Message Exchange games, meanings and blindness

Nicholas Asher*, Soumya Paul[1], Antoine Venant[1]

*CNRS, Institut de Recherche en Informatique de Toulouse (IRIT)
[1]IRIT, Université de Toulouse
Research supported by ERC Grant 269427

August 2015

What we saw first lecture

- Gricean communication
- Problems with non-cooperative settings
- troublesome backwards induction argument

What we saw second lecture

- Basics of infinitary games over infinitary languages
- BM games, GS games, ME games
- winning conditions in the Borel Hierarchy

This lecture: coming back to conversation

- Meanings
- games over sequences of models
- analysis of examples
- commitment based semantics
- problems with determinacy

Consistency and meaning

- we saw last time that consistency is a desirable even necessary constraint on a winning condition.
- but supposing consistency is enforced over all possible strings. Then,
- the consistent strings in $(V_0^+ \cup V_1^+)^\omega$ determine a set of models (consistent sets of sentences or discourse moves).
- so this constraint determines meaning for $(V_0^+ \cup V_1^+)^\omega$.
- we could maybe even extend this to gauge when speakers make moves that are contrary to expectations (implicatures)

What sort of meanings?

- notice the gulf here between credible meaning (believability) and what speakers put forward as public commitments.
- the prosecutor may not believe Bronston at all, if he says he never had a bank account. (Maybe he has gotten records from the bank).
- meaning of discourse moves and conversations as public commitments by the dialogue's interlocutors.
- dialogue moves become actions that transform the commitments of the interlocutors. Transforms models

Example

i : I have a bank account in $M \rightsquigarrow C_i^M + C_i$bank) ($M \oplus$ *I have a bank account*)

Why commitment semantics isn't so simple

- Deepak's comment in the Dialogue Workshop: an agent is not a static object but evolves.
- When one player makes a new commitment, other players' commitments shift too.

Example

C: N. isn't coming to the meeting (π_1). It's been cancelled (π_2).

A: That's not why N. isn't coming. He's sick.

C: I didn't say that N. wasn't coming because the meeting was cancelled. The meeting is cancelled because N. isn't coming.

What's going on?

- C makes an ambiguous discourse commitment. Explanation vs. Result
- What does A need to take on board to object? $C_A C_C \text{Explanation}(\pi_1, \pi_2)$
- So A's commitments change when C's commitments do on the first turn.
- Similarly when C responds he needs to make the commitment $C_C C_A C_C \text{Explanation}(\pi_1, \pi_2)$
- it looks like we can get very deep nestings of commitments. More on this in lecture 4.

An analysis of one of our examples: Bronston

- The conversational goal: Eventually, $C^*_{P,B}C_B\text{bank} \vee C^*_{P,B}C_B\neg\text{bank}$
- P first asks a question; he commits to a partition of the state space
- B responds by committing to one of the elements of the partition.
- Update the model first with: $C_BC_P(\{\text{bank}, \neg\text{bank}\})$ and then update with $C_B\neg\text{bank}$.
- P's follow up question implies: $C_PC_B\text{QAP}(\{\text{bank}, \neg\text{bank}\}, \neg\text{bank})$
- and updates the model with a new commitment: $C_P(\{\mathit{Past}\ \text{bank}, \mathit{Past}\ \neg\text{bank}\})$.
- Now *B* makes an ambiguous discourse move: he could be answering the question, or he could be evading it.

Exploiting semantics: a more general notion of consistency

- A play ρ of an ME game in $(V_0^+ \cup V_1^+)^\omega$, with V a language of discourse moves with a well-defined semantics is consistent for player i iff for every model $\mathfrak{A}$ there is no finite prefix x of ρ such that $\mathfrak{A} \oplus \|x\| \models C_i\bot$. .
- This constraint is easily seen to be Π^0_2 but more general than what we had before.
- It is adaptable to any logic associated with the semantics for the base language of discourse moves that is compact.
- can also explore ω consistency in this framework: $\{\exists x \in \mathbb{N} F(x), \neg F(0), \neg F(1), ...\}$ is an ω inconsistent set.

Some tricky questions about ω consistency

- how do we define the appropriate limit point of the update sequence? If the process is monotonic, then $\mathfrak{A} \oplus \rho = \bigcap_{n \in \omega} \mathfrak{A}\rho_n$.
- A play ρ of an ME game in $(V_0^+ \cup V_1^+)^\omega$ is ω-consistent for player i iff for every model $\mathfrak{A}$ $\mathfrak{A} \oplus \|\rho\| \models \neg C_i \bot$.
- open question: is the update process always monotonic? Depends on what discourse moves you allow and what semantics you have for them!
- it's unclear, for instance, that the semantics of acknowledgment and especially corrections can be monotonic, in which case our definition of the limit point of the update sequence is not correct.

Consistent but ω inconsistent formulas

Example

N. I'm sleepy π_1

S. OK. π_2

- acknowledgments are tricky.
- Given a weak semantics for assertions, $[\pi_1]_N \top \models C_N \text{sleepy}$
- π_2 actually is a complex discourse move: it's $\pi_3 : \text{Acknowledge}(\pi_1, \pi_2)$
- $[\text{Acknowledge}(\pi_1, \pi_2)]_S [\pi_1]_N \top \models C_S C_N \text{sleepy}$
- $\{\neg C^*_{N,S} C_N \text{sleepy}\} \cup \bigcup_{n \in \omega} \{C^n_{N,S} C_N \text{sleepy}\}$ is consistent but ω inconsistent. Every finite subset of this set is satisfiable but the whole is not.
- A continuation of our example with an infinite string of alternating acknowledgments by N and S on this semantics yields $C^*_{N,S} C_N \text{sleepy}$.

Correspondence results

- with our semantics, we can now see an ME game as played. not over a set of linguistic strings $x_1.y_1.x_2.y_2\ldots$, but over ω sequences of structures:
- $\langle \mathfrak{A} \oplus \|x_1\|, (\mathfrak{A} \oplus \|x_1\|) \oplus \|y_1\|, ((\mathfrak{A} \oplus \|x_1\|) \oplus \|y_1\|) \oplus \|x_2\|, \ldots \rangle$
- fix the structures to be a set by fixing a common (super) domain for all possible structures.
- the game space is now a quotient structure: $(V_0^+ \cup V_1^+)^\omega / \|.\|$ Strings that yield the same sequences of model updates form an equivalence class.
- Every interpretable ω sequence of discourse moves by 0 and 1 yields a ω sequence of structures. Not every sequence of structures is expressible by an interpretable ω length sequence.

Determinacy

- Leaning on Martin's proof, we can show that ME games with Borel determinable winning conditions are determined.
- So why should the loser ever play?
- Maybe he doesn't know that he's going to lose (chess).
- Or maybe there are unknown unknowns.

Two ways to deal with a no-win situation

- Captain Kirk when facing a no-win game (Kobayashi Maru) did the right thing: he changed the game. (K: I don't believe in a no-win situation)
- ME games are particularly simple objects: the parameters of variation are the vocabulary and the winning condition.
- We look at two transformations of games each one dealing with one of these parameters.

Misdirection

goals are hard to figure out

- What's Bronston doing with his indirect reply to P's question: *the company had one there for about 6 months*?
- He's looking cooperative. But he's actually fishing (phishing?).
- In fact, it's not easy to know exactly what the Prosecutor's goal is in the conversation.
- E.g. is it (i): $\Diamond(C_B\text{bank} \vee C_B\neg\text{bank} \vee C_B\text{evade-question})$ for *any structure* that admits an update with the conversational sequence in Example 2?
- Or is it (ii): $\Diamond(C_B\text{bank} \vee C_B\neg\text{bank} \vee C_B\text{evade-question})$ in *preferred* structures that admit an update with the conversational sequence?
- difference between QAP and IQAP in SDRT (Asher & Lascarides, 2003).
- difference between entailment and defeasible entailment.

Misdirection

Bronston's adjustments

- Suppose *P* was playing (ii).
- If *B* was playing the complement of (ii), then he's doomed.
- But if *P* is playing (ii), *B* can shift the game to a non zero sum game by adopting the complement of (i). And he still has a chance to avoid conviction: *I never said that*.
- $\overline{\text{(ii)}} \cap \text{(i)} \neq \emptyset$; Bronston can implicate an answer without that answer holding in all possible updates.
- So *B* can shift from a "no win" situation for him to a situation in which *P* is happy by having attained his goal and *B* has attained his.

Misdirection

Adding incompleteness

- Let's go back to the point about goals being hard to discern.
- In strategic situations, it can behoove a player to hide his true winning conditions.
- So *B* might have (at least) 2 types for *P*, "entailment *P*" and "implicature *P*".
- It is rational to try to see if *P* is "implicature *P*", so play that game first.
- If not, well... *B* was in trouble anyway.
- Luckily for *B*, *P* was of the softer type. Given that *P* ultimate goal was to make sure that *B*'s ultimate winning condition (avoid conviction) was unattainable, that was a mistake.

Playing with different moves (Asher & Paul, 2013)

Theorem

Let X and Y be two alphabets such that $X \subsetneq Y$. We have the following in the Borel hierarchy:

$$
\begin{array}{cccccccccccc}
 & & \circlearrowright & & & & \circlearrowright & & \circlearrowright & \circlearrowright & & \circlearrowright \\
\Sigma^0_1 & \longrightarrow & \Sigma^0_2 & & \Sigma^0_3 & \longrightarrow & \Sigma^0_4 & \cdots & \Sigma^0_\omega & \Sigma^0_{\omega+1} & \cdot\cdot & \Sigma^0_{\omega_1} \\
\circlearrowright & & & & \circlearrowright & & & & \circlearrowright & \circlearrowright & & \circlearrowright \\
\Pi^0_1 & & \Pi^0_2 & \longrightarrow & \Pi^0_3 & & \Pi^0_4 & \cdots & \Pi^0_\omega & \Pi^0_{\omega+1} & \cdot & \Pi^0_{\omega_1}
\end{array}
$$

Proof idea: look at encodings of winning conditions in the reduced vocabulary within the expanded vocabulary. The result is a Cantor-set like construction of Win_X within Y^ω. E.g. consider $\{a,b\}^\omega$ within $\{a,b,c\}^\omega$

What the encoding does

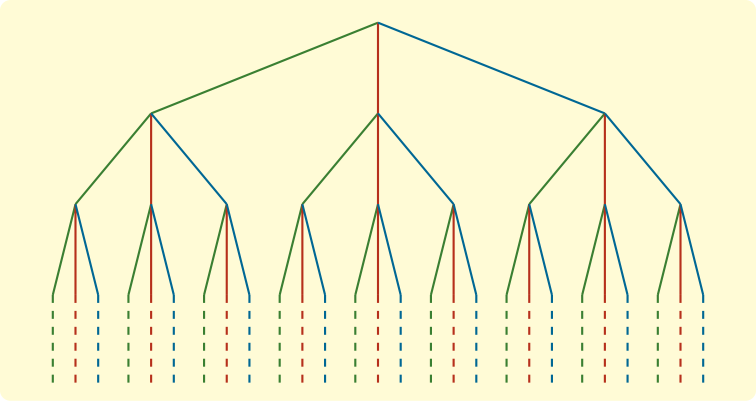

Figure: the Cantor set

Applications

- "unknown unknowns". Quayle is unaware of some moves by Bentsen, and so can neither anticipate them or have had the time to devise an adequate retort.
- this is a sort of epistemic blindness or unawareness about conversational moves.
- Bentsen is playing in a larger game arena $Y_0 \cup Y_1$ while Quayle is playing in the smaller game arena $X_0 \cup X_1$.
- Unbeknownst to Quayle, Bentsen has "shifted" the conversational game on him.

Back to Quayle

Dan Quayle's famous gaffe and Lloyd Bentsen's even more famous response again

Quayle: ... What kind of qualifications does Dan Quayle have to be president," "What kind of qualifications do I have,"... I have as much experience in the Congress as Jack Kennedy did when he sought the presidency.
Bentsen: Senator, I served with Jack Kennedy. I knew Jack Kennedy. Jack Kennedy was a friend of mine. Senator, you're no Jack Kennedy.

- The theorem says his winning objective was more complex because it had to be coded in the vocabulary of the larger game. $\Pi^0_2 \mapsto \Pi^0_3$, a much more complex set.

Comparing BM to ME games again

The situation in a BM game

- Σ_1 sets are always co-meager. Π_2 sets are countable intersections of co-meager sets and under some conditions are themselves co-meager. So by the BM theorem, a player can have a winning strategy for some Π_2 winning conditions.
- Quayle's winning condition sure looks co-meager. So by the BM theorem he has a winning strategy.
- On the other hand Π_3 sets are countable unions of Σ_2 that the construction in the Asher & Paul theorem guarantees to be meager. (Cantor sets have measure 0.)
- So Quayle's *Win* in the larger game is meaguer, and by the BM theorem, a player cannot have a winning strategy for such Π_3 winning conditions.

Comparing BM to ME games again

The situation in an ME game

- However, in ME games, meagerness of *Win* is not sufficient for the opponent to win.
- Π_3 conditions can have winning strategies. Quayle did have a way out.
- Details in Asher, Paul & Venant 2015b (available from authors). At least say, that you didn't mean to give the implicature. Thus you deny the presuppositions of the attack so that it cannot be felicitously made.
- But if you're caught off guard, you may not be able to respond to the attack
- Bensten shifted the game from a no-win situation for him to a winning situation in the ME game.

Conclusions on game shifts

- game shifts are to be expected, as players can not only dynamically update their commitments but also change their goals, if things aren't panning out as they thought.
- given that players may conceal moves (arguments), it may be that both players have shifted games and also moves they are unaware of.
- arguments and strategic conversations become a complicated business

Some benefits for linguists from games

- games furnish a new foundation for implicatures, independent of Gricean cooperativity
- games help linguists explore implicatures in new ways.
- contributions of discourse moves include implicatures like rhetorical relations
- sometimes the contribution is arguable or unclear (cf Valentino). We call such contributions (implicatures) *unsafe*.
- we model safety via certain sorts of equilibria in local games (roughly a pair of turns involving both players).

What's left to do

- looking more closely at discourse moves
- bridging the gap between the large and fine structure of dialogue