An Epistemic Interpretation of Bidirectional Optimality Based on Signaling Games

Michael Franke

Universiteit van Amsterdam

Institute for Logic, Language and Computation

Amsterdam, The Netherlands

To some, the relation between bidirectional optimality theory and game theory seems obvious: strong bidirectional optimality corresponds to Nash equilibrium in a strategic game (Dekker and van Rooij 2000). But in the domain of pragmatics this formally sound parallel is conceptually inadequate: the sequence of utterance and its interpretation cannot be modelled reasonably as a strategic game, because this would mean that speakers choose formulations independently of a meaning that they want to express, and that hearers choose an interpretation irrespective of an utterance that they have observed. Clearly, the sequence of utterance and interpretation requires a dynamic game model. One such model, and one that is widely studied and of manageable complexity, is a signaling game. This paper is therefore concerned with an epistemic interpretation of bidirectional optimality, both strong and weak, in terms of beliefs and strategies of players in a signaling game. In particular, I suggest that strong optimality may be regarded as a process of internal self-monitoring and that weak optimality corresponds to an iterated process of such self-monitoring. This latter process can be derived by assuming that agents act rationally to (possibly partial) beliefs in a self-monitoring opponent.

---

1 Bidirectional Optimality in Pragmatics

Optimality theory (OT) has its origin in phonology (Prince and Smolensky 1997), but has been readily applied to other linguistic subdisciplines such as syntax, semantics (Hendriks and de Hoop 2001), and pragmatics (c.f. the contributions in Blutner and Zeevat 2004). Abstractly speaking, OT is a model of how input and output representations are associated with each other based on grammatical preferences on input-output matching. More concretely, for models of pragmatic interpretation we are interested in how a set M of (input) forms and a set T of (output) meanings are matched by language users in production and interpretation. An OT-SYSTEMS $\langle \text{Gen}, \geq \rangle$ is then just a pair $\langle \text{Gen}, \geq \rangle$ consisting of a GENERATOR $\text{Gen} \subseteq M \times T$ that gives us the initially possible form-meaning pairs and an ordering $\geq$ on elements of Gen that measures how well the elements of the generator satisfy certain standards of grammaticality, normality, efficiency, or whatever might be at stake for a particular explanation of pragmatic language use.[1]

Based on the ordering $\geq$, an OT-system specifies the preferred input-output associations in several ways. Since $\geq$ is an ordering on a set of input-output pairs, we can either take a production perspective and ask which output is best when we fix the input dimension, or we can take a comprehension perspective and ask which input is best when we fix the output dimension. The former production perspective is taken by OT-syntax, the latter comprehension perspective is taken by OT-semantics. Abstractly, we can define the set of UNIDIRECTIONALLY OPTIMAL PAIRS as follows:

$$\begin{aligned} \text{OT}_{\text{syn}} &= \{\langle m,t\rangle \in \text{Gen} \mid \neg\exists t' : \langle m,t'\rangle \in \text{Gen} \wedge \langle m,t'\rangle \succ \langle m,t\rangle\} \\ \text{OT}_{\text{sem}} &= \{\langle m,t\rangle \in \text{Gen} \mid \neg\exists m' : \langle m',t\rangle \in \text{Gen} \wedge \langle m',t\rangle \succ \langle m,t\rangle\}. \end{aligned}$$

Optimization along both dimensions at the same time is also possible, of course. This is BIDIRECTIONAL OPTIMALITY and it comes in two varieties, a strong notion and a weak notion (Blutner 1998, 2000). We say that an input-output pair is STRONGLY OPTIMAL iff it is unidirectionally optimal for both production and comprehension:

$$\text{BIOT}_{\text{str}} = \text{OT}_{\text{syn}} \cap \text{OT}_{\text{sem}}$$

is the set of all strongly optimal pairs. Adopting Jäger's reformulation of Blutner's original definition (Jäger 2002), we say that a pair $\langle m,t\rangle$ is WEAKLY OPTIMAL iff

(i) there is no weakly optimal $\langle m,t'\rangle$ such that $\langle m,t'\rangle \succ \langle m,t\rangle$; and

(ii) there is no weakly optimal $\langle m',t\rangle$ such that $\langle m',t\rangle \succ \langle m,t\rangle$;

and we denote the set of all weakly optimal pairs with $\text{BIOT}_{\text{weak}}$. It is obvious

[1] Normally, the ordering $\geq$ would be derived from a set of ranked constraints, but for the purposes of this paper we can safely abstract from that.

that all strongly optimal pairs are also weakly optimal, but it may be the case that there are weakly optimal pairs which are not strongly optimal.

How should we interpret the various optimality notions for applications to linguistic pragmatics? What exactly does it mean when an OT-system selects a given form-meaning pair as weakly optimal but not strongly optimal, or as unidirectionally optimal but not strongly optimal? These are the general questions that this paper seeks to address.

Proponents of OT-pragmatics are not unanimous about this issue. Some propose to think of unidirectional and strong optimality as measures of online pragmatic competence, but reject the notion that weak optimality has anything to do with actual pragmatic reasoning (Blutner and Zeevat 2004, 2008). Weak optimality is rather viewed from a diachronic, evolutionary perspective as giving the direction into which semantic meaning of expressions will most likely shift over time by pragmatic pressures.

Opposed to this view, others treat also weak optimality as a model of pragmatic reasoning competence. Under this interpretation different notions of optimality express different levels of *perspective taking*: whereas unidirectional optimization does not require to take the interlocutor's perspective into account, bidirectional optimization does (cf. Hendriks et al. 2007, chapter 5). More strongly even, optimality theory in pragmatics is often related to theory of mind (ToM) reasoning (Premack and Woodruff 1978): unidirectional optimization is taken to involve no ToM reasoning (or zero-order ToM), strong optimization would correspond to first-order, and weak optimization would involve second-order ToM reasoning (see, for instance, Flobbe et al. 2008, p. 424).

Given the controversy about its conceptual interpretation, what would be required is, in a manner of speaking, an interpretation of the basic notions of optimality theory that clarifies (some of) its intended use in pragmatic applications. For this purpose, it would be most welcome to supply in particular an *epistemic interpretation* of OT, i.e., an interpretation that links OT's basic notions to more familiar features of human cognition such as beliefs and preferences. Comparison to a related game theoretic model can help achieve this, especially when we focus on an epistemic characterization of player behavior. This is what this paper tries to achieve by linking pragmatic OT-systems to particular kinds of signaling games and by linking notions of optimality to particular player types of varying degree of sophistication.

The paper is structured as follows. I will first review critically the most commonly adopted characterization of OT-pragmatics in terms of strategic games in section 2. It will transpire that a strategic game is inadequate to capture the sequential nature of speech and its uptake and interpretation. Section 3 explores a different characterization of OT in terms of signaling games, and section XYZ finally links optimality notions to iterated best responses.

2 BiOT and Strategic Games

Bidirectional optimization is simultaneous optimization of both the production and the comprehension perspective. At first glance, this looks very similar to an equilibrium state in which the speaker's and the hearer's preferences are balanced. And, indeed, there is a *prima facie* very plausible link between BIOT and game theory. Dekker and van Rooij (2000) (henceforth D&vR) show that the notion of strong optimality corresponds one-to-one to the notion of Nash equilibrium in an *optimality game*.[2] An optimality game is a straightforward translation of an OT-system into a strategic game. D&vR continue to show that weak optimality corresponds with the outcome of a process that we could call *iterated Nash-selection*. Let's first look at the analysis of D&vR in more detail and then reflect critically.

2.1 Strong Optimality as Nash Equilibrium

Formally a strategic game is a triple $\langle N, (A)_{i\in N}, (\succeq)_{i\in N}\rangle$ where N is a set of players, A_i are the actions available to player i and $\succeq_i$ is player i's preference relation over action profiles $\times_{j\in N} A_j$, i.e., possible outcomes of the game. A Nash equilibrium of a strategic game is an action profile a^* such that for all $i \in N$ there is no $a_i \in A_i$ for which:[3]

$$(a^*_{-i}, a_i) \succ_i a^*.$$

In words, a Nash equilibrium is an action profile which no player would like to deviate from given that all other players conform.

Take an OT-system with forms M, meanings T —assuming for simplicity

[2] D&vR use the term "interpretation game" for what I call "optimality game." I would like to reserve the former term for a particular kind of signaling game to be introduced later.

[3] Here, (a^*_{-i}, a_i) is the action profile which is derived from a^* by replacing player i's action in a^* with a_i.

that Gen $= M \times T$— and some ordering $\geq$ over form-meaning pairs. An OPTIMALITY GAME, as defined by D&vR, is a strategic game between a speaker S and a hearer H such that the speaker selects a form, $A_S = M$, the hearer selects a meaning, $A_H = T$, and the players' preferences are just equated with the ordering of the OT-system, $\geq_S = \geq_H = \geq$.

An action profile $\langle m, t \rangle$ is a Nash equilibrium of an optimality game iff

(i) there is no $m' \in M$ such that $\langle m', t \rangle >_S \langle m, t \rangle$; and

(ii) there is no $t' \in T$ such that $\langle m, t' \rangle >_H \langle m, t \rangle$.

But since $\geq_S = \geq_H = \geq$ this is the case just when $\langle m, t \rangle \in \text{BIOT}_{\text{str}}$. Consequently, every Nash equilibrium of an optimality game is a strongly optimal pair in the corresponding OT-system, and every strongly optimal pair of an OT-system is a Nash equilibrium of the corresponding optimality game. D&vR's result in slogan form: strong optimality is Nash equilibrium (in an optimality game).

2.2 Weak Optimality as Iterated Nash Selection

In order to understand D&vR's characterization of weak optimality, we should first notice that the recursive definition of weak optimality given above is rather cumbersome to apply. In practice, therefore, most often weakly optimal pairs are computed via a manageable algorithm which iteratively computes optimal pairs. D&vR's characterization of weak optimality is inspired by this iterative computation process, so that we should first revisit the BIOT-algorithm.

2.2.1 The BIOT*-Algorithm*

The BIOT-algorithm, which is due to Jäger (2002) and given in figure 1, iteratively computes three disjoint sets of form-meaning pairs (Jäger 2002):

(i) the set Pool_n of form-meaning pairs still in competition for optimality after n rounds of iteration;

(ii) the set Opt_n of form-meaning pairs that have been identified as optimal after round n;

(iii) the set Blo_n of form-meaning pairs that are blocked by an optimal pair and therefore removed from the pool.

$\text{Pool}_0 \leftarrow \text{Gen}$
$\text{Opt}_0 \leftarrow \emptyset$
$\text{Blo}_0 \leftarrow \emptyset$
$n \leftarrow 0$
while $\text{Pool}_n \neq \emptyset$ **do**
$\quad \text{Opt}_{n+1} \leftarrow \text{Opt}_n \cup \{\langle m,t\rangle \in \text{Pool}_n \mid$
$\qquad \neg\exists \langle m',t\rangle \in \text{Pool}_n\ \langle m',t\rangle > \langle m,t\rangle \wedge$
$\qquad \neg\exists \langle m,t'\rangle \in \text{Pool}_n\ \langle m,t'\rangle > \langle m,t\rangle\}$
$\quad \text{Blo}_{n+1} \leftarrow \text{Blo}_n \cup \{\langle m,t\rangle \in \text{Pool}_n \mid$
$\qquad \exists \langle m',t\rangle \in \text{Opt}_{n+1}\ \langle m',t\rangle > \langle m,t\rangle \vee$
$\qquad \exists \langle m,t'\rangle \in \text{Opt}_{n+1}\ \langle m,t'\rangle > \langle m,t\rangle\}$
$\quad \text{Pool}_{n+1} \leftarrow \text{Pool}_0 \setminus (\text{Opt}_{n+1} \cup \text{Blo}_{n+1})$
$\quad n \leftarrow n+1$
end while

Figure 1: The BIOT-algorithm

Initially, Pool_0 is the set Gen and there are no optimal or blocked forms. The algorithm then iteratively computes optimal pairs based on a comparison of forms left in the pool and removes optimal and blocked pairs from the pool until every form-meaning pair is removed from the pool as either optimal or blocked. We could think of the pool at round n as a reduced OT-system. The BIOT-algorithm thus repeatedly checks for strong optimality in ever more reduced OT-systems and thus selects all and only weakly optimal pairs (see Jäger 2002; Franke 2009, for more formal detail).

2.2.2 Iterated Nash Selection

The main idea of D&vR's characterization of weak optimality is now this. Firstly, we saw that the BIOT-algorithm iteratively computes strongly optimal pairs, based on a shrinking pool of candidate pairs. Secondly, we also saw that strong optimality can be likened to Nash equilibrium in optimality games. Hence, the workings of the BIOT-algorithm can be recast in game theoretic terms as a process of iteratively removing action profiles from competition for Nash equilibrium that are, in a way of speaking, dominated by a Nash equilibrium.

In order to make this idea more precise, D&vR allow strategic games to have partial preferences. For games with partial preferences, not every definition of Nash equilibrium will do, but the one given above applies. The process of ITERATED NASH-SELECTION on a strategic game $I_0 = \langle N, (A)_{i\in N}, (\geq_0)_{i\in N}\rangle$ is defined inductively as follows: let NE_n be the set of Nash equilibria of game I_n;

I_{n+1} is derived from I_n by restricting the preferences $\succeq_{n,i}$ to:

$$\succeq_{n+1,i} = \{\langle x, y\rangle \in \succeq_{n,i} \mid \neg\exists z \in \mathrm{NE}_n : z \succ_{n,i} x\}.$$

If for some index n we have $I_n = I_{n+1}$, we consider the process to be terminated, and call NE_n the *outcome* of the process of iterated Nash-selection. D&vR show that this process corresponds to the BIOT-algorithm if applied to optimality games: if I is the optimality game corresponding to an OT-system, then the outcome of iterated Nash-selection on I contains all and only the weakly optimal pairs of the OT-system.

2.3 Critique

The characterization of strongly optimal pairs as Nash equilibria in an optimality game has some *prima facie* plausibility and seems unanimously endorsed as *the* link between OT and game theory. But on closer look the suggested parallel turns out not to be very sensible. To model communication as a strategic game is to assume that speakers choose formulations *independently* of a meaning that they want to express, and that hearers choose an interpretation *irrespective* of an utterance that they have observed. But this is clearly inadequate for pragmatic explanations. Obviously, speakers choose forms *conditional* on a meaning to be expressed and hearers choose interpretations of a *given* form not interpretations *per se*.

Here is a concrete example to make my argument more tangible. Let us consider the OT-system that Hendriks and Spenader (2005) use in order to explain the preferred interpretations of sentences (1) and (2).

(1) Bert washed himself.

(2) Bert washed him.

Clearly, for (most) adult speakers of English the sentence (1) has only a coreferential reading for the reflexive pronoun, i.e., (1) means that Bert washed Bert. In contrast, sentence (2) has *no* coreferential reading for the non-reflexive pronoun, i.e., to (most) adult speakers (2) means that Bert washed someone other than himself. In order to model not only adult interpretation, but also a peculiar pattern in the acquisition of this piece of pragmatic competence, Hendriks and Spenader (2005) adopt a simple OT-system with two forms $m_{\texttt{himself}}$ for (1) and

$m_{\texttt{him}}$ for (2), and two meanings $t_{\texttt{BB}}$ for a situation in which Bert washed Bert and $t_{\texttt{BE}}$ for a situation in which Bert washed Ernie. All possible form-meaning combinations are generated in this system and the ordering can be visualized as follows:[4]

$$\begin{tikzcd} & t_{\texttt{BB}} & t_{\texttt{BE}} \\ m_{\texttt{himself}} & \bullet & \bullet \arrow[l] \arrow[d] \\ m_{\texttt{him}} & \bullet \arrow[u] \arrow[r, leftrightarrow] & \bullet \end{tikzcd}$$

This gives rise to the following sets of optimal pairs:

$$\begin{aligned} \text{Opt}_{\text{syn}} &= \{\langle m_{\texttt{himself}}, t_{\texttt{BB}}\rangle, \langle m_{\texttt{him}}, t_{\texttt{BE}}\rangle\} \\ \text{Opt}_{\text{sem}} &= \{\langle m_{\texttt{himself}}, t_{\texttt{BB}}\rangle, \langle m_{\texttt{him}}, t_{\texttt{BB}}\rangle, \langle m_{\texttt{him}}, t_{\texttt{BE}}\rangle\} \\ \text{BIOT}_{\text{str,weak}} &= \{\langle m_{\texttt{himself}}, t_{\texttt{BB}}\rangle, \langle m_{\texttt{him}}, t_{\texttt{BE}}\rangle\} \end{aligned}$$

We should now ask what it means to say that the strongly optimal pairs are $\langle m_{\texttt{himself}}, t_{\texttt{BB}}\rangle$ and $\langle m_{\texttt{him}}, t_{\texttt{BE}}\rangle$ and whether this squares with what it means to be a Nash equilibrium in an optimality game.

Suppose the speaker (Alice) and the hearer (Bob) are playing the corresponding strategic optimality game. In this game, both Alice and Bob make effectively simultaneous and *independent* decision. We may imagine that this is achieved, e.g., by writing down and passing to a judge the choice between either $m_{\texttt{him}}$ or and $m_{\texttt{himself}}$ for Alice, and between $t_{\texttt{BB}}$ and $t_{\texttt{BE}}$ for Bob. A Nash equilibrium is then a pair of actions $\langle m, t\rangle$ such that, firstly, given Bob's choice t, Alice would not strictly prefer a message different from m, and, secondly, given Alice's choice m, Bob would not strictly prefer an interpretation different from t. This means that if this game is played repeatedly, and if, for example, Bob shows a tendency to play $t_{\texttt{BB}}$ however slightly more frequently, then Alice would start to play $m_{\texttt{himself}}$ more frequently, and the whole process would start reinforcing itself until we reach the *steady state* in which Alice always plays $m_{\texttt{himself}}$ and Bob always plays $t_{\texttt{BB}}$. This steady state is a Nash equilibrium, and to think of Nash equlibria as steady states in this way is indeed the most prevalent textbook interpretation of this solution concept (e.g. Osborne and Ru-

[4] An arrow from one form-meaning pair to another indicates that the form-meaning pair to which the arrow points is strictly more preferred according to $\geq$. It is not essential here that the ordering is derived from particular constraints, each with its own independent motivation (see Hendriks and Spenader 2005, for details).

binstein 1994; Osborne 2004; Heap and Varoufakis 2004).

But this has nothing to do with either the way that we imagine communication to proceed if online pragmatic reasoning is concerned, or with a reasonable model of language evolution under pragmatic pressures. It is also not the way we would commonly interpret a set of (strongly) optimal form-meaning pairs. The above set of strongly optimal pairs, which contains $\langle m_{\texttt{himself}}, t_{\texttt{BB}}\rangle$ and $\langle m_{\texttt{him}}, t_{\texttt{BE}}\rangle$, is commonly taken to describe *conditional* production and interpretation behavior (that is in a certain sense optimal). In particular, from a production point of view this set captures that if the speaker wants to express the meaning $t_{\texttt{BB}}$ then it is optimal to use message $m_{\texttt{himself}}$ and that if the the speaker wants to express the meaning $t_{\texttt{BE}}$ then it is optimal to use message $m_{\texttt{him}}$. Similarly, from an interpretation point of view the set captures that if the hearer observes message $m_{\texttt{himself}}$, he should optimally interpret this as meaning $t_{\texttt{BB}}$, and if the hearer observes message $m_{\texttt{him}}$, he should optimally interpret this as meaning $t_{\texttt{BE}}$.

In effect, that means that Nash equilibrium is an inadequate characterization of strong optimality, because optimality games are, qua *strategic* game, the inadequate game model for pragmatic OT-systems. The critique then carries over to Dekker and van Rooij's characterization of weak optimality in terms of iterated Nash selection. Phrased polemically, if Nash equilibrium is an inadequate characterization of strong optimality, then if you repeatedly link Nash equilibrium and strong optimality in a reduced system (be it OT-system or optimality game), then this is not making things better, but worse.[5]

3 BiOT and Signaling Games

The above considerations suggest that the natural way of interpreting a set of form-meaning pairs —be they optimal or not— is not as a set of Nash equilibria, but rather as a (possibly partial) specification of *conditional* production and interpretation behavior. Speech production proceeds from a thought or intention that needs to be expressed to a choice of form to express the desired content with. Interpretation of an utterance starts only after a message that needs to be interpreted has been observed. This is all natural, I believe, but it does call for a different game model to match pragmatic OT-systems: we need at least a

[5] For more detailed criticism also of the concept of iterated Nash selection see Franke (2009).

dynamic game in which the speaker chooses a message conditional on a to-be-expressed meaning, and the hearer subsequently chooses an interpretation given that he has observed a form.

The perhaps most manageable and (for that reason) most widely studied kind of game that fits this description is a *signaling game*. A signaling game is a special kind of dynamic game with incomplete information that has been studied extensively in philosophy (Lewis 1969), economics (Spence 1973), biology (Zahavi 1975; Grafen 1990) and linguistics (Parikh 1991, 1992, 2001; van Rooij 2004). Informally speaking, the idea is that the sender (the agent modelling the speaker) knows the true state of affairs t, but the receiver (the agent modelling the hearer) does not. Given the true state t the sender then chooses a message m which the receiver observes. Subsequently, the receiver chooses an action a as his proper response. An outcome of such a game is given as the triple $\langle t, m, a \rangle$. Naturally, sender and receiver may prefer some outcomes more than others and these preferences may select for a particular class of sender and receiver behavior under a given solution concept.

Formally, a signaling game (with meaningful signals) is a tuple

$$\langle \{S, R\}, T, \mathrm{Pr}, M, [\![\cdot]\!], A, \mathrm{U}_S, \mathrm{U}_R \rangle$$

where sender S and receiver R are the players of the game; T is a set of states of the world; $\mathrm{Pr} \in \Delta(T)$ is a probability distribution over T, which represents the receiver's uncertainty which state in T is actual;[6] M is a set of messages that the sender can send; $[\![\cdot]\!] : M \to \mathcal{P}(T) \setminus \emptyset$ is a denotation function that gives the predefined semantic meaning of a message as the set of all states where that message is true (or otherwise semantically acceptable); A is the set of response actions available to the receiver; and $\mathrm{U}_{S,R} : T \times M \times A \to \mathbb{R}$ are utility functions for both sender and receiver that give a numerical value for, roughly, the desirability of each possible play of the game.[7]

In general, behavior of players in dynamic games is represented in terms

6 As for notation, $\Delta(X)$ is the set of all probability distributions over set X, Y^X is the set of all functions from X to Y, $X : Y \to Z$ is alternative notion for $X \in Z^Y$, and $\mathcal{P}(X)$ is the power set of X.

7 To rule out certain irrelevant and aberrant cases, I will assume throughout that for each state t there is at least one message m such that $t \in [\![m]\!]$ and that Pr has full support, i.e., that $\mathrm{Pr}(t) > 0$ for all $t \in T$.

of STRATEGIES which select possible moves for each agent for any of their choice points in the game. For signaling games, a PURE SENDER STRATEGY $s \in M^T$ is a function from states to messages which specifies which message the sender will or would send in each state that might become actual. A PURE RECEIVER STRATEGY $r \in A^M$ is a function from messages to actions which similarly specifies which action the receiver will or would choose as a response to each message he might observe. (Obviously, the receiver knows only what message has been sent, but not what state is actual, so he has to choose an action for each message he might observe and cannot condition his choice on the actual state of affairs). A PURE STRATEGY PROFILE $\langle s, r\rangle$ is then a characterization of the players' *joint behavior* in a given signaling game.

3.1 Optimal Pairs as Partial Strategies

If my previous argument is correct, and a set of optimal pairs, is to be interpreted as a specification of *conditional* production or comprehension behavior, then we should generally link sets of form-meaning pairs, be they optimal or not, to strategies in a suitable signaling game. In particular, a set of form-meaning pairs partially defines a sender or receiver strategy in a SIGNALING GAME WITH INTERPRETATION ACTIONS where

(i) the set of states in the signaling game are the meanings T of the OT-system; these are the meanings that the speaker might want to express;

(ii) the set of messages in the signaling game are the forms M of the OT-system; these are the messages the speaker can choose to express a meaning when she wants to; and

(iii) the set of receiver actions in the signaling game are interpretations, i.e., the meanings T of the OT-system.

In general, we can read off a (partial) description of a sender and receiver strategy for such a game from any set $O \subseteq M \times T$. If we agree to write

$$O(t) = \{m \in M \mid \langle m, t\rangle \in O\} \quad \text{and} \quad O(m) = \{t \in T \mid \langle m, t\rangle \in O\}, \qquad (3.1)$$

the set of pure sender strategies in a signaling game with interpretation actions compatible with O is:

$$\mathsf{S}(O) = \{s \in \mathsf{S} \mid O(t) \neq \emptyset \rightarrow s(t) \in O(t)\}\,;$$

and the set of pure receiver strategies compatible with O is:

$$\mathsf{R}(O) = \{r \in \mathsf{R} \mid O(m) \neq \emptyset \rightarrow r(m) \in O(m)\}\,.$$

Obviously, an arbitrary set O need not specify a full strategy. For instance, there may be states t for which $O(t)$ is empty, so that when taken as a description of a sender strategy O is only a *partial* description. I suggest that this is really how we should set the link between OT and game theory in pragmatics: sets of form-meaning pairs —no matter whether any notion of optimality has selected these— are specifications of strategies in a corresponding signaling game with interpretation actions.

3.2 OT-Systems and Signaling Games

Linking form-meaning pairs to strategies may be a natural idea, but this much does not yet fix a complete translation between OT-systems and signaling games. Some correspondences are hardly worth mentioning: speakers correspond to senders and hearers correspond to receivers, of course, and the generator places restrictions on the set of possible form-meaning associations and this naturally finds its expression in the semantic denotation function

$$\langle m, t \rangle \in \text{Gen iff } t \in [\![m]\!]$$

if we assume that the corresponding signaling game makes truthful signaling obligatory, i.e., that the sender can only ever send a true message in a given state. But all this still does not fix an interpretation of the ordering $\geq$ of the OT-system. Also the prior probabilities $\Pr(\cdot)$ and the utilities $\mathrm{U}_{S,R}$ for both sender and receiver are still unspecified.

Formally, there are many possibilities of translation between OT-systems and signaling games. I have explored one such formal parallel in Franke (2009), where I link optimality notions with the behavior of strategic types in a sequence of iterated best responses. An iterated best response model, or IBR model for short, is an epistemic solution concept in which different strategic types of players are defined in terms of their beliefs about opponent player behavior (cf. Jäger 2008; Jäger and Ebert 2009). The beginning of the sequence is given by naïve strategic types of level 0 who do not take their opponent's perspective into account, but who may be susceptible to certain focal framing effects in the game structure (cf. Schelling 1960, for focality). Players of level $k+1$ then believe that they are facing a level-k opponent and play a best response to that belief. It is then possible to identify in particular the behavior of naïve level-0 receivers with unidirectionally optimal interpretation, the behavior of level-1 senders with unidirectionally optimal production and the interpretation of level-2 receivers with strong optimality if we assume that the receiver uses a particular, simplistic (and strictly speaking incorrect) belief formation process when computing his posterior beliefs after receiving a message (see Franke 2009, for details).

This characterization of optimality notions in terms of IBR reasoning assumed an independently motivated IBR model and tried to match optimality notions with as little amendment as possible onto the strategic types of this model. It turned out, however, that especially a characterization of weak optimality is rather difficult, because the game-theoretic idea of a rational best response to a belief in an opponent strategy is *holistic* in the sense that it takes into account the whole of an opponents strategy (see also section 4.2). This makes it possible that certain form-meaning associations appear optimal in early stages, but are dismissed as optimal later on, because every possible form-meaning association is always reconsidered at every iteration step. Opposed to that, the BIOT algorithm, which selects for weak optimality, is rather *myopic* in that the set of optimal form-meaning associations grows monotonically. The upshot of this is that Bayesian rationality, if based on a standard belief in opponent strategy, does not always match the fast-and-frugal form-meaning selection process modelled by the BIOT algorithm. In Franke (2009) I therefore give a restriction on agent's belief formation, which is admittedly rather severe, but which guarantees a match between rationalistic IBR and weak optimality.

In the following, I would like to go a different route, one that is closer to OT-pragmatics and parts from the idea of staying as close as possible to the rationalistic norms of standard game theory. I would like to start out from the assumption that sets of form-meaning pairs describe partial strategies of senders and receivers in a signaling game. Based on this, it is possible to simply reconstruct the BIOT-algorithm as a *behavioral definition* of strategic types of players. Finally, we can then look back at this behavioral characterization and ask which *epistemic assumptions*, e.g., about belief formation, rationality or preferences, would give rise to this behavior and how these assumptions square with the common interpretation of optimality notions in the pragmatic OT-community on the one hand, and the accepted standards of game theory on the other. The epistemic interpretation of optimality that I end up suggesting is that bidirectional optimality is a process of, if necessary iterative, self-monitoring for congruence between form-meaning associations in production and interpretation.

4 Iterated Self-Monitoring

In order to match sets of form-meaning pairs to strategies of senders and receivers, we should assume that the set of receiver actions equals the set of states $T = A$. Going a step further, let us also assume that the signaling game corresponding to a given OT-system has a particular payoff structure, namely that the signaling game models a situation in which sender and receiver would like to communicate the true state of affairs successfully. This is achieved by setting:

$$\mathrm{U}_S(t, m, a) = \mathrm{U}_R(t, m, a) = \begin{cases} 1 & \text{if } t = a \\ 0 & \text{otherwise.} \end{cases}$$

Let us call a signaling game with this payoff structure an INTERPRETATION GAME.

Recall that according to the standard interpretation of unidirectional optimality, as outlined in section 1, we want to link unidirectional optimality to the behavior of senders and receivers who do *not* take their opponent's strategy into account but only follow their own preferences on form-meaning associations as specified by a given OT-system. This can be achieved if we assume that there are *naïve strategic types* which do not take a belief about their opponent into account, but merely play a rational best response given their preferences about form-meaning associations.

4.1 Unidirectional Optimality and Naïve Players

For the sender this is easily achieved by assuming that messages have *state-dependent costs*. We model this by a function $\mathrm{C} : T \times M \rightarrow \mathbb{R}$ that associates for every state t and message m the costs $\mathrm{C}(t, m)$ that sending m in state t incurs for the sender.[8] To translate the speaker's preferences, as captured in $\geq$ into the signaling game, we simply assume that for all $\langle m, t\rangle$ and $\langle m', t\rangle$ in Gen:

$$\langle m, t\rangle \geq \langle m', t\rangle \;\text{ iff }\; \mathrm{C}(t, m) \leq \mathrm{C}(t, m').$$

We may then assume that a naïve, but rational sender type S_0, who does not take interpretation behavior into account but otherwise cares for her preferences, will choose a message that minimizes costs in each state. We can represent this sender type by a set of pure strategies as follows:

$$S_0 = \left\{ s \in \mathsf{S} \mid \forall t \in T \;\; s(t) \in \arg\min_{m \in M} \mathrm{C}(t, m) \right\}.$$

By construction, it is trivially so that:[9]

$$\langle m, t\rangle \in \textsc{ot}_{\text{syn}} \;\text{ iff }\; m \in S_0(t).$$

In words, our naïve sender type corresponds behaviorally to unidirectional optimality along the production dimension.

For the receiver a similar move is possible. Since in an interpretation game states correspond one-to-one to actions, and, moreover, the receiver would like to match his response action to the true state, we find that a receiver who does not take his opponent's strategy into account would maximize for the most likely state in which a given message could have been sent (given the restrictions on truthful signaling). That is to say that *prima facie* we would like to construct a naïve receiver type similar to S_0 who takes into account only his preferences as represented in his prior probabilities $\Pr(\cdot)$. The problem with this is that not all OT-orderings can be translated in this way because, obviously,

[8] I will follow standard practice and assume that these costs are nominal, i.e., that they apply only when expected utilities based on U_S reach a tie.

[9] As for notation, a set of pure sender strategies like S_0 can equivalently be represented as a set of form-meaning pairs. With this, $S_0(t)$ is defined by the notational convention in (3.1).

$\Pr(\cdot)$ only specifies a *global* ordering on T independent of the message that the receiver observes. However, there are independent arguments for thinking of the receiver's prior probabilities merely as a simplistic and convenient way of specifying those global form-meaning associations that do not vary with the message (see Franke 2009, section 3.1). If we then want to be able to translate any arbitrary OT-ordering into a signaling game via prior probabilities, we should adapt the definition of a signaling game to include an ASSOCIATION FUNCTION $\text{Ass} : M \times T \to \mathbb{R}$, such that for all $\langle m, t\rangle$ and $\langle m, t'\rangle$ in Gen we have:[10]

$$\langle m, t\rangle \geq \langle m, t'\rangle \text{ iff } \text{Ass}(m, t) \geq \text{Ass}(m, t').$$

Based on his associative preferences, a naïve receiver R_0, who is rational but does not take into account his opponent's behavior, will maximize for each observed message the likelihood of matching the true state by selecting a maximally associated state:

$$R_0 = \left\{ r \in \mathsf{R} \mid \forall m \in M \; r(m) \in \arg\max_{t \in T} \text{Ass}(m, t) \right\}.$$

Again, by construction, this corresponds with unidirectional optimality along the comprehension dimension:

$$\langle m, t\rangle \in \text{OT}_{\text{sem}} \text{ iff } t \in R_0(m).$$

In line with the common idea that unidirectional optimization does not involve taking the opponent's behavior into account, the above definition of naïve players offers a straightforward behavioral implementation of unidirectional optimality in a signaling game. Moreover, this characterization also allows to draw further conclusions about a possible epistemic interpretation of unidirectional optimality. We should think of preferences, as captured in the OT-ordering $\geq$, as the strength of associating form-meaning pairs. This is given by grammar, in a wide sense of the term, and may involve contextual association biases, depending on the intended application of the OT-system. But, crucially, building on these basic grammatical preferences, unidirectional optimality is supplied by

[10] A prior probability function $\Pr(\cdot)$ is then just a special case of an association function: constant over all m and scaled to the interval $[0; 1]$.

Bayesian rationality in the absence of any conjecture about opponent behavior.

4.2 Strong Optimality as Self-Monitoring

Strong optimality is defined as the intersection of unidirectional optimization along the comprehension and the production dimension, and is therefore often considered an operation that takes into account the opponent's strategy (e.g. Hendriks and Spenader 2005; Flobbe et al. 2008). However, there is a fundamental difference between the way game theory models such perspective taking from the way this notion is present in strong and weak optimality. This section therefore suggests to look at strong optimality as a mere self-monitoring, not as genuine perspective taking in the strong game-theoretic sense.

Let us begin by looking more closely at the idea of perspective taking in game theory. If a rational agent takes the behavior of an opponent into account, game theorists assume that the agent plays a rational best response to the belief that her opponent is behaving in the specified way. Take, for instance, the behavior of a naïve receiver R_0. A belief in this behavior is a belief that message m is interpreted as some state in $R_0(m)$. If a sender plays a best response to this belief, she optimizes her behavior, based on her preferences, by taking into account the complete interpretation behavior of R_0, i.e., the way *all* messages are interpreted according to R_0. In other words, perspective taking in game theory is *holistic* in the sense that the *whole* strategy of the opponent is taken into account when making a choice.

This is not what strong optimality implements. In order to adhere to strong optimality, it is usually not necessary to take the whole strategy of the opponent into account. For example, a sender only has to do two things if she wants to conform to strong optimality (when this is possible): firstly, given a state t, she needs to check her production preferences to compute $\mathrm{OT}_{\mathrm{syn}}(t) \subseteq M$; secondly, she has to check whether some message in $\mathrm{OT}_{\mathrm{syn}}(t)$ would also be interpreted as t given the receiver's interpretative preferences. It becomes clear thus that strong optimization merely implements a simple *associative feedback-loop*, but not full perspective-taking in the standard game-theoretic sense. In other words, under this interpretation strong optimality is mere SELF-MONITORING to check for *association congruence* between production and comprehension.

4.3 Weak Optimality as Iterated Self-Monitoring

This idea of monitoring production by self-interpretation and monitoring interpretation by self-production also carries over to an interpretation of weak optimality. Remember that the BIOT-algorithm repeatedly checks for strong optimality in reduced OT-systems where optimal and blocked form-meaning pairs are removed in every step. This process can be mirrored by defining more sophisticated player types of level $n > 0$ whose behavior corresponds to the n-th round of computation of the BIOT-algorithm.

For this purpose, let AC_0 be the set of *level*-0 *association congruent from-meaning pairs*: $\mathrm{AC}_0 = \textsc{biot}_{\mathrm{str}}$. Let us then define level-$(n+1)$ players as playing in conformity with level-n association congruence where possible:[11]

$$S_{n+1}(t) = \begin{cases} \mathrm{AC}_n(t) & \text{if } \mathrm{AC}_n(t) \neq \emptyset \\ \arg\min_{m \in M \setminus \mathrm{AC}_n(T)} \mathrm{C}(t, m) & \text{otherwise} \end{cases}$$

$$R_{n+1}(m) = \begin{cases} \mathrm{AC}_n(m) & \text{if } \mathrm{AC}_n(m) \neq \emptyset \\ \arg\max_{t \in T \setminus \mathrm{AC}_n(M)} \mathrm{Ass}(m, t) & \text{otherwise.} \end{cases}$$

To complete the inductive construction, we also need to define *level*-$(n + 1)$ *association congruence* as: $\mathrm{AC}_{n+1} = S_n \cap R_n$. This construction, call it *iterated self-monitoring*, quite obviously replicates exactly the workings of the BIOT-algorithm.

Iterated self-monitoring is not only a rephrasing of the BIOT-algorithm, but actually helps interpreting weak optimality. For we can now ask and answer the question which assumptions about the psychology of agents give rise to the above behavior of sophisticated players. The obvious answer is that agents perform self-monitoring iteratively, but only when necessary, and believe that their opponents do too. More concretely, the behavior of a sophisticated level-$(n + 1)$ sender follows from two simple assumptions:

(i) the player performs self-monitoring based on the behavior of level-n players and plays accordingly when this gives a result;

(ii) where this gives no result, the player plays rationally given the *partial*

[11] I write $\mathrm{AC}_n(T)$ as the set of all m for which there is some t such that $\langle m, t\rangle \in \mathrm{AC}_n$, and similarly for $\mathrm{AC}_n(M)$.

belief that the opponent adheres to (i).

Let us first validate that these two assumptions indeed give rise to the behavior of sophisticated players as defined above and reflect on the conceptual implications afterwards.

Take, for instance, a sender of level $(n + 1)$ who wishes to express the state t. (The argument for the receiver is parallel.) Firstly, S_{n+1} would perform self-monitoring based on level-n behavior and thus compute level-n association congruence. If some message satisfies level-n association congruence for t, any message with this property would be used. This way, the first assumption directly assures that $S_{n+1}(t) = \mathrm{AC}_n(t)$ whenever $\mathrm{AC}_n(t) \neq \emptyset$.

The second assumption is just a little bit more complicated. It kicks in when the sender wants to express some $t \notin \mathrm{AC}_n(M)$. In that case, S_{n+1} is required to play rationally to the belief that her opponent's behavior is characterized by the (possibly partial) strategy $R_{n+1}(m) = \mathrm{AC}_n(m)$ for all m such that $\mathrm{AC}_n(m) \neq \emptyset$.[12] Given such a partial conjecture, it would always yield an expected utility of 0 (possibly minus some nominal cost, of course) to try to express a state in $t \notin \mathrm{AC}_n(M)$ with a message $m \in \mathrm{AC}_n(T)$. But in the absence of a definite conjecture about how messages in $M \setminus \mathrm{AC}_n(T)$ are interpreted, any such message has at least a positive chance of obtaining the right interpretation, so that the expected utility of sending a message from the $M \setminus \mathrm{AC}_n(T)$ in t will be strictly bigger than zero, and, in fact, equal for all messages in this set. Consequently, a rational level-$(n + 1)$ sender will choose any cost-minimal message in $M \setminus \mathrm{AC}_n(T)$ in each state $t \notin \mathrm{AC}_0(M)$. It turns out that the second assumption effectively gives, via partiality of belief in self-monitoring, a rationalistic explanation of the blocking mechanism of the biot-algorithm.

4.4 Relflection on Iterated Self-Monitoring

Taken together, this suggests that we should think of weak optimality as a process of self-monitoring to the maximal depth necessary to express or interpret a form. Since $\mathrm{AC}_n \subseteq \mathrm{AC}_{n+1}$ for all n, there is no need to compute more sophisticated play than the minimal k for which $\mathrm{AC}_k(t) \neq \emptyset$, when expressing t, or $\mathrm{AC}_k(m) \neq \emptyset$, when interpreting m. Only when necessary, further iteration

[12] Notice that this belief may be *partial*, for it may mean that S_{n+1} has no belief about how her opponent will interpret a message m for which $\mathrm{AC}_n(m) = \emptyset$ if such messages exist.

of self-monitoring takes place, by adopting a belief that the opponent also performs such iterated self-monitoring. At each step of this procedure, however, the conjecture about opponent behavior is not the full-fledged perspective taking that is standard in game theory, but only an associative feedback and the assumption that the opponent also performs such self-monitoring.

Interestingly enough, Bayesian rationality features in this interpretation of optimality only as an explication of preference maximization *in the absence* of a conjecture about opponent behavior. In other words, unlike in the structurally similar IBR models of, for instance, Jäger and Ebert (2009) and Franke (2009), the more sophisticated types do not rely on deeper and deeper nestings of belief in rationality. The sophisticated types that match the BIOT-algorithm only require ever more nested beliefs in self-monitoring. This is in a sense a weaker requirement, but it may nonetheless explain why weak optimality is often too strong a theoretical prediction to be borne out in reality (cf. the arguments by Beaver and Lee 2004): that agents can coordinate successfully on weakly optimal communication behavior becomes dubious proportional to the number of iteration steps in self-monitoring, due to natural restrictions on cognitive resources.

However, to say that nested belief in self-monitoring, as found in BIOT under the interpretation favored here, is weaker than nested belief in rationality under full-fledged perspective taking, as found in recent IBR models, is not necessarily an argument *for* BIOT and against IBR. In order to be an argument for BIOT we would have to motivate why exactly this kind of self-monitoring should occur in pragmatic language use. It is fairly standard to assume monitoring by internal self-interpretation (cf. Levelt 1989), but this is not necessarily so for comprehension. This points favorably into the direction of an asymmetric approach to BIOT, as advanced by (see Zeevat 2000).

Finally, it is also not implausible to accept simple self-monitoring as a reasonable mental operation, be it in production alone or also in interpretation, yet to reject nested beliefs in self-monitoring opponents as a natural cognitive process. This would corroborate the position of, for instance, Blutner and Zeevat (2008) that only strong optimality is reasonable as an online mechanism, while weak optimality is not.

5 Conclusion

To take stock, I have argued that it useful and desirable to match optimality theory with game theory in order to supply a characterization of OT's basic notions in terms of agents' mental states and behavioral disposition. I have tried to show that the analogy between OT-systems and strategic optimality games suggested by Dekker and van Rooij (2000) is conceptually flawed, and does not achieve this end. Therefore, I have suggested to work out a connection between signaling games and OT-systems, and between optimality notions and different kinds of more or less sophisticated player types. From this point of view, unidirectional optimality is Bayesian rationality that takes into account only preferences on form-meaning associations in the absence of a conjecture about opponent behavior. Strong optimality turned out to be best described as a simple self-monitoring feedback process, not as full strategic perspective taking. Weak optimality then presents itself as an iterated process of such self-monitoring which is defined in terms of beliefs in self-monitoring and rational responses to these (possibly partial) beliefs.

References

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.

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 and Henk Zeevat, eds. (2004). *Optimality Theory and Pragmatics*. Palgrave MacMillan.

— (2008). "Optimality-Theoretic Pragmatics". To appear in: Claudia Maienborn, Klaus von Heusinger and Paul Portner (eds.) *Semantics: An International Handbook of Natural Language Meaning*.

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

Flobbe, Liesbeth et al. (2008). "Children's Application of Theory of Mind in Reasoning and Language". In: *Journal of Logic, Language and Information* 17. Pp. 417–442.

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

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

Heap, Shaun P. Hargreaves and Yanis Varoufakis (2004). *Game Theory — A Critical Text (Second Edition)*. Routledge.

Hendriks, Petra and Helen de Hoop (2001). “Optimality Theoretic Semantics”. In: *Linguistics and Philosophy* 24. Pp. 1–32.

Hendriks, Petra and Jennifer Spenader (2005). “When Production Precedes Comprehension: An Optimization Approach to the Acquisition of Pronouns”. In: *Language Acquisition* 13.4. Pp. 319–348.

Hendriks, Petra et al. (2007). “Conflicts in Interpretation”. Unpublished book manuscript, Groningen, Nijmegen, Utrecht.

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). “Game Theory in Semantics and Pragmatics”. Unpublished manuscript, University of Bielefeld.

Jäger, Gerhard and Christian Ebert (2009). “Pragmatic Rationalizability”. In: *Proceedings of Sinn und Bedeutung 13*. Ed. by Arndt Riester and Torgrim Solstad. Pp. 1–15.

Levelt, Willem J. M. (1989). *Speaking: From Intention to Articulation*. MIT Press.

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

Osborne, Martin J. (2004). *An Introduction to Game Theory*. New York: Oxford University Press.

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

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*. San Francisco: Morgan Kaufmann Publishers Inc. Pp. 85–94.

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

Premack, David and Guy Woodruff (1978). “Does the Chimpanzee have a Theory of Mind”. In: *Behavioral and Brain Sciences* 1.4. Pp. 515–526.

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

van Rooij, Robert (2004). "Signalling Games Select Horn-Strategies". In: *Linguistics and Philosophy* 27. Pp. 493–527.

Schelling, Thomas C. (1960). *The Strategy of Conflict*. Cambridge, Massachusetts: Harvard University Press.

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

Zahavi, Amotz (1975). "Mate Selection — A Selection for a Handicap". In: *Journal of Theoretical Biology* 53. Pp. 205–214.

Zeevat, Henk (2000). "The Asymmetry of Optimality Theoretic Syntax and Semantics". In: *Journal of Semantics* 17.3. Pp. 243–262.