Game Theoretic Pragmatics

Day 4
Aspects of Bounded Rationality

Anton Benz

Centre for General Linguistics (ZAS), Berlin

19 August 2010

Communication and Speaker's intentions

Widely held:

1. Speaker's intentions define communicated meaning.
2. Interpretation involves recognition of speaker's intentions

Grice:

[S] means$_{nn}$ something by x' is roughly equivalent to '[S] uttered x with the intention of inducing a belief by means of the recognition of this intention. [Grice, 1957, p. 385]

Relevance Theory
[Sperber and Wilson, 2004]

Ostensive–Inferential Communication

1. **The informative intention**:
   The intention to inform an audience of something.
2. **The communicative intention**:
   The intention to inform an audience of one's informative intention.

[Sperber and Wilson, 2004, p. 611]

⇒ Understanding is achieved when the communicative intention is fulfilled.

⇒ Audience has to recognise the informative intention.

Contrast: Two Communicating Computers

1. Computer 1 encodes some command *p*.
2. The signal is transmitted to computer 2.
3. Computer 2 decodes signal again and executes command.

$\Rightarrow$ No recognition of computer 1's *intentions* involved.

$\Rightarrow$ Human communication principally different.

Cognitive Costs

Intuitively:

1. The simple picture of communicating is very efficient.
2. Reasoning about speaker's intentions is costly.

⇒ Is reasoning about speaker's intentions really necessary for hunman communication?

Communication as an intentional activity

Approach:

- Address questions indirectly: How much reasoning about each other is necessary?
- Question can be studied in game theoretic model:
  1. If calculating best response by hearer involves reasoning about speaker's strategy, reasoning about speaker's intentions **may** take place.
  2. If calculating best response by hearer involves no reasoning about speaker's strategy, reasoning about speaker's intentions **can not** take place.

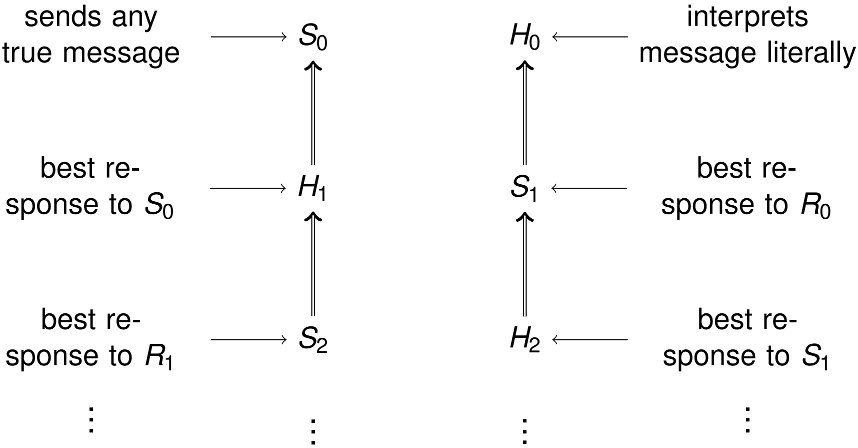
The Role of the IBR Model

1. IBR model makes reasoning about each other explicit.
2. It provides a measure of complexity and predicts a lower limit.
3. Lower limit predicts reasoning about speaker's strategy must take place.

$\Rightarrow$ Is there a method which undercuts the IBR lower limit?

The Iterated Best Response (IBR) Model

[Franke, 2009, Jäger and Ebert, 2009]



Undercutting IBR Limit

We will see:

1. For normal communication backward induction is sufficient.
2. Follows: Speaker reasons once about hearer.
3. Recognition of the speaker's intentions follows rather than precedes interpretation.

Justification:

1. Backward induction is, in general, not appropriate for signalling games.
2. Provide justification by investigating:
   i. Notion of Bayesian perfect equilibria.
   ii. Natural Information.

Role of Optimal answer Model

1. Basic structures: support problems.
2. Solved by backward induction.

$\Rightarrow$ No reasoning about speaker's strategy involved.

$\Rightarrow$ Considerations also provide justification of OA model.

Models of Signalling Behaviour

GT Models

non–evolutionary

online

- immediate effects
- two interlocutors

competence

- several round of interaction
- two interlocutors

evolutionary

- replication involving several generations
- large populations (infinite)

Models of Signalling Behaviour

GT Models

- non–evolutionary
  - online
    - immediate effects
    - two interlocutors
  - competence
    - several round of interaction
    - two interlocutors
- evolutionary
  - replication involving several generations
  - large populations (infinite)

Outline

1. Baysian Perfect Equilibria and Backward Induction
2. Natural Information
3. Natural Information and Implicatures
4. Implicatures and Reasoning about Each Other

Section 1

Baysian Perfect Equilibria and Backward Induction

[Fudenberg and Tirol, 1991]

Baysian Perfect Equilibrium

1. The textbook equilibrium concept for signalling games is the concept of a Bayesian Perfect (Nash) Equilibrium.
2. Idea: Once the hearer receives a signal F from the speaker, he learns new information about the state of the world.

⇒ The information state after receiving F corresponds to the conditional probability distribution $P(v|\theta, F)$ defined by:

$$P(v|\theta_H, F) = \frac{\sum_{\theta_S} P(v, \theta_S, \theta_H)\, S(F|\theta_S)}{\sum_v \sum_{\theta_S} P(v, \theta_S, \theta_H)\, S(F|\theta_S)}. \quad (1.1)$$

3. Hearer's response to signal has to be a best response taking the conditional probability distribution $P(v|\theta, F)$ into account.

Baysian Perfect Equilibrium

Definition 1

A strategy pair (S, H) is a (weak) Bayesian Perfect Nash equilibrium of a signalling game $\mathscr{G}$, iff (S, H) is Nash equilibrium which satisfies:

1. If $P(F|\theta_H) > 0$, i.e. if $\sum_{v,\theta_S} P(v, \theta_S, \theta_H)\, S(F|\theta_S) > 0$, then

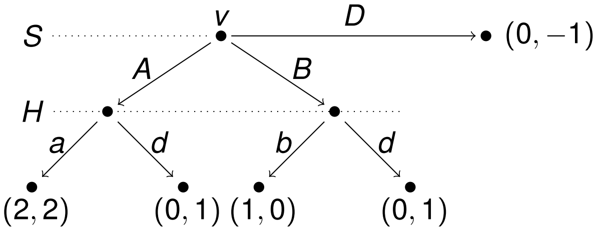
$$H(a|\theta_H, F) > 0 \Rightarrow a \in \mathscr{B}_H(F|\theta_H). \quad (1.2)$$

2. If $P(F|\theta_H) = 0$, then there exists a probability distribution $\mu(v|\theta_H, F)$ such that $H(a|\theta_H, F) > 0$ implies that for all $a' \in \mathcal{A}$:

$$\sum_v \mu(v|\theta_H, F)\, u(v, F, a') \leq \sum_v \mu(v|\theta_H, F)\, u(v, F, a). \quad (1.3)$$

   This means: there must be some probability distribution μ which makes a **a best response** to F **relative to** μ.

- Every Bayesian Perfect Equilibrium is a Nash equilibrium by definition; but
- Not every Nash equilibrium is Bayesian perfect.



Equilibria

- Two types of Nash equilibria: (S_1, H_1) and (S_2, H_2) with

	v
S_1	A
S_2	D

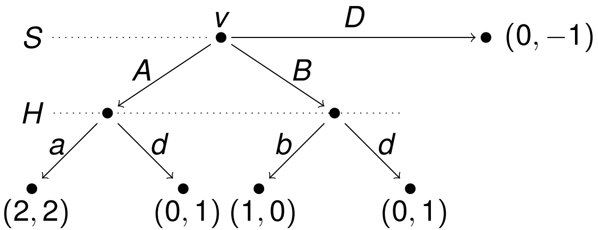
	A	B	D
H_1	a	x	y
H_2	d	d	d

$x, y \in \{a, b, d\}$

- But only one type of Baysian perfect equilibrium: (S_1, H_1') with

	A	B	D
H_1'	a	d	y

$y \in \{a, b, d\}$



1. Nash equilibrium does not take out–of–equilibrium signals into account, i.e. move which shouldn't occur if the players follow their equilibrium strategies.
2. Baysian perfection takes into account the information available to an agent when acting.
3. It also considers the states after out–of–equilibrium signals.
   - If an out–of–equilibrium signal F is sent, it must be possible to assign a belief to the hearer which makes his choice of action rational.
   - That is: there must be a probability distribution μ such that his choice a is a best response to F given μ.

It can be shown:

$$\mathscr{E}_H(H|S) = \sum_{\theta_H, F} P(\theta_H, F)\, \mathscr{E}_H(H|\theta_H, F). \tag{1.4}$$

Follows: If one ignores out–of–equilibrium signals = surprise messages, then all Nash equilibria are Bayesian perfect.

Signalling Games of Pure Coordination

1. From now on we only consider signalling games of pure coordination:

   $$\langle \Omega, \Theta_S, \Theta_H, P, \mathcal{F}, \mathcal{A}, u \rangle$$

   This means, $u = u_S = u_H$.

2. Furthermore, we assume that u can be divided into a sum

   $$u(v, F, a) = u(v, a) - c(F) \tag{1.5}$$

   with

   i. a payoff $u(v, a)$ function only depending on the world and the action performed by the hearer;
   ii. A cost function $c : \mathcal{F} \longrightarrow \mathbb{R}_0^+$.

A Criterion for Nash Equilibria

Lemma 2

Let $\langle \Omega, \Theta_S, \Theta_H, P, \mathcal{F}, \mathcal{A}, u \rangle$ *be a signalling game. Assume that for all* $\theta_S \in \Theta_S$ $P(\theta_S) > 0$. *Then* (S, H) *is a Nash equilibrium if, and only if:*

$$\forall \theta_S, \theta_H \, (S(F|\theta_S) > 0 \wedge P(\theta_H|\theta_S) > 0 \Rightarrow H(\mathscr{B}(\theta_S, \theta_H)|F) = 1). \quad (1.6)$$

with

$$\mathscr{B}(\theta_S, \theta_H) = \{a \in \mathcal{A} \mid \forall b \, \mathscr{E}(b|\theta_S, \theta_H) \leq \mathscr{E}(a|\theta_S, \theta_H)\}, \quad (1.7)$$

and

$$\mathscr{E}(a|\theta_S, \theta_H) = \sum_v P(v|\theta_S, \theta_H) \, u(v, a). \quad (1.8)$$

Weak Pareto Dominance

Given the conditions of Lemma 2, it holds for a Nash equilibrium (S, H) satisfying (1.6):

1. If signals are cheap, i.e. if they have no influence on payoffs, then (S, H) weakly Pareto dominates all other signalling pairs (S', H').
2. If costs of signals are nominal, then (S, H) weakly Pareto dominates all other signalling pairs (S', H') up to nominal costs.

Bayesian Perfect Equilibria and Interpreted Signalling Games

Interpreted Signalling Games of Pure Coordination

We concentrate on:

Definition 3

An **interpreted signalling game of pure coordination** is a tuple $\langle \Omega, \Theta_S, \Theta_H, P, \mathcal{F}, \mathcal{A}, u, [\![\,.\,]\!] \rangle$ such that:

1. $\langle \Omega, \Theta_S, \Theta_H, P, \mathcal{F}, \mathcal{A}, u \rangle$ is a signalling game of pure coordination.
2. $[\![\,.\,]\!]$ is an interpretation function for signals $F \in \mathcal{F}$; i.e.

$$[\![\,.\,]\!] : \mathcal{F} \longrightarrow \mathcal{P}(\Omega), \; F \mapsto [\![F]\!]. \tag{1.9}$$

We write signalling game instead of signalling game of **pure coordination** if it is clear from context what is meant.

Expected Utility and Best Actions Based on Semantics

This does not take the speaker's signalling strategy into account!

1. $\langle \Omega, \Theta_S, \Theta_H, P, \mathcal{F}, \mathcal{A}, u, [\![\,.\,]\!] \rangle$: an interpreted signalling game.
2. Hearer's belief after learning F given type θ_H:

$$P(X|\theta_H, [\![F]\!]) = \frac{\sum_{\theta_S} P(X \cap [\![F]\!], \theta_S, \theta_H)}{\sum_{\theta_S} P([\![F]\!], \theta_S, \theta_H)} \text{ for } [\![F]\!] \neq \emptyset. \quad (1.10)$$

3. Hearer's best expected utility after learning F given type θ_H:

$$\mathscr{E}_H(a|\theta_H, [\![F]\!]) = \sum_v P(v|[\![F]\!], \theta_H)\, u_H(v, F, a). \quad (1.11)$$

4. Hearer's best actions after learning F given type θ_H:

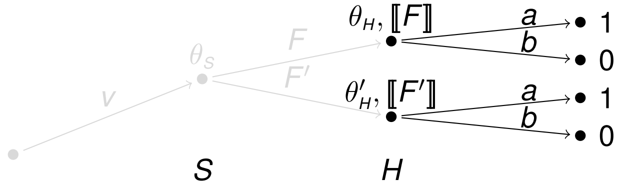
$$\mathscr{B}_H(\theta_H, [\![F]\!]) = \{a \in \mathcal{A} \mid \forall a'\ \mathscr{E}_H(a'|\theta_H, [\![F]\!]) \leq \mathscr{E}_H(a|\theta_H, [\![F]\!])\}. \quad (1.12)$$

Backward Induction

Backward induction proceeds as follows:

1. First backward induction step:

$$H(\,.\,|F,\theta_H) \in \mathscr{B}_H(\theta_H, [\![F]\!]). \tag{1.13}$$



Backward Induction

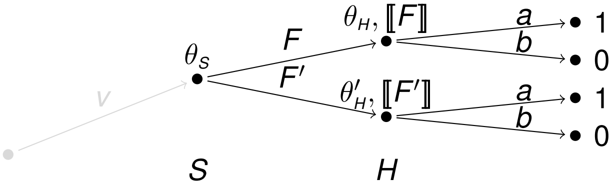
Backward induction proceeds as follows:

1. First backward induction step:

$$H(\,.\,|F,\theta_H) \in \mathscr{B}_H(\theta_H, [\![F]\!]). \tag{1.13}$$

2. Second backward induction step:

$$S(\,.\,|\theta_S) \in \mathscr{B}_S(H). \tag{1.14}$$



Why is Backward Induction of Interest?

1. Backward induction is the most simple method for finding Nash equilibria of sequential games.
   i. Hearer only takes semantic information into account.
   ii. Speaker takes hearer's perspective into account.
   $\Rightarrow$ No extended reasoning about each other.
2. To be fully justified, each participant has to exactly know the node of the game tree in which he decides.
   $\Rightarrow$ Not satisfied in signalling games!

$\Rightarrow$ When is backward induction a reliable means leading to Bayesian perfect equilibria in signalling games?

When Backward Induction is not Optimal

Backward induction cannot be reliable if

1. Signalling games are not games of pure coordination.
2. Certain states of the world cannot be expressed.
3. Speaker knowledge is limited and expectations about the world differ greatly.

Zero Sum game

No private hearer information: $P(v|\theta_H) = P(v)$.

- Example: zero sum game, i.e. a game with strictly opposing payoffs.
- Backward induction would allow the speaker to manipulate the hearer.

Ω	$P(w_i\|\theta_S)$	$P(w_i)$	a	b
w_1	1	$1/4$	1 ; 0	0 ; 1
w_2	0	$1/4$	0 ; 1	1 ; 0
w_3	0	$1/4$	1 ; 0	0 ; 1
w_4	0	$1/4$	0 ; 1	1 ; 0

Answer $\{w_1, w_2, w_4\}$ would induce H to choose a.

Ineffability

No private hearer information: $P(v|\theta_H) = P(v)$.

Assume:

- There are w_1, w_2 such that for all F: $w_1 \in [\![F]\!] \Leftrightarrow w_2 \in [\![F]\!]$, and
- $[\![F_1]\!] = [\![F_2]\!] = \{w_1, w_2\}$.

$\Rightarrow$ Strategy (S, H) found by backward induction cannot lead to optimal success in:

Ω	$P(w_i \mid \theta_S)$	$P(w_i)$	a	b
w_1	0	$1/2$	1	0
w_2	1	$1/6$	0	1
w_3	0	$1/3$	1	0

Strategy pair (S', H') with $S'(w_i) = F_i$, $H'(F_i) = w_i$ is more successful.

Differing Expectations

No private hearer information: $P(v|\theta_H) = P(v)$.

Backward induction: Speaker cannot induce hearer to choose action *b*:

Ω	$P(w_i|\theta_S)$	$P(w_i)$	a	b
w_1	$1/4$	$3/8$	1	0
w_2	$3/4$	$1/8$	0	1
w_3	0	$1/2$	1	0

This can be avoided if e.g. for all $\theta \in \Theta_S$:

1. Speaker is expert: $\exists a \in \mathcal{A}\ P(O(a)|\theta) = 1$; or
2. Common prior: $\forall v \in K_\theta\ P(v|\theta) = P(v|K_\theta)$ with $K_\theta = \{v \mid P(v|\theta) > 0\}$.

Backward Induction and Bayesian Perfection

Lemma 4

Let $\langle \Omega, \Theta_S, \Theta_H, P, \mathcal{F}, \mathcal{A}, u, [\![\,.\,]\!]\rangle$ *be an interpreted signalling game. Assume that for* $K_\theta = \{v \mid P(v|\theta) > 0\}$*:*

$$\exists F \in \mathcal{F}\, \forall \theta_H \in \Theta_H : \ (K_\theta \subseteq [\![F]\!] \,\&\, \mathscr{B}_H(\theta_H, [\![F]\!]) \subseteq \mathscr{B}_H(\theta_S, \theta_H)). \qquad (1.15)$$

Then the strategy pair (S, H) *found by backward induction is a Bayesian perfect Nash equilibrium which weakly dominates all other strategy pairs* (S', H')*.*

If cost of signals are **nominal**, then (S, H) weakly dominates all other strategy pairs up to nominal costs.

Special Cases I

Let $\langle \Omega, \Theta_S, P, \mathcal{F}, \mathcal{A}, u, [\![\, . \,]\!] \rangle$ be an interpreted signalling game for which every state is expressible, signals are cheap, and beliefs are derived from a common prior; i.e.

1. $\forall X \subseteq \Omega \, \exists F \in \mathcal{F} \; X = [\![F]\!]$,
2. $\forall F, F' \in \mathcal{F} \, \forall v \in \Omega \, \forall a \in \mathcal{A} \; u(v, F, a) = u(v, F', a)$.
3. $\forall v \in K_\theta \; P(v|\theta) = P(v|K_\theta)$ with $K_\theta = \{v \mid P(v|\theta) > 0\}$.

Then:

$$\forall \theta \in \Theta_S \, \exists F \in \mathcal{F} : \; (K_\theta \subseteq [\![F]\!] \wedge \mathscr{B}_H([\![F]\!]) \subseteq \mathscr{B}(\theta)) \qquad (1.16)$$

Special Cases II

Let $\langle \Omega, \Theta_S, \Theta_H, P, \mathcal{F}, \mathcal{A}, u, [\![\,.\,]\!] \rangle$ be an interpreted signalling game for which every state is expressible, signals are cheap, and beliefs are derived from a common prior; i.e.

1. $\forall X \subseteq \Omega\, \exists F \in \mathcal{F}\; X = [\![F]\!]$,
2. $\forall F, F' \in \mathcal{F}\, \forall v \in \Omega\, \forall a \in \mathcal{A}\; u(v, F, a) = u(v, F', a)$.
3. $\forall \theta_S \in \Theta_S\, \exists a \in \mathcal{A}\; P(O(a)|\theta_S) = 1$ (Expert).

Then:

$$\forall \theta_H, \theta_S\, \exists F \in \mathcal{F} : \; (P([\![F]\!]|\theta_S) = 1 \wedge \mathscr{B}_H(\theta_H, [\![F]\!]) \subseteq \mathscr{B}(\theta_S, \theta_H)) \qquad (1.17)$$

$O(a)$: set of all worlds in which a is an optimal action.

Signalling Games and Support Problems

Constructing Support problems from Signalling Games

1. Let $\mathscr{G}$ be an interpreted signalling game $\langle \Omega, \Theta_S, P, \mathcal{F}, \mathcal{A}, u, [\![\,.\,]\!] \rangle$ without private knowledge for the hearer.
2. Let $\mathcal{S}$ be a set of support problems $\sigma = \langle \Omega, P_S, P_H, \mathcal{F}, \mathcal{A}, u, [\![\,.\,]\!] \rangle$. Let there be a bijection $\Theta_S \longrightarrow \mathcal{S}$, $\theta \mapsto \sigma_\theta$ such that for all $\theta \in \Theta_S$:

$$P_S^{\sigma_\theta}(v) = P(v|\theta).$$
$$P_H^{\sigma_\theta}(v) = P(v).$$

3. Then, for all strategy pairs (S, H):
   - (S, H) is justified by backward induction for the signalling game $\mathscr{G}$

     if, and only if
   - $(S(\,.\,|\sigma_\theta), H)$ is justified by backward induction for σ_θ.
4. If, in addition, for all $\theta \in \Theta_S$:

$$P(\theta|v) = \frac{P(\theta)}{P(K_\theta)} \text{ for } v \in K_\theta := \{v \in \Omega \mid P(v|\theta) > 0\}. \qquad (1.18)$$

   Then, σ_θ is a support problem which satisfies:

$$\forall X \subseteq \Omega : P_S(X) = P_H(X|K_S^{\sigma_\theta});$$

   with $K_S^{\sigma_\theta} = \{v \in \Omega \mid P_S^{\sigma_\theta}(v) > 0\}$.

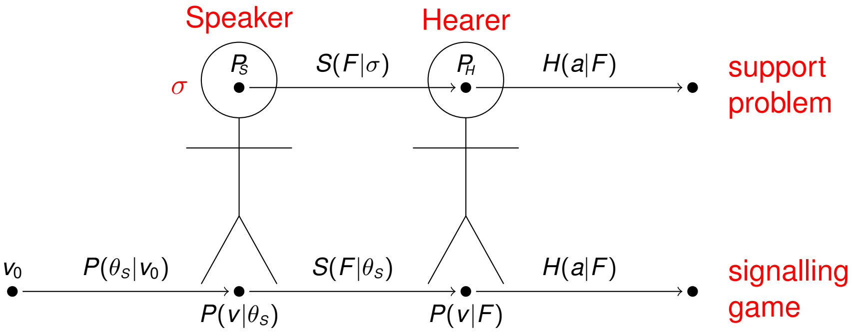
Constructing Signalling Games from Sets of Support Problems

Let $\mathcal{S}$ be a set of interpreted support problems $\sigma = \langle \Omega, P_S, P_H, \mathcal{F}, \mathcal{A}, u, [\![\,.\,]\!] \rangle$ which may only differ with respect to P_S^σ.

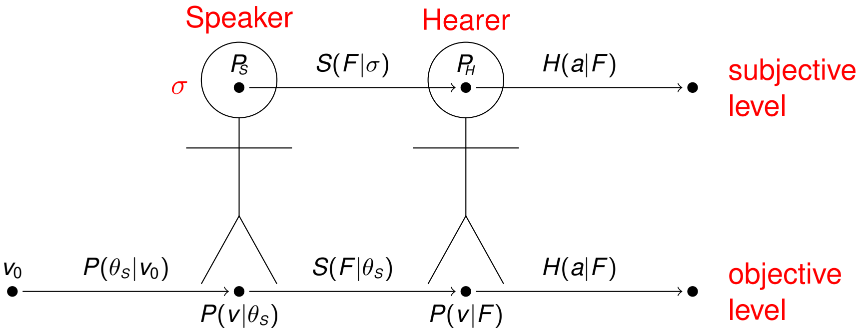
1. P be any probability measure on $\mathcal{S}$ for which $P(\sigma) > 0$ for all $\sigma \in \mathcal{S}$,
2. Let $\Theta_S := \mathcal{S}$, and $P(v, \sigma) := P(\sigma)\, P_S^\sigma(v)$.

Then P is a probability measure on $\Omega \times \Theta_S$, and $\mathscr{G} = \langle \Omega, \Theta_S, P, \mathcal{F}, \mathcal{A}, u, [\![\,.\,]\!] \rangle$ is a signalling game.

Subjective and Objective Level



Subjective and Objective Level



Section 2

Natural Information

Natural Meaning: The Gricean Picture

[Grice, 1957]

Natural Meaning

Natural meaning:

Example 5

1. Those spots mean measles.
2. Those spots didn't mean anything to me, but to the doctor they meant measles.

1. In both sentences, the word *meaning* refers to natural meaning.
2. Meaning of spots does not depend on someone using them with the intention to communicate.

Non–Natural Meaning

Non–natural meaning:

Example 6

1. Those three rings on the bell (of the bus) mean that the bus is full.
2. That remark, ‘Smith couldn’t get on without his trouble and strife,’ meant that Smith found his wife indispensable.

1. In both sentences, the word *meaning* refers to non–natural meaning.
2. Meaning of rings or the remark depends on someone using them with the intention to communicate.

Communicated meaning

Grice: Communicated meaning = *non-natural* meaning (meaning_{nn}):

Definition 7

[S] means_{nn} something by x' is roughly equivalent to '[S] uttered x with the intention of inducing a belief by means of the recognition of this intention. [Grice, 1957, p. 385]

Grice' picture example

Example 8 (Mr. X)

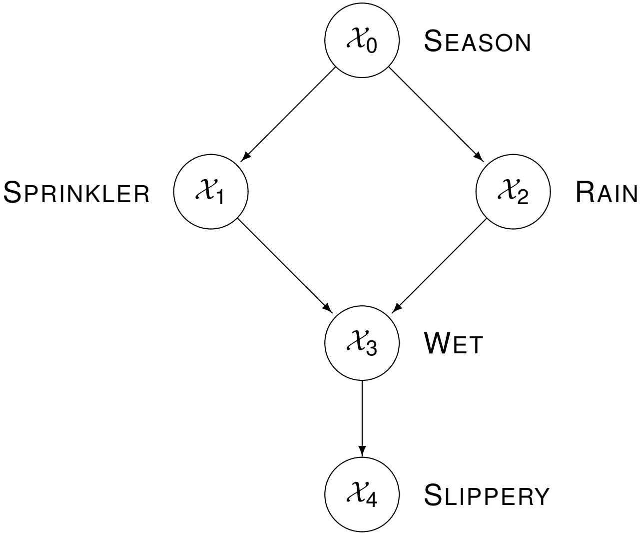
1. I show Mr. X a photograph of Mr. Y displaying undue familiarity to Mrs. X.
2. I draw a picture of Mr. Y behaving in this manner and show it to Mr. X.

1. In both communicated: Mr. Y had been unduly familiar with Mrs. X.
2. The drawing achieves this only by recognition of the senders intentions.

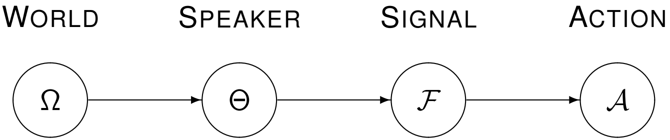
Natural Meaning in Causal Networks

[Benz, 2009]

Causal Networks

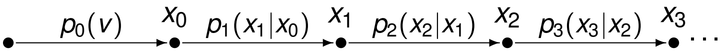


Causal Network of Communication



The causal network associated to a signalling game.

A Markovian Process



Joint probability:

$$P(\langle x_0, \ldots, x_n \rangle) := \prod_{i=0}^{n} p_i(x_i|x_{i-1}). \tag{2.19}$$

Marginal probability:

$$P_i(X) = P({\pi_i}^{-1}[X]), \text{ for } X \subseteq \mathcal{X}_i. \tag{2.20}$$

Conditional marginal probability:

$$P_{i|j}(X|Y) = P({\pi_i}^{-1}[X]|{\pi_j}^{-1}[Y]). \tag{2.21}$$

Natural Meaning

Definition 9

Let $(\mathcal{X}_i, p_i)_{i=0,\dots,n}$ be a linear causal network. Then, for $X \subseteq \mathcal{X}_i$ and $Y \subseteq \mathcal{X}_j$ with $P_j(Y) > 0$, we set

$$(\mathcal{X}_i, p_i) \models Y \Rrightarrow X :\Longleftrightarrow \; P_{i|j}(X|Y) = 1. \tag{2.22}$$

We say that event Y *naturally means* that X.

Natural Meaning in Signalling Games

Definition 10

Let $\mathscr{G} = \langle \Omega, \Theta_s, P, \mathcal{F}, \mathcal{A}, u, [\![\,.\,]\!] \rangle$ be a signalling game and (S, H) a set of signalling strategies. Let $F \in \mathcal{F}$ such that $\sum_{v,\theta_S} P(v, \theta_s)\, S(F|\theta_s) > 0$, and $R \subseteq \Omega$ or $R \subseteq \Theta_s$. Then we write

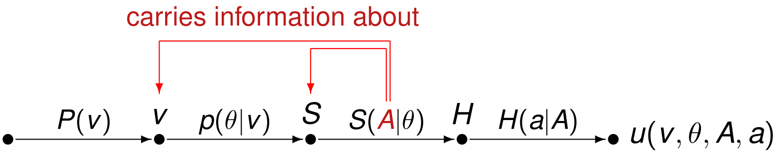
$$\langle \mathscr{G}, S, H \rangle \models F \Rrightarrow R \tag{2.23}$$

iff

$$(\mathcal{X}_i, p_i) \models \{F\} \Rrightarrow R \tag{2.24}$$

for the causal network defined by P, S, H and $\Omega, \Theta_s, \mathcal{F}$.

Natural Information in Signalling Games



- The order of actions defines a linear causal process.

Section 3

Natural Information and Implicatures

Implicatures

[Grice, 1989, p. 86]

What is an implicature?

"... what is implicated is what is required that one assume a speaker to think in order to preserve the assumption that he is observing the Cooperative Principle (and perhaps some conversational maxims as well), ..."

Definition of Implicatures in a Set if Support Problems

Definition 11 (Implicature)

Let $\mathcal{S}$ be a given set of support problems with joint decision problem $\langle(\Omega, P_H), \mathcal{A}, u\rangle$. Let $A, R \subseteq \Omega$ be two propositions with $A \in \mathrm{Op}_\sigma$ for some $\sigma \in \mathcal{S}$. Then we set:

$$A \mathrel{+\!>} R \Leftrightarrow \forall\sigma \in \mathcal{S}\,(A \in \mathrm{Op}_\sigma \rightarrow P_S^\sigma(R) = 1), \tag{3.25}$$

- Op_σ: Set of optimal answers in situation σ.

From Support Problems to Signalling Games

1. Let $\mathscr{G}$ be an interpreted signalling game $\langle \Omega, \Theta_S, P, \mathcal{F}, \mathcal{A}, u, [\![\,.\,]\!] \rangle$ for which $P(\theta) > 0$ for all $\theta \in \Theta_S$.
2. Let $\mathcal{S}$ be a set of support problems $\sigma = \langle \Omega, P_S, P_H, \mathcal{F}, \mathcal{A}, u, [\![\,.\,]\!] \rangle$ for which there is a bijection $\Theta_S \longrightarrow \mathcal{S}$, $\theta \mapsto \sigma_\theta$ such that for all $\theta \in \Theta_S$:

$$P_S^{\sigma_\theta}(v) > 0 \textit{ iff } P(v|\theta) > 0.$$

3. Then, the canonical solution (S, H) to $\mathcal{S}$ satisfies: If $R \subseteq \Omega$ and $F \in \mathrm{Op}_\sigma$ for some $\sigma \in \mathcal{S}$, then

$$\langle \mathcal{S}, S, H \rangle \models F +\!> R \textit{ iff } \langle \mathscr{G}, S, H \rangle \models F \Rrightarrow R. \qquad (3.26)$$

Remarks

1. $\langle \mathscr{G}, S, H \rangle \models A \Rrightarrow R$ does not depend on H.
2. If we interpret probabilities in support problems as subjective probabilities, then the result says:

$\Rightarrow$ Even if backward induction leads to objectively sub–optimal actions, the implicatures nevertheless are true!

Generalisation of the Notion of Implicatures in Signalling Games

Definition 12

Implicature

$$\langle \mathscr{G}, S, H \rangle \models Y +> X \iff \langle \mathscr{G}, S, H \rangle \models Y \Rrightarrow X, \tag{3.27}$$

i.e. *Y implicates X* iff *Y naturally means X*.

Section 4

Implicatures and Reasoning about Each Other

[Benz, 2010]

Defining Non–natural Meaning

[Grice, 1989, p. 92]

Definition 13 (meaning$_{nn}$)

"[*S*] meant something by uttering *x*" is true, iff for some audience *A*, [*S*] uttered *x* intending:

1. *A* to produce a particular response *r*
2. *A* to think (recognize) that [*S*] intends **(1)**
3. *A* to fulfil **(1)** on the basis of his fulfilment of **(2)**.

On the basis of is to be understood:

1. The addressee's thinking that *S* intends him to respond with *r* is at least part of his reason to produce *r*;
2. It is not merely a cause for his producing *r*.

Defining Non–natural Meaning

[Grice, 1989, p. 92]

Definition 13 (meaning$_{nn}$)

"[S] meant something by uttering x" is true, iff for some audience A, [S] uttered x intending:

1. A to produce a particular response r
2. A to think (recognize) that [S] intends **(1)**
3. A to fulfil **(1)** on the basis of his fulfilment of **(2)**.

On the basis of is to be understood:

1. The addressee's thinking that S intends him to respond with r is at least part of his reason to produce r;
2. It is not merely a cause for his producing r.

Causation not enough

[Grice, 1989, p. 92]

Example 14 (Amusement)

Weaker causality condition entails that a speaker *S* would $mean_{nn}$ something by doing something *x* with the intended effect of:

1. *A* to be amused
2. *A* to think that [*S*] intended him to be amused
3. *A* to be amused (at least partly) as a result of his thinking that [*S*] intended him to be amused.

But: to cause somebody to be amused by making him recognises that one tries to make him amused is not a case of communication.

Varieties of Communicated Meaning

- Scalar implicature generally assumed not to be context sensitive.
- Can possibly be inferred by default rule.

$\Rightarrow$ Likely not to involve recognition of speaker's intention.

- Relevance implicature generally assumed to be context sensitive.
- Can not be inferred by default rule.

$\Rightarrow$ Likely to involve recognition of speaker's intention.

Implicatures

[Grice, 1989, p. 86]

What is an implicature?

"... what is implicated is what is required that one assume a speaker to think in order to preserve the assumption that he is observing the Cooperative Principle (and perhaps some conversational maxims as well), ..."

Seems to imply by definition that the hearer has to recognise the speaker's intentions!

The Standard Scalar Implicatures

[Franke, 2009]

Example 15

Some of the boys came to the party. $+>$ not all came

$$H_0 = \left\{ \begin{array}{l} \textit{some} \mapsto \theta_{\exists\neg\forall}, \theta_{\forall} \\ \textit{all} \mapsto \theta_{\forall} \end{array} \right\} \tag{4.28}$$

$$S_1 = \left\{ \begin{array}{l} \theta_{\exists\neg\forall} \mapsto \textit{some} \\ \theta_{\forall} \mapsto \textit{all} \end{array} \right\} \tag{4.29}$$

$$H_2 = \left\{ \begin{array}{l} \textit{some} \mapsto \theta_{\exists\neg\forall} \\ \textit{all} \mapsto \theta_{\forall} \end{array} \right\} \tag{4.30}$$

$$S_3 = S_1 \tag{4.31}$$

$$H_4 = H_2. \tag{4.32}$$

IBR Lower Limit

1. Shortest IBR path to stability:
    1. R_0–S_1–R_2–S_3 if reasoning starts with hearer.
    2. S_0–R_1–S_2–R_3–S_4 if reasoning starts with speaker.

⇒ Hearer has to take speaker strategy into account at least once!

Undercutting IBR Limit

- We show that relevance implicatures can be calculated in a sequence shorter than that predicted by the IBR model.
- Derivation of implicature does not involve reasoning about speaker expectations or strategies!
- We use:

Lemma 16

Let $\mathcal{S}$ be a set of support problems with joint decision problem $\langle(\Omega, P_H), \mathcal{A}, u\rangle$. Assume furthermore that

1. *S is an expert for every $\sigma \in \mathcal{S}$,*
2. $\forall v \in \Omega\, \exists \sigma \in \mathcal{S}\; P_S^\sigma(v) = 1$.

Let $\sigma \in \mathcal{S}$ and $A, R \subseteq \Omega$ be two propositions with $A \in \mathrm{Op}_\sigma$. Then, it follows that:

$$A +> R \text{ iff } A^* \cap A^+ \subseteq R.$$

Out-of-Petrol

Example 17

H: I am out of petrol.
S: There is a garage round the corner. (*A*)
$+>$ The garage is open. (*R*)

Model:

Ω	G	R	go-to-g	search
w_1	$+$	$+$	1	ε
w_2	$+$	$-$	0	ε
w_3	$-$	$-$	0	ε

Assumption: $EU_H(\text{go-to-g}|G) > \varepsilon$. (go to garage better than randon search)

Recognition of Intentions

1. $A^* := \{v \in A \mid P_H(v) > 0\}$ only depends on answer A and P_H.
2. $A^+ = \bigcap\{O(a) \mid a \in \mathscr{B}(A)\}$ only depends on A, P_H, and u.

Follows:

- The implicature $A^* \cap A^+$ only depends on A, P_H, and utilities u.

⇒ The implicature can be calculated without considering the speaker's intentions.

- That the answer has to be interpreted as recommendation follows from the context created by background question: 'Where to look for petrol?'

Reasoning

1. $A^* = \{w_1, w_2\}$ (garage round corner)
2. $\mathscr{B}(A) = \{\text{go-to-g}\}$. (set of optimal actions)
3. $O(\text{go-to-g}) = \{w_1\} = R$.
   (set of worlds in which go-to-g is optimal)
4. $\Rightarrow A^+ = \bigcap\{O(a) \mid a \in \mathscr{B}(A)\} = \{w_1\}$.
5. $\Rightarrow A^* \cap A^+ = \{w_1\} = R$.
6. $\Rightarrow G +> R$. (Lem. 16)
7. Hence $P_E^\sigma(R) = 1$ (speaker believes R)

Comparison with IBR model

1. Shortest IBR path to stability:
   1 R_0–S_1–R_2–S_3 if reasoning starts with hearer.
   2 S_0–R_1–S_2–R_3–S_4 if reasoning starts with speaker.
2. Path in OA model: R_0–S_1.

Conclusion

Answer to questions about the recognition of intentions:

- For normal communication, the hearer does not need to take into account the speaker's state.
- Implicatures can be calculated on the basis of literal meaning, hearer's expectations, and joint purpose of talk exchange.

$\Rightarrow$ It does not have to be explicit.

- It is following rather than preceding interpretation.

The End

Literature:

Benz, A. (2009).
Outline of the foundations for a theory of implicatures.
In Benz, A. and Blutner, R., editors, *Papers on Pragmasemantics*, volume 51 of *ZAS Papers in Linguistics*, pages 153–201. Center for General Linguistics, Berlin.

Benz, A. (2010).
Implicatures and the recognition of speaker intentions in a game theoretic model.
In Miranda, L. A. P. and Madariaga, A. I., editors, *Advances in cognitive science: learning, evolution, and social action. Proceedings of the ILCLI international workshop on cognitive science*, pages 15–34. University of the Basque Country Press, Bilbao.

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

ILLC Dissertation Series DS-2009-11.

Fudenberg, D. and Tirol, J. (1991).
Game Theory.
The MIT Press, Cambridge, MA.

Grice, H. P. (1957).
Meaning.
Philosophical Review, 66:377–388.

Grice, H. P. (1989).
Studies in the Way of Words.
Harvard University Press, Cambridge MA.

Jäger, G. and Ebert, C. (2009).
Pragmatic rationalizability.
In Riester, A. and Solstad, T., editors, *Proceedings of Sinn und Bedeutung*, volume 13, pages 1–15.

Sperber, D. and Wilson, D. (2004).
Relevance theory.

In Horn, L. and Warden, G., editors, *The Handbook of Pragmatics*, pages 607–631. Blackwell Publishing, Oxford.