still, it is also possible to assume that each level-$k+1$ agent has a belief in the form of a non-trivial probability distribution over strategic levels. We will, for simplicity, stick to the simpler myopic versions in the following.[2]

Given a sequence of strategic types for the speaker and the listener, with their associated behavior, the last crucial thing is to specify what the model's overall behavioral prediction is. Again, there is room for application-specific design choices. We might be interested in the most sophisticated behavior included in the sequences, if such exists, or we might be interested in an average, somehow weighted, over the behavior of many strategic types. The former usually happens when we look at general explanations for general pragmatic facts; the latter is more relevant when we want to account for concrete, perhaps even numerical data from, say, psycholinguistic experiments. It is a benefit of the IxR-approaches, as compared to, e.g., equilibrium analyses, to be flexible enough to provide predictions for both ideally rational unlimited back-and-forth reasoning (which could be the outcome of learning when playing the relevant game repeatedly), as well as depth-limited reasoning in one-shot cases.

To illustrate these differences, the following sections will elaborate on some of the sketched possibilities of how to fill in concrete instances of IxR reasoning schemes. First, we will look at a model in which agents respond rationally to their behavioral beliefs (the variations IBR and ICR). Then we

[2] Similar myopic models are discussed as *level-k models* in the literature on behavioral economics (e.g. Crawford, 2003; Crawford and Iriberri, 2007). Like-minded models without myopicity are usually called *cognitive hierarchy models* (e.g. Camerer, 2003; Camerer, Ho, and Chong, 2004; Ho, Camerer, and Weigelt, 1998; Rogers, Palfrey, and Camerer, 2009). Our IxR-models are essentially specialized adaptations of these economic models to the case of natural language interpretation.

will look at a probabilistic variant (labeled IQR) that dispenses with the belief of full rationality and instead assumes that agents make choices with a probability proportional to how preferable a given option is relative to its alternatives. In the latter model, agents not only *are* boundedly rational in their decision making, but also believe that their interlocutors are as well.

3 Iterated Best & Cautious Response

Preliminaries. If A is an (m, n)-matrix and $\mathbf{p}$ a vector of length n, then we write $A \times \mathbf{p}$ and $A - \mathbf{p}$ to denote row-wise multiplication and subtraction. We also use a normalization operator $\text{Norm}(A)$ that maps matrix A onto another (m, n)-matrix such that $\text{Norm}(A)_i \propto A_i$ if $\sum_j (A_{ij}) > 0$ and $\text{Norm}(A)_{ij} = 1/n$ otherwise. We write $T(A)$ for the transpose of A.

Strategies. A sender strategy σ is a row-stochastic (n_T, n_M)-matrix, mapping each state onto a probability distribution over messages. A sender strategy describes how likely each message is chosen in each state. Likewise, since the receiver chooses states as interpretations in reaction to an observed message, a receiver strategy ρ is a row-stochastic (n_M, n_A)-matrix, mapping each message onto a probability distribution over actions. Here are two random examples for sender and receiver strategies for the some-all game introduced earlier:

$$\sigma = \begin{array}{ccc} & m_{\text{some}} & m_{\text{all}} \\ t_{\exists\neg\forall} & .4 & .6 \\ t_{\forall} & .9 & .1 \end{array} \qquad \rho = \begin{array}{ccc} & t_{\exists\neg\forall} & t_{\forall} \\ m_{\text{some}} & .8 & .2 \\ m_{\text{all}} & .3 & .7 \end{array}$$

Rows in these strategy matrices represent the situations in which agents need to make a choice, and each row then gives the respective choice probabilities. The ordering of rows and columns follows the specification of the game. (It's given in gray here for illustration, but will be left out henceforth.) The above sender strategy, for instance, expresses that the sender chooses m_{some} when in state $t_{\exists\neg\forall}$ with probability .4. Strategies with a 1 in each row are called *pure strategies*. Let S and R be the set of all pure sender and receiver strategies. Sender strategies can represent both the sender's behavior, as well as the receiver's beliefs about the sender's behavior. Likewise for receiver strategies.

Naïve Types. Naïve level-0 types need not represent models of actual behavior, but rather represent the literal meanings of the messages involved, coerced into the format of a (possibly mixed) sender strategy or receiver strategy respectively.

$$S_0 = \{\text{Norm}(B)\}$$
$$R_0 = \{\text{Norm}(T(B))\}\,.$$

When constructing a game, only such types are being considered where at least one message is true. Likewise, all messages considered are consistent, i.e., they are true of at least one type. Therefore neither B nor $T(B)$ has rows consisting only of 0 entries.

Sophisticated Types. All strategic types are defined as sets of strategies. From these we want to derive a set of possible beliefs of higher-order types about the interlocutor's possible behavior. There is room for interesting conceptual variation here, and we consider three obvious possibilities. Let X

be an ordered set of strategies, be it receiver or sender strategies. Let $|X| = d$. In the most unrestricted case, we could allow *any* possible belief as to which strategy the opponent plays:

$$\Pi(X) = \left\{ \sum_{x_i \in X} p_i x_i \mid \langle p_1, \ldots, p_d \rangle \in \Delta(d) \right\} .$$

This, however, is often rather too unconstrained and allows agents to endorse too many unjustified biases. If agents blend out the possibility of some opponent strategy entirely, they might miss crucial information. Intuitively speaking, a careful reasoner would better not rule out any strategy in X entirely. We therefore consider the following set of all *cautious beliefs* based on set X:

$$\Pi^c(X) = \left\{ \sum_{x_i \in X} p_i x_i \mid \langle p_1, \ldots, p_d \rangle \in \Delta^+(d) \right\} .$$

Cautious beliefs can still be rather biased, almost completely ruling out some strategies in favor of others. That's why yet another salient possible assumption about belief formation is to consider entirely *unbiased beliefs* that consider each option equally likely:

$$\Pi^u(X) = \left\{ \sum_{x \in X} \frac{1}{d} X \right\} .$$

Based on a set of beliefs Π we define the set of best responses to these as: $\mathrm{BR}(\Pi) = \bigcup \{\mathrm{BR}(\pi) \mid \pi \in \Pi\}$. The best response of the sender to a belief ρ about the receiver's behavior is defined as:

$$\mathrm{BR}_S(\rho) = \left\{ s \in \mathsf{S} \mid s_{ij} = 1 \Rightarrow j \in \arg_k \max(\mathrm{U}_S T(\rho) - \mathbf{c})_{ik}) \right\} .$$

In plain English, a best response to ρ is any pure strategy that maps each type to a message maximizing the utility to be expected under ρ.

The definition of the receiver's best response to a sender strategy σ is more subtle because there may be messages m_j with $\sigma_{ij} = 0$ for all i, i.e., messages that a sender playing σ would never use. As receiver strategies are total functions from messages to actions, the best response to σ has to specify how to react to such so-called *surprise messages*. A fairly simple solution to this problem is to fall back to the literal meaning of the message in question. In other words, a best response to a surprise message m_j is any action a_i with $B_{ij} = 1$. This leads to the following definition:

$$\begin{aligned} \mathrm{BR}_R(\sigma) \quad = \quad \{ r \in \mathsf{R} \mid \quad & (r_{ij} = 1 \wedge \max T(\sigma)_j > 0 \Rightarrow j \in \arg_k \max((T(\sigma) \times \mathbf{p}) \mathrm{U}_R)_{ik}) \wedge \\ & (r_{ij} = 1 \wedge \max T(\sigma)_j = 0 \Rightarrow B_{ji} = 1) \} . \end{aligned}$$

Higher-order types are then defined as either:

$$S_{k+1} = \mathrm{BR}(\Pi^u(R_k)) \qquad\qquad R_{k+1} = \mathrm{BR}(\Pi^u(S_k)) \tag{1}$$

in which case we obtain an IBR-model using unbiased beliefs (e.g. Franke, 2009, 2011), or as:

$$S_{k+1} = \mathrm{BR}(\Pi^c(R_k)) \qquad\qquad R_{k+1} = \mathrm{BR}(\Pi^c(S_k)) \tag{2}$$

in which case we obtain an ICR-model based on cautious beliefs (e.g. Jäger, 2013).[3]

[3] Note that the present version of the ICR-model differs from the one in (e.g. Jäger, 2013) insofar as only a single, possibly mixed, strategy is assumed here for naïve types, while (e.g. Jäger, 2013) uses a, possibly non-singleton, set of pure strategies.

Overall Predictions. So far, the above definitions give us two sequences of beliefs and behavior for speakers and listeners of varying theory-of-mind capacities. It remains to be fixed what then actually the overall prediction of such models is. Again, there is room for different implementations. When we are looking for an explication of idealized reasoning about language use, we would be most interested in the most sophisticated behavior present in the sequence. Indeed, if the underlying signaling game is finite (i.e., there are finitely many states, messages and actions), there will only be finitely many sets of pure strategies, so that at some point or other the inductive definition of strategic types (as sets of pure strategies) will start to loop. Since everything is deterministic, these loops will be repeated infinitely. Therefore, any strategic type that occurs in a loop is compatible with also an unbounded theory-of-mind capacity (even if we enter the loop already after 1 round, for instance). Consequently, the *limit prediction* of IBR- and ICR-type models is:

$$S_\omega = \left\{ s \in \mathsf{S} \mid \forall i \, \exists j > i \, : \, s \in \mathsf{S}_j \right\} \qquad\qquad R_\omega = \left\{ r \in \mathsf{R} \mid \forall i \, \exists j > i \, : \, r \in \mathsf{R}_j \right\} .$$

In many cases, the limit prediction of IBR-type models is an equilibrium of the signaling game (see Franke, 2011; Jäger, 2013, for details).

On the other hand, we might not be interested in the predictions of pragmatic back-and-forth reasoning under common belief in rationality. For instance, we might believe that, e.g., the subjects of an experiment on pragmatic reasoning, perform strategic reasoning as described by the IBR/ICR-scheme, but do not reason themselves all the way to a fixed point, so to speak (c.f. Degen and Franke, 2012; Degen, Franke, and Jäger, 2013). In that case, we could formulate a belief about how likely we think each strategic type is and derive an overall prediction from that. If $f \in \Delta(\mathbb{N})$ is a probability distribution over strategic types $0, 1, 2, \dots$, the overall prediction of IBR-style models is the f-weighted average over (unbiased) averages of strategies at each level k:

$$\sigma = \sum_k f(k) \, \text{Norm}\left(\sum_{s \in \mathsf{S}_k} s \right) \qquad\qquad \rho = \sum_k f(k) \, \text{Norm}\left(\sum_{r \in \mathsf{R}_k} r \right) .$$

Related models from behavioral game theory often use a Poisson distribution for this purpose, mostly for practical reasons (e.g. Camerer, 2003; Camerer, Ho, and Chong, 2004). The Poisson distribution is a discrete probability distribution with a single parameter τ that regulates its shape. The probability $f_\tau(k)$ of strategic type k under a Poisson distribution is:

$$f_\tau(k) = \text{Pois}_\tau(k) = \frac{\tau^k \, e^{-\tau}}{k!}$$

Examples of Poisson distributions for various τ are given in Figure 1.

Examples. As a first example, let's compute the limit predictions of an unbiased IBR-model for the some-all game introduced earlier. Remember that we had $T = \{t_{\exists\neg\forall}, t_\forall\}$, $M = \{m_{\text{some}}, m_{\text{all}}\}$, $\mathbf{p} = \langle 1/2, 1/2 \rangle$, $\mathbf{c} = \langle 0, 0 \rangle$ and a matrix B that captured the logical semantics as usual. The behavior of naïve types is:

$$\mathsf{S}_0 = \left\{ \begin{pmatrix} 1 & 0 \\ .5 & .5 \end{pmatrix} \right\} \qquad\qquad \mathsf{R}_0 = \left\{ \begin{pmatrix} .5 & .5 \\ 0 & 1 \end{pmatrix} \right\} .$$

In words, when a truthful speaker talks about state $t_{\exists\neg\forall}$ there is only one option, namely to use m_{some}. But in state $t_\forall$, he could use either m_{some} or m_{all} as both are true. Similarly, when the receiver hears message m_{all} there is only one possible interpretation, namely that the actual state is

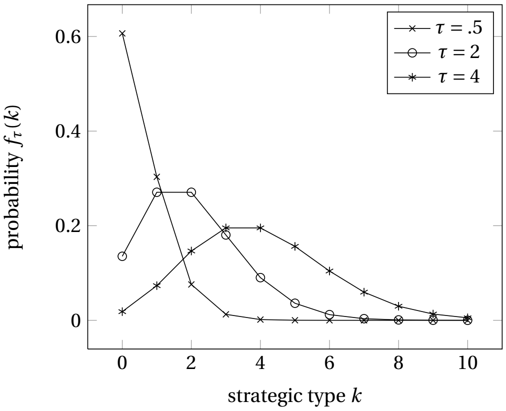

Figure 1: Examples of Poisson-distributed strategic types. For instance, $\tau = .5$ yields an expectation of ca. 60% level-0 players in the population of experimental subjects, ca. 30% level-1 players etc. The higher the value for τ, the deeper the expected strategic reasoning depth.

$t_\forall$. But when hearing m_{some}, the receiver's beliefs, obtained from updating his prior beliefs with the semantic meaning of message m_{some}, are maximally undecided between states.

The unbiased beliefs that level-1 sender and receiver have are, respectively:

$$\Pi^u(\mathsf{R}_0) = \left\{\begin{pmatrix}1 & 0\\ .5 & .5\end{pmatrix}\right\} \qquad \Pi^u(\mathsf{S}_0) = \left\{\begin{pmatrix}.5 & .5\\ 0 & 1\end{pmatrix}\right\}$$

For instance, the sender believes that when he utters m_{some} the receiver will adopt either interpretation with equal probability. The set of best responses to these beliefs are singleton and constitute a fixed point of the reasoning sequence already:

$$\mathsf{S}_1 = \left\{\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\right\} \qquad \mathsf{R}_1 = \left\{\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\right\}. \tag{3}$$

According to these strategies, the sender uses m_{some} only in state $t_{\exists\neg\forall}$ and the receiver interprets m_{some} to mean $t_{\exists\neg\forall}$.

As we are only dealing with belief sets that are singletons in this example, using cautious beliefs instead of unbiased beliefs yields the same prediction.

$$\Pi^c(\mathsf{R}_0) = \left\{\begin{pmatrix}1 & 0\\ .5 & .5\end{pmatrix}\right\} \qquad \Pi^c(\mathsf{S}_0) = \left\{\begin{pmatrix}.5 & .5\\ 0 & 1\end{pmatrix}\right\}$$

The set of best responses to all of these beliefs are the ones in Equation (3).

Another interesting example is reasoning towards Horn's division of pragmatic labor. Consider the Horn game introduced earlier, where we had $T = \{t, t^*\}$, $M = \{m, m^*\}$, $\mathbf{p} = \langle 1/2 + \epsilon, 1/2 - \epsilon\rangle$, $\mathbf{c} = \langle 0, \delta\rangle$ and a trivial semantic matrix B with value 1 everywhere. The naïve types are then:

$$\mathsf{S}_0 = \left\{\begin{pmatrix}.5 & .5\\ .5 & .5\end{pmatrix}\right\} \qquad \mathsf{R}_0 = \left\{\begin{pmatrix}.5 & .5\\ .5 & .5\end{pmatrix}\right\}$$

So neither the naïve sender strategy nor the naïve receiver strategy establish a correlation between messages and meanings. Consequently, in the first round the sophisticated players will simply use prior information and message costs, which leads to

$$\mathsf{S}_1 = \left\{\begin{pmatrix}1 & 0\\1 & 0\end{pmatrix}\right\} \qquad \mathsf{R}_1 = \left\{\begin{pmatrix}1 & 0\\1 & 0\end{pmatrix}\right\}$$

Under S_1, m^* is a surprise message. Therefore we get

$$\mathsf{S}_2 = \left\{\begin{pmatrix}1 & 0\\1 & 0\end{pmatrix}\right\} \qquad \mathsf{R}_2 = \left\{\begin{pmatrix}1 & 0\\1 & 0\end{pmatrix}, \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\right\}$$

As we have a non-singleton set in R_2, IBR and ICR diverge at this point. For IBR, we continue with

$$\Pi^u(S_2) = \left\{\begin{pmatrix}1 & 0\\1 & 0\end{pmatrix}\right\} \qquad \Pi^u(R_2) = \left\{\begin{pmatrix}1 & 0\\.5 & .5\end{pmatrix}\right\},$$

and thus

$$\mathsf{S}_3 = \left\{\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\right\} \qquad \mathsf{R}_3 = \left\{\begin{pmatrix}1 & 0\\1 & 0\end{pmatrix}, \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\right\}$$

$$\mathsf{S}_4 = \left\{\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\right\} \qquad \mathsf{R}_4 = \left\{\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\right\}$$

At this point the sequence has reached a fixed point. In the fixed point, Horn's division of pragmatic labor emerges, i.e., the unmarked type is associated with the unmarked message and the marked type with the marked message.

Using cautious beliefs, we have

$$\Pi^c(S_2) = \left\{\begin{pmatrix}1 & 0\\1 & 0\end{pmatrix}\right\} \qquad \Pi^c(R_2) = \left\{\begin{pmatrix}1 & 0\\\alpha & 1-\alpha\end{pmatrix} \mid \alpha \in (0,1)\right\},$$

Depending on the value of α, there are two possible best responses to strategies in $\Pi^c(R_2)$:

$$\mathsf{S}_3 = \left\{\begin{pmatrix}1 & 0\\1 & 0\end{pmatrix}, \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\right\} \qquad \mathsf{R}_3 = \left\{\begin{pmatrix}1 & 0\\1 & 0\end{pmatrix}, \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\right\}$$

This non-determinism continues for the sender at the next round:

$$\mathsf{S}_4 = \left\{\begin{pmatrix}1 & 0\\1 & 0\end{pmatrix}, \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\right\} \qquad \mathsf{R}_4 = \left\{\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\right\}$$

The fixed point, which is identical to the fixed point of the IBR-sequence, is reached in the next round:

$$\mathsf{S}_5 = \left\{\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\right\} \qquad \mathsf{R}_5 = \left\{\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\right\}$$

There are cases where the assumption of cautious or unbiased beliefs yields different predictions. One is an extended Horn game in which there are $n > 2$ states and messages. States are strictly linearly ordered by pior probabilities and messages by costs. For concreteness sake, let us say there are three types, three messages and three actions, $\mathbf{p} = \langle .5, .3, .2\rangle$, and $\mathbf{c} = \langle 0, .05, .01\rangle$. B is a $(3,3)$-matrix with the entry 1 everywhere. The IBR-sequence then comes out as:

$$\mathsf{S}_0 = \left\{\begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}\right\} \qquad \mathsf{R}_0 = \left\{\begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}\right\}$$

$$\mathsf{S}_1 = \left\{\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}\right\} \qquad \mathsf{R}_1 = \left\{\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}\right\}$$

$$\mathsf{S}_2 = \mathsf{S}_1 \qquad \mathsf{R}_2 = \left\{\begin{pmatrix} 1 & 0 & 0 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix} \mid \begin{array}{l} \mathbf{x}, \mathbf{y} \in \Delta(3), \\ \max \mathbf{x} = \max \mathbf{y} = 1 \end{array}\right\}$$

$$\mathsf{S}_3 = \left\{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}\right\} \qquad \mathsf{R}_3 = \mathsf{R}_2$$

$$\mathsf{S}_4 = \mathsf{S}_3 \qquad \mathsf{R}_4 = \left\{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ x_1 & x_2 & x_3 \end{pmatrix} \mid \begin{array}{l} \mathbf{x} \in \Delta(3), \\ \max \mathbf{x} = 1 \end{array}\right\}$$

$$\mathsf{S}_5 = \left\{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right\} \qquad \mathsf{R}_5 = \mathsf{R}_4$$

$$\mathsf{S}_6 = \mathsf{S}_5 \qquad \mathsf{R}_6 = \mathsf{R}_5$$

The ICR-sequence, however, comes out as:

$$\mathsf{S}_0 = \left\{\begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}\right\} \qquad \mathsf{R}_0 = \left\{\begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}\right\}$$

$$\mathsf{S}_1 = \left\{\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}\right\} \qquad \mathsf{R}_1 = \left\{\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}\right\}$$

$$\mathsf{S}_2 = \mathsf{S}_1 \qquad \mathsf{R}_2 = \left\{\begin{pmatrix} 1 & 0 & 0 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix} \mid \begin{array}{l} \mathbf{x}, \mathbf{y} \in \Delta(3), \\ \max \mathbf{x} = \max \mathbf{y} = 1 \end{array}\right\}$$

$$\mathsf{S}_3 = \mathsf{R}_2 \qquad \mathsf{R}_3 = \mathsf{R}_2$$

$$\mathsf{S}_4 = \mathsf{S}_3 \qquad \mathsf{R}_4 = \left\{\begin{pmatrix} 1 & 0 & 0 \\ 0 & x_1 & x_2 \\ 0 & y_2 & y_3 \end{pmatrix} \mid \begin{array}{l} \mathbf{x}, \mathbf{y} \in \Delta(2), \\ \max \mathbf{x} = \max \mathbf{y} = 1 \end{array}\right\}$$

$$\mathsf{S}_5 = \mathsf{R}_4 \qquad \mathsf{R}_5 = \mathsf{R}_4$$

The point of divergence between the IBR-sequence and the ICR-sequence is S_3. To see why this is, let us compare the different belief sets that are derived from R_2:

$$\Pi^u(\mathsf{R}_2) = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix} \right\} \qquad \Pi^c(\mathsf{R}_2) = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \end{pmatrix} \mid \alpha, \beta \in \Delta^+(3) \right\}.$$

For an agent holding the belief in $\Pi^u(\mathsf{R}_2)$, messages m_2 and m_3 have the same chances of inducing the desired action in t_2 and t_3, namely $1/3$. As m_2 is less costly than m_3, she will choose m_2. Now suppose an agent is in state t_2, and she holds one of the beliefs in $\Pi^c(\mathsf{R}_2)$. Her expected utilities for the three messages are $-\mathbf{c}_1$, $\alpha_2 - \mathbf{c}_2$, and $\beta_2 - \mathbf{c}_3$ respectively. Depending on the distributions α and β, either of these values may be maximal. Therefore each of the three messages may be rational in t_2, depending on the agent's specific belief. The same holds *ceteris paribus* for state t_3. Therefore there are 9 cautious responses to R_2, as opposed to a single best response under IBR. This difference leads to further divergences in later reasoning steps and ultimately to different fixed points.

ICR leads to the somewhat counter-intuitive prediction that the least marked type is always expressed by the least marked message, but that the more marked types can be expressed by any message, while the least marked message is always interpreted as the least marked type and each more marked message can be interpreted as any type except the least marked one. Whether the added predictive ability of unbiased — as opposed to cautious — beliefs is an advantage is not as clear as it may seem though. Beaver and Lee (2004), for instance, argue that we don't see cases of generalized Horn's division of pragmatic labor for more than 2 meanings and forms in natural languages.

Another example where IBR and ICR lead to different outcomes is a variant of the *some-all game* discussed above. In the *some-but-not-all game* we still have two states $T = \{t_{\exists\neg\forall}, t_\forall\}$, but we have three messages now, $M = \{m_{\text{some}}, m_{\text{all}}, m_{\text{sbna}}\}$, where m_{sbna} represents the utterance "Joe ate some but not all of the cookies." The only state where m_{sbna} is true is $t_{\exists\neg\forall}$. So we have:

$$B = \begin{array}{c} \\ t_{\exists\neg\forall} \\ t_\forall \end{array} \begin{array}{c} \begin{array}{ccc} m_{\text{some}} & m_{\text{all}} & m_{\text{sbna}} \end{array} \\ \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \end{array}.$$

Both states are *a priori* equally likely, i.e., $\mathbf{p} = \langle .5, .5 \rangle$. Message m_{sbna} is slightly more complex than the other two messages, so let us say that $\mathbf{c} = \langle 0, 0, .1 \rangle$. The IBR-sequence for this game comes out as:

$$\mathsf{S}_0 = \left\{ \begin{pmatrix} .5 & 0 & .5 \\ .5 & .5 & 0 \end{pmatrix} \right\} \qquad \mathsf{R}_0 = \left\{ \begin{pmatrix} .5 & .5 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}$$

$$\mathsf{S}_1 = \left\{ \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \right\} \qquad \mathsf{R}_1 = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}$$

$$\mathsf{S}_2 = \mathsf{S}_1 \qquad \mathsf{R}_2 = \mathsf{R}_1$$

The presence of the specific if costly message m_{sbna} here prevents the emergence of a scalar implicature for the less specific cheap message m_{some}.[4]

[4] It should be noted that the version of IBR developed by Franke (2009) uses a more sophisticated protocol for belief revision, which lets IBR converge to the same fixed point as ICR.

The ICR-sequence reaches a different fixed point:

$$\mathsf{S}_0 = \left\{ \begin{pmatrix} .5 & 0 & .5 \\ .5 & .5 & 0 \end{pmatrix} \right\} \qquad \mathsf{R}_0 = \left\{ \begin{pmatrix} .5 & .5 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}$$

$$\mathsf{S}_1 = \left\{ \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \right\} \qquad \mathsf{R}_1 = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}$$

$$\mathsf{S}_2 = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \right\} \qquad \mathsf{R}_2 = \mathsf{R}_1$$

$$\mathsf{S}_3 = \mathsf{S}_2 \qquad \mathsf{R}_3 = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}$$

$$\mathsf{S}_4 = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \right\} \qquad \mathsf{R}_4 = \mathsf{R}_3$$

$$\mathsf{S}_5 = \mathsf{S}_4 \qquad \mathsf{R}_5 = \mathsf{R}_4$$

So according to ICR, the scalar implicature for m_{some} does emerge here as well, and the more costly m_{sbna}, being superfluous, will never be used but preserves its literal meaning in the fixed point. Here the point of divergence between IBR and ICR is S_2. The unbiased and cautious belief sets respectively induced by R_1 are:

$$\Pi^u(\mathsf{R}_1) = \left\{ \begin{pmatrix} .5 & .5 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \right\} \qquad \Pi^c(\mathsf{R}_1) = \left\{ \begin{pmatrix} \alpha & 1-\alpha \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \mid \alpha \in (0,1) \right\}.$$

For a sender in state $t_{\exists\neg\forall}$ holding the belief in $\Pi^u(\mathsf{R}_1)$, the optimal message is m_{sbna} because it leads to a utility of .9, while m_{some} only has the expected utility .5. If the sender holds one of the beliefs in $\Pi^c(\mathsf{R}_1)$, however, the expected utilities of m_{some} and m_{sbna} are .9 and α respectively. Depending on the value of α, either m_{some} or m_{sbna} may be the optimal choice. Therefore m_{some} is not a surprise message under S_2, which enables the emergence of the scalar implicature.

Finally, let's have a brief look at the lifted game introduced at the end of Section 1 accounts for the epistemic inferences in (4) that are triggered when the hearer attends to the epistemic state of the speaker explicitly. Remember that we had three states $T = \{t_{[\exists\neg\forall]}, t_{[\forall]}, t_{[\exists\neg\forall,\forall]}\}$ that captured different knowledge states of the speaker. The hearer's assumptions about the speaker's competences are captured by different prior probabilities $\mathbf{p} = \{p_1, p_2, p_3\}$. We'd like to check three cases:

(i) uncertainty about competence $p_1 = p_2 = p_3$;

(ii) assumed competence (and no other bias) $p_1 = p_2 > p_3$;

(iii) assumed incompetence (and no other bias) $p_1 = p_2 < p_3$.

For illustration of how IxR-reasoning deals with assumptions about competence, let's focus on the IBR-reasoning scheme starting from a naïve speaker. The naïve speaker's behavior is independent of the priors:

$$\mathsf{S}_0 = \left\{ \begin{pmatrix} 1 & 0 \\ .5 & .5 \\ 1 & 0 \end{pmatrix} \right\}.$$

But depending on the receiver's priors, different best responses to this strategy ensue (the three cases from above are given here from left to right):

$$\mathsf{R}_1^{(i)} = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \right\} \qquad \mathsf{R}_1^{(ii)} = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \right\} \qquad \mathsf{R}_1^{(iii)} = \left\{ \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \right\}.$$

So, when the hearer beliefs that the speaker is as likely competent as not, he will not favor either semantically possible interpretation of m_{some}, but if he is biased in either way, his best response will follow suit. These receiver strategies are actually part of the fixed point behavior of this sequence. Other variants have the same fixed points.

Another relevant example which cannot be spelled out here for reasons of space are free-choice inferences as in (5b) associated with utterances as in (5a) where disjunctions occur under existential models (c.f. Kamp, 1973, 1978).

(5) a. You may take an apple or a pear.

b. ⇝ You may take an apple and you may take a pear.

Both IBR and ICR as spelled out here can account for these inferences in a way that is technically (but not conceptually) very similar to the solution of Fox (2007) (c.f. Franke, 2009, 2011, for details).

4 Iterated Quantal Response

Models of the IBR-variety assume that agents play fully rational best responses at each step of the pragmatic reasoning sequence. When we would like to explain behavioral data from psycholinguistic experiments with game theoretic models, we might wish to drop this idealistic assumption. This section therefore describes a probabilistic variant from the IxR family, where we iterate *quantal* responses. We will introduce this notion presently. The model we obtain in this way will be called iterated quantal response (IQR) model (see Degen, Franke, and Jäger, 2013, for a concrete application to experimental data).

Let us first get acquainted with the notion of a quantal response function by contrasting it with the classical best response function. Generally speaking, a response function takes expected utilities to choice probabilities. Classical, rational best responses maximize expected utility. In case of ties, agents are indifferent. So if U is an arbitrary expected utility matrix (with rows as choice points and columns as choices), then the classical best response function is simply $\text{BR}(\text{U}) = \text{Norm}(\texttt{max row}(\text{U}))$, where $\texttt{max row}(\text{U})$ returns a matrix of the same size as U where the maxima in each row are replaced by a 1 and all other values by 0. Opposed to that, the quantal response function is motivated by the idea that decision makers may make mistakes in calculating their expected utilities, or, equivalently, make mistakes in implementing the $\text{BR}(\cdot)$ function. These mistakes are small trembles so that choices with a similar expected utility receive similar choice probabilities. More concretely, quantal response assumes that choices are proportional to their expected utility. If U is an arbitrary matrix of expected utilities, then $\text{QR}_\lambda(\text{U})$ is the unique row-stochastic matrix with $\text{QR}_\lambda(\text{U})_{ij} \propto \exp(\lambda \text{U}_{ij})$. Here λ is a rationality parameter. We obtain entirely random choices for $\lambda = 0$; the higher λ, the more rational the modelled agent is, with $\lim_{\lambda\to\infty} \text{QR}_\lambda(\text{U}) = \text{BR}(\text{U})$. The quantal response function is also known as *logit choice rule* (because it reduces to the logistic function for binary choice), as *soft-max* function (Sutton and Barto, 1998) or, if $\lambda = 1$ as *Luce's choice rule* (Luce, 1959).

The quantal response function is all we need to define a simple IQR model on top of what we have defined before in the context of IBR models. Unlike the latter, IQR is defined entirely in terms of probabilistic strategies. In the context of a probabilistic model like IQR, probabilistic strategies of agent X are both (i) descriptions of X's probabilistic behavior (that capture aggregate data from a behavioral experiment, for example), and (ii) beliefs of agent Y about what X is doing.

As before we assume that naïve agents only consider the semantic meaning of messages and the immediate part of the game that concerns them (i.e., costs for senders, and priors and payoffs for receivers). The naïve sender behavior is then characterized by a quantal response to imaginary utilities resulting from a desire to speak truthfully and minimize costs. The naïve receiver behavior is given by a quantal response to imaginary expected utilities from beliefs in the truth of messages and a desire to maximize payoffs. Sophisticated types of level-$k+1$ play quantal responses to expected utilities derived from the belief that their opponent is of level k. We parameterize the whole model with a, for simplicity, single parameter λ for both sender and receiver at all levels. With this define:

$$\begin{aligned} \sigma_0 &= \mathrm{QR}_\lambda(B - \mathbf{c}) & \rho_0 &= \mathrm{QR}_\lambda(\mu_0\, \mathrm{U}_R) \\ \sigma_{k+1} &= \mathrm{QR}_\lambda(\mathrm{U}_S\, T(\rho_k) - \mathbf{c}) & \rho_{k+1} &= \mathrm{QR}_\lambda(\mathrm{Norm}(T(\sigma) \times \mathbf{p})\, \mathrm{U}_R)\,. \end{aligned}$$

Notice that by this definition agents not only play quantal responses but also believe that their interlocutor does: agents believe in the interlocutor's bounded rationality, believe that the interlocutor believes in it and so on.

Consider as an example once more the some-all game. There are no message costs and only flat priors in this simple case, so, by semantic meaning, we get the following sender strategies for different values of $\lambda \in \{0, 0.5, 1, 5\}$:

$$\mathrm{QR}_0(B) = \begin{pmatrix} .5 & .5 \\ .5 & .5 \end{pmatrix} \qquad \mathrm{QR}_{0.5}(B) \approx \begin{pmatrix} .622 & .378 \\ .5 & .5 \end{pmatrix}$$

$$\mathrm{QR}_1(B) \approx \begin{pmatrix} .731 & .269 \\ .5 & .5 \end{pmatrix} \qquad \mathrm{QR}_5(B) \approx \begin{pmatrix} .993 & .007 \\ .5 & .5 \end{pmatrix}$$

With $\lambda = 0$ the quantal response function returns purely arbitrary choices. As λ grows we approach the best response function. In case of payoff ties, quantal response returns equal choice probabilities.

For further illustration, let's look at the IQR-sequences that start with a naïve sender. It suffices to keep track of the diagonal, instead of the full speaker strategy matrix, since rows in each matrix sum to one. We look at two concrete values of λ, namely .5 (on the left) and 5 (on the right).

$$\begin{aligned} \sigma_0 &\approx \langle 0.622, .5\rangle & \sigma_0 &\approx \langle 0.993, .5\rangle \\ \rho_1 &\approx \langle 0.514, .517\rangle & \rho_1 &\approx \langle 0.839, .992\rangle \\ \sigma_2 &\approx \langle 0.503, 0.504\rangle & \sigma_2 &\approx \langle 0.984, .984\rangle \\ \rho_3 &\approx \langle 0.5, 0.5\rangle & \rho_3 &\approx \langle 0.992, .992\rangle \end{aligned}$$

We see that for $\lambda = .5$ the IQR-sequence converges rapidly to unbiased random choices. But when $\lambda = 5$, it converges to a probabilistic strategy where the probability of the scalar inference, i.e., the probability that the receiver interprets "some" as "some but not all" converges to approximately .993.

In general, the value of λ crucial and allows fitting the model to empirical data. The plot in Figure 2 shows the probability of a scalar inference in the sequence starting with a naïve sender for different values of λ. For $\lambda \leq 2$ the sequence converges to .5. For λ sufficiently bigger than 2, we get ever higher probabilities of scalar inferences in the limit. In summary, unlike models from the IBR-family, IQR models give us a parameterized probabilistic prediction about how likely listeners draw scalar inferences and about how likely speakers would conform to the Gricean Quantity prediction. These

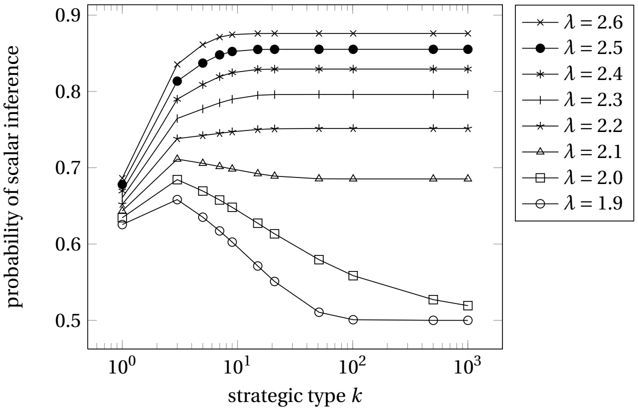

Figure 2: Probability with which receivers of strategic type k (in the sequence starting with a naïve sender) chooses the interpretation $t_{\exists\neg\forall}$ in response to m_{some} in the some-all game under IQR for different values of λ.

probabilities can be used to predict quantitative data from experiments on pragmatic language use (c.f. Degen and Franke, 2012; Degen, Franke, and Jäger, 2013; Frank and Goodman, 2012; Goodman and Stuhlmüller, 2013)

When applied to the Horn game, the IQR model unfortunately gives peculiar predictions. Consider a Horn game with prior probabilities $\mathbf{p} = \langle 3/4, 1/4 \rangle$ and costs $\mathbf{c} = \langle 0, .2 \rangle$. Since both messages are true in both states, the naïve sender's probabilities for choosing messages only depend on the costs, which are identical in each state. Consequently, the choice probabilities of unmarked and marked messages are the same in both states. For $\lambda = 5$, for instance, we obtain:

$$\sigma_0 = \begin{pmatrix} .731 & .269 \\ .731 & .269 \end{pmatrix}.$$

The posterior beliefs of a level-1 receiver will then be the same, irrespective of the message that was sent. Hence, the receiver's choice probabilities will be the same for each message. Concretely, for $\lambda = 5$ we get:

$$\rho_1 = \begin{pmatrix} .924 & .076 \\ .924 & .076 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} .731 & .269 \\ .731 & .269 \end{pmatrix}.$$

This sequence then quickly converges to a fixed point of mutual quantal responses, which, if rounded, yields exactly the numbers above. The problem, then, is that IQR simply does not predict Horn's division of pragmatic labor.

The reason why IBR/ICR-style models were able to predict (non-generalized) Horn's division of pragmatic labor was tightly connected to the occurrence of surprise messages. Since the speaker believed that the listener, at some point in the sequence, would not expect the marked message to be sent, the speaker believed that the listener's reaction to the surprising marked message would not be the choice of the unmarked state with certainty. That allowed the inference to be lifted off the ground and to break the symmetry in the reasoning chain, so to speak. But there are no surprise messages when agents believe in quantal responses because for any utility matrix U and any λ the matrix $\text{QR}_\lambda(\text{U})$ will have only strictly positive entries (due to the exponential function used to compute quantal responses). In other words, quantal response allows for all choices, no matter how bad,

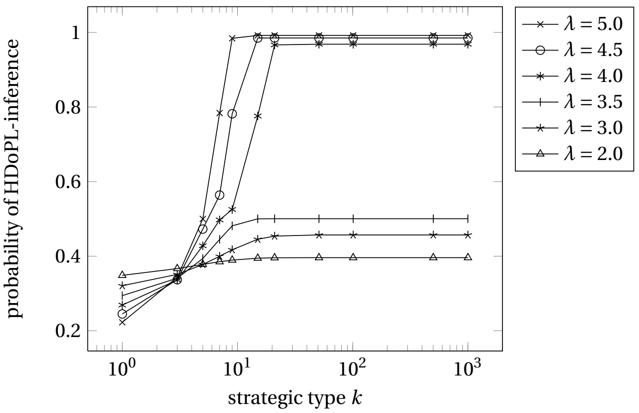

Figure 3: Probablity with which receivers of different strategic depth k (in the sequence starting with a naïve sender) select the marked interpretation for the marked expression in a Horn game for different values of λ.

to occur with some positive probability. This then excludes the kind of explanation for Horn's division of pragmatic labor that IBR/ICR-models could rely on.

Fortunately, there are several ways in which this problem can be solved (see Bergen, Levy, and Goodman, 2012, for a solution in a closely related model). One is to allow agents to put different emphasis on exploitation and exploration at different choice points. The intuition is this: if agents expect to be able to achieve high payoffs by some action at a given choice point, the chance that they make random mistakes should be lower than when no action promises a satisfactory reward; in the latter case, agents might be indifferent and "explore" more alternative actions, whereas when the road to success is clear, they would "exploit" that fact without much shopping around. This rough intuition can be implement in a number of ways, for instance, by allowing different adjustments of λ for different choice points (i.e., rows in the utility matrix) depending on the maximum expected utility value in that choice point. To capture this, we can define an amended quantal response rule like so: $\text{QR}^+_\lambda(\text{U})_{ij} \propto \lambda\,\text{U}_{ij}\,\max_{j'}\text{U}_{ij'}$. (This simple definition works nicely for the Horn game where the maximal expected utility of either sender or receiver in any choice point is always between 0 and 1. To be applicable to other games, payoffs would need to be scaled first.)

An IQR-model with this amended quantal response function predicts Horn's division of pragmatic labor, including the generalized case. It does so rather quickly even. Figure 3 shows the probability with which the receiver selects the marked interpretation when observing the marked message for different values of λ at different depth of strategic reasoning. As before, we see that λ has to exceed a certain threshold, after which the probability in question converges to values substantially bigger than chance.

5 Comparison with other Approaches

Some members of the IxR-family are superficially similar to other prominent models of pragmatic language use and interpretation. But superficial resemblance notwithstanding, there are often noteworthy conceptual divergences. We take a closer look at *Gricean* and *Neo-Gricean* theories (e.g. Atlas and Levinson, 1981; Gazdar, 1979; Grice, 1975; Horn, 1984; Levinson, 2000), *bidirectional optimality* *theory* (Blutner, 1998, 2000), the *intentions-first* approach of Geurts (2010), and the Bayesian *rational speech-act model* (Bergen, Levy, and Goodman, 2012; Frank and Goodman, 2012; Goodman and Stuhlmüller, 2013).

Gricean and Neo-Gricean Theories. The game theoretic approach to pragmatic inference computation presented here is very much in the spirit of Grice's (1975) original proposal. Grice saw his account of implicature computation as a special case of reasoning about goal-oriented rational agency. Pragmatic back-and-forth reasoning of the kind presented here takes this idea seriously by modeling pragmatic inferences as rational explanations of speaker behavior. This also implies a clear conceptual difference with work in the so-called Neo-Gricean tradition (e.g. Atlas and Levinson, 1981; Gazdar, 1979; Horn, 1984; Levinson, 2000), as well as relevance theory (Sperber and Wilson, 1995, 2004), which is sometimes called a Post-Gricean approach. The main conceptual difference is that the game theoretic approach presented here aim does not hinge on considering better or alternative formulations of Grice Maxim's of Conversation. Rather, it does without maxims entirely. The assumptions it operates with are familiar from Grice's work, most importantly: a shared interest between speaker and hearer of transmitting true information from the former to the latter. But the main motor of explanation is rationality, and beliefs about rationality etc. The approach presented here therefore is more basic and also more general, because it makes predictions also for non-cooperative contexts, or when full information transfer is not relevant (see Franke, de Jager, and van Rooij, 2012; de Jaegher and van Rooij, 2013; Stalnaker, 2006, for some applications).

Finally, another noteworthy difference between Neo-Gricean and the present game theoretic approach concerns epistemic inferences. We assumed here that epistemic inferences arise when the hearer explicitly attends to the speaker's epistemic condition. If he doesn't, the approach sketched here predicts that inferences could operate on a factual layer without any representation of the speaker's epistemic state at all. In contrast, Neo-Gricean approaches usually assume that base-level inferences, such as in (1) are always mediated by reasoning about the speaker's epistemic state. Yet a different route is taken by theories that propose to see base-level implicatures as part of the grammatical system and invoke genuine Gricean reasoning only for epistemic inferences (e.g. Fox, 2007). The game theoretic approach sketched here seems to walk a middle path: it acknowledges a distinguished level of factual inferences (which may, for instance, imply different behavior of base-level and epistemic inferences in fossilization of pragmatic inferences over time), but invokes the exact same general-purpose machinery for the calculation of base-level and epistemic inferences nonetheless.

Bidirectional Optimality Theory. Optimality theory is a widely applicable framework for studying mappings between different systems of representation (Prince and Smolensky, 1997). Bidirectional optimality theory is an extension of optimality theory, first proposed by Blutner (1998, 2000), that was designed specifically to take pragmatic back-and-forth reasoning into account in an alternative formalization of Neo-Gricean pragmatics (c.f. Atlas and Levinson, 1981; Gazdar, 1979; Horn, 1984). Bidirectional optimality proved to be widely applicable to many interesting problems along the semantics/pragmatics interface (c.f. Blutner, de Hoop, and Hendriks, 2006; Hendriks et al., 2010). The central notion of bidirectional optimality requires a form-meaning mapping to be optimal for the speaker and the listener. While the original definition was static with a close resemblance to Nash equilibrium (Dekker and van Rooij, 2000), Jäger (2002) gave an algorithmic procedure for determining bidirectionally optimal form-meaning mappings. This algorithmic procedure coincides with an IBR/ICR-reasoning scheme almost perfectly, but there are some key divergences, in particular, as Franke and Jäger (2012) showed, bidirectional optimality cannot handle pure Quantity reasoning, as needed to compute scalar inferences, without further stipulation.

The Intentions-First Approach. The IxR-reasoning schemes bear a close resemblance also to Geurts' (2010) intentions-first approach to Quantity reasoning. Geurts suggests this approach in particular as an explanation of free-choice inferences (mentioned above in connection with example (5)). While the general Gricean scheme for Quantity implicatures focuses on reasoning about alternative expressions the speaker could have used, the intentions-first approach focuses on the speaker's possible intentions behind a given utterance. If the speaker utters "you may take an apple or a pear", the listener may ask himself whether the speaker might be intending to convey that taking an apple is okay, but not taking a pear. If that was the speaker's intention, the best thing to say for him would have been "you may take and apple", not the actual utterance. A parallel argument leads the listener to conclude that the speaker also doesn't intend to convey that taking a pear is okay, but not taking an apple. Like the intentions-first approach, IxR-models also explicitly represent the speaker's possible epistemic states. Different from the former, IxR-reasoning is in a sense holistic, weighing at each stage in the sequence *all* possible epistemic states of the speaker with *all* possible expressions. The holistic and formal nature of IxR-reasoning schemes yields unambiguous and principled predictions, but can become quite cumbersome to compute. Where reasoning about only parts of the alternative expressions and potential epistemic states of the speaker is relevant, this can be formally modelled as IxR-reasoning with the help of so-called *awareness structures*, which capture diverging representations of the context of utterance in different (even counterfactual) information states speaker and listener may find themselves in (see Franke, 2013b).

The Rational Speech-Act Model. The rational speech-act (RSA) model is a Bayesian model of pragmatic language use and its interpretation (Bergen, Levy, and Goodman, 2012; Frank and Goodman, 2012; Goodman and Stuhlmüller, 2013). Bayesian reasoning is an integral part of classical game theoretic reasoning, and indeed the RSA model is closely related to IxR-models, but there are also interesting conceptual divergences. The RSA model of Frank and Goodman (2012) assumes that the speaker's behavior is given by a quantal best response to the assumption that the listener interprets expressions literally, and that the listener's behavior is given by the listener's posterior beliefs given a belief in this behavior of the speaker. That is very close to an IxR-style sequence involving a naïve receiver, a level-1 sender and a level-2 receiver. In fact, this is exactly the sequence of reasoning steps that the *optimal assertions* model of Benz and van Rooij (2007) assumed, which is a direct predecessor of the IxR-family of models. However, there are some differences in the way the behavior of speakers and listeners is defined in RSA. Firstly, the speaker's utilities in RSA are defined with respect to an information theoretic measure for the distance between his own and the belief of a naïve listener after hearing a message. In other words, while RSA assumes that the speaker cares about the *beliefs* of the listener only, IxR-models assume that the speaker cares foremost about the *action* the listener performs in response to an utterance. Secondly, the level-2 listener in RSA is not assumed to choose an optimal action based on his beliefs (e.g., by a best or quantal response), but chooses interpretation actions with a probability linearly proportional to their expected utility. Again, RSA is mainly concerned with the listener's beliefs, while IxR-models look for the listener's concrete actions based on his beliefs. Finally, the most obvious difference between IxR-reasoning and the RSA model is that the former considers more strategic types than the latter. As shown, for instance, by Degen and Franke (2012), interlocutors occasionally engage in higher-order reasoning than assumed by RSA (this is also assumed by the extended RSA model of Bergen, Levy, and Goodman, 2012). But whether this legitimates the full-blown type hierarchies of IxR-models or just a subset thereof remains an open empirical issue, just as the question how to decide between each of the above differences between RSA and IxR-models is ultimately an empirical question.

Models of pragmatic back-and-forth reasoning of the IxR-variety offer explicit accounts of the speaker's and the listener's beliefs and action choices. Under the idealized assumption of rationality and common belief therein, we can account for general pragmatic reasoning patterns that may have been acquired by repeatedly playing the interpretation game in question. But IxR-models are also flexible enough to account for pragmatic reasoning that is limited in its strategic depth, such as needed to account for subjects' reasoning in psycholinguistic experiments. Some of the most exciting future work, as we see it, will consist in refining game theoretic models of pragmatic reasoning alongside a growing body of empirical, indeed quantitative, data from psycholinguistic experiments.

References

Atlas, Jay David and Stephen Levinson (1981). "It-clefts, Informativeness, and Logical Form". In: *Radical Pragmatics*. Ed. by Peter Cole. Academic Press, pp. 1–61.

Beaver, David and Hanjung Lee (2004). "Input-Output Mismatches in Optimality Theory". In: *Optimality Theory and Pragmatics*. Ed. by Reinhard Blutner and Henk Zeevat. Palgrave MacMillan. Chap. 6, pp. 112–153.

Benz, Anton, Gerhard Jäger, and Robert van Rooij, eds. (2006). *Game Theory and Pragmatics*. Hampshire: Palgrave MacMillan.

Benz, Anton and Robert van Rooij (2007). "Optimal Assertions and what they Implicate". In: *Topoi* 26, pp. 63–78.

Bergen, Leon, Roger Levy, and Noah D. Goodman (2012). "That's what she (could have) said: How alternative utterances affect language use". In: *Proceedings of the 34th Annual Meeting of the Cognitive Science Conference*.

Blutner, Reinhard (1998). "Lexical Pragmatics". In: *Journal of Semantics* 15, pp. 115–162.

— (2000). "Some Aspects of Optimality in Natural Language Interpretation". In: *Journal of Semantics* 17, pp. 189–216.

Blutner, Reinhard, Helen de Hoop, and Petra Hendriks (2006). *Optimal Communication*. Stanford: CSLI Publications.

Camerer, Colin F. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press.

Camerer, Colin F., Teck-Hua Ho, and Juin-Kuan Chong (2004). "A Cognitive Hierarchy Model of Games". In: *The Quarterly Journal of Economics* 119.3, pp. 861–898.

Crawford, Vincent P. (2003). "Lying for Strategic Advantage: Rational and Boundedly Rational Misrepresentation of Intentions". In: *American Economic Review* 93.1, pp. 133–149.

Crawford, Vincent P. and Nagore Iriberri (2007). "Fatal Attraction: Salience, Naïveté, and Sophistication in Experimental "Hide-and-Seek" Games". In: *The American Economic Review* 97.5, pp. 1731–1750.

Crawford, Vincent P. and Joel Sobel (1982). "Strategic Information Transmission". In: *Econometrica* 50, pp. 1431–1451.

Degen, Judith and Michael Franke (2012). "Optimal Reasoning About Referential Expressions". In: *Proceedings of SemDial 2012 (SeineDial): The 16th Workshop on the Semantics and Pragmatics of Dialogue*. Ed. by Sarah Brown-Schmidt, Jonathan Ginzburg, and Staffan Larsson, pp. 2–11.

Degen, Judith, Michael Franke, and Gerhard Jäger (2013). “Cost-Based Pragmatic Inference about Referential Expressions”. In: *Proceedings of CogSci*.

Dekker, Paul and Robert van Rooij (2000). “Bi-Directional Optimality Theory: An Application of Game Theory”. In: *Journal of Semantics* 17, pp. 217–242.

Fox, Danny (2007). “Free Choice and the Theory of Scalar Implicatures”. In: *Presupposition and Implicature in Compositional Semantics*. Ed. by Uli Sauerland and Penka Stateva. Hampshire: Palgrave MacMillan, pp. 71–120.

Frank, Michael C. and Noah D. Goodman (2012). “Predicting Pragmatic Reasoning in Language Games”. In: *Science* 336.6084, p. 998.

Franke, Michael (2009). “Signal to Act: Game Theory in Pragmatics”. PhD thesis. Universiteit van Amsterdam.

— (2011). “Quantity Implicatures, Exhaustive Interpretation, and Rational Conversation”. In: *Semantics & Pragmatics* 4.1, pp. 1–82.

— (2013a). “Game Theoretic Pragmatics”. In: *Philosophy Compass* 8.3, pp. 269–284.

— (2013b). “Pragmatic Reasoning about Unawareness”. In: *Erkenntnis*.

Franke, Michael and Gerhard Jäger (2012). “Bidirectional Optimization from Reasoning and Learning in Games”. In: *Journal of Logic, Language and Information* 21.1, pp. 117–139.

Franke, Michael, Tikitu de Jager, and Robert van Rooij (2012). “Relevance in Cooperation and Conflict”. In: *Journal of Logic and Computation* 22.1, pp. 23–54.

Gazdar, Gerald (1979). *Pragmatics: Implicature, Presupposition, and Logical Form*. New York: Academic Press.

Geurts, Bart (2010). *Quantity Implicatures*. Cambridge, UK: Cambridge University Press.

Goodman, Noah D. and Andreas Stuhlmüller (2013). “Knowledge and Implicature: Modeling Lanuage Understanding as Social Cognition”. In: *Topics in Cognitive Science* 5, pp. 173–184.

Grafen, Alan (1990). “Biological Signals as Handicaps”. In: *Journal of Theoretical Biology* 144, pp. 517–546.

Grice, Paul Herbert (1975). “Logic and Conversation”. In: *Syntax and Semantics, Vol. 3, Speech Acts*. Ed. by Peter Cole and Jerry L. Morgan. Academic Press, pp. 41–58.

Hendriks, Petra et al. (2010). *Conflicts in Interpretation*. London: Equinox Publishing.

Ho, Teck-Hua, Colin Camerer, and Keith Weigelt (1998). “Iterated Dominance and Iterated Best Response in Experimental ‘p-Beauty Contests’ ”. In: *The American Economic Review* 88.4, pp. 947–969.

Horn, Laurence R. (1984). “Towards a New Taxonomy for Pragmatic Inference: Q-based and R-based Implicature”. In: *Meaning, Form, and Use in Context*. Ed. by Deborah Shiffrin. Washington: Georgetown University Press, pp. 11–42.

de Jaegher, Kris and Robert van Rooij (2013). “Game-Theoretic Pragmatics Under Conflicting and Common Interests”. In: *Erkenntnis*.

Jäger, Gerhard (2002). “Some Notes on the Formal Properties of Bidirectional Optimality Theory”. In: *Journal of Logic, Language and Information* 11.4, pp. 427–451.

— (2008). “Applications of Game Theory in Linguistics”. In: *Language and Linguistics Compass* 2/3, pp. 406–421.

— (2013). “Rationalizable Signaling”. In: *Erkenntnis*.

Kamp, Hans (1973). “Free Choice Permission”. In: *Proceedings of the Aristotelian Society* 74, pp. 57–74.

— (1978). “Semantics versus Pragmatics”. In: *Formal Semantics and Pragmatics for Natural Languages*. Ed. by Franz Guenthner and Siegfried Josef Schmidt. Dordrecht: Reidel, pp. 255–287.

Levinson, Stephen C. (2000). *Presumptive Meanings. The Theory of Generalized Conversational Implicature.* Cambridge, Massachusetts: MIT Press.

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

Luce, Duncan R. (1959). *Individual Choice Behavior: A Theoretical Analysis.* New York: Wiley.

Nowak, Martin A. and David C. Krakauer (1999). “The Evolution of Language”. In: *PNAS* 96, pp. 8028–8033.

Parikh, Prashant (1991). “Communication and Strategic Inference”. In: *Linguistics and Philosophy* 473–514.14, p. 3.

— (1992). “A Game-Theoretic Account of Implicature”. In: *TARK ’92: Proceedings of the 4th conference on Theoretical aspects of reasoning about knowledge.* Ed. by Yoram Moses. San Francisco: Morgan Kaufmann Publishers Inc., pp. 85–94.

— (2001). *The Use of Language.* Stanford University: CSLI Publications.

Prince, Alan and Paul Smolensky (1997). “Optimality: From Neural Networks to Universal Grammar”. In: *Science* 275, pp. 1604–1610.

Rogers, Brian W., Thomas R. Palfrey, and Colin Camerer (2009). “Heterogeneous Quantal Response Equilibrium and Cognitive Hierarchies”. In: *Journal of Economic Theory* 144.4, pp. 1440–1467.

Sauerland, Uli (2004). “Scalar Implicatures in Complex Sentences”. In: *Linguistics and Philosophy* 27, pp. 367–391.

Schulz, Katrin and Robert van Rooij (2006). “Pragmatic Meaning and Non-monotonic Reasoning: The Case of Exhaustive Interpretation”. In: *Linguistics and Philosophy* 29, pp. 205–250.

Spector, Benjamin (2006). “Scalar Implicatures: Exhaustivity and Gricean Reasoning”. In: *Questions in Dynamic Semantics.* Ed. by Maria Aloni, Alistair Butler, and Paul Dekker. Amsterdam, Singapore: Elsevier, pp. 229–254.

Spence, Andrew Michael (1973). “Job market signaling”. In: *Quarterly Journal of Economics* 87, pp. 355–374.

Sperber, Dan and Deirde Wilson (1995). *Relevance: Communication and Cognition (2nd ed.)* Oxford: Blackwell.

— (2004). “Relevance Theory”. In: *Handbook of Pragmatics.* Ed. by Laurence R. Horn and Gregory Ward. Oxford: Blackwell, pp. 607–632.

Stalnaker, Robert (2006). “Saying and Meaning, Cheap Talk and Credibility”. In: *Game Theory and Pragmatics.* Ed. by Anton Benz, Gerhard Jäger, and Robert van Rooij. Hampshire: Palgrave MacMillan, pp. 83–100.

Sutton, Richard S. and Andrew G. Barto (1998). *Reinforcement Learning.* MIT Press.