

# Illusory inferences and their relationship to framing

LC2 Squib

Adèle Mortier

10 décembre 2018

## 1 Introduction

“Illusory inferences” 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 :

$$\left| \begin{array}{l} \text{non-atomic logical formula} \\ \text{atomic cue} \end{array} \right.$$

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 :

$$\begin{aligned} A \vee B \vee C \vee D &\iff (A \vee B) \vee (C \vee D) \iff (A \vee B \vee C) \vee D \iff D \vee B \vee C \vee A \\ (A \rightarrow B) \vee (C \rightarrow D) &\iff (\neg A \vee B) \vee (\neg C \vee D) \iff (\neg A \vee B \vee \neg C) \vee D \iff D \vee B \vee \neg C \vee \neg A \end{aligned}$$

In what follows, we would like to develop a way to compare the illusory potential of several well-known 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.

## 2 Problems

We present here several puzzles we would like to compare. Some of them seem trigger intuitive answers that turn out to be wrong ; some other trigger answer that appear to be right.

$$\begin{aligned} (\text{XOR}/\rightarrow) &\left| \begin{array}{l} \text{If there is a king then there is an ace or else, if there is not a king then there is an ace} \\ \text{There is a king} \end{array} \right. \\ \implies &\text{There is an ace (there is not an ace!)} \end{aligned}$$

$$\begin{aligned} (\text{XOR}/\wedge) &\left| \begin{array}{l} \text{There is a king and there is an ace or else, there is not a king and there is an ace} \\ \text{There is a king} \end{array} \right. \\ \implies &\text{There is an ace (indeed)} \end{aligned}$$

$$\begin{aligned} (\text{XOR}/\vee) &\left| \begin{array}{l} \text{There is a king or there is an ace or else, there is not a king or there is an ace} \\ \text{There is a king} \end{array} \right. \\ \implies &\text{There is an ace (there is not an ace!)} \end{aligned}$$

$$\begin{aligned} (\text{XOR}/\text{XOR}) &\left| \begin{array}{l} \text{There is a king or else there is an ace or else, there is not a king or else there is an ace} \\ \text{There is a king} \end{array} \right. \\ \implies &\text{There is not an ace (it should be impossible to conclude!)} \end{aligned}$$

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. This last puzzle questions the influence of scope and commutativity. We can also present 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  $\neg Q$ )

### 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 :

- The rate of wrong answers for a given puzzle ;
- The average confidence in false answers ;
- The average time needed to come to a final conclusion ;

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

We detail three possible explanations of what happens when people reason about such problems.

#### 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. We then build models of the properties that make a formula true, they only represent what has to be fulfilled, and not what should not be fulfilled.

#### 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 :

$$Or_{lazy}(A)(B) = \text{if } A \text{ then } A \text{ else } B \quad (1)$$

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. We facing two alternatives, reasoners chose the one that fits the best what they already know. If the problem is as follows :

$$\begin{array}{|l} X \text{ XOR } Y \\ Z \end{array}$$

Then the alternative that will be considered as true is :

$$\operatorname{argmax}_{W \in \{X, Y\}} \mathbb{P}[W|Z] \quad (2)$$

(XOR/ $\rightarrow$ ) :

$$\mathbb{P}[K \rightarrow A|K] = \mathbb{P}[A|K, K] = \mathbb{P}[A|K] \quad (3)$$

$$\mathbb{P}[\neg K \rightarrow A|K] = \mathbb{P}[A|\neg K, K] \simeq \mathbb{P}[A] \quad (4)$$

(XOR/ $\wedge$ ) :

$$\mathbb{P}[K \wedge A|K] = \mathbb{P}[A, K|K] = \mathbb{P}[A|K] \quad (5)$$

$$\mathbb{P}[\neg K \wedge A|K] = \mathbb{P}[A, \neg K|K] = 0 \quad (6)$$

(XOR/ $\vee$ ) :

$$\mathbb{P}[K \vee A|K] = \mathbb{P}[A \cup K|K] = \mathbb{P}[A|K] \quad (7)$$

$$\mathbb{P}[\neg K \vee A|K] = \mathbb{P}[A \cup \neg K|K] = \mathbb{P}[A|K] \quad (8)$$

(XOR/XOR) :

$$\mathbb{P}[KXORA|K] = \mathbb{P}[\neg A] \quad (9)$$

$$\mathbb{P}[\neg KXORA|K] = \mathbb{P}[A] \quad (10)$$

### 4.4 Graph-theoretic approach