Combining TTR and game theory in dialogue modelling

Ellen Breitholtz, Robin Cooper, and Bill Noble

9 October, 2019

Outline

Games in a theory of language as action

Games in TTR

Social meaning games in GT

Argument games using topoi

Topoi and personae

A probabilistic model of topoi as social signals

Outline

Language as action

- Language as action (Austin, 1962; Lewis, 1969; Clark, 1996; Barwise and Perry, 1983)
- Agents need to coordinate action: coordination games (Lewis, 1969)

Two kinds of games

- Dialogue games build on techniques used in coordination games involving non-linguistic agents
- *Interaction games* in TTR, a type theory with records (Cooper, 2014; Breitholtz, 2014; Cooper, in prep)
- *Social meaning games* Burnett (2019), drawing on techniques from Game Theory (GT) Lewis (1969)
- Combining these types of games in terms of a theory of dialogue involving *Information State Update*: Ginzburg's *KoS* (Ginzburg, 2012)

Outline

Games in TTR

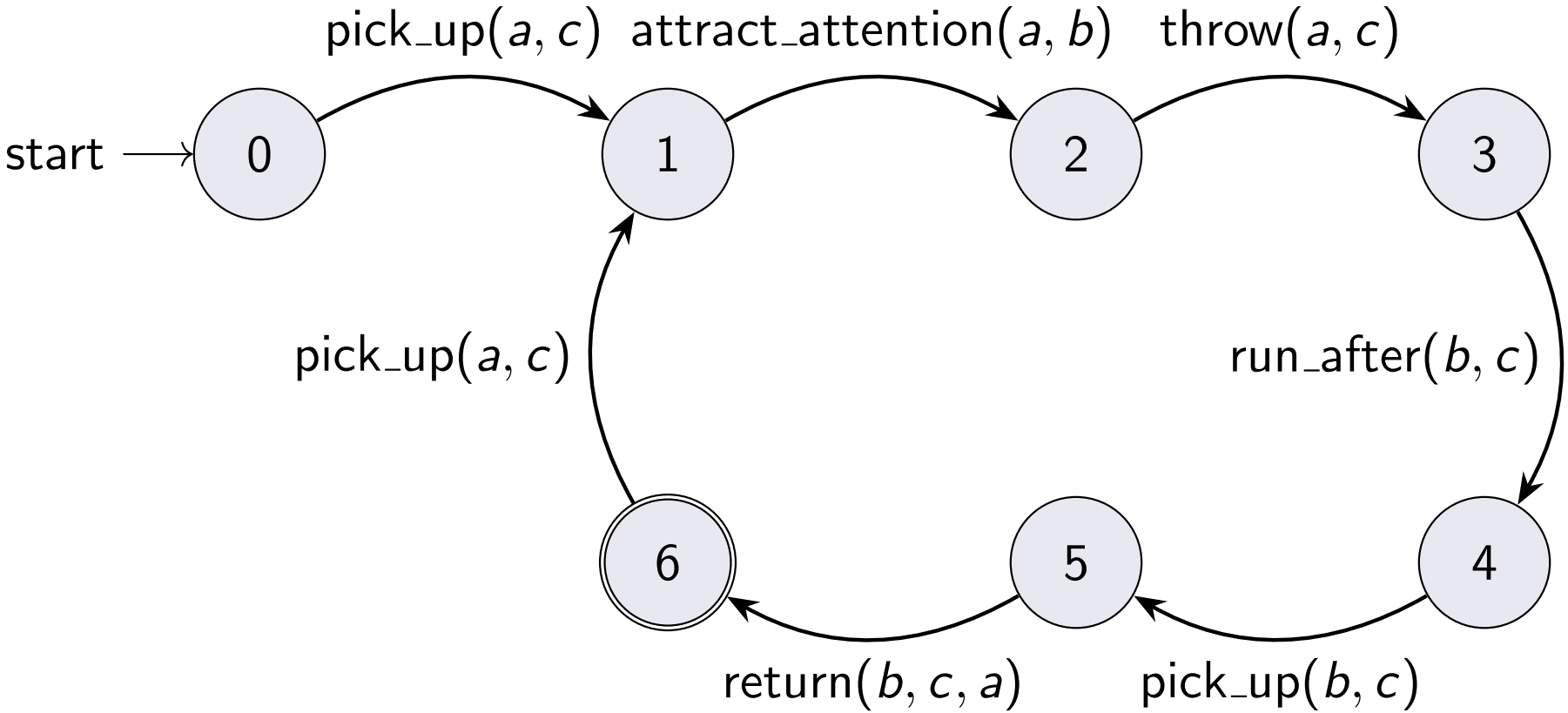
- Cooper (in prep), Ch. 1 (discussed here)
- Breitholtz (2014) in relation to enthymematic reasoning
- related to Ginzburg on genre and conversation types

String types

cf. work by Tim Fernando, e.g. Fernando (2015)

1. if T_1, $T_2 \in$ **Type**, then $T_1 ^\frown T_2 \in$ **Type**
   $a : T_1 ^\frown T_2$ iff $a = x ^\frown y$, $x : T_1$ and $y : T_2$
2. if $T \in$ **Type** then $T^+ \in$ **Type**.
   $a : T^+$ iff $a = x_1 ^\frown \ldots ^\frown x_n$, $n > 0$ and for i, $1 \leq i \leq n$, $x_i : T$
   ...

A game of fetch



$$(\text{pick\_up}(a,c)^\frown\text{attract\_attention}(a,b)^\frown\text{throw}(a,c)^\frown\text{run\_after}(b,c)^\frown \text{pick\_up}(b,c)^\frown\text{return}(b,c,a))^+$$

Information states and gameboards

- Information states (gameboards) are used by agents to keep track of where they are in the creation of an event belonging to a certain type
- each agent has their own view of the state of the game
- plays an essential role in coordination
- *information state* (Larsson, 2002) and *gameboard* (Ginzburg, 1994, 2012, originally Lewis, 1979) are adopted from the literature on dialogue
- we shall model information states as records and use 'gameboard' to refer to types of information states

The types *InfoState* and *InitInfoState*

$$\begin{array}{rl} \mathit{InfoState} & \left[\ \text{agenda} \quad : \quad [\mathit{RecType}]\ \right] \\ \mathit{InitInfoState} & \left[\ \text{agenda}{=}[] \quad : \quad [\mathit{RecType}]\ \right] \end{array}$$

Game of fetch (human, a, dog, b, and stick, c)

- game as a set of update functions corresponding to transitions in a finite state automaton
- an initial update function
  $\lambda r{:}\left[\text{agenda}{=}[]{:}[\mathit{RecType}]\right]$.
  $\left[\text{agenda}{=}[\left[\text{e:pick\_up}(a,c)\right]]{:}[\mathit{RecType}]\right]$
- a non-initial, non-final update function
  $\lambda r{:}\left[\text{agenda}{=}[\left[\text{e:pick\_up}(a,c)\right]]{:}[\mathit{RecType}]\right]$
  $\lambda e{:}\left[\text{e:pick\_up}(a,c)\right]$.
  $\left[\text{agenda}{=}[\left[\text{e:attract\_attention}(a,b)\right]]{:}[\mathit{RecType}]\right]$
- a final update function
  $\lambda r{:}\left[\text{agenda}{=}[\left[\text{e:return}(b,c,a)\right]]{:}[\mathit{RecType}]\right]$
  $\lambda e{:}\left[\text{e:return}(b,c,a)\right]$.
  $\left[\text{agenda}{=}[]{:}[\mathit{RecType}]\right]$

Game of fetch (with roles abstracted)

$$\lambda r^*{:}\begin{bmatrix} \mathsf{h} & : & \mathit{Ind} \\ \mathsf{c}_{\mathrm{human}} & : & \mathsf{human(h)} \\ \mathsf{d} & : & \mathit{Ind} \\ \mathsf{c}_{\mathrm{dog}} & : & \mathsf{dog(d)} \\ \mathsf{s} & : & \mathit{Ind} \\ \mathsf{c}_{\mathrm{stick}} & : & \mathsf{stick(s)} \end{bmatrix} .$$

$$\{\quad \lambda r{:}\big[\mathsf{agenda}{=}[]{:}[\mathit{RecType}]\big] \ .$$
$$\big[\mathsf{agenda}{=}[\big[\mathsf{e{:}pick\_up}(r^*.\mathsf{h},r^*.\mathsf{s})\big]]{:}[\mathit{RecType}]\big],$$
$$\lambda r{:}\big[\mathsf{agenda}{=}[\big[\mathsf{e{:}pick\_up}(r^*.\mathsf{h},r^*.\mathsf{s})\big]]{:}[\mathit{RecType}]\big]$$
$$\lambda e{:}\big[\mathsf{e{:}pick\_up}(r^*.\mathsf{h},r^*.\mathsf{s})\big] \ .$$
$$\big[\mathsf{agenda}{=}[\big[\mathsf{e{:}attract\_attention}(r^*.\mathsf{h},r^*.\mathsf{d})\big]]{:}[\mathit{RecType}]\big],$$
$$\ldots,$$
$$\lambda e{:}\big[\mathsf{e{:}return}(r^*.\mathsf{d},r^*.\mathsf{s},r^*.\mathsf{h})\big] \ .$$
$$\big[\mathsf{agenda}{=}[]{:}[\mathit{RecType}]\big]$$

$$\}$$

Type acts

judgements

specific $o :_A T$ "agent A judges object o to be of type T"

non-specific $:_A T$ "agent A judges that there is some object of type T"

queries

specific $o :_A T?$ "agent A wonders whether object o is of type T"

non-specific $:_A T?$ "agent A wonders whether there is some object of type T"

creations

non-specific $:_A T!$ "agent A creates something of type T"

Action rules

- also known as: licensing conditions, affordances (Gibson, 1979)
- $\underbrace{\varphi_1 \quad \cdots \quad \varphi_n}_{\psi}$
- $\varphi_1, \ldots, \varphi_n$ license/afford ψ
- $\varphi_1, \ldots, \varphi_n$ and ψ are characterized by type acts
- Note: ψ does not *follow* from $\varphi_1, \ldots, \varphi_n$. ψ is just something that is licensed or afforded by $\varphi_1, \ldots, \varphi_n$.

Action rules for Fetch

- $s_{i,A}$ represents A's current information state
- "Execute (contribute to the creation of a witness for) the type on the top of the agenda"
- $$\frac{s_{i,A} :_A \left[\text{agenda:} \begin{bmatrix}\text{fst:}\mathit{RecType} \\ \text{rst:}\mathrm{list}(\mathit{RecType})\end{bmatrix}\right]}{:_A s_{i,A}.\text{agenda.fst!}}$$

Action rules for Fetch, *contd*

- $s_{i+1,A}$ represents A's updated information state
- e^* represents a current event
- f is an update function of the game Fetch.
- "if a move of the game has just been executed put the type of an allowable next move on the agenda"
- $$\frac{f : (T_1 \to (T_2 \to Type)) \quad s_{i,A} :_A T_1 \quad e^* :_A T_2}{s_{i+1,A} :_A f(s_{i,A})(e^*)}$$
- "if you are in a state that can be updated by one of the games update functions without a triggering event, update accordingly"
- $$\frac{f : (T \to Type) \quad s_{i,A} :_A T}{s_{i+1,A} :_A f(s_{i,A})}$$

A problem

What do we do when games are non-deterministic, there is more than one update function that can be applied?

Outline

-ing/-in' variation as social cue

- Use of *-ing/-in'* verbal morphology (Labov, 2012, p. 22, cited by Burnett)
- use of *-ing/-in'* varies depending on context
- Burnett:*-ing/-in'* associated with social and individual characteristcs
  - *-in'* indicates 'friendly', but also possibly 'incompetent'
  - *-ing* indicates 'competent', but also possibly 'aloof'
- combinations of such (perceived) characteristics make up different social *personae*
- key concept in third wave sociolinguistics (Eckert, 2012)

Social meaning games

Burnett (2019)

Definition 4.1. *A **Social Meaning Game** is a tuple* $\langle \{S, L\}, \langle \mathbb{P}, > \rangle, M, C, [\cdot], Pr \rangle$ *where:*

1. *S and L are the players.* Two players
2. *$\langle \mathbb{P}, > \rangle$ is the **universe** (a relational structure), where*
   - *$\mathbb{P} = \{p_1, \ldots, p_n\}$ is a finite set of properties.* Properties such as 'friendly'
   - *$>$ is a relation on $\mathbb{P}$ that is irreflexive.*
3. *M is a finite set of **messages**.* ing/'in
4. *C is a measure function on M describing the **cost** of each message.*
5. *$[\cdot]$ is the **indexation** relation (to be described below).* e.g. 'in is friendly
6. *Pr is a probability distribution over sets of properties describing L's **prior beliefs** about S.* e.g. to what extent does L think Obama is friendly

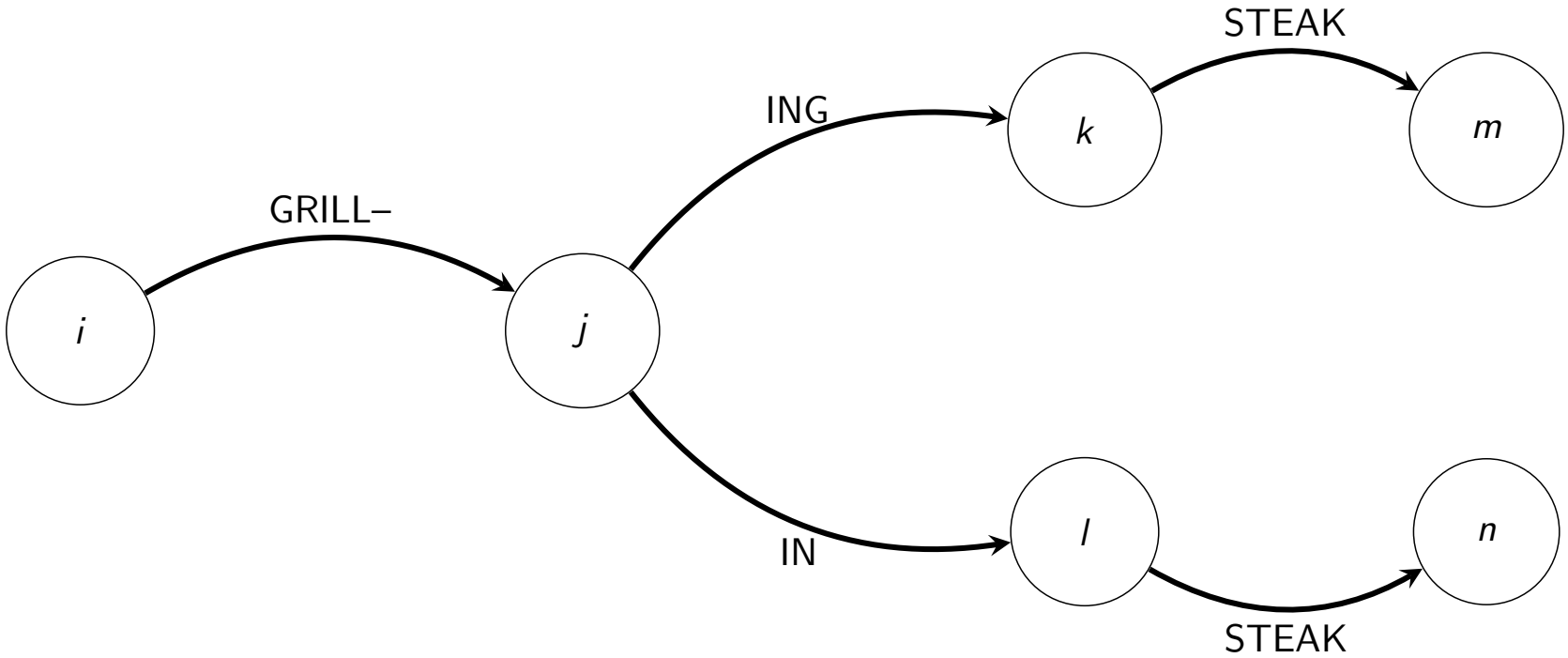
A problem

- Not immediately obvious how such games should be integrated into a general theory of dialogue.
- **Solution:** Embed the games in the kind of information state update/dialogue gameboard approach associated with TTR (Ginzburg, 2012; Cooper and Ginzburg, 2015)

One way of putting TTR and GT together

- For each non-deterministic transition in a TTR game there is a Burnett game to help you make the choice
- That is, if you have more than one update function defined for the current state of the game you need a GT game to choose between them
- The probabilities associated with the different options are computed by a game referring to the mental states of the speaker and addressee as discussed by Burnett.
- Congenial with an information state update (gameboard) approach to dialogue
- *cf.* also HMMs

A simple example: Grilling steak



Outline

Argumentation in dialogue

- Estimating attitudes of addressee when choosing how to make an argument
- Involves estimating prior likelihood of addressee being convinced by a given argument

Our Corpus

- 40 triadic dialogues where participants have been asked to discuss a moral dilemma (Lavelle *et al.*, 2012)
- 20 of these conversations involves a patient diagnosed with schizophrenia

The balloon task

- Subjects asked to discuss a moral dilemma: Four people in a hot air balloon about to crash killing all four unless one of the four is thrown out
- Pilot, 7 months pregnant woman (his wife), doctor (about to find a cure for cancer) and a child prodigy (new Mozart)

Part of a dialogue

- 42 A So I mean the person it seems like the person with least value is the pregnant woman.
- 48 B [she's] pregnant.
- 51 B [So you're] killing two people instead of one.
- 52 C Yhh and another thing is would he be able to pilot the balloon if his wife is overboard?

Two arguments

- if you throw out the pregnant woman, you are killing two people
- if the pregnant woman is thrown out, the pilot (her husband) may not be able to operate the balloon

Enthymemes and Topoi

- Enthymemes = (logically) incomplete arguments
  - the conclusion does not necessarily follow from the premises
  - rely on what is "in the mind" of the listener
- The speaker expects the listener to have access to (and to acknowledge) a particular *topos* (or set of topoi) which warrants the argument. (Aristotle)
- The topoi chosen affect whether the listener will be persuaded or not.
- Enthymemes and/or topoi in conversation (Jackson and Jacobs, 1980; Ducrot, 1988; Anscombre, 1995; Breitholtz, 2014)

Two topoi

- τ_1 there is a choice between sacrificing n and sacrificing m people $m > n \rightarrow$ sacrifice n people
- τ_2 someone is upset $\rightarrow$ they will not be able to perform demanding tasks

Part of a dialogue

- 42 A So I mean the person it seems like the person with least value is the pregnant woman.
- 48 B [she's] pregnant.
- 51 B [So you're] killing two people instead of one. τ_1
- 52 C Yhh and another thing is would he be able to pilot the balloon if his wife is overboard? τ_2

Argument game

- A TTR game (*cf.* suggestion games in Breitholtz (2014))
- Main moves: speaker makes an argument, listener accepts or rejects it
- In order to make an argument you have to first choose an appropriate topos
- Need a GT game

Argument game: choose topos

A tuple $\langle \{S, L\}, T_{cg}, \mathbb{T}, C, \mathfrak{I}, Pr \rangle$ where:

1. S and L are the *players* Two players
2. T_{cg} is a record type representing the *common ground* (*universe*) Type of the balloon situation
3. $\mathbb{T}$ is a finite set of *topoi* which S regards as relevant to the common ground Topoi on which arguments may be based
4. C_S is a measure function on $\mathbb{T}$ Cost of presenting topoi for S
   C_L is a measure function on $\mathbb{T}$ Cost of accepting topoi for L
5. $\mathfrak{I}$ is a relation between members of $\mathbb{T}$ and enthymemes *instantiating* them based on objects introduced in T_{cg}
6. Pr is probability distribution over $\mathbb{T}$ What S regards as topoi most likely to be accepted by L

Calculating the potential utility of using a topos

For $\tau \in \mathbb{T}$, S estimates potential utility of τ

$$\mathit{utility}_S(\tau) = \max(0, \mathit{Pr}(\tau) - C_S(\tau))$$

Payoffs: Actual payoff of τ for both players depending on whether L accepts or rejects

	Accept		Reject	
τ	$1 - C_S(\tau)$	$1 - C_L(\tau)$	0	$C_L(\tau)$

Updating expected probability of L being convinced

Let $\alpha \geq 2$ Temperature constant regulating learning rate

L accepts τ:

$$Pr(\tau) := Pr(\tau) + \frac{1-Pr(\tau)}{\alpha}$$ Increase probability that τ is convincing

$$\forall \tau' \neq \tau Pr(\tau') := Pr(\tau') - \frac{1-Pr(\tau)}{\alpha(|\mathbb{T}|-1)}$$ Decrease probability on other topoi

L rejects τ:

$$Pr(\tau) := Pr(\tau) - \frac{Pr(\tau)}{\alpha}$$ Decrease probability that τ is convincing

$$\forall \tau' \neq \tau Pr(\tau') := Pr(\tau') + \frac{Pr(\tau)}{\alpha(|\mathbb{T}|-1)}$$ Increase probability on other topoi

An example

$\mathbb{T} = \{\tau_1, \tau_2\}, \alpha = 2$
$C_S(\tau_1) = 0, C_S(\tau_2) = .2;\ C_L(\tau_1) = .8, C_L(\tau_2) = .3$
$Pr(\tau_1) = .75, Pr(\tau_2) = .25$

	Accept		Reject	
τ_1	$1 - C_S(\tau_1) = 1$	$1 - C_L(\tau_1) = .2$	0	$C_L(\tau_1) = .8$
τ_2	$1 - C_S(\tau_2) = .8$	$1 - C_L(\tau_2) = .7$	0	$C_L(\tau_2) = .3$

$\text{Utility}_S(\tau_1) = Pr(\tau_1) - C_S(\tau_1) = .75$
$\text{Utility}_S(\tau_2) = Pr(\tau_2) - C_S(\tau_2) = .05$

S chooses τ_1 based on estimated utility, L rejects based on actual payoff.

Update: $Pr(\tau_1) = .75 - \frac{.75}{2} = .375, Pr(\tau_2) = .25 + \frac{.75}{2\times 1} = .625$
$\text{Utility}_S(\tau_1) = Pr(\tau_1) - C_S(\tau_1) = .375$
$\text{Utility}_S(\tau_2) = Pr(\tau_2) - C_S(\tau_2) = .425$

S chooses τ_2 based on new estimated utilities, L accepts based on actual payoff.

Do topoi have social meaning?

- We have suggested a way of choosing argumentational strategies based on social considerations
- The way linguistic cues are related to social meaning in sociolinguistics is by means of persona
- Can we relate personae to topoi?

Outline

Personae in terms of topoi

Returning to the balloon corpus...

Available topoi:

- τ_1: x is a child $\rightarrow$ don't sacrifice x
- τ_2: x may achieve great things $\rightarrow$ don't sacrifice x
- τ_3: There is a choice between sacrificing n people and $n+1$ people $\rightarrow$ sacrifice n people

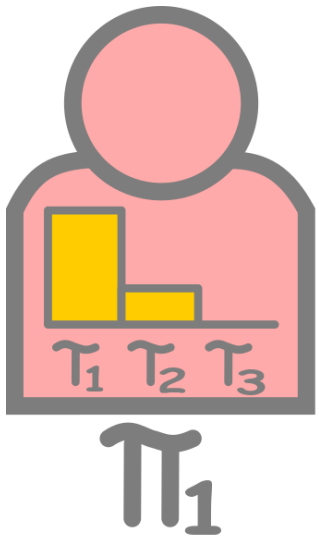
Personae in terms of topoi

Returning to the balloon corpus...

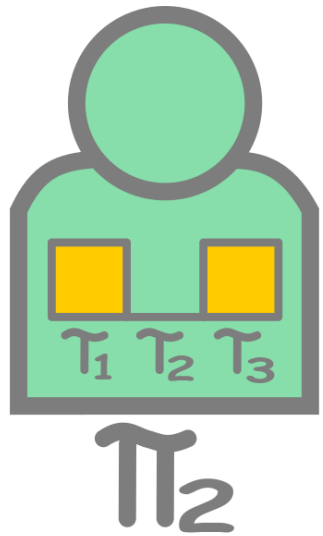
Available topoi:

- τ_1: x is a child $\rightarrow$ don't sacrifice x
- τ_2: x may achieve great things $\rightarrow$ don't sacrifice x
- τ_3: There is a choice between sacrificing n people and $n+1$ people $\rightarrow$ sacrifice n people

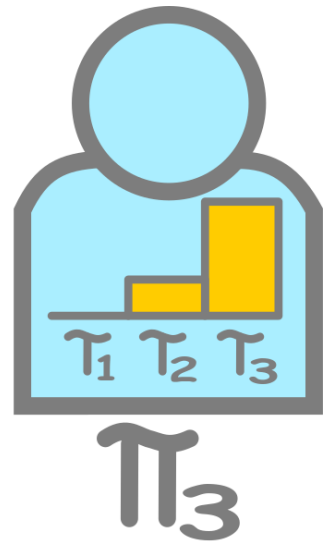
Relevant personae:



the virtue ethicist



the humanist



the cold rationalist

Outline

A probabilistic model of topoi as social signals

Some goals:

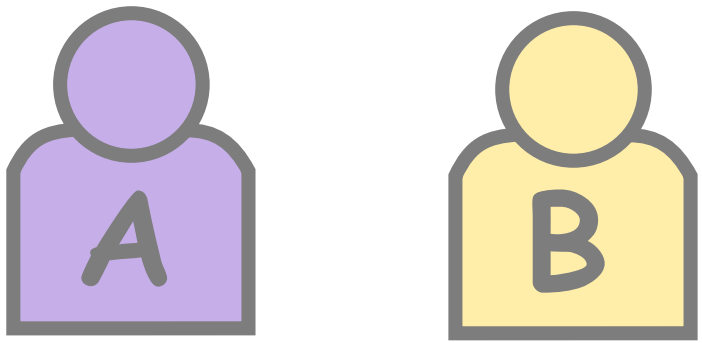
- Define the relationship between *topoi* and *personae*
- Formalize a notion of *social meaning* for topoi
- Model updates to the *social context* resulting from social signals, such as topoi.

A probabilistic model of topoi as social signals

Some goals:

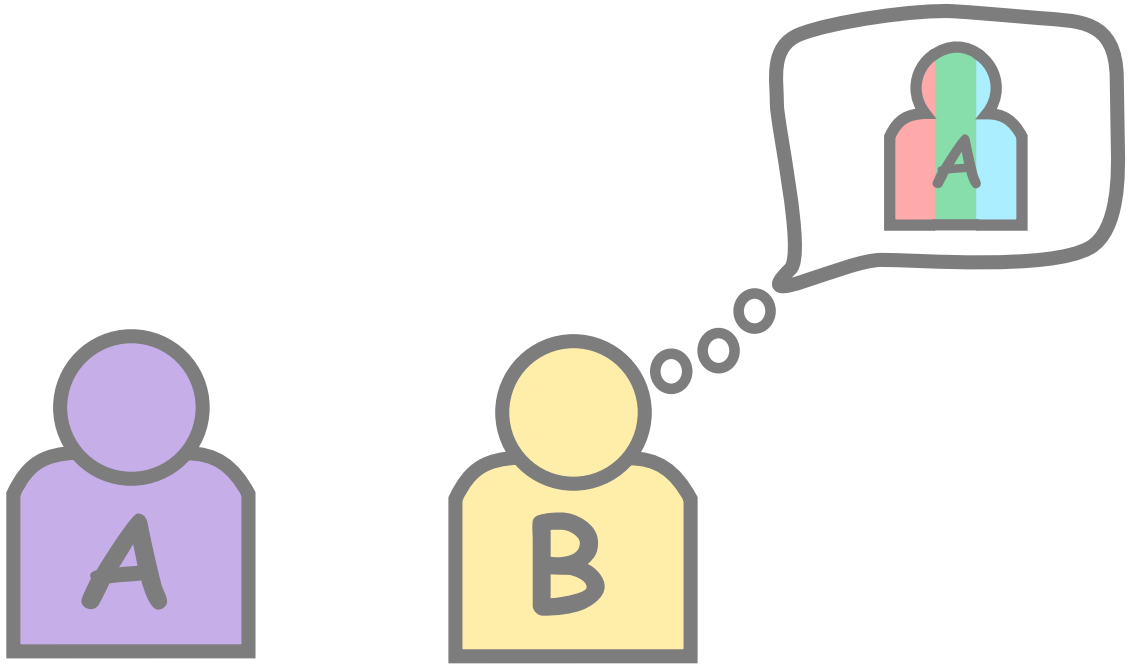
- Define the relationship between *topoi* and *personae*
- Formalize a notion of *social meaning* for topoi
- Model updates to the *social context* resulting from social signals, such as topoi.
- Lay the groundwork for *Bayesian social meaning games*
- Formulate some questions:
  - Do patients with schizophrenia use personae (via topoi) differently from non-patients?
  - How does social uncertainty contribute to the interpretation of social signals?

The setup...



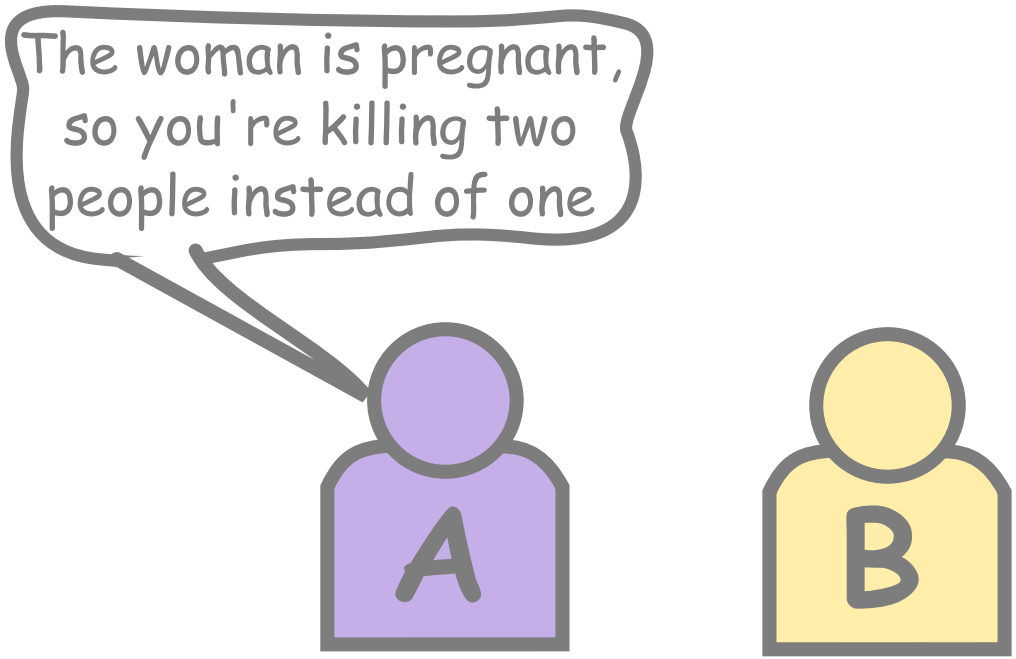
- Assume we have two speakers: A and B

The setup...



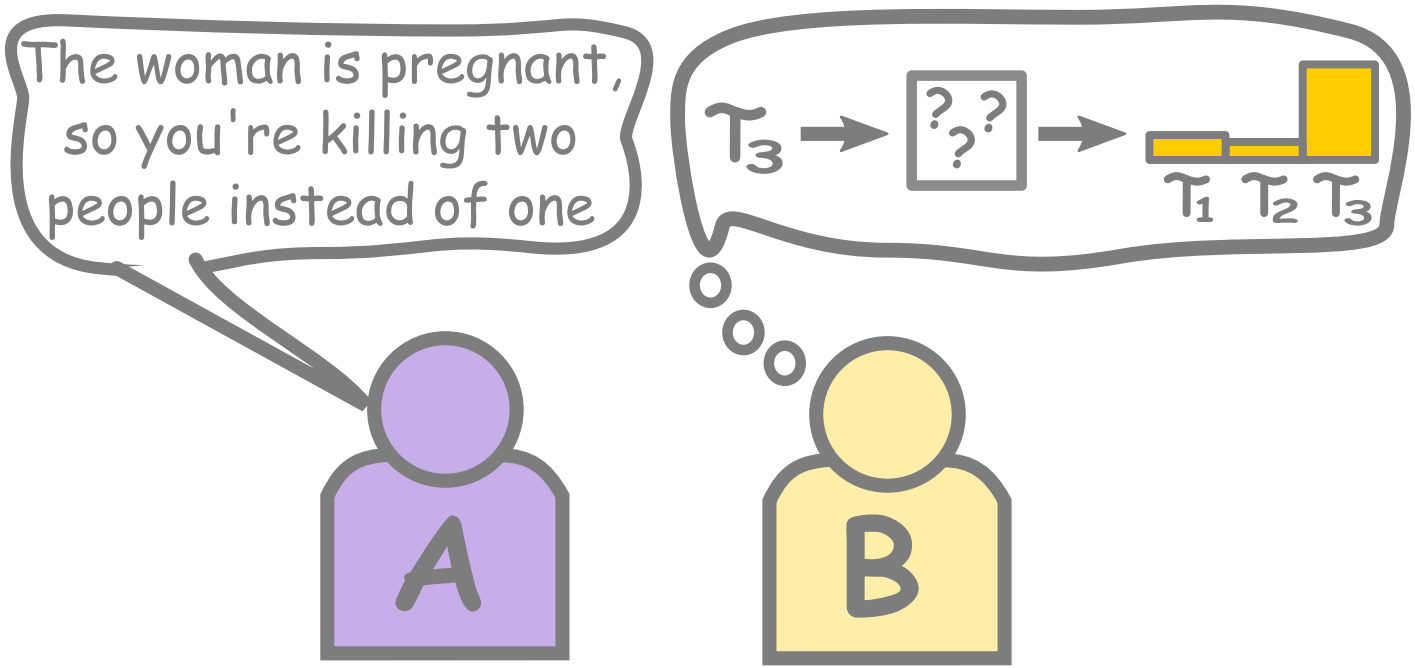
- Assume we have two speakers: A and B
- B's model of A is a probability distribution over personae, according to how likely B finds each as a persona for A

A wild utterance appears!



- Which topos does the utterance *evoke*?

A wild utterance appears!



- Which topos does the utterance *evoke*?
- What is the social meaning of that topos?
  - We define the social meaning of the topos in terms of *ideologically related* topoi.
  - This relatedness goes through the personae it projects.

Let's take a minute to justify this. . .

Social meaning as an indexical field

> *The the meanings of variables are not precise or fixed but rather constitute a field of potential meanings – an indexical field, or constellation of ideologically related meanings, any one of which can be activated in the situated use of the variable.*
>
> Eckert (2008)

- The social meaning of a topos is a probability distribution of idealogically related topoi:

$$[\![\tau^*]\!]_\Delta(\tau) = Pr(\tau \mid \tau^*)$$

Social meaning as an indexical field

> *The the meanings of variables are not precise or fixed but rather constitute a field of potential meanings – an indexical field, or constellation of ideologically related meanings, any one of which can be activated in the situated use of the variable.*
>
> Eckert (2008)

- The social meaning of a topos is a probability distribution of idealogically related topoi:

$$[\![\tau^*]\!]_\Delta(\tau) = Pr(\tau \mid \tau^*)$$

- *Idealogically related* means related **through personae**:

$$Pr(\tau \mid \tau^*) = \sum_{\pi \in \Pi} Pr(\tau \mid \pi) \cdot Pr(\pi \mid \tau^*)$$

The category adjustment effect

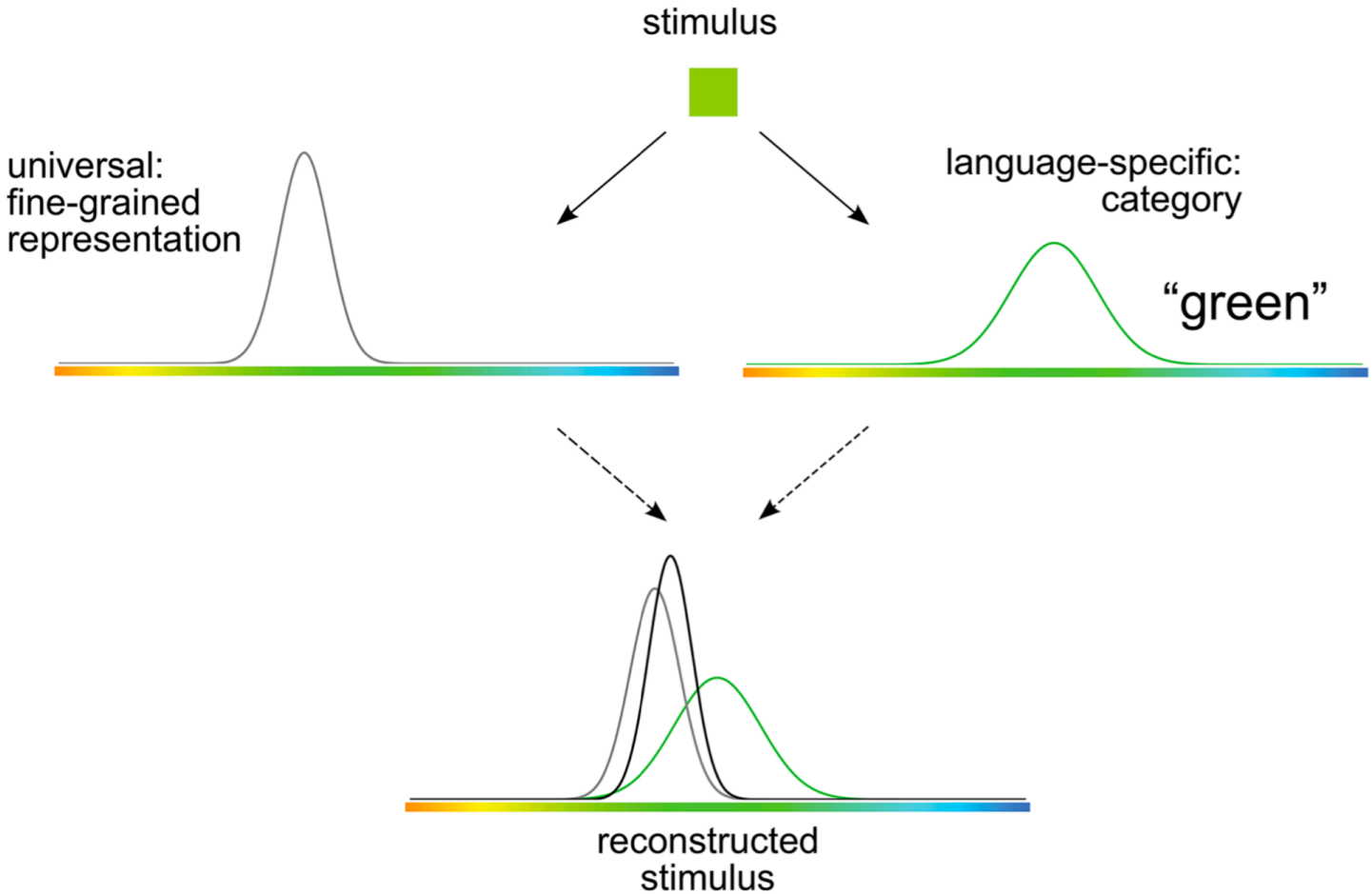


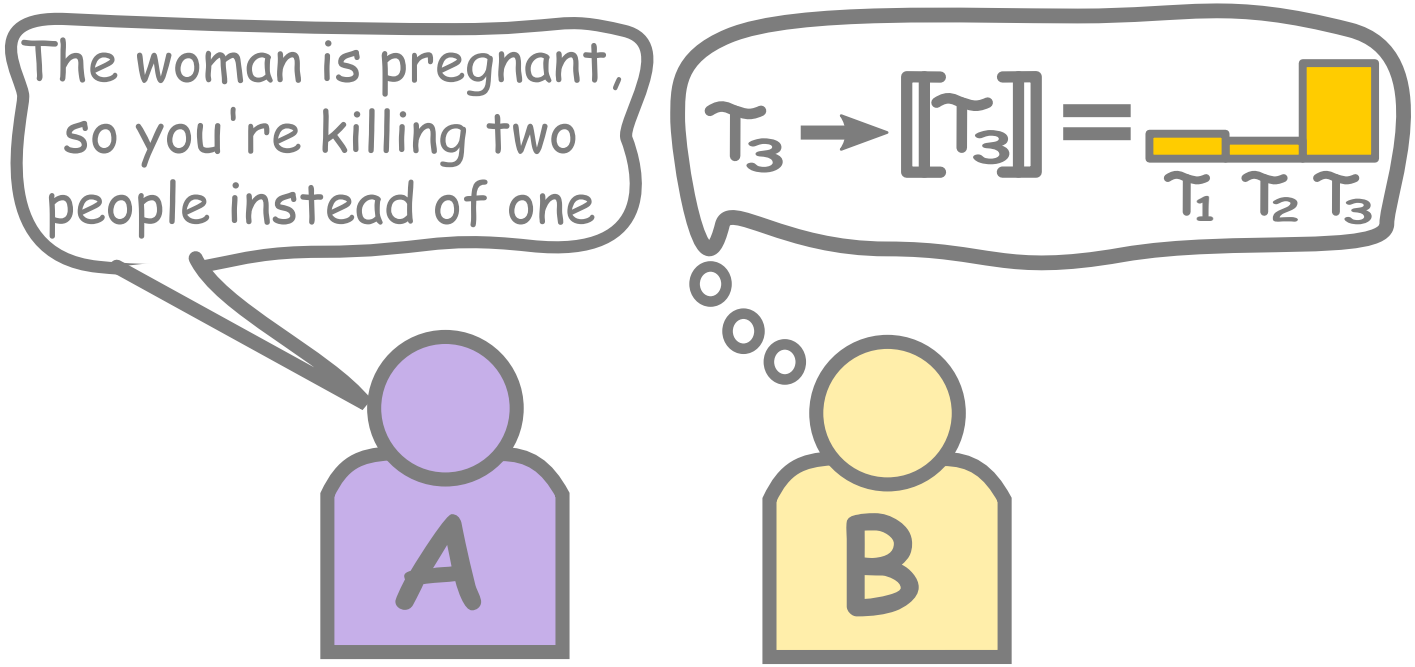
Figure: Figure 3 from Regier and Xu (2017)

The category adjustment effect

- Stimulus = The topos
- Category = Personae
- *Reconstructed stimulus* = A distribution over topoi (i.e., the indexical field)

$$
\begin{aligned}
[\![\tau^*]\!]_\Delta(\tau) &= Pr(\tau \mid \tau^*) \\
&= \sum_{\pi \in \Pi} Pr(\tau \mid \pi) \cdot Pr(\pi \mid \tau^*) \\
&= \sum_{\pi \in \Pi} Pr(\tau \mid \pi) \cdot \frac{Pr(\tau^* \mid \pi) \cdot Pr(\pi)}{Pr(\tau^*)}
\end{aligned}
$$

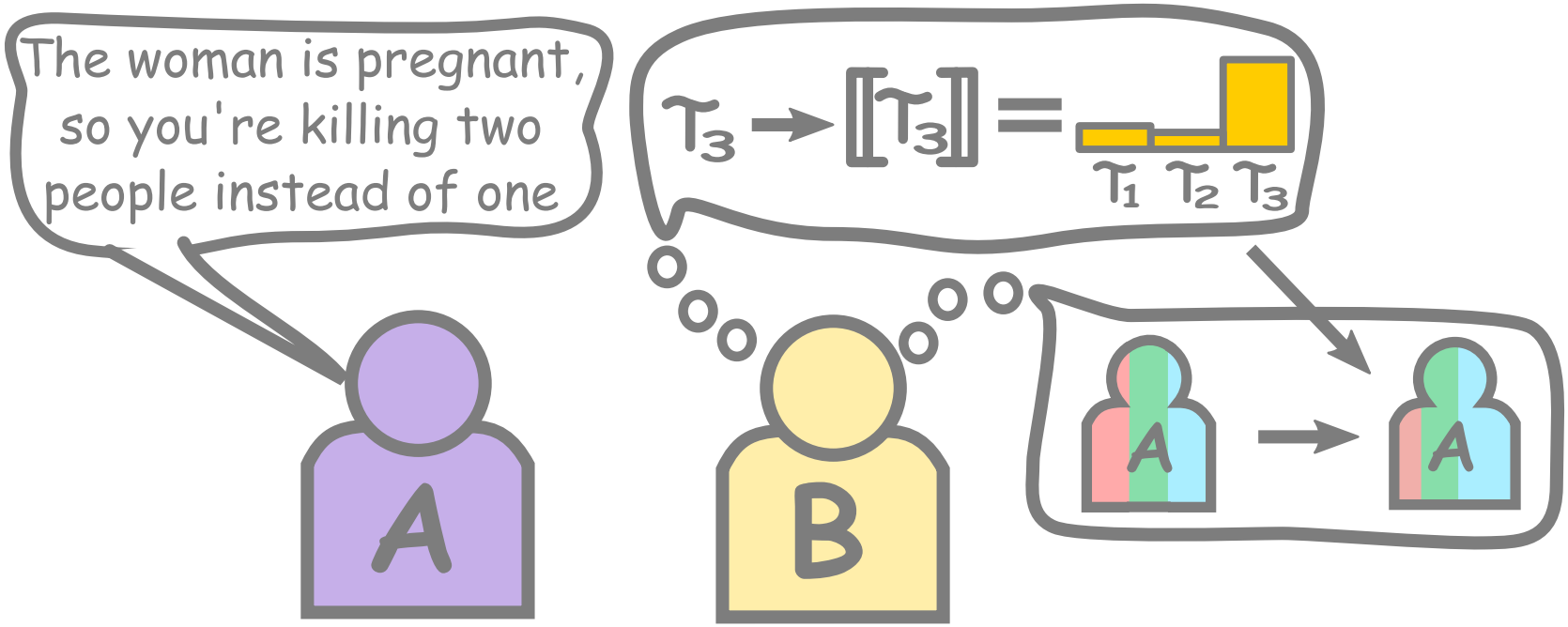
Once again: The social signal produced by τ_3



B interprets τ_3 as a distribution over other topoi it evokes:

$$[\![\tau^*]\!]_\Delta(\tau) = \sum_{\pi \in \Pi} Pr(\tau \mid \pi) \cdot \frac{Pr(\tau^* \mid \pi) \cdot Pr(\pi)}{Pr(\tau^*)}$$

Updating the social context



B updates her model of A's persona (as a result of A's use of τ_3):

$$\hat{Pr}(\pi) = \sum_{\tau} Pr(\pi \mid \tau) \cdot [\![\tau^*]\!]_{\Delta}(\tau)$$

Conclusions

- Games in TTR — no strategy for non-determinism
- Game theory — no integration into a general dialogue theory
- Combine the two kinds of game
- Games for choosing topoi
- Personae characterized in part by distributions over topoi

Bibliography I

Anscombre, Jean-Claude (1995) La Théorie des topoi: Sémantique ou rhétorique?, *Hermés*, Vol. 15, .

Austin, J. (1962) *How to Do Things with Words*, Oxford University Press, ed. by J. O. Urmson.

Barwise, Jon and John Perry (1983) *Situations and Attitudes*, Bradford Books, MIT Press, Cambridge, Mass.

Breitholtz, Ellen (2014) *Enthymemes in Dialogue: A micro-rhetorical approach*, PhD dissertation, University of Gothenburg.

Burnett, Heather (2019) Signalling Games, Sociolinguistic Variation and the Construction of Style, *Linguistics and Philosophy*, Vol. 42, No. 5, pp. 419–450.

Clark, Herbert (1996) *Using Language*, Cambridge University Press, Cambridge.

Bibliography II

Cooper, Robin (2014) How to do things with types, in V. de Paiva, W. Neuper, P. Quaresma, C. Retoré, L. S. Moss and J. Saludes (eds.), *Joint Proceedings of the Second Workshop on Natural Language and Computer Science (NLCS 2014) & 1st International Workshop on Natural Language Services for Reasoners (NLSR 2014) July 17-18, 2014 Vienna, Austria*, pp. 149–158, Center for Informatics and Systems of the University of Coimbra.

Cooper, Robin (in prep) Type theory and language: from perception to linguistic communication. Draft of book chapters available from `https://sites.google.com/site/typetheorywithrecords/drafts`.

Cooper, Robin and Jonathan Ginzburg (2015) Type Theory with Records for Natural Language Semantics, in Lappin and Fox (2015), pp. 375–407.

Bibliography III

Ducrot, Oswald (1988) Topoï et Formes Topique, *Bulletin d'études de la linguistique française*, Vol. 22, pp. 1–14.

Eckert, Penelope (2008) Variation and the Indexical Field, *Journal of Sociolinguistics*, Vol. 12, No. 4, pp. 453–476.

Eckert, Penelope (2012) Three waves of variation study: The emergence of meaning in the study of variation, *Annual Review of Anthropology*, Vol. 41, pp. 87–100.

Fernando, Tim (2015) The Semantics of Tense and Aspect: A Finite-State Perspective, in Lappin and Fox (2015).

Gibson, James J. (1979) *The Ecological Approach to Visual Perception*, Houghton, Mifflin and Company, Also available in a Classic Edition published by Psychology Press, 2015.

Ginzburg, Jonathan (1994) An update semantics for dialogue, in H. Bunt (ed.), *Proceedings of the 1st International Workshop on Computational Semantics*, Tilburg University.

Bibliography IV

Ginzburg, Jonathan (2012) *The Interactive Stance: Meaning for Conversation*, Oxford University Press, Oxford.

Jackson, Sally and Scott Jacobs (1980) Structure of conversational argument: Pragmatic bases for the enthymeme, *Quarterly Journal of Speech*, Vol. 66, No. 3, pp. 251–265.

Labov, William (2012) *Dialect diversity in America: The politics of language change*, University of Virginia Press.

Lappin, Shalom and Chris Fox, eds. (2015) *The Handbook of Contemporary Semantic Theory*, second edition, Wiley-Blackwell.

Larsson, Staffan (2002) *Issue-based Dialogue Management*, PhD dissertation, University of Gothenburg.

Bibliography V

Lavelle, Mary, Patrick GT Healey and Rosemarie McCabe (2012) Is nonverbal communication disrupted in interactions involving patients with schizophrenia?, *Schizophrenia bulletin*, Vol. 39, No. 5, pp. 1150–1158.

Lewis, David (1969) *Convention*, Harvard University Press.

Lewis, David (1979) Scorekeeping in a Language Game, *Journal of Philosophical Logic*, Vol. 8, pp. 339–359.

Regier, Terry and Yang Xu (2017) The Sapir-Whorf hypothesis and inference under uncertainty, *Wiley Interdisciplinary Reviews: Cognitive Science*, Vol. 8, No. 6, p. e1440.