
Abstract

This thesis offers a general game theoretic model of language use and interpretation and applies it to linguistic pragmatics in the vein of Grice (1989). The model presented here —called the IBR model— explains pragmatic phenomena, such as conversational implicatures, as arising from a sequence of iterated best responses: starting from the literal, semantic meaning as a psychologically salient attractor of attention, speaker and hearer initially compute the rational best responses to a literal use or interpretation of expressions; subsequently, agents continue computing best responses to best responses, for as long as this is reasonable and their cognitive resources permit.

This algorithmic solution procedure is simple and intuitively appealing. But more importantly, it has a clear epistemic interpretation as modelling so-called “level- k thinking,” which has gained recent popularity in behavioral game theory (Stahl and Wilson 1995; Ho et al. 1998; Camerer et al. 2004). Laboratory data supports the assumption that human reasoners are cognitively biased and possibly resource-bounded in the sense that they are susceptible to focal framing effects and perform theory of mind reasoning possibly only to a given depth k . Thus conceived, the IBR model formally implements a number of empirically attested assumptions about the cognitive architecture of human reasoners. The IBR model then effectively provides a novel non-equilibrium solution concept as a form of strong rationalizability (Battigalli 2006) in which these psychological assumptions have been implemented. The thesis aims to show how this turn towards psychological realism solves outstanding conceptual problems with game theoretic approaches to communication, and, moreover, improves on predictions in linguistic applications.

Firstly, by implementing semantic meaning as a focal attractor of atten-

tion, the IBR model singles out those strategies that conform to our intuitions about credible communication without altogether precluding the possibility, and even occasional optimality, of lying, misleading and distrust (see Farrell and Rabin 1996; Stalnaker 2006). Secondly, the model explicitly represents agents with absent or only limited capacity of taking opponent behavior and reasoning into account. This sheds light on higher-order theory of mind reasoning in language use and especially in the pattern of acquisition of pragmatic competence by young children (see Noveck 2001; Papafragou and Musolino 2003). An in-depth comparison of the IBR model with bidirectional optimality theory (Blutner 2000) suggests that the former is the better tool for modelling limitations of theory of mind reasoning in interpretation and acquisition. Finally, the IBR model unifies and extends a series of recent work in game theoretic pragmatics (see especially Benz 2006; Stalnaker 2006; Benz and van Rooij 2007; Jäger 2007). It yields formidable predictions for, among other phenomena, complex and nested cases of scalar implicatures, generalized M-implicatures and free-choice readings. The model also backs up natural accounts of conditional perfection, and unconditional readings of conditionals.

Contents

Contents · vii

Acknowledgments · xi

Overview · xv

1 · What is Game Theoretic Pragmatics? · 1

1.1 · Gricean Pragmatics · 2

1.1.1 · The Gricean Programme · 4

1.1.2 · Conversational Implicatures · 6

1.1.3 · Brands of Griceanism · 9

1.2 · Game Theory for Gricean Pragmatics · 13

1.2.1 · Static Games & Their Solutions · 13

1.2.2 · Signaling Games in Gricean Pragmatics · 17

1.2.3 · Solving Signaling Games · 24

1.2.4 · Implementing Semantic Meaning · 34

2 · The Iterated Best Response Model · 43

2.1 · Focal Points & Iterated Best Response · 44

2.1.1 · Semantic Meaning as a Focal Point · 46

2.1.2 · Iterated Best Response Reasoning as Pragmatic Inference · 48

2.1.3 · Strategic-Type Models · 52

2.2 · The Vanilla Model · 53

2.2.1 · Strategic Types and the IBR Sequence · 54

2.2.2 · Examples: Scalar & M-Implicatures · 59

2.2.3 · Consistency: Naïve & Sophisticated Updates · 66	66
2.2.4 · Truth Ceteris Paribus & Skewed Priors · 71	71
2.3 · Forward Induction · 76	76
2.3.1 · Trouble-Maker “Some But Not All” · 77	77
2.3.2 · Forward Induction and Strong Belief in Rationality · 78	78
2.3.3 · Restrictions on Counterfactual Beliefs · 81	81
2.4 · Overview and Comparison · 89	89
2.4.1 · Versions of the IBR Model · 89	89
2.4.2 · Related Models · 91	91
2.4.3 · IBR vs. Rationalizability · 96	96
2.5 · Semantic Meaning and Credibility · 107	107
2.5.1 · Message Credibility · 108	108
2.5.2 · Credibility-Based Refinements · 113	113
3 · Games and Pragmatic Interpretation · 123	123
3.1 · Game Models Revisited · 124	124
3.1.1 · Interpretation Games · 124	124
3.1.2 · Construction of Interpretation Games · 126	126
3.1.3 · Examples: Multiple Scalar Items · 134	134
3.2 · Epistemic Lifting of Signaling Games · 138	138
3.2.1 · The Epistemic Status of Scalar Implicatures · 138	138
3.2.2 · The Implicatures of Plain Disjunctions · 140	140
3.2.3 · Lifted Signaling Games · 141	141
3.2.4 · Examples · 145	145
3.3 · Free Choice Inferences · 156	156
3.3.1 · Free Choice from Anti-Exhaustivity · 158	158
3.3.2 · Anti-Exhaustivity from Iteration · 161	161
3.3.3 · Simplification of Disjunctive Antecedents · 169	169
3.4 · Games at the Semantics-Pragmatics Interface · 175	175
4 · Perspective, Optimality & Acquisition · 181	181
4.1 · Optimality Theory in Pragmatics · 182	182
4.1.1 · OT-Systems · 182	182
4.1.2 · Uni- and Bidirectional Optimality · 184	184
4.1.3 · Example: M-Implicatures in BiOT · 187	187
4.1.4 · BiOT as a Model of Pragmatic Interpretation · 189	189

4.2 · BiOT and Game Theory · 190
4.2.1 · BiOT and Strategic Games · 191
4.2.2 · Critique · 192
4.2.3 · BiOT and Signaling Games · 194
4.3 · An Epistemic Interpretation of Optimality · 196
4.3.1 · Comprehension Lags in Language Acquisition · 196
4.3.2 · Unidirectional Optimality · 199
4.3.3 · Bidirectional Optimality · 203
4.4 · Scalar Implicatures in Language Acquisition · 213
4.4.1 · Overview of Some Recent Studies · 213
4.4.2 · Tolerance vs. Conceptualization · 220
5 · The Pragmatics of Conditionals · 225
5.1 · Meaning and Use of Conditionals · 226
5.1.1 · Semantics for Conditionals · 227
5.1.2 · Kinds of Conditionals · 231
5.2 · Conditional Perfection · 234
5.2.1 · Approaching Perfection · 236
5.2.2 · Two Sources of Perfection · 239
5.2.3 · Rationalizing Indirectness · 249
5.2.4 · Forward Induction under Awareness Dynamics · 252
5.3 · Unconditional Readings · 257
5.3.1 · Biscuit Conditionals · 258
5.3.2 · Unconditionality Beyond Biscuits · 262
5.3.3 · Conditional Independence · 265
5.3.4 · Biscuits in Discourse · 274
5.3.5 · Projection and a Big Fat Lie · 277
6 · Conclusions & Outlook · 279
References · 281
Notation, Symbols & Abbreviations · 299
Samenvatting · 301
Abstract · 303

Overview

This thesis is interdisciplinary in nature. Its main contribution is an application of game theory to linguistic pragmatics. Since perhaps not many people will be familiar with both subjects at once, the need arises to introduce the basics of both fields. Although admittedly the thesis spends more effort on explaining the relevant concepts of game theory to the linguist than on explaining the relevant concepts of pragmatics to the game theorist, I would sincerely hope that the text is accessible, at least in its gist, to anybody proficient in some adjacent academic field who is interested in the topic. Be that as it may, it would certainly be forlorn optimism to expect that all of my possible readers are equally interested in all issues addressed here. I would therefore like to give a brief overview of the content of this thesis, together with an indication which parts belong to either the *linguist's track* or the *game theorist's track*. The linguist's track contains all linguistic applications and only the absolutely necessary information on game theory. The game theorist's track, on the other hand, contains the game theoretic details and only the absolutely necessary information on linguistic pragmatics.

There are five main chapters. Chapter 1 introduces the basics of both Gricean pragmatics and game theory. Chapter 2 spells out the central iterated best response (IBR) model of pragmatic reasoning. Chapter 3 is dedicated to linguistic applications of the IBR model. Chapter 4 compares the IBR model to bidirectional optimality theory and discusses data from language acquisition. Finally, chapter 5 is mainly linguistic and deals with use and interpretation of conditionals. (A more thorough abstract of the thesis can be found at the end, on page 303.)

The linguist's track obligatorily contains sections 1.2, where basic concepts

of game theory are introduced, as well as sections 2.1 and 2.2 to understand the basic IBR model. Reading section 2.3, which discusses a refinement of the IBR model, is also recommended. After that any part or portion of chapters 3, 4 and 5 that seems relevant to the reader's concern should be intelligible. In fact, chapter 5 is nearly independent of the game theoretic framework (with the exception of sections 5.2.4 and 5.3.4).

The game theorist's track obligatorily contains section 1.1, which introduces the basic ideas of Gricean pragmatics, as well the whole of chapter 2. Here especially sections 2.4 and 2.5 are relevant, which compare the IBR model to related game theoretic approaches. The game theorist might furthermore take interest in section 3.1 where I discuss my preferred interpretation of signaling games in a linguistic context. Finally, a cursory glance at some of the applications in sections 3.2 and 3.3 will help understand better the linguistic motivation behind the present approach.