Game Theory and Gricean Pragmatics Lesson IV

Anton Benz

Zentrum für Allgemeine Sprachwissenschaften
ZAS Berlin

Course Overview

- Lesson 1: Introduction
  - From Grice to Lewis
  - Relevance Scale Approaches
- Lesson 2: Signalling Games
  - Lewis‘ Signalling Conventions
  - Parikh‘s Radical Underspecification Model
- Lesson 3: The Optimal Answer Approach I
- Lesson 4: The Optimal Answer Approach II
  - Decision Contexts with Multiple Objectives
  - Comparison with Relevance Scale Approaches

The Optimal Answer Approach II

Lesson 4 – April, 5th

Overview of Lesson IV

- Implicatures in Decision Problems with Multiple Objectives
- Relevance Scale Approaches
- Three Negative Results
  - RSA can‘t avoid misleading answers
  - RSA can‘t avoid unintended implicatures
  - Optimisation of relevance not a conversational maxim

Implicatures in Decision Problems with Multiple Objectives

Main Examples - Answers

Peter: I have to buy wine for our dinner banquette. I get into trouble with our secretary if I spend too much money on it. We still have some French wine. Where can I buy Italian wine?

Bob: At the Wine Centre.

+> Peter can buy Italian wine at a low price at the Wine Centre.

Main Examples – Embedded Questions

In the afternoon Ann tells Bob that Peter bought some Italian wine but it was obviously completely overpriced. Bob gets very angry about it.

Ann: “Maybe, it was not his fault.”

Bob: “Oh, Peter, knows where he can buy Italian wine.”

+> Peter knows where he can buy Italian wine at a low price.

- Observation:
  - Implicatures depend on contextually salient preferences.
  - Preferences are not introduced by question.
- Goal:
  - Explain impicatures for both examples.
  - Derive explanation for embedded questions from model for answers to direct questions.
- Methodology:
  - Optimal Answer Approach

Hip Hop at Roter Salon

J: Is the Music in Roter Salon **ok**?

Direct reference to speaker‘s preferences.

Italian Wine

Peter: Where can I buy Italian wine?

No reference to speaker‘s preferences.

Multiple Attributes

- **Observation:** Often, preferences depend only on a finite number of attributes a_i of outcomes s.

$$u(s) = f(a_1(s),\ldots,a_n(s))$$

- **Idea** (Italian wine)**:**
  - The question predicate defines an attribute.
  - Other attributes may be added from context.
  - f must be inferred from world knowledge and context.
- **Optimal Answers:** Calculated as before.

Italian Wine (Price)

- $a_1(s) = 0$: $s \models$ Peter didn‘t buy It. W.
- $a_1(s) = 1$: $s \models$ Peter bought It. W.
- $a_2(s) = 0$: $s \models$ Price was high.
- $a_2(s) = 1$: $s \models$ Price was low.
- $f(0,i) < f(1,0) < f(1,1)$
- Assumption: $\forall i,j\ \exists s\ a_j(s) = i$

Variations on Italian Wine

Peter, the office assistant, was sent to buy Italian wine for an evening dinner.

1. In the afternoon Ann tells Bob that Peter went shopping but that he returned without wine. Bob gets very angry about it.

   Ann: “Maybe, it was not his fault.”

   Bob: “Oh, Peter, knows where he can buy Italian wine.”

1. In the afternoon Ann tells Bob that Peter bought some Italian wine but it was obviously completely overpriced. Bob gets very angry about it.

   Ann: “Maybe, it was not his fault.”

   Bob: “Oh, Peter, knows where he can buy Italian wine.”

1. In the afternoon Ann tells Bob that Peter bought some Italian wine but it took a long time because he went to one of the wine shops in the centre and he was caught in the city traffic. Bob gets very angry about it.

Ann: “Maybe, it was not his fault.”

Bob: “Oh, Peter, knows where he can buy Italian wine.

Example (attribute unrelated to buying event)

Peter visits Ann and Bob. He is obviously very excited and has to tell Ann and Bob about it. In the metro he sat opposite of a very nice and attractive Italian woman. She talked with her girl friend. So Peter learned that her name is Maria and that she jobs at an Italian wine shop near the station. He immediately got excited about her but he had to leave the subway and there was no chance to get her attention. They talk quite some while about this event and Peter’s chances to get this girl. After he left, Ann says to Bob: ‘Poor Peter, he will not meet her again!’ Bob: ‘Peter knows where he can buy Italian wine.’

Intuition

„X knows QUESTION“ is true

iff

X is an expert who can answer QUESTION.

Knowing an Answer

E knows (in an absolute sense) an optimal answer in world w iff

1. $P_E(w)>0$
1. $\exists\, a \in \mathcal{A}\ P_E(O(a))=1$

with $O(a) := \{v \in \Omega |\ \forall b \in \mathcal{A}\ u(b,v) \leq u(a,v)\}$

Towards an Interpretation of Embedded Questions

E knows where/when/ E can do φ. (A)

$$\Rightarrow$$

$$\exists\, a \in A\ P_E(\Omega^* \cap O(a))=1$$

Ω^*: common ground between speaker and hearer.

Example
(with partial information)

Bob ordered Peter, the office assistant, to buy Italian wine for an evening dinner. In a break Ann tells him that Peter came back from town but without wine. Bob gets very angry about it, such that Ann replies: “You know that the transportation union is on strike for weeks now. Maybe, he just didn’t find a shop which still has Italian wine.” Bob answers: “No, Peter, knows where he can buy Italian wine. I told him this morning that the Wine Centre received foreign wine, he just has to cycle a bit further. I was there at 11 o’clock. They have Italian wine.”

Relevance Scale Approaches

Game and Decision Theory

- **Decision theory:** Concerned with decisions of **individual** agents
- **Game theory:** Concerned with interdependent decisions of **several** agents.

Basic Issue

If Gricean Pragmatics can be modelled in:

- Decision Theory: Non-interactional view sufficient.
- Game Theory but not Decision Theory: Interactional view necessary!
  - H.H. Clark‘s Interactional Approach
  - Alignment Theory (Pickering, Garrod)
  - Conversational Analysis

Relevance Scale Approach
(with real valued relevance measure)

- Let $\mathcal{M}$ be a set of propositions.
- R : $\mathcal{M} \rightarrow \Re$ real valued function with

R(A) $\leq$ R(B) $\Leftrightarrow$ B is at least as relevant as A.

➢ then A +> $\neg$B iff R(A) < R(B).

Two Types of Relevance Scale Approaches

- Argumentative view: Arthur Merin
- Non-Argumentative view: Robert van Rooij
  - Relevance Maximisation
  - Exhaustification
- We concentrate on van Rooij's early (2003, 2004) relevance scale approach.
- All results apply to van Rooij-Schultz (2006) exhaustification as well.

General Situation

We consider situations where:

- A person *I*, called inquirer, has to solve a decision problem $((\Omega, P), A, u)$.
- A person *E*, called expert, provides *I* with information that helps to solve *I*'s decision problem.
- P_E represents E's expectations about Ω at the time when she answers.

Support Problems

$$\begin{array}{ccccc}
\text{Expert E} & & \text{I decides} & & \text{Evaluation} \\
\text{answers} & & \text{for action} & & \\
\downarrow & & \downarrow & & \downarrow \\
\bullet & \xrightarrow{A} & \bullet & \xrightarrow{a} & \bullet \\
\uparrow & & \uparrow & & \uparrow \\
\text{expectations} & & \text{expectations} & & \text{utility} \\
\text{of E} & & \text{of I} & & \text{function} \\
(\Omega, P_E) & & (\Omega, P_I) & & u(v, a)
\end{array}$$

Assumptions

- The answering expert E tries to maximise the relevance of his answer.
- Relevance is defined by a real valued function R: $\wp(\Omega) \rightarrow \Re$.
- R only depends on the decision problem $((\Omega, P), A, u)$.
- E can only answer what he believes to be true.

Sample Value of Information

(Measures of Relevance I)

New information A is relevant if

- it leads to a different choice of action, and
- it is the more relevant the more it increases thereby expected utility.

Sample Value of Information

- Let ((Ω, P),A,u) be a given decision problem.
- Let a* be the action with maximal expected utility before learning *A*.

Possible definition of **Relevance** of *A*:

$$UV(A) = \max_{a \in \mathcal{A}} EU(a|A) - EU(a^*|A).$$

(Sample Value of Information)

Utility Value
(Measures of Relevance II)

Possible alternative e.g.:

New information A is relevant if

- it increases expected utility.
- it is the more relevant the more it increases it.

$$UV'(A) = \max_{a \in \mathcal{A}} EU(a|A) - \max_{a \in \mathcal{A}} EU(a).$$

The Italian Newspaper Example

Somewhere in the streets of Amsterdam...

a) J: Where can I buy an Italian newspaper?
b) E: At the station and at the Palace but nowhere else. (SE)
c) E: At the station. (A) / At the Palace. (B)

Answers

Assumptions:

1. $P_I(A) > P_I(B)$
2. *E* knows that $A \wedge B$, i.e. $P_E(A \cap B)=1$.

Then:

- With sample value of information: Only B is relevant.
- With utility value: A, B, and $A \wedge B$ are equally relevant.

- Assume now that *E* learned that:

  (¬A) there are no Italian newspapers at the station.

- With sample value of information: ¬A is relevant.
- With utility value: the uninformative answer is the most relevant answer.

Need: Uniform definition of relevance that explains all examples.

In order to get a better intuition about *relevance*, we present a non-linguistic example of a decision problem.

We will see that *desired* information and *relevant* information are two different concepts.

A Decision Problem

- An oil company has to decide where to build a new oil production platform.
  - Given the current information it would invest the money and build the platform at a place off the shores of Alaska.
  - An alternative would be to build it off the coast of Brazil.
- Build a platform off the shores of Alaska. (act a)
- Build it off the shores of Brazil. (act b)

- The company decides for exploration drilling.
- Using sample value of information means:
  - Only if the exploration drilling gives hope that there is a larger oil field off the shores of Brazil, the company got relevant information.
- Using utility value of information:
  - Only if the exploration drilling rises the expectations about the amount of oil, the company got relevant information.

- Desired: Information that leads to the best decision.
  - Information is desired as long as it leads to optimal decision even if it confirms current decision or decreases expectations.

Relevant information $\neq$ desired information

Finally, we reconsider the Out of Petrol Example and the two opposing inferences of implicatures.

We will see later, that no relevance scale approach can explain the implicatures and non-implicatures of the Out of Petrol example.

Implicatures and Relevance Scales

The Out of Patrol Example

- A stands in front of his obviously immobilised car.
  - A: I am out of petrol.
  - B: There is a garage around the corner. (G)

+> The garage is open (H)

An Explanation of the Out of Petrol Example

Set H^*:= The negation of H

1. B said that G but not that H^*.
2. H^* is relevant and $G \wedge H^* \Rightarrow G$.
3. Hence if $G \wedge H^*$, then B should have said $G \wedge H^*$ (Quantity).
4. Hence H^* cannot be true, and therefore H.

Problem: We can exchange H and H* and still get a *valid* inference:

1. B said that G but not that H.
2. H is relevant and $G \wedge H \Rightarrow G$.
3. Hence if $G \wedge H$, then B should have said $G \wedge H$ (Quantity).
4. Hence H cannot be true, and therefore H*.

- Let $\mathcal{M}$ be the set of admissible answers.
- Let R : $\mathcal{M} \longrightarrow \Re$ be either utility value or sample value of information.

Then:

$$A +> B \text{ iff } R(A) < R(B)$$

Makes the second inference true, i.e. G implikates that the garage is **closed**!

Three Negative Results

Basic Issue

Is there any relevance measure R such that:

- Optimisation of relevance leads to optimal answers.
- The criterion A +> B iff R(A) < R(B) makes correct predictions?

Main Results

- **Answerhood:** No relevance scale approach can avoid predicting misleading answers.
- **Implicatures:** No relevance scale approach can avoid predicting certain unintended implicatures.
- The notion of relevance that predicts correctly in the **Out-of-Patrol** example does not define a conversational maxim.

In the following, we present principled examples that cannot be explained by any relevance scale approach.

Relevance and Optimal Answers

First Negative Result

Strike in Amsterdam I

There is a strike in Amsterdam and therefore the supply with foreign newspapers is a problem. The probability that there are Italian newspapers at the station is slightly higher than the probability that there are Italian newspapers at the Palace, and it might be that there are no Italian newspapers at all. All this is common knowledge between I and E.

- Now E learns that

  (N) the Palace has been supplied with foreign newspapers.

- In general, it is known that the probability that Italian newspapers are available at a shop increases significantly if the shop has been supplied with foreign newspapers.

We describe the epistemic states by:

$P_I(A) > P_I(B)$ and $P_X(B \cap N) > P_X(A \cap N)$ for $x = I, E$.

It follows that going to the Palace (*b*) is preferred over going to the station (*a*):

$$EU_I(a, N) = \sum_{v \in N} P_I(v|N) \times u(v, a) = P_I(A \cap N);$$

$$EU_I(b, N) = \sum_{v \in N} P_I(v|N) \times u(v, b) = P_I(B \cap N).$$

E.g. Sample Value of Information predicts:

***N* is relevant.**

Strike in Amsterdam II

- We assume the same scenario as before but E learns this time that

(M) the Palace has been supplied with British newspapers.

Due to the fact that the British delivery service is rarely affected by strikes and not related to newspaper delivery services of other countries, this provides no evidence whether or not the Palace has been supplied with Italian newspapers.

- M provides no evidence whether or not there are Italian newspaper at the station (A) or the Palace (B)
- We assume therefore:

$$P_E(A) = P_E(M \cap A) > P_E(M \cap B) = P_E(B).$$

- $M \subseteq N$: Hence E knows N. Is N still a good answer?
- I's epistemic state hasn't changed

E.g. Sample Value of Information predicts:

N **is still relevant.**

Support problems

Definition 3.2 *A* support problem *is a five–tuple* $\langle \Omega, P_E, P_I, (\mathcal{A}, <), u \rangle$ *where* (Ω, P_E) *is a finite probability space and* $\langle (\Omega, P_I), (\mathcal{A}, <), u \rangle$ *a decision problem with tie break rule. We assume:*

$$\forall X \subseteq \Omega \; P_E(X) = P_I(X|K) \text{ for } K = \{v \in \Omega \,|\, P_E(v) > 0\}.$$

Italian Newspaper Properties

Let $K := \{v \in \Omega \mid P_E(v) > 0\}$, $EU = EU_l$

- $\forall a \in \mathcal{A}$ $EU(a|A) = EU(a|B) \Rightarrow R(A) = R(B)$
- $EU(a_\Omega|K) < EU(a_K|K) \Rightarrow R(\Omega) < R(K)$
- $R(K) = R(\Omega) \Rightarrow \forall C\ (K \subseteq C \subseteq \Omega \Rightarrow R(C) \leq R(\Omega))$

If R a relevance measure has properties 1-3, then we call R *monotone*.

- For a support problem σ the set of maximally relevant answers is given by:

$$\mathrm{MR}_\sigma := \{A \in Adm_\sigma \mid \forall B \in Adm_\sigma\ R(D_\sigma, B) \leq R(D_\sigma, A)\}$$

  - The set of optimal answers Op_σ is identical to the set of non-misleading answers.

First Negative Result

- Relevance scale approaches can't avoid misleading answers:

Lemma 5.2 *For each support problem* $\sigma \in \mathcal{S}$ *let* $R(D_\sigma, .) : \mathcal{P}(\Omega_\sigma) \rightarrow \mathbf{R}$ *be a monotone relevance measure. Then, for some* $\sigma \in \mathcal{S}$*:*

$$\mathrm{MR}_\sigma \cap \mathrm{Op}_\sigma = \emptyset.$$

Relevance and Implicatures

Second Negative Result

Relevance Scale Approach

- Let **M** be a set of propositions.
- Let $\leq$ be a linear well-founded pre-order on **M** with interpretation:

  $A \leq B \Leftrightarrow$ B is at least as relevant as A.

- then A +> B iff $A < B$.

Lemma

No relevance scale approach can satisfy the following set of implicatures:

1. $A_1 +\!> H_1$;
2. $A_2 +\!> H_2$;
3. not $A_1 +\!> \neg A_2$;
4. not $A_2 +\!> \neg A_1$;
5. not $A_1 +\!> H_2$.

Where it is assumed that A_1, A_2, H_1, and H_2 are pair-wise distinct propositions.

Proof:

1. $A_1 +\!> H_1$ implies $A_1 \prec \neg H_1$;
2. not $A_1 +\!> \neg A_2$ implies $A_1 \not\prec A_2$;
3. not $A_2 +\!> \neg A_1$ implies $A_2 \not\prec A_1$;
4. hence, $A_1 \approx A_2$ from the last two lines;
5. $A_2 +\!> H_2$ implies $A_2 \prec \neg H_2$;
6. hence, $A_1 \prec \neg H_2$ from lines 4 and 5;
7. not $A_1 +\!> H_2$ implies $A_1 \not\prec \neg H_2$, in contradiction to line 6.

An Example

(Argentine wine)

- Somewhere in Berlin... Suppose J approaches the information desk at the entrance of a shopping centre.
- He wants to buy Argentine wine. He knows that staff at the information desk is very well trained and know exactly where you can buy which product in the centre.
- E, who serves at the information desk today, knows that there are two supermarkets selling Argentine wine, a Kaiser’s supermarket in the basement and an Edeka supermarket on the first floor.
- J: I want to buy some Argentine wine. Where can I get it?
- E: Hm, Argentine wine. Yes, there is a Kaiser’s supermarket downstairs in the basement at the other end of the centre.

Propositions

1. A_1: There is a Kaiser's supermarket in the shopping centre.
2. A_2: There is an Edeka supermarket in the shopping centre.
3. H_1: The Kaiser's supermarket sells Argentine wine.
4. H_1: The Edeka supermarket sells Argentine wine.

No Relevance scale approach can explain this example.

The Argentine Wine Example is just a special case of the Out of Petrol Example.

Relevance and Conversational Maxims

Third Negative Result

The Out of Patrol Example

A stands in front of his obviously immobilised car.

A: I am out of petrol.

B: There is a garage around the corner. (G)

+> The garage is open (H)

The “correct” explanation

Set H*:= The negation of H

- B said that G but not that H*.
- H* is relevant and $G \wedge H^* \Rightarrow G$.
- Hence if $G \wedge H^*$, then B should have said $G \wedge H^*$ (Quantity).
- Hence H* cannot be true, and therefore H.

Is there a relevance measure that makes the argument valid?

The previous result shows that this is not possible if the relevance measure defines a linear pre-order on propositions.

The Posterior Sample Value of Information

Let O(*a*) be the set of worlds where action *a* is optimal.

If

1. the speaker said that A;
2. it is common knowledge that $\exists a\ P_E(O(a)) = 1$
3. for all $X \subseteq H^*$: $UV_I(X|A) > 0$,

then H is true.

Here, $UV_I(X|A)$ is the **sample value** of information **posterior** to learning A:

$$UV_I(X|A) := EU_I(a_{A\cap X}|A\cap X) - EU_I(a_A|A\cap X)$$

Application to Out-of-Petrol Example

- Let $X \subseteq H^*$ = ‘the garage is closed’
- A: ‘there is a garage round the corner’
- We assume that the inquirer has a better alternative than going to a closed garage.
- It follows then that $UV_I(X|A) > 0$, and our criterion predicts that

  H: ‘the garage is open’

  is true.

Standard expectations about *Relevance*:

Relevance

- is presumed to be maximised by the answering person.
- defines a linear pre-order on the set of possible answers.
- is definable from the receivers perspective.
- makes the ‘standard’ explanation in the out-of-patrol example valid.

Violated by Posterior Sample Value of Information

Relevance

- **is presumed to be maximised by the answering person.**
- **defines a linear pre-order on a set of possible answers.**
- **is definable from the receivers perspective.**
- **makes the 'standard' explanation in the out-of-patrol example valid.**

Relevance and Conversational Maxim

Conversational Maxim:

- presumed to be followed by the speaker.
- Necessary for calculating appropriate answers and implicatures.

$\Rightarrow$ The relevance measure defined by the posterior sample value of information does not define a conversational maxim.

The End