The role of linguistic interpretation in human failures of reasoning

Salvador Mascarenhas
The University of Oxford
Ecole Normale Supérieure

ESSLLI 2016, week 2

1 Introduction

Failures of Reasoning [Slide 4]

A bat and ball cost $1.10 together. The bat costs one dollar more than the ball. *How much does the ball cost?*

Failures of Reasoning [Slide 5]

Mary has met every king or every queen of Europe. Mary has met the king of the Netherlands. Does it follow that *Mary has met the king of Spain*?

Failures of Reasoning [Slide 6]

Jack is looking at Anne, but Anne is looking at George. Jack is married, and George is not. *Is a married person looking at an unmarried person?* Yes No Cannot tell

Anne is either married or she isn't.

Failures of Reasoning [Slide 7]

The Scandinavian peninsula is the European area with the greatest percentage of people with blond hair and blue eyes. This is the case even though (as in Italy) every possible combination of hair and eye color occurs. Suppose we choose at random an individual from the Scandinavian population.

Which do you think is the most probable?

1. The individual has blond hair.
2. The individual has blond hair and blue eyes.

Conceptual house-cleaning [Slide 8]

Compelling fallacies
are (classically) *invalid* inference patterns that we often *accept*.

Repugnant validities
are (classically) *valid* inference patterns that we often *reject*.

Failures of Reasoning? [Slide 9]

- The reasoning problems seen so far have all received accounts in terms of *general purpose reasoning mechanisms* from within psychology and cognitive science
- But it's also possible that some (perhaps even *many*) of these alleged failures of reasoning are in fact the result of *sound* reasoning acting on *non-obvious but perfectly reasonable and predictable* interpretations of the premises
- If we don't contenance this possibility in a *systematic* and *sophisticated* way, we run the risk of *misdiagnosing* interesting interpretive processes as failures of reasoning

Affirming the consequent [Slide 10]

P_1: If the card is long then the number is even. P_2: The number is even. Conclusion: The card is long.

+ pragmatics
P_1: If the card is long then the number is even. $P_{1'}$: *Only if* the card is long is the number even. P_2: The number is even. Conclusion: The card is long.

Disjunction introduction [Slide 11]

P_1: The card is long. Conclusion: The card is long or the number is even.

- *A or B* suggests strongly that *one doesn't know* which one of *A* and *B* is in fact the case

2 Toward convergence between the psychology of reasoning and linguistic semantics

Two themes in semantics [Slide 13]

- Semantics is about *entailment*
  Capturing what sentences follow from a sentence *S* is part of what characterizes the meaning of *S*.
- Semantics is about mapping structures to *representations*
  Modeling and constraining content is part of understanding meaning.

Standard view

- Entailment is real entailment, *what follows classically*
- The right representations are *truth conditional*

A project for semantics as cognitive science [Slide 14]

- To focus on the overlap between semantics and the psychology of reasoning

Reasoning

- Reasoning is about *naive entailment* — whatever people deem to follow from a sentence or set of sentences
- Reasoning is about manipulating *mental representations* of sentences (among other things) to find new mental representations.

Semantics as cognitive science ought to produce representations of content appropriate to feed into a reasoning module and concern itself with what humans take to follow from sentences.

Two notions of "what follows" [Slide 15]

(1) John speaks German and Mary speaks French, or Bill speaks Italian.

Classical entailment (disjunctive syllogism)
If Bill doesn't speak Italian, then John speaks German.

Naive "entailment" (20/20 acceptance, Koralus & Mascarenhas, 2013)
If John speaks German, then Mary speaks French.

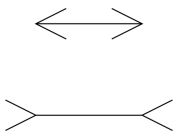
Two notions of representation of content [Slide 16]

John speaks German and Mary speaks French, or Bill speaks Italian.

Truth conditions
$(J \wedge M) \vee B$
(worlds where John speaks German *intersect* worlds where Mary speaks French) *union* worlds where Bill speaks Italian

Inquisitive semantics / mental models
$\{J \wedge M, B\}$
two *alternatives*: *1.* John speaks German and Mary speaks French, *2.* Bill speaks Italian

An analogy with vision [Slide 17]



John speaks German
and Mary speaks French,
or Bill speaks Italian

REPRESENTATION TYPE	**vision**	**semantics**	DESIDERATA FOR REP. **vision**	**semantics**
External	what is true of a piece of visual data	which worlds make a sentence true	horizontal lines have $=$ lengths	disjunctions are sets of worlds
Internal	what the human visual system constructs	what the human faculty of language constructs	horizontal lines have $\neq$ lengths	disjunctions raise alternatives

3 Illusory inferences from disjunction

Illusory inference from disjunction [Slide 19]

(2) P_1: Either Jane is kneeling by the fire and she is looking at the TV or otherwise Mark is standing at the window and he is peering into the garden.
P_2: Jane is kneeling by the fire.
Concl.: Jane is looking at the TV.

Does it follow that *Jane is looking at the TV*?

A fallacy [Slide 20]

(3) Illusory inference from disjunction, schematically:
P_1: $(a \wedge b) \vee (c \wedge d)$
P_2: a
Conclusion: b

- About 85% of subjects accept the conclusion (Walsh and Johnson-Laird, 2004)
- There is no significant effect of whether a, b, c, and d have distinct subjects

Falsified at a model where a, c, and d are true, but b is false.

Not a trivial issue of exclusive 'or'
$(a \wedge b \wedge \neg(c \wedge d)) \vee (c \wedge d \wedge \neg(a \wedge b))$

More on the illusory inference [Slide 21]

- Walsh and Johnson-Laird's illusory inferences can be simplified substantially while getting the same acceptance rates (around 90%)
- Multiple ways of formulating disjunction do not produce a significant difference

(4) John is watching TV and Mary is playing tennis, or Bill is doing homework.
John is watching TV.

(5) John is watching TV and Mary is playing tennis, or else Bill is doing homework.
John is watching TV.

(6) Either John is watching TV and Mary is playing tennis, or Bill is doing homework.
John is watching TV.

Mental models account [Slide 22]

Mental model theory account of the illusory inference from disjunction (combining elements from Johnson-Laird (1983) and Koralus and Mascarenhas (2013)

- Reasoners build mental representations (mental models) that verify each of the premises.
- Disjunctive premises are represented as sets of alternative mental models.
- P_1 gives rise to a set of two alternative models: a minimal model of $a \wedge b$ and a minimal model of $c \wedge d$.
- **Upon hearing P_2, a, reasoners notice that it is related to the first alternative model for P_1, but not the second.** This makes them ignore the second model.
- The combined representation of the premises is therefore only one mental model: $a \wedge b$. From here, b follows.

4 A reasoning-based account: the erotetic theory of reasoning

The erotetic theory of reasoning [Slide 24]

The erotetic principle

- *Part I* — Our natural capacity for reasoning proceeds by treating successive premises as questions and maximally strong answers to them.
- *Part II* — Systematically asking a certain type of question as we interpret each new premise allows us to reason in a classically valid way.

Commitment on interpretation
Disjunctions raise alternatives and put pressure toward *choosing* an alternative — *disjunctions are like questions* in this regard (Inquisitive Semantics: Groenendijk, 2008, Mascarenhas, 2009)

"Disjunctions are like questions?" [Slide 25]

> A standard (if not *the* standard) employment of 'or' is in the specification of possibilities (one of which is supposed by the speaker to be realized, although *he does not know which one*).
>
> Grice, *Indicative Conditionals*, p. 68

- Questions are modeled as sets of propositions...
  {it's raining, it's not raining}
- ...so are disjunctions in many modern approaches to free choice, counterfactuals, exceptional scope-taking, a.o.
  {John is at home, Mary is at work}
- Inquisitive Semantics: disjunctions are at the core of inquisitiveness — disjunctions are the building blocks of questions

Hamblin (1958), Kratzer & Shimoyama (2002), Fine (2012), Charlow (2014)

An argument from morphology [Slide 26]

A(n in)famous fact
Very many natural languages have the same morphemes for the *interrogative complementizer* and *disjunction operator* (and indefinites, more on which later)

- Malayalam is a good example (Jayaseelan, 2004)

(7) John-oo Bill-oo wannu.
John-or Bill-or came
'John or Bill came.'

(8) Mary wannu-oo?
Mary came-or
'Did Mary come?'

(cf. also Japanese 'ka', Korean 'ka', several variations of Slavic 'li', Polish 'czy', and so on)

Illusory inference on the erotetic theory [Slide 27]

(9) P_1: John is watching TV and Mary is playing tennis, or Bill is doing homework.
P_2: John is watching TV.
C: Mary is playing tennis.

Question
Are we in a John-watching-TV and Mary-playing-tennis situation, or in a Bill-doing-homework situation?

Incomplete answer
We are in a John-watching-TV situation.

Jumping to conclusions
I see, so the first answer to the question is the true answer.

References

Fine, Kit (2012). A difficulty for the possible world analysis of counterfactuals. *Synthese*.

Hamblin, Charles L. (1958). Questions. *Australasian Journal of Philosophy*, 36(3):159–168.

Horn, Laurence (2000). From *if* to *iff*: conditional perfection as pragmatic strengthening. *Journal of Pragmatics*, 32:289–326.

Jayaseelan, K.A. (2004). Comparative morphology of quantifiers. Ms. The English and Foreign Languages University (Hyderabad).

Johnson-Laird, Philip N. (1983). *Mental models: towards a cognitive science of language, inference, and consciousness*. Cambridge: Cambridge University Press.

Katzir, Roni (2007). Structurally-defined alternatives. *Linguistics and Philosophy*, 30:669–690.

Koralus, Philipp and Salvador Mascarenhas (2013). The erotetic theory of reasoning: bridges between formal semantics and the psychology of deductive inference. *Philosophical Perspectives*, 27:312–365.

Kratzer, Angelika and Junko Shimoyama (2002). Indeterminate pronouns: the view from Japanese. In *Third Tokyo Conference on Psycholinguistics*.

Mascarenhas, Salvador (2009). *Inquisitive Semantics and Logic*. Master's thesis, ILLC.

Mascarenhas, Salvador (2011). Licensing by modification: the case of positive polarity pronouns. In Ana A. Guevara, Anna Chernilovskaya and Rick Nouwen, editors, *Proceedings of* Sinn und Bedeutung *16*, pages 417–429.

Mascarenhas, Salvador (2013). An interpretation-based account of illusory inferences from disjunction. Talk given at *Sinn und Bedeutung 18*.

Mascarenhas, Salvador (2014a). *Formal Semantics and the Psychology of Reasoning: Building new bridges and investigating interactions*. Ph.D. thesis, New York University.

Mascarenhas, Salvador (2014b). A note on the cardinalities of sets of scalar alternatives. Under review.

Mascarenhas, Salvador and Philipp Koralus (2015). Reasoning with quantifiers beyond syllogisms: illusory inferences and the erotetic theory of reasoning. To be presented at CogSci 2015.

Tversky, Amos and Daniel Kahneman (1983). Extensional versus intuitive reasoning: the conjunction fallacy in probability judgment. *Psychological Review*, 90:293–315.

Walsh, Clare and Philip N. Johnson-Laird (2004). Coreference and reasoning. *Memory and Cognition*, 32:96–106.