Illusory inferences and their relationship to framing

LC2 Squib

Adèle Mortier

Introduction

"Illusory inferences" [4] are illogical conclusions that people tend to draw in a systematic way when faced with certain classes of logical puzzles. Therefore, these inferences represent a real challenge for logicians and psychologists, because they question the ability of humans to reason in a sound way. The puzzles we are interested in have the following form :

> non-atomic logical formula
> atomic cue

Where the logical expressions are somewhat "translated" into natural language. This translation may be problematic ; since natural language is by nature ambiguous, it is difficult to ensure that all people understand the logical operators of natural language as they should be understood within the "pure" logic. Moreover, there can be different possible translations for the same logical reality ; and as we will see, these translations may be interpreted differently. That is what we call "framing" in our context.

In what follows, we would like to develop a way to compare the illusory potential of several puzzles that have been well-documented over the last 30 years. Some of them convey the same logical truth-conditions, but are presented in a different way. If these problems and the conclusions they trigger are now well-known, trying to rank them with respect to their difficulty appears as a quite novel approach to the general problem of illusory inferences.

1 Problems

1.1 Presentation

We present here several puzzles we would like to compare. Some of them seem to trigger intuitive answers that turn out to be wrong ; some other trigger the right answers (w.r.t. the logical "gold standard").

(XOR/$\rightarrow$) :
> - If there is a king then there is an ace or else,
> if there is not a king then there is an ace
> - There is a king
$\implies$ There is an ace (there is not an ace !)

(XOR/$\wedge$) :
> - There is a king and there is an ace or else,
> there is not a king and there is an ace
> - There is a king
$\implies$ There is an ace (indeed)

(XOR/$\vee$) :
> - There is a king or there is an ace or else,
> there is not a king or there is an ace
> - There is a king
$\implies$ There is not an ace (indeed)

(XOR/XOR)
> - There is a king or else there is an ace or else,
> there is not a king or else there is an ace
> - There is a king
$\implies$ There is not an ace (impossible to conclude !)

1.2 Interpretations

The first puzzle, (XOR/$\rightarrow$) seems to strongly suggest that there is an ace ; yet, the opposite conclusion should be drawn. The second puzzle (XOR/$\wedge$) also suggests that there is an ace, and indeed, this condition must be met (it is a control). The third problem is maybe less easy than the last ones, but we hypothesize that it should lead to the inference that there is not an ace, which is a correct inference. Note that logically speaking, the first problem and the third problem are the same, but only one of them

(the first) is expected to lead to an illusory inference. The last problem (XOR/XOR) is also debatable ; but it may lead to the conclusion that there is not as ace ; whereas it should be impossible to conclude.

1.3 Issues

1.3.1 Disjunction

The problems we studied above involve the expression "or else", that has a disjunctive force. The main issue is that the precise semantics people give to this operator it quite unclear :

— some of them might "miss the point" and interpret this as a conjunction – although some control experiments involving a "think-aloud" tasks have shown that it is quite unlikely [4] ;

— some of them may see this as an inclusive disjunction – even though the word "else" should inhibit this interpretation, "or else" could also mean something like $\lambda P.\ \lambda Q.\ P \vee \neg P \wedge Q$ which is itself equal to $\lambda P.\ \lambda Q.\ P \vee Q = \vee$;

— some of them may see this as an exclusive disjunction – which was what we expected at the beginning ;

Solutions to this problem would be to try to control the interpretations people do by presenting the problem in a very explicit way. Another possibility would be to ignore this problem, and consider that the subjects may do several exhaustification operations on top of the reasoning process.

1.3.2 Entailment

Another problem lies in the use of if-clauses, that are very controversial in the field of reasoning [4, 1]. Although if-clauses are defined unequivocally in classical logic ($\rightarrow$ is indeed equivalent to $\lambda P.\ \lambda Q.\ \neg P \vee Q$), the conclusions that people tend to draw are often quite different :

— either they consider "if A then B" as a function that takes a proposition A and returns a proposition B (note that it is an intuitionistic interpretation very close to the BHK interpretation of conditionals, see for instance [6]) ;

— or they consider that "if A then B" is an expression akin to "A and B", but that is only defined when A is true (hypothesis of the defective truth table see for instance [2]).

Both interpretations agree on the fact that natural language conditionals behave weirdly when the antecedent is false ; this also explains why people perform poorly at Wason selection task and other tasks involving *modus tollens.* The first interpretation makes the consequent necessary, while the second interpretation *depends* on the consequent.

2 Experiment

2.1 Measures

We aim at ranking the different illusory problem with respect to their "illusory potential". We hypothesize that the first problem should have the strongest illusory potential, while the others would have lower illusory potentials.

To do so, we may consider three different metrics (for a given puzzle, and a given answer) :

— the degree of wrongness of the answer (*e.g.* which proportion of the logically valid models draw the same conclusion) ;

— the frequency of the answer (among the participants) ;

— the average confidence in the answer ;

— the average time needed to come up with the answer ;

Puzzles with a strong illusory potential should lead to quick and very definitive answers, for a majority af participants.

2.2 Design

To design the experiment, we have to control how people interpret the main disjunctive operator; we are interested in both inclusive and exclusive interpretations but want them to be clearly delineated. We use a protocol similar to the ones designed in [4, 5]. For the inclusive or, we introduce the two disjuncts by telling the participant that "at least one of the following properties is true; and both of them can be true". For the exclusive or, we introduce the two disjuncts by telling the participant that "one and only one of the following properties is true". This formulation may help to avoid misinterpretations of the operator. The question asked to the participant after the presentation of the puzzle will have the form of a multiple choice question (basically between A/¬A/cannot say).

To avoid learning effect and any possible interaction between the XOR and th OR conditions, we assign each participant to only one of the two conditions (OR or XOR). Each participant will be confronted with one control problem (where the intuitive answer given by our model is congruent with the logical answer) and two illusory illusory inferences, in a randomized order. The reaction time (time between the presentation of the problem and the definitive answer) would also be measured; the participant will be told at the beginning of the experiment to answer as quickly but also as reliably as possible.The participant will be invited to rate their level of confidence after they have given their answer.

3 Models

Different models of human reasoning may explain the phenomena we expect to measure. But actually, some of our conditions are quite unpredictable (especially the XOR/OR, XOR/XOR and OR/XOR cases). Therefore, we would like to see whether the existing models will capture what happens experimentally with these puzzles. We may also investigate how these models allow to measure the "illusory potential" of our different problems, theoretically.

3.1 Mental model theory

MM theory is a very popular theory introduced by Johnson-Laird and that covers a lot of problems like the ones we have seen so far. It is based on the strong assumption that people fail to represent "what is false". More precisely, they build mental representations of the "models" that make a formula true, but only in terms of what has to be fulfilled, not in terms of what has to be invalid. This models are thus inspired from model theory, but they are in a way underspecified. Computing a mental model for a given formula consists in :

  (i) building the explicit truth table of the problem;

 (ii) identifying the models (or valuations) such that the problem is valid;

(iii) simplifying these models by only keeping a list of the atomic formulae whose valuation is 1.

The step that generates illusory inferences (and deviates from model theoretic logic) is the last one.

3.2 Probabilistic approach

This approach considers reasoners as hypothesis testers. When facing two alternatives, reasoners chose the one that fits the best what they already know. If the problem is as follows :

$$\left| \begin{array}{l} \text{A (X)OR B} \\ \text{C} \end{array} \right.$$

Then the alternative that will be considered as true is :

$$argmax_{X \in \{A,B\}} \mathbb{P}[X|C]$$

If we consider that $\mathbb{P}[X|C]$ is a good approximation of $\mathbb{P}[C \to X]$, then $\mathbb{P}[X|C]$ is the degree to which the alternative X and the cue C "fit together". This goes along with the fact that people prefer "consistent" scenarios to non consistent one. However, other related measures could be considered, for instance :

$$\begin{array}{ccc} \mathbb{P}[X|C] & - & \mathbb{P}[X] \\ \log(\mathbb{P}[X|C]) & - & \log(\mathbb{P}[X]) \end{array}$$

These measures are differential measures. The first measure is basically the increment of probability mass that is added to X by the fact that C is known. The rationale behind this formula is that people tend to prefer the option whose probability is "boosted" the most by the cue they are given – they want the cue to be optimally useful. The second measure is also a differential measure, but a multiplicative one (because of the log). This approach is consistent with the fact that people perceive intervals between numerical quantities logarithmically.

3.3   Graph-theoretic approach

Here, we develop an alternative approach that is inspired by proof nets, a graphical way of representing proofs. We divide the reasoning process into 5 steps :

(i) building of the proof tree corresponding to the non-atomic formula ;

(ii) computation of the truth values on the leaves using the context ;

(iii) bottom-up propagation of the truth values ;

(iv) deduction of additional truth values ;

(v) top-down propagation of the new truth values.

The reasoning process amounts to a traversal of the tree, that is successively bottom-up and top-down. Note that steps (iii), (iv), (v) may be repeated more than once, until a truth value ($\top$) has been found for the root. We assume that the branches of the tree are labeled by a pair (context ;truth value), where the context corresponds to what the reasoner has inferred – what he knows to be the case.

Références

[1] Pierre Barrouillet and J F Lecas. Illusory inferences from a disjunction of conditionals : A new mental models account. *Cognition*, 76 :167–73 ; discussion 175, 08 2000.

[2] Ruth M.J. Byrne and P.N. Johnson-Laird. 'if' and the problems of conditional reasoning. *Trends in Cognitive Sciences*, 13(7) :282 – 287, 2009.

[3] Bart Geurts. Entertaining alternatives : Disjunctions as modals. *Natural Language Semantics*, 13(4) :383–410, dec 2005.

[4] P.N Johnson-Laird and Fabien Savary. Illusory inferences : a novel class of erroneous deductions. *Cognition*, 71(3) :191 – 229, 1999.

[5] Sangeet Khemlani and P. N. Johnson-Laird. Disjunctive illusory inferences and how to eliminate them. *Memory & Cognition*, 37(5) :615–623, jul 2009.

[6] Anne S. Troelstra. Proof theory and constructive mathematics. In Jon Barwise and H. Jerome Keisler, editors, *Handbook of Mathematical Logic*, pages 973–1052. North-Holland Pub. Co., 1977.