Game theoretic pragmatics

Gerhard Jäger
gerhard.jaeger@uni-tuebingen.de

(based on joint work with Michael Franke)

February 18, 2011

MIT

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Overview

Structure of the talk

- Signaling games
- Literal meaning and rationality
- Justification of pragmatic decisions
- Examples
- Conclusion

Signaling games

- sequential game:
  1. **nature** chooses a world w
     - out of a pool of possible worlds W
     - according to a certain probability distribution p^*
  2. nature shows w to sender **S**
  3. S chooses a message m out of a set of possible signals M
  4. S transmits m to the receiver **R**
  5. R chooses an action a, based on the sent message.
- Both S and R have preferences regarding R's action, depending on w.
- S might also have preferences regarding the choice of m (to minimize signaling costs).

Tea or coffee?

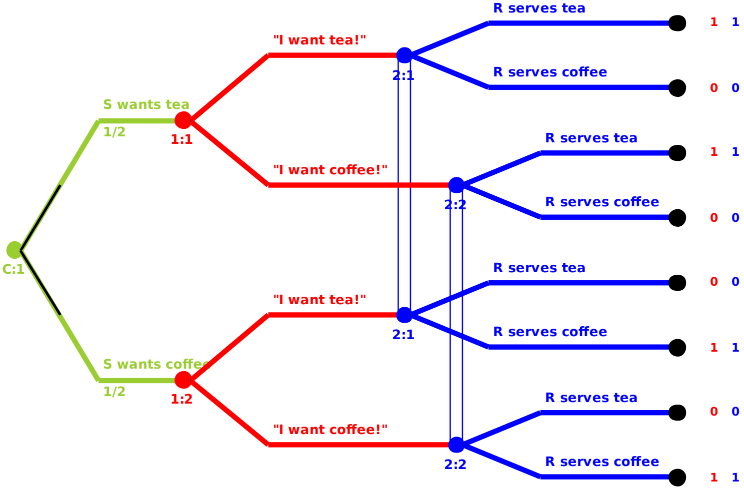
An example

- Sally either prefers tea (w_1) or coffee (w_2), with $p^*(w_1) = p^*(w_2) = \frac{1}{2}$.
- Robin either serves tea (a_1) or coffee (a_2).
- Sally can send either of two messages:
  - m_1: *I prefer tea.*
  - m_2: *I prefer coffee.*
- Both messages are costless.

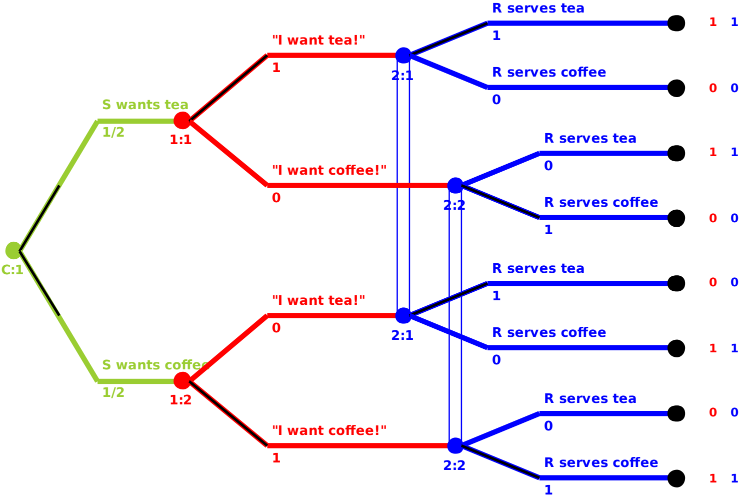
	a_1	a_2
w_1	$1, 1$	$0, 0$
w_2	$0, 0$	$1, 1$

Table: utility matrix

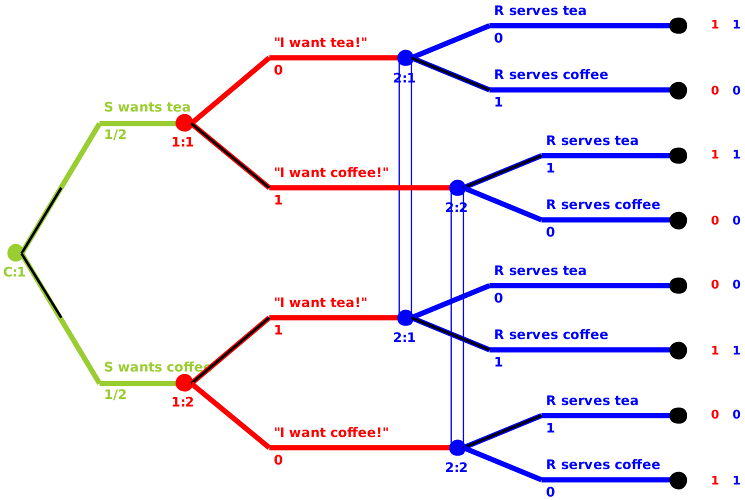
Extensive form



Extensive form



Extensive form



A coordination problem

- two strict Nash equilibria
  - S always says the truth and R always believes her.
  - S always says the opposite of the truth and R interprets everything ironically.
- Both equilibria are equally rational.
- Still, first equilibrium is more reasonable because it employs exogenous meanings of messages for equilibrium selection.
- Criterion for equilibrium selection:

> *Always say the truth, and always believe what you are told!*

- What happens if it is not always rational to be honest/credulous?

Partially aligned interests

Rabin's (1990) example

- In w_1 and w_2, S and R have identical interests.[a]
- In w_3, S would prefer R to believe in w_2.
- The propositions $\{w_1\}$ and $\{w_2, w_3\}$ are *credible*.
- The propositions $\{w_2\}$ and $\{w_3\}$ are *not credible*.

[a] Unless mentioned otherwise, I always assume a uniform distribution p^* over W.

	a_1	a_2	a_3
w_1	$10, 10$	$0, 0$	$0, 0$
w_2	$0, 0$	$10, 10$	$5, 7$
w_3	$0, 0$	$10, 0$	$5, 7$

Table: Partially aligned interests

Partially aligned interests

Rabin's (1990) example

- Suppose there are three messages:
  - m_1: We are in w_1.
  - m_2: We are in w_2.
  - m_3: We are in w_3.
- reasonable S will send m_1 if and only if w_1
- reasonable R will react to m_1 with a_1
- nothing else can be inferred

	a_1	a_2	a_3
w_1	$10, 10$	$0, 0$	$0, 0$
w_2	$0, 0$	$10, 10$	$5, 7$
w_3	$0, 0$	$10, 0$	$5, 7$

Table: Partially aligned interests

Revised maxim

Always say the truth,
and always believe what you are told,
unless you have reasons to do otherwise!

But what does this mean?

Justification and best responses

Jusification of decisions

- decisions must be **justifiable**
- two kinds of justification:

  I use/interpret a message the way I do because:

  - *this is what the literal meaning of the message dictates, or*
  - *because this is the best I can do, given my justifiable belief about the decisions of the other player.*

Justification and best responses

Sally's belief states

- (first order) belief of the sender:
  - function ρ from messages to probability distribution of actions
  - $\rho(a|m)$: S's subjective probability that R performs action a if S sends message m

Justification and best responses

Robin's belief states

- (first order) belief of the receiver has two components:
  - function σ from worlds to probability distributions over messages
  - $\sigma(m|w)$: R's subjective probability that S sends messages m if she is in world w
  - function σ^* from messages to probability distributions over worlds
  - $\sigma^*(w|m)$: R's posterior probability that w is the case after observing message m
  - σ and σ^* are connected via **Bayes' Rule**

$$\sigma^*(w|m) = \frac{\sigma(m|w)p^*(w)}{\sum_{w'} \sigma(m|w')p^*(w')}$$

Justification: Rabin's example again

Utility matrix

	a_1	a_2	a_3
w_1	$10, 10$	$0, 0$	$0, 0$
w_2	$0, 0$	$10, 10$	$5, 7$
w_3	$0, 0$	$10, 0$	$5, 7$

literal meanings

$$\|m_1\| = \{w_1\}$$
$$\|m_2\| = \{w_2\}$$
$$\|m_3\| = \{w_3\}$$

Justification: Rabin's example again

The Honest Sender

- Suppose (Robin supposes that) Sally is simply honest, and non-deliberating.
- This means she sends a true message in each world.
- No further criteria about message selection are known.
- Best model σ_0 of such a sender is the one that is consistent with the assumptions and **maximizes entropy**.
- This means that in each world w, $\sigma_0(\cdot|w)$ is a uniform distribution:

$$\sigma_0(m|w) = \begin{cases} \frac{1}{|\{m|w \models m\}|} & \text{if } w \models m, \\ 0 & \text{else.} \end{cases}$$

Justification: Rabin's example again

Can be represented in a matrix with worlds as rows and messages as columns.

σ_0	m_1	m_2	m_3
w_1	1	0	0
w_2	0	1	0
w_3	0	0	1

This is justifiable by the literal meaning of the messages.

Justification: Rabin's example again

Bayesian reasoning

- Robin assumes σ_0.
- He actually needs σ_0^* (posterior probablities of worlds given messages)
- $\sigma_0^*(\cdot|m)$:
  - take column m in σ_0
  - multiply each entry with corresponding value of p^*
  - normalize the column: divide it by its sum, such that it becomes a probability distribution
  - if m is a zero-column (i.e. m is a surprise message): assume uniform distribution over $\|m\|$
  - result is the row in σ^*

σ_0^*	w_1	w_2	w_3
m_1	1	0	0
m_2	0	1	0
m_3	0	0	1

Justification: Rabin's example again

Expected utility

- Robin's utility depends on w and a
- σ_0^* gives him a probability distribution over W
- this enables him to assess the exected utility for each action, conditional on each message:

$$EU(a|m) = \sum_w \sigma_0^*(w|m) u_r(w, a)$$

- comes down to matrix multiplication:

$$EU_r = \sigma_0^* \cdot u_r$$

Justification: Rabin's example again

$$
\begin{aligned}
EU_r &= \sigma_0^* \cdot u_r \\
&= \begin{pmatrix} \sigma_0^* & w_1 & w_2 & w_3 \\ \hline\hline m_1 & 1 & 0 & 0 \\ m_2 & 0 & 1 & 0 \\ m_3 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} & a_1 & a_2 & a_3 \\ \hline\hline w_1 & 10 & 0 & 0 \\ w_2 & 0 & 10 & 7 \\ w_3 & 0 & 0 & 7 \end{pmatrix} \\
&= \begin{pmatrix} & a_1 & a_2 & a_3 \\ \hline\hline m_1 & 10 & 0 & 0 \\ m_2 & 0 & 10 & 7 \\ m_3 & 0 & 0 & 7 \end{pmatrix}
\end{aligned}
$$

Justification: Rabin's example again

Best response

- If Robin is rational he will maximize his expected utility after each message
- If he believs in σ_0, he will — for each message m — pick an action that is maximal within m's row in EU_r
- If Sally assumes that this is how Robin thinks, and if she has no further information, her model ρ_0 of Robin's behavior will the conditional probability distribution that is consistent with these assumptions and maximizes entropy
- Hence: ρ_0 puts equal probability on each action that maximizes EU_r, and 0 probability elsewhere
- ρ_0 is the **best response** to σ_0:

$$\rho_0 = BR_r(\sigma_0)$$

Justification: Rabin's example again

Expected utility

	a_1	a_2	a_3
m_1	10	0	0
m_2	0	10	7
m_3	0	0	7

Best response

ρ_0	a_1	a_2	a_3
m_1	1	0	0
m_2	0	1	0
m_3	0	0	1

Justification: Rabin's example again

Iterated Best Response

- Suppose Sally believes that Robin plays according to ρ_0.
- Then Sally (if she is rational) will play the best response $\sigma_1 = BR_s(\rho_0)$ to ρ_0.
- Calculation goes as follows:
  - Sally's Expected Utility:

$$EU_s(m|w) = \sum_a \rho_0(a|m)u_s(w,a)$$

  - boils down to matrix multiplication:

$$EU_s = u_s \cdot \rho_0^T$$

  - σ_1 places row-wise equal probability on each row-maximal value in EU_s, and 0 elsewhere

Justification: Rabin's example again

$$EU_s = u_s \cdot \rho_0^T$$

$$= \left(\begin{array}{c|ccc} & a_1 & a_2 & a_3 \\ \hline w_1 & 10 & 0 & 0 \\ w_2 & 0 & 10 & 5 \\ w_3 & 0 & 10 & 5 \end{array}\right) \cdot \left(\begin{array}{c|ccc} \rho_0 & a_1 & a_2 & a_3 \\ \hline m_1 & 1 & 0 & 0 \\ m_2 & 0 & 1 & 0 \\ m_3 & 0 & 0 & 1 \end{array}\right)^T$$

$$= \left(\begin{array}{c|ccc} & m_1 & m_2 & m_3 \\ \hline w_1 & 10 & 0 & 0 \\ w_2 & 0 & 10 & 5 \\ w_3 & 0 & 10 & 5 \end{array}\right)$$

$$BR(\rho_0) = \left(\begin{array}{c|ccc} \sigma_1 & m_1 & m_2 & m_3 \\ \hline w_1 & 1 & 0 & 0 \\ w_2 & 0 & 1 & 0 \\ w_3 & 0 & 1 & 0 \end{array}\right)$$

Justification: Rabin's example again

Iterated Best Response

- If Robin anticipates this, he will play according to $\rho_1 = BR_r(\sigma_1)$.
- Computation is exactly as for ρ_0, but using σ_1 instead of σ_0.
- As the third column of σ_1 only contains 0s, the principle

  TRUTH CETERIS PARIBUS

  applies: assume a uniform distribution over $\|m_3\|$ in σ_1^*.

$\sigma_1^* =$

	w_1	w_2	w_3
m_1	1	0	0
m_2	0	$\frac{1}{2}$	$\frac{1}{2}$
m_3	0	0	1

Justification: Rabin's example again

$$
\begin{aligned}
EU_r &= \sigma_1^* \cdot u_r \\
&= \left(\begin{array}{c|ccc} \sigma_1^* & w_1 & w_2 & w_3 \\ \hline\hline m_1 & 1 & 0 & 0 \\ m_2 & 0 & \frac{1}{2} & \frac{1}{2} \\ m_3 & 0 & 0 & 1 \end{array}\right) \cdot \left(\begin{array}{cccc} & a_1 & a_2 & a_3 \\ \hline\hline w_1 & 10 & 0 & 0 \\ w_2 & 0 & 10 & 7 \\ w_3 & 0 & 0 & 7 \end{array}\right) \\
&= \left(\begin{array}{cccc} & a_1 & a_2 & a_3 \\ \hline\hline m_1 & 10 & 0 & 0 \\ m_2 & 0 & 5 & 7 \\ m_3 & 0 & 0 & 7 \end{array}\right) \\
BR_r(\sigma_1) &= \left(\begin{array}{cccc} \rho_1 & a_1 & a_2 & a_3 \\ \hline\hline m_1 & 1 & 0 & 0 \\ m_2 & 0 & 0 & 1 \\ m_3 & 0 & 0 & 1 \end{array}\right)
\end{aligned}
$$

Justification: Rabin's example again

Iterated Best Response

- This procedure can be iterated indefinitely.
- General pattern:

$$
\begin{aligned}
\rho_n &= BR_r(\sigma_n) \\
\sigma_{n+1} &= BR_s(\rho_n)
\end{aligned}
$$

Justification: Rabin's example again

$$
\begin{aligned}
EU_s &= u_s \cdot \rho_1^T \\
&= \left(\begin{array}{c|ccc} & a_1 & a_2 & a_3 \\ \hline\hline w_1 & 10 & 0 & 0 \\ w_2 & 0 & 10 & 5 \\ w_3 & 0 & 10 & 5 \end{array}\right) \cdot \left(\begin{array}{c|ccc} \rho_1 & a_1 & a_2 & a_3 \\ \hline\hline m_1 & 1 & 0 & 0 \\ m_2 & 0 & 0 & 1 \\ m_3 & 0 & 0 & 1 \end{array}\right)^T \\
&= \left(\begin{array}{c|ccc} & m_1 & m_2 & m_3 \\ \hline\hline w_1 & 10 & 0 & 0 \\ w_2 & 0 & 5 & 5 \\ w_3 & 0 & 5 & 5 \end{array}\right) \\
BR(\rho_1) &= \left(\begin{array}{c|ccc} \sigma_2 & m_1 & m_2 & m_3 \\ \hline\hline w_1 & 1 & 0 & 0 \\ w_2 & 0 & \frac{1}{2} & \frac{1}{2} \\ w_3 & 0 & \frac{1}{2} & \frac{1}{2} \end{array}\right)
\end{aligned}
$$

Justification: Rabin's example again

$$EU_r = \sigma_2^* \cdot u_r$$

$$= \begin{pmatrix} \sigma_2^* & w_1 & w_2 & w_3 \\ \hline\hline m_1 & 1 & 0 & 0 \\ m_2 & 0 & \frac{1}{2} & \frac{1}{2} \\ m_3 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} & a_1 & a_2 & a_3 \\ \hline\hline w_1 & 10 & 0 & 0 \\ w_2 & 0 & 10 & 7 \\ w_3 & 0 & 0 & 7 \end{pmatrix}$$

$$= \begin{pmatrix} & a_1 & a_2 & a_3 \\ \hline\hline m_1 & 10 & 0 & 0 \\ m_2 & 0 & 5 & 7 \\ m_3 & 0 & 5 & 7 \end{pmatrix}$$

$$BR_r(\sigma_2) = \begin{pmatrix} \rho_2 & a_1 & a_2 & a_3 \\ \hline\hline m_1 & 1 & 0 & 0 \\ m_2 & 0 & 0 & 1 \\ m_3 & 0 & 0 & 1 \end{pmatrix}$$

Justification: Rabin's example again

Fixed point of IBR sequence

- $\rho_2 = \rho_1$
- Hence for all $n > 2$:

$$\begin{aligned} \sigma_n &= \sigma_2 \\ \rho_n &= \rho_2 \end{aligned}$$

- (σ_2, ρ_1) is a fixed point for best response calculation
- If Sally and Robin only consider justifiable strategies and are both sufficiently sophisticated — and these facts are common knowledge — they will play according to these fixed point strategies

IBR sequence for Rabin's example

σ_0	m_1	m_2	m_3
w_1	1	0	0
w_2	0	1	0
w_3	0	0	1

ρ_0	a_1	a_2	a_3
m_1	1	0	0
m_2	0	1	0
m_3	0	0	1

σ_1	m_1	m_2	m_3
w_1	1	0	0
w_2	0	1	0
w_3	0	1	0

ρ_2	a_1	a_2	a_3
m_1	1	0	0
m_2	0	0	1
m_3	0	0	1

σ_2	m_1	m_2	m_3
w_1	1	0	0
w_2	0	$\frac{1}{2}$	$\frac{1}{2}$
w_3	0	$\frac{1}{2}$	$\frac{1}{2}$

ρ_1	a_1	a_2	a_3
m_1	1	0	0
m_2	0	0	1
m_3	0	0	1

$$F = (\sigma_2, \rho_1)$$

Interpretation games

- How does this relate to linguistic examples?
- There is a quasi-algorithmic procedure (due to Franke 2009) how to construct a game from an example sentence.

What is given?

- example sentence
- set of expression alternatives
- jointly form set of messages
- question under discussion QUD
- set of complete answers to QUD is the set of possible worlds

What do we need?

- interpretation function $\| \cdot \|$
- prior probability distribution p^*
- set of actions
- utility functions

Interpretation games

QUD

- often QUD is not given explicitly
- procedure to construct QUD from expression m and its alternatives $ALT(m)$:
  - Let ct be the context of utterances, i.e. the maximal set of statements that is common knowledge between Sally and Robin.
  - any subset w of $ALT(m) \cup \{\neg m' | m' \in ALT(m)\}$ is a possible world iff
    - w and ct are consistent, i.e. $w \cup ct \not\vdash \bot$
    - for any set $X : w \subset X \subseteq ALT(m) \cup \{\neg m' | m' \in ALT(m)\}$, $ct \cup X$ is inconsistent

Interpretation games

Game construction

- interpretation function:

$$\|m'\| = \{w | w \vdash m\}$$

- p^* is uniform distribution over W
- justified by principle of insufficient reason
- set of actions is W
- intuitive idea: Robin's task is to figure out which world Sally is in
- utility functions:

$$u_{s/r}(w, a) = \begin{cases} 1 & \text{iff } w = a \\ 0 & \text{else} \end{cases}$$

- both players want Robin to succeed

Example: Quantity implicatures

(1)
a. Who came to the party?
b. SOME: Some boys came to the party.
c. NO: No boys came to the party.
d. ALL: All boys came to the party.

Game construction

- $ct = \emptyset$
- $W = \{w_{\neg\exists}, w_{\exists\neg\forall}, w_{\forall}\}$
- $w_{\neg\exists} = \{\text{NO}\}, w_{\exists\neg\forall} = \{\text{SOME}\}, w_{\forall} = \{\text{SOME}, \text{ALL}\}$
- $p^* = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$

- interpretation function:

$$\begin{aligned} \|\text{SOME}\| &= \{w_{\exists\neg\forall}, w_{\forall}\} \\ \|\text{NO}\| &= \{w_{\neg\exists}\} \\ \|\text{ALL}\| &= \{w_{\forall}\} \end{aligned}$$

- utilities:

	$a_{\neg\exists}$	$a_{\exists\neg\forall}$	$a_{\forall}$
$w_{\neg\exists}$	$1,1$	$0,0$	$0,0$
$w_{\exists\neg\forall}$	$0,0$	$1,1$	$0,0$
$w_{\forall}$	$0,0$	$0,0$	$1,1$

Interpretation games

- utility functions are identity matrices
- therefore the step *multiply with utility matrix* can be omitted in best response computation
- also, restriction to uniform priors makes simplifies computation of posterior distribution
- simplified IBR computation:

Interpretation games

Sally

1. flip ρ along diagonal
2. place a 0 in each cell that is non-maximal within its row
3. normalize each row

Robin

1. flip σ along diagonal
2. if a row contains only 0s, fill in a 1 in each cell corresponding to a true world-message association
3. place a 0 in each cell that is non-maximal within its row
4. normalize each row

Example: Quantity implicatures

σ_0	NO	SOME	ALL
$w_{\neg\exists}$	1	0	0
$w_{\exists\neg\forall}$	0	1	0
$w_{\forall}$	0	$\frac{1}{2}$	$\frac{1}{2}$

ρ_0	$w_{\neg\exists}$	$w_{\exists\neg\forall}$	$w_{\forall}$
NO	1	0	0
SOME	0	1	0
ALL	0	0	1

σ_1	NO	SOME	ALL
$w_{\neg\exists}$	1	0	0
$w_{\exists\neg\forall}$	0	1	0
$w_{\forall}$	0	0	1

ρ_1	$w_{\neg\exists}$	$w_{\exists\neg\forall}$	$w_{\forall}$
NO	1	0	0
SOME	0	1	0
ALL	0	0	1

$$F = (\rho_0, \sigma_1)$$

In the fixed point, SOME is interpreted as entailing $\neg$ALL, i.e. exhaustively.

Lifted games

- So far, it is hard-wired in the model that Sally has complete knowledge (or, rather, complete belief — whether or not she is right is inessential for IBR) about the world she is in.
- corresponds to strong version of **competence assumption**
- Sometimes this assumption is too strong:

Lifted games

1. **a.** Ann or Bert showed up. (= OR)
   b. Ann showed up. (= A)
   c. Bert showed up. (= B)
   d. Ann and Bert showed up. (= AND)

- w_a: Only Ann showed up.
- w_b: Only Bert showed up.
- w_{ab}: Both showed up.

Utility matrix

	a_a	a_b	a_{ab}
w_a	1	0	0
w_b	0	1	0
w_{ab}	0	0	1

Lifted games

IBR sequence

σ_0	OR	A	B	AND
w_a	$\frac{1}{2}$	$\frac{1}{2}$	0	0
w_b	$\frac{1}{2}$	0	$\frac{1}{2}$	0
w_{ab}	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

ρ_0	w_a	w_b	w_{ab}
OR	$\frac{1}{2}$	$\frac{1}{2}$	0
A	1	0	0
B	0	1	0
AND	0	0	1

σ_1	OR	A	B	AND
w_a	0	1	0	0
w_b	0	0	1	0
w_{ab}	0	0	0	1

ρ_1	w_a	w_b	w_{ab}
OR	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
A	1	0	0
B	0	1	0
AND	0	0	1

OR comes out as a message that would never be used!

Lifted games

- full competence assumption is arguably too strong
- weaker assumption (Franke 2009):
  - Sally's information states are **partial answers to QUD**, ie. **sets** of possible worlds
  - Robin's task is to figure out which information state Sally is in.
  - *ceteris paribus*, Robin receives slightly higher utility for smaller (more informative) states

Costs

- Preferences that are independent from correct information transmission are captured via *cost functions* for sender and receiver.
- For the sender this might be, *inter alia*, a preference for simpler expressions.
- For the receiver, the *Strongest Meaning Hypothesis* is a good candiate.

Lifted games

Formally

- cost functions c_s, c_r: $c_s : (POW(W) - \{\emptyset\}) \times M \mapsto \mathbb{R}^+$
- costs are **nominal**:

$$0 \leq c_s(i,m), c_r(i,m) < \min(\frac{1}{|POW(W) - \emptyset|^2}, \frac{1}{|ALT(m)|^2})$$

- guarantees that cost considerations never get in the way of information transmission considerations
- new utility functions:

$$\begin{aligned} u_s(i,m,a) &= -c_s(i,m) + \begin{cases} 1 & \text{if } i = a, \\ 0 & \text{else,} \end{cases} \\ u_r(i,m,a) &= -c_r(a,m) + \begin{cases} 1 & \text{if } i = a, \\ 0 & \text{else.} \end{cases} \end{aligned}$$

Modified IBR procecure

Sally

- flip ρ along the diagonal
- subtract c_s
- place a 0 in each cell that is non-maximal within its row
- normalize each row

Robin

- flip σ along diagonal
- if a row contains only 0s,
  - fill in a 1 in each cell corresponding to a true world-message association
- else
  - subtract c_r^T
- place a 0 in each cell that is non-maximal within its row
- normalize each row

The Strongest Meaning Hypothesis

- if in doubt, Robin will assume that Sally is competent
- captured in following cost function:

$$c_r(a, m) = \frac{|a|}{\max(|M|, 2^{|W|})^2}$$

$$\begin{aligned}
c_r(\{w_a\}, \cdot) &= \tfrac{1}{49} & c_r(\{w_a, w_{ab}\}, \cdot) &= \tfrac{2}{49} \\
c_r(\{w_b\}, \cdot) &= \tfrac{1}{49} & c_r(\{w_b, w_{ab}\}, \cdot) &= \tfrac{2}{49} \\
c_r(\{w_{ab}\}, \cdot) &= \tfrac{1}{49} & c_r(\{w_a, w_b, w_{ab}\}, \cdot) &= \tfrac{3}{49} \\
c_r(\{w_a, w_b\}, \cdot) &= \tfrac{2}{49}
\end{aligned}$$

Lifted games

IBR sequence: 1

σ_0	OR	A	B	AND
$\{w_a\}$	$\frac{1}{2}$	$\frac{1}{2}$	0	0
$\{w_b\}$	$\frac{1}{2}$	0	$\frac{1}{2}$	0
$\{w_{ab}\}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
$\{w_a, w_b\}$	1	0	0	0
$\{w_a, w_{ab}\}$	$\frac{1}{2}$	$\frac{1}{2}$	0	0
$\{w_b, w_{ab}\}$	$\frac{1}{2}$	0	$\frac{1}{2}$	0
$\{w_a, w_b, w_{ab}\}$	1	0	0	0

Lifted games

IBR sequence: flipping and subtracting costs

ρ_0	$\{w_a\}$	$\{w_b\}$	$\{w_{ab}\}$	$\{w_a, w_b\}$	$\{w_a, w_{ab}\}$	$\{w_b, w_{ab}\}$	$\{w_a, w_b, w_{ab}\}$
OR	0.48	0.48	0.23	**0.96**	0.46	0.46	0.94
A	**0.48**	−0.02	0.23	−0.04	0.46	−0.04	−0.06
B	−0.02	**0.48**	0.23	−0.04	−0.04	0.46	−0.06
AND	−0.02	−0.02	**0.23**	−0.04	−0.04	−0.04	−0.06

Lifted games

IBR sequence: 2

ρ_0	$\{w_a\}$	$\{w_b\}$	$\{w_{ab}\}$	$\{w_a, w_b\}$	$\{w_a, w_{ab}\}$	$\{w_b, w_{ab}\}$	$\{w_a, w_b, w_{ab}\}$
OR	0	0	0	1	0	0	0
A	1	0	0	0	0	0	0
B	0	1	0	0	0	0	0
AND	0	0	1	0	0	0	0

Lifted games

IBR sequence: 3

σ_1	OR	A	B	AND
$\{w_a\}$	0	1	0	0
$\{w_b\}$	0	0	1	0
$\{w_{ab}\}$	0	0	0	1
$\{w_a, w_b\}$	1	0	0	0
$\{w_a, w_{ab}\}$	$\frac{1}{2}$	$\frac{1}{2}$	0	0
$\{w_b, w_{ab}\}$	$\frac{1}{2}$	0	$\frac{1}{2}$	0
$\{w_a, w_b, w_{ab}\}$	1	0	0	0

Lifted games

- OR is only used in $\{w_a, w_b\}$ in the fixed point
- this means that it carries two implicatures:
  - exhaustivity: Ann and Bert did not both show up
  - ignorance: Sally does not know which one of the two disjuncts is true

Sender costs

2.
   a. Ann or Bert or both showed up. (= AB-OR)
   b. Ann showed up. (= A)
   c. Bert showed up. (= B)
   d. Ann and Bert showed up. (= AND)
   e. Ann or Bert showed up. (= OR)
   f. Ann or both showed up. (= A-OR)
   g. Bert or both showed up. (= B-OR)

- Message (e) is arguably more efficient for Sally than (a)
- Let us say that $c_s(\cdot, \text{AB-OR}) = \frac{1}{50}, c_s(\cdot, \text{A-OR}) = c_s(\cdot, \text{B-OR}) = \frac{1}{75}, c_s(\cdot, \text{OR}) = c_s(\cdot, \text{AND}) = \frac{1}{100}$), and $c_s(\cdot, \text{A}) = c_s(\cdot, \text{B}) = 0$.

More ignorance implicatures

IBR sequence: 1

σ_0	AB-OR	A	B	AND	OR	A-OR	B-OR
$\{w_a\}$	$\frac{1}{4}$	$\frac{1}{4}$	0	0	$\frac{1}{4}$	$\frac{1}{4}$	0
$\{w_b\}$	$\frac{1}{4}$	0	$\frac{1}{4}$	0	$\frac{1}{4}$	0	$\frac{1}{4}$
$\{w_{ab}\}$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$
$\{w_a, w_b\}$	$\frac{1}{2}$	0	0	0	$\frac{1}{2}$	0	0
$\{w_a, w_{ab}\}$	$\frac{1}{4}$	$\frac{1}{4}$	0	0	$\frac{1}{4}$	$\frac{1}{4}$	0
$\{w_b, w_{ab}\}$	$\frac{1}{4}$	0	$\frac{1}{4}$	0	$\frac{1}{4}$	0	$\frac{1}{4}$
$\{w_a, w_b, w_{ab}\}$	$\frac{1}{2}$	0	0	0	$\frac{1}{2}$	0	0

More ignorance implicatures

IBR sequence: 1

ρ_0	$\{w_a\}$	$\{w_b\}$	$\{w_{ab}\}$	$\{w_a, w_b\}$	$\{w_a, w_{ab}\}$	$\{w_b, w_{ab}\}$	$\{w_a, w_b, w_{ab}\}$
AB-OR	0	0	0	1	0	0	0
A	1	0	0	0	0	0	0
B	0	1	0	0	0	0	0
AND	0	0	1	0	0	0	0
OR	0	0	0	1	0	0	0
A-OR	1	0	0	0	0	0	0
B-OR	0	1	0	0	0	0	0

More ignorance implicatures

IBR sequence: 2

σ_1	AB-OR	A	B	AND	OR	A-OR	B-OR
$\{w_a\}$	0	1	0	0	0	0	0
$\{w_b\}$	0	0	1	0	0	0	0
$\{w_{ab}\}$	0	0	0	1	0	0	0
$\{w_a, w_b\}$	0	0	0	0	1	0	0
$\{w_a, w_{ab}\}$	0	1	0	0	0	0	0
$\{w_b, w_{ab}\}$	0	0	1	0	0	0	0
$\{w_a, w_b, w_{ab}\}$	0	0	0	0	1	0	0

More ignorance implicatures

IBR sequence: 2

ρ_1	$\{w_a\}$	$\{w_b\}$	$\{w_{ab}\}$	$\{w_a, w_b\}$	$\{w_a, w_{ab}\}$	$\{w_b, w_{ab}\}$	$\{w_a, w_b, w_{ab}\}$
ORBOTH	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$
A	1	0	0	0	0	0	0
B	0	1	0	0	0	0	0
AND	0	0	1	0	0	0	0
OR	0	0	0	1	0	0	0
A-OR	$\frac{1}{3}$	0	$\frac{1}{3}$	0	$\frac{1}{3}$	0	0
B-OR	0	$\frac{1}{3}$	$\frac{1}{3}$	0	0	$\frac{1}{3}$	0

More ignorance implicatures

IBR sequence: 3

σ_2	AB-OR	A	B	AND	OR	A-OR	B-OR
$\{w_a\}$	0	1	0	0	0	0	0
$\{w_b\}$	0	0	1	0	0	0	0
$\{w_{ab}\}$	0	0	0	1	0	0	0
$\{w_a, w_b\}$	0	0	0	0	1	0	0
$\{w_a, w_{ab}\}$	0	0	0	0	0	1	0
$\{w_b, w_{ab}\}$	0	0	0	0	0	0	1
$\{w_a, w_b, w_{ab}\}$	1	0	0	0	0	0	0

More ignorance implicatures

IBR sequence: 3

ρ_2	$\{w_a\}$	$\{w_b\}$	$\{w_{ab}\}$	$\{w_a, w_b\}$	$\{w_a, w_{ab}\}$	$\{w_b, w_{ab}\}$	$\{w_a, w_b, w_{ab}\}$
ORBOTH	0	0	0	0	0	0	1
A	1	0	0	0	0	0	0
B	0	1	0	0	0	0	0
AND	0	0	1	0	0	0	0
OR	0	0	0	1	0	0	0
A-OR	0	0	0	0	1	0	0
B-OR	0	0	0	0	0	1	0

Conclusion

- IBR model formalizes neo-Gricean program
- Principle of cooperativity: identical preferences of sender and receiver
- Quality: Honesty is default strategy
- Quantity, Relevance: captured in utility function
- Manner: captured in cost function
- further applications
    - free choice implicatures
    - conditional perfection
    - I-implicatures, M-implicatures
    - pragmatics of measure terms
- next project: presuppositions

I-implicatures

(2) a. John opened the door. $(= \text{OPEN})$
b. John opened the door using the handle. $(= \text{OPEN-H})$
c. John opened the door with an axe. $(= \text{OPEN-A})$

formally

- $W = \{w_h, w_a\}$
- $p^*(w_1) = \frac{2}{3}, p^*(w_2) = \frac{1}{3}$
- $\|\text{OPEN-H}\| = \{w_h\}$, $\|\text{OPEN-A}\| = \{w_a\}$, and $\|\text{OPEN}\| = \{w_h, w_a\}$
- $c(m_1) = c(m_2) \in \frac{1}{20}$, $c(m_3) = 0$

	a_h	a_a
w_h	$1, 1$	$0, 0$
w_a	$0, 0$	$1, 1$

I-implicatures

σ_0	OPEN	OPEN-H	OPEN-A
w_h	$\frac{1}{2}$	$\frac{1}{2}$	0
w_a	$\frac{1}{2}$	0	$\frac{1}{2}$

ρ_0	w_h	w_a
OPEN	1	0
OPEN-H	1	0
OPEN-A	0	1

σ_1	OPEN	OPEN-H	OPEN-A
w_h	1	0	0
w_a	0	0	1

ρ_1	w_h	w_a
OPEN	1	0
OPEN-H	1	0
OPEN-A	0	1

$$F = (\sigma_1, \rho_0)$$

Measure terms

Krifka (2002,2007) notes that measure terms can be used in a precise or in a vague way, and that more complex expressions are less likely to be used in a vague way. Here is a schematic analysis:

- w_1, w_3: 100 meter, w_2, w_4: 101 meter
- m_{100}: "one hundred meter"
  m_{101}: "one hundred and one meter"
  m_{ex100}: "exactly one hundred meter"
- $\|m_{100}\| = \|m_{ex100}\| = \{w_1, w_3\}$,
  $\|m_{101}\| = \{w_2, w_4\}$
- $c(m_{100}) = 0$,
  $c(m_{101}) = c(m_{ex100}) = 0.15$
- a_1, a_3: 100, a_2, a_4: 101

- in w_1, w_2 precision is important
- in w_3, w_4 precision is not important

	a_1	a_2	a_3	a_4
w_1	1	0.5	1	0.5
w_2	0.5	1	0.5	1
w_3	1	0.9	1	0.9
w_4	0.9	1	0.9	1

Measure terms

σ_0	m_{100}	m_{101}	m_{ex100}
w_1	$\frac{1}{2}$	0	$\frac{1}{2}$
w_2	0	1	0
w_3	$\frac{1}{2}$	0	$\frac{1}{2}$
w_4	0	1	0

ρ_0	a_1	a_2	a_3	a_4
m_{100}	$\frac{1}{2}$	0	$\frac{1}{2}$	0
m_{101}	0	$\frac{1}{2}$	0	$\frac{1}{2}$
m_{ex100}	$\frac{1}{2}$	0	$\frac{1}{2}$	0

σ_1	m_{100}	m_{101}	m_{ex100}
w_1	1	0	0
w_2	0	1	0
w_3	1	0	0
w_4	1	0	0

ρ_1	a_1	a_2	a_3	a_4
m_{100}	$\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{1}{3}$
m_{101}	0	1	0	0
m_{ex100}	$\frac{1}{2}$	0	$\frac{1}{2}$	0

σ_2	m_{100}	m_{101}	m_{ex100}
w_1	0	0	1
w_2	0	1	0
w_3	1	0	0
w_4	1	0	0

ρ_2	a_1	wa_2	a_3	a_4
m_{100}	0	0	$\frac{1}{2}$	$\frac{1}{2}$
m_{101}	0	1	0	0
m_{ex100}	1	0	0	0

M-implicatures

3. **a.** John stopped the car. (= STOP)
   b. John made the car stop. (= MAKE-STOP)

- w_1: John used the foot brake.
- w_2: John drove the car against a wall.
- $\|\text{STOP}\| = \|\text{MAKE-STOP}\| = \{w_1, w_2\}$
- $c(\text{STOP}) = 0$; $c(\text{MAKE-STOP} = 0.1$
- $p^*(w_1) = .8$; $p^*(w_2) = .2$.

Utility matrix

	a_1	a_2
w_1	1	0
w_2	0	1

M-implicatures

IBR sequence

σ_0	STOP	MAKE-STOP
w_1	$\frac{1}{2}$	$\frac{1}{2}$
w_2	$\frac{1}{2}$	$\frac{1}{2}$

ρ_0	a_1	a_2
STOP	1	0
MAKE-STOP	1	0

σ_1	STOP	MAKE-STOP
w_1	1	0
w_2	1	0

ρ_1	a_1	a_2
STOP	1	0
MAKE-STOP	$\frac{1}{2}$	$\frac{1}{2}$

σ_2	STOP	MAKE-STOP
w_1	1	0
w_2	0	1

ρ_2	a_1	a_2
STOP	1	0
MAKE-STOP	0	1