

# Illusory inferences and their relationship to framing

LC2 Squib

Adèle Mortier

16 décembre 2018

## 1 Introduction

“Illusory inferences” [3] 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. For instance, the following (equivalent) logical formulae may have different translations :

$$\begin{array}{llll} A \vee B \vee C \vee D & \iff & \text{associativity} & (A \vee B) \vee (C \vee D) \\ & \iff & \text{associativity} & (A \vee B \vee C) \vee D \\ & \iff & \text{commutation} & D \vee B \vee C \vee A \\ (A \rightarrow B) \vee (C \rightarrow D) & \iff & \text{equivalence} & (\neg A \vee B) \vee (\neg C \vee D) \\ & \iff & \text{associativity} & (\neg A \vee B \vee \neg C) \vee D \\ & \iff & \text{commutation} & D \vee B \vee \neg C \vee \neg A \end{array}$$

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 appears as a quite novel approach to the general problem of illusory inferences.

## 2 Problems

### 2.1 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 answer that appear to be right.

(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

$\Rightarrow$  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

$\Rightarrow$  There is an ace (there is not an ace!)

(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

$\Rightarrow$  There is not an ace (it should be impossible to conclude!)

## 2.2 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 convincing than the first one, but we hypothesize that it should also lead to the illusory inference that there is an ace. Note that logically speaking, the first problem and the third problem are the same. 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.

We could also think of natural generalizations of these problems, where  $\neg K$  (“there is not a king”) is replaced by another atomic formula (for instance  $Q$ , for “there is a queen”, or  $\neg Q$ ). But this is a bit beyond the scope of this squib.

## 2.3 Issues

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

- some of them could sort of “miss the point” and interpret this as a conjunction (although some control experiments have shown that it is quite unlikely [3]) ;
- some of them could see this as an inclusive disjunction (even though the word “else” should inhibit this interpretation) ;
- some of them could see this as an exclusive disjunction (that is favored by “else”) ;
- but more interestingly, some of them could see this as an asymmetric “or”, of the form  $\lambda P. \lambda Q. P \vee \neg P \wedge Q$ . This interpretation is linked to an algorithmic reading of “or” where the second disjunct is evaluated only if the first one is false.

Solutions to this problem would be to try to control the interpretations people do by presenting the problem in a very explicit way. But the last interpretation we have considered may be very difficult to remove totally. Another solution would be to ignore this problem, and consider that the subjects may do several exhaustification operations on top of the reasoning process.

Another problem lies in the use of if-clauses, that are very controversial in the field of reasoning. 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) ;
- 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.

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.

A	B	$A \rightarrow B$	A	B	$A \rightarrow B$
0	0	1	0	0	-
0	1	1	0	1	-
1	0	0	1	0	0
1	1	1	1	1	1

TABLE 1 – Truth tables corresponding to “material” implication (left) and to “real-world” conditionals used by reasoners (right)

### 3 Experience proposals

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 of participants.

For the experimental setting, we present the puzzles a bit differently. To avoid misinterpretations of the operator “or else” (intended to be an exclusive or), we will present the two alternative separately in a randomized order, and specify that one and only one of the alternatives is true. This allows to observe the behavior of the participant with respect to the inner operators only. Each participant will face the control puzzle (XOR/ $\wedge$ ) plus one additional illusory puzzle, to avoid learning effect.

## 4 Possible explanations

### 4.1 Mental model theory

It is a very popular theory 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. These models are thus inspired from model theory, but they are in a way underspecified. Computing a mental model for a given formula consists in :

- building the explicit truth table of the problem ;
- identifying the models (or valuations) such that the problem is valid ;
- simplifying these models by only keeping a list of the atomic formulae whose valuation is 1.

Problem	Mental model			Mental model conclusion	Intuitive conclusion	Logical conclusion
(XOR/ $\rightarrow$ )	$\overline{K}$ — $\overline{A}$	$A$ $\neg K$		A	A	$\neg A$
(XOR/ $\wedge$ )	$\overline{K}$ — $\overline{A}$	$A$ $\neg K$		A	A	A
(XOR/ $\vee$ )	$\overline{K}$ — $\overline{A}$ — $\neg K$	$\neg A$ $\neg K$		$\neg A$	$\neg A$ (?)	$\neg A$
(XOR/XOR)	$\overline{K}$ — $\neg K$	$\neg A$ $\neg K$		$\neg A$	$\neg A$ (?)	nothing
(OR/ $\rightarrow$ )	$\overline{K}$ — $\overline{A}$ $K$ $A$	$\neg K$ $\neg K$		A	A	nothing
(OR/ $\wedge$ )	$\overline{K}$ — $\overline{A}$ $K$ $A$	$\neg K$ $\neg K$		A	A	A
(OR/ $\vee$ )	$\overline{K}$ — $\overline{A}$ $K$ — $\neg K$ — $A$	$\neg K$ $\neg K$		nothing	A (?)	nothing
(OR/XOR)	$\overline{K}$ — $\overline{A}$ $K$ — $\neg K$ — $A$	$\neg K$ $\neg K$		nothing	$\neg A$ (?)	nothing

TABLE 2 – Mental model predictions *vs* logical conclusions. The question mark means that the judgment may be not very sharp, and that we should test these framings.

## 4.2 Algorithmic approach

This approach assumes that people tend to reason like algorithms ; and like algorithms, they may be prone to “lazy” evaluation : they compute truth values only when it is needed. The most simple example of what is “lazy” would be the following :

$$LOR_{lazy}(A)(B) = \text{if } A \text{ then } A \text{ else } B$$

This amounts to saying that first we examine the left disjunct (Left OR), and if it is true we consider the sentence true. If it is not, we examine the right disjunct and check whether it is true.

Here, we use a variant of this kind of lazy processing, where we assume that people continue the evaluation if they cannot find enough evidence for a given answer :

X XOR Y : first suppose X is true and conclude what are the truth values of X’s subformulae. If X cannot be true or does not allow to conclude, then try on Y.

(XOR/ $\rightarrow$ ) : we suppose that K is true. We begin by evaluating  $K \rightarrow A$ , and this returns A, so A must be true. If the two XORed alternatives had been swapped, we would have begun by evaluating  $\neg K \rightarrow A$ , which gave us nothing, and then we would have evaluated  $K \rightarrow A$ , with the same result as before.

(XOR/ $\wedge$ ) : we suppose that K is true. We begin by evaluating  $K \wedge A$ , which forces A to be true, so A must be true. If the two XORed alternatives had been swapped, we would have begun by evaluating  $\neg K \wedge A$  which gave us nothing, and then we would have evaluated  $K \wedge A$ , with the same result as before.

(XOR/ $\vee$ ) : we suppose that K is true. We begin by evaluating  $K \vee A$ , which gives us nothing special, so we evaluate  $\neg K \vee A$ , which gives us A (by disjunction elimination). So A is true. If the two XORed alternatives had been swapped, we would have begun by evaluating  $\neg K \vee A$ , with the same result as before.

(XOR/XOR) : we suppose the K is true. We begin by evaluating  $KXORA$ , and within  $KXORA$ , we begin by evaluating K , which is true. Then, we move to the second alternative, and evaluate  $\neg KXORA$ . We first begin by evaluating  $\neg K$ , which contradicts our hypothesis. We then move to A, which must be true. If the two XORed alternatives had been swapped, we would have begun by evaluating  $\neg KXORA$ ,

which would have given A too.

In that framework, the illusory potential of an inference could be measured as the number of steps that have to be performed to find an answer. The first three problems appears to be equally costly, while the fourth one demands more evaluations.

### 4.3 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 :

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

Then the alternative that will be considered as true is :

$$\operatorname{argmax}_{X \in \{A, B\}} \mathbb{P}[X|C]$$

If we consider that  $\mathbb{P}[X|C]$  is a good approximation of  $\mathbb{P}[C \rightarrow X]$ , 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 :

$$\begin{aligned} \mathbb{P}[X|C] &= \mathbb{P}[X] \\ \log(\mathbb{P}[X|C]) &= \log(\mathbb{P}[X]) \end{aligned}$$

(XOR/ $\rightarrow$ ) :

$$\begin{aligned} \mathbb{P}[K \rightarrow A|K] &= \mathbb{P}[A|K, K] = \mathbb{P}[A|K] \\ \mathbb{P}[\neg K \rightarrow A|K] &= \mathbb{P}[A|\neg K, K] \simeq \mathbb{P}[A] \end{aligned}$$

(XOR/ $\wedge$ ) :

$$\begin{aligned} \mathbb{P}[K \wedge A|K] &= \mathbb{P}[A, K|K] = \mathbb{P}[A|K] \\ \mathbb{P}[\neg K \wedge A|K] &= \mathbb{P}[A, \neg K|K] = 0 \end{aligned}$$

(XOR/ $\vee$ ) :

$$\begin{aligned} \mathbb{P}[K \vee A|K] &= \mathbb{P}[A \cup K|K] = \mathbb{P}[A|K] \\ \mathbb{P}[\neg K \vee A|K] &= \mathbb{P}[A \cup \neg K|K] = \mathbb{P}[A|K] \end{aligned}$$

(XOR/XOR) :

$$\begin{aligned} \mathbb{P}[KXORA|K] &= \mathbb{P}[\neg A] \\ \mathbb{P}[\neg KXORA|K] &= \mathbb{P}[A] \end{aligned}$$

### 4.4 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 (by elimination) ;
- (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). A branch labeled with  $(K; )$  has an underspecified truth value. The context corresponds to what the reasoner has inferred – what he knows to be the case. Initially (at the root of the tree), the context consists of the cue that has been given to the reasoner as part of the problem. The context can be enriched when the reasoner crosses conditionals, that basically add the consequent to the context if the context matches the antecedent. The truth value of a branch corresponds to the truth value of the subformula that corresponds to this branch. Truth values are computed incrementally, bottom-up or top-down (using the regular truth-functional semantics of the operators crossed)

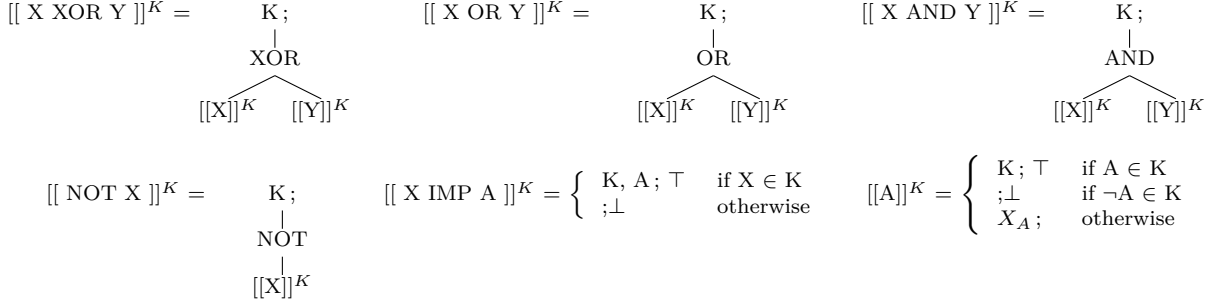


TABLE 3 – Tree building rules

The bottom-up propagation rules for  $\top$  and  $\perp$  correspond to the classical rules associated with the different binary operators. For instance, a AND node whose two incoming branches are labeled  $\top$  will propagate  $\top$ ; a XOR node that has one  $\perp$ -branch and one  $\top$ -branch will propagate  $\top$  too, etc. When the propagation reaches the root of the tree, we consider that we are done, and we can conclude that all the formulae that can be read on the leaves of the  $\top$ -branches hold.

If the propagation stops before the root is reached, we have to do a so-called deduction step. The deduction step must be performed preferentially in very specific node of the tree, namely the nodes that already have an incoming  $\top$ -branch. If no such node exists, we perform the deduction step on a node that has an incoming  $\perp$ -branch. Basically, the deduction step infers additional propositions that must hold for the whole formula to be satisfied :

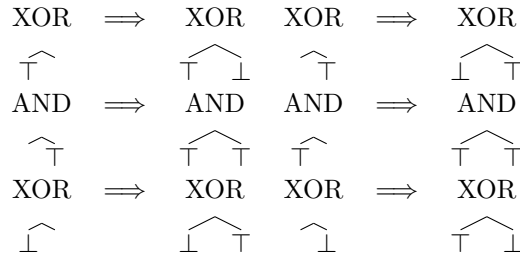


TABLE 4 – Deduction rules

Then, it becomes possible to perform top-down propagation to deduce truth values for yet unknown atomic formulae (noted  $X_A$ ). The top-down propagation is performed quite naturally using the classical rules; for instance, a XOR that has a  $\perp$  top branch and a  $\top$  incoming branch will have another  $\top$  incoming branch. The only “special” rule is the following :

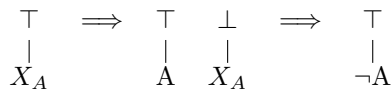


TABLE 5 – Top-down propagation rule at the leaf-level

## 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] Bart Geurts. Entertaining alternatives : Disjunctions as modals. *Natural Language Semantics*, 13(4) :383–410, dec 2005.
- [3] P.N Johnson-Laird and Fabien Savary. Illusory inferences : a novel class of erroneous deductions. *Cognition*, 71(3) :191 – 229, 1999.
- [4] Sangeet Khemlani and P. N. Johnson-Laird. Disjunctive illusory inferences and how to eliminate them. *Memory & Cognition*, 37(5) :615–623, jul 2009.

## Appendix

A XOR B	A OR B	A AND B	A IMP B
A $\neg$ B	A $\neg$ B	A B	A B
$\neg$ A B	$\neg$ A B		...

TABLE 6 – Mental models for basic operators

