Paradoxical inferences, biconditional interpretation, and exclusivity

Išvadų paradoksualumas, dvilypė interpretaciją ir nesuderinamumąs

MIGUEL LÓPEZ-ASTORGA, University of Talca, Chile

Abstract

Two inferences correct in classical logic are controversial in cognitive science. The reason is that people do not always deem them as valid inferences. One of them is the rule to introduce a conditional. The other one is the rule to introduce a disjunction. The theory of mental models has an account for them. Their conclusions refer to models, and, in both cases, one of those models is inconsistent with the premise. When semantics modulates and removes the incoherent model, the inferences are accepted as correct. The present paper tries to describe those phenomena within the framework of first-order predicate logic. It proposes that the rule to introduce a conditional is not admitted when the conclusion is not a conditional, but a biconditional. It also claims that the rule to introduce a disjunction is not accepted when the disjunction is exclusive. These latter points are the novelty of the paper. People do not actually reject the two mentioned inferences correct in classical logic. What individuals reject is to introduce a biconditional taking as one of its clauses just a proposition already presented in the inference (which is also forbidden in classical logic) and to infer an exclusive disjunction from just a proposition, which is taken as one of the disjuncts (which is also forbidden in classical logic).

KEYWORDS: biconditional, conditional, exclusive disjunction, first-order predicate logic, inclusive disjunction.

Introduction

There are inferences that are correct in first-order predicate logic. However, people many times do not accept them. Strictly speaking, they are not paradoxes (they are just inferences correct in classical logic that people do not consider to be right). But they are often called ‘paradoxical inferences’ (e.g., Orenes & Johnson-Laird, 2012). (1) is an example of one of them.

Ex 1

These people do not eat apples
---------------------------------------------------------------------------
Therefore, if these people study geography, these people do not eat apples

Inference (1) is correct in first-order predicate calculus. (2) holds in that calculus.
Ex 2 \[ \exists x \neg A x \therefore \exists x (G x \rightarrow \neg A x) \]

Where ‘\( \exists \)’ represents the existential quantifier, ‘\( \neg \)’ is negation, ‘\( A \)’ is a predicate indicating to eat apples, ‘\( \therefore \)’ denotes logical deduction, ‘\( G \)’ is a predicate meaning to study geography, and ‘\( \rightarrow \)’ stands for the logical conditional. Something similar occurs with (3). It is also accepted in first-order predicate calculus.

Ex 3

<table>
<thead>
<tr>
<th>These people eat fruits</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ \text{Therefore, either these people eat cake or these people eat fruits} ]</td>
</tr>
</tbody>
</table>

Deduction (4) is assumed in first-order predicate logic, too.

Ex 4 \[ \exists x F x \therefore \exists x (C x \lor F x) \]

Where ‘\( F \)’ is a predicate indicating to eat fruits, ‘\( C \)’ is a predicate signifying to eat cake, and ‘\( \lor \)’ refers to the logical disjunction.

The problem is that, following the literature, human beings do not always embrace inferences such as (1) and (3) (Orenes & Johnson-Laird, 2012). This suggests that first-order predicate logic cannot explain what happens in those cases. The theory of mental models (e.g., Khemlani, Byrne, & Johnson-Laird, 2018) can account for why inferences such as (1) and (3) are spurned. In fact, it can explain even why other inferences with the same logical structure are, in general, admitted. For example, if (1) is transformed into (5),

Ex 5

<table>
<thead>
<tr>
<th>These people do not eat apples</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ \text{Therefore, if these people eat fruits, these people do not eat apples} ]</td>
</tr>
</tbody>
</table>

The literature also shows that the inference is often approved (Orenes & Johnson-Laird, 2012). From logic, this is hard to understand. The formal structure of (5) is (2) as well. To note it, it is only necessary to modify the meaning of ‘\( G \)’, which would denote to eat fruits now.

The same happens with (3). If its conclusion is that in (6),

Ex 6

<table>
<thead>
<tr>
<th>These people eat fruits</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ \text{Therefore, either these people eat apples or these people eat fruits} ]</td>
</tr>
</tbody>
</table>

It tends to be validated (Orenes & Johnson-Laird, 2012). Again, the logical structure is identical in both inferences. (4) corresponds to both (3) and (6). To note that, it is only necessary to assume that what ‘\( C \)’ means now is to eat apples, that is, the same as ‘\( A \)’.

This paper is not intended to present objections against the explanations of the theory of mental models. It just tries to describe how these phenomena could be expressed by means of first-order predicate logic. Hence, without rejecting the account of the human cognition the theory of mental models gives, the paper will seek to capture these phenomena from well-formed formulae and derivations in first-order predicate calculus. This is the contribution of this paper, since it appears that the account the theory of mental models offers is incompatible with an account based on classical logic. The arguments below will try to show the opposite.

The first section will describe how the theory of mental models considers people to understand the conditional and disjunction. Then, the paper will offer an account of the difficulties linked to (1), (2), and (5) from the theory of mental models. The third section will address (3), (4), and (6) based on that very theory. The following section will show a possible way to deal with (1), (2), and (5) from the perspective of first-order predicate calculus. The last section will also assume the framework of first-order predicate logic, but it will be devoted to the problems of (3), (4), and (6).
According to the theory of mental models, when individuals make inferences, they review the models corresponding to every sentence. The models of every sentence represent the possible situations when the particular sentence is true. A conditional such as (7) has the models in (8) (see also, e.g., Byrne & Johnson-Laird, 2020; an example based on the thematic content in (1) is given below).

\[
\begin{align*}
\text{Ex 7} & \quad \text{If } p, \text{ then } q \\
\text{Ex 8} & \quad \Diamond(p \land q) \land \Diamond(\neg p \land q) \land \Diamond(\neg p \land \neg q)
\end{align*}
\]

Where ‘\(\Diamond(x)\)’ represents that x is possible and ‘\(&\)’ works as a conjunction (this way to express models is that used in papers such as that of López-Astorga, 2022a; in fact, the models in (8) are expressed as they are in López-Astorga, 2022a, p. 30). Symbol ‘\(\Diamond\)’ is not the operator indicating possibility in modal logic. It provides that what is between brackets is possible. However, the theory of mental models has many characteristics differentiating it from logic (see also, e.g., Johnson-Laird & Ragni, 2019).

On the other hand, the theory also assigns models to an exclusive disjunction such as (9).

\[
\begin{align*}
\text{Ex 9} & \quad \text{Either } p \text{ or } q, \text{ but not both of them} \\
\text{Ex 10} & \quad \Diamond(p \land \neg q) \land \Diamond(\neg p \land q)
\end{align*}
\]

Likewise, the theory of mental models addresses inclusive disjunctions such as (11).

\[
\begin{align*}
\text{Ex 11} & \quad \text{Either } p \text{ or } q, \text{ or both of them} \\
\text{Ex 12} & \quad \Diamond(p \land q) \land \Diamond(p \land \neg q) \land \Diamond(\neg p \land q)
\end{align*}
\]

As said, the theory of mental models is not classical logic. One of the differences is that people do not always detect all the possibilities in (8), (10), and (12). If effort does not suffice, they do not succeed in noting all the possibilities (see also, e.g., Quelhas, Rasga, & Johnson-Laird, 2019). Nevertheless, the most relevant difference here is that the theory predicts that the models can be modulated by virtue of the content of the clauses (semantics) and the situation in which the sentences are used (pragmatics) (see also, e.g., Goodwin & Johnson-Laird, 2018).

The next sections show the way modulation works for the conditional and disjunction.

**Introducing a conditional and its problems**

Based on this, it is easy to understand why people do not often endorse inferences such as (1). If ’p’ means that ‘these people study geography’ and ‘q’ represents that ‘these people eat apples’, the models of the conclusion in (1) are those in (13).

\[
\begin{align*}
\text{Ex 13} & \quad \Diamond(p \land \neg q) \land \Diamond(\neg p \land \neg q) \land \Diamond(\neg p \land q)
\end{align*}
\]

This causes a problem: one of the models in (13) shows a scenario in which the premise is not the case. It is the third model in (13), that is, the model in which p is not true and q is not false (\(\neg p \land q\)). If that scenario is possible when the conclusion holds, the conclusion cannot be supported: the conclusion implies the possibility of the premise (\(\neg q\)) being false (this section is literature review; so, an explanation similar to this one, but resorting to different examples, is to be found in Orenes & Johnson-Laird, 2012).
However, the case of (5), whose conclusion is ‘if these people eat fruits, these people do not eat apples’, is different. Its content triggers modulation processes (see also, e.g., Quelhas, Johnson-Laird, & Juhos, 2010). In principle, it seems that, if the meaning of ‘p’ is modified assuming that it designs that ‘these people eat fruits’, nothing changes. But this is not correct. The contents of the clauses make it impossible for the models of the conclusion in (5) to be those in (13). They are those in (14).

\[
\Box(p \& \neg q) \& \Box(\neg p \& \neg q)
\]

The difference between (13) and (14) is that (14) does not include the third model in (13). Apples are fruit. So, it is not possible to eat apples without eating fruits. This eliminates the problems, since that was the model inconsistent with the premise (not-q). In (14) there is no case in which \(q\) is true. That explains why individuals many times think that inferences such as (5) are valid (this section is literature review; so, an account akin to this one with different examples is to be found in Orenes & Johnson-Laird, 2012).

**Introducing a disjunction and its problems**

The approach of the theory of mental models helps understand the cases of (3) and (6), too. If the equivalences are now that ‘p’ indicates that ‘these people eat cake’ and ‘q’ provides that ‘these people eat fruits’, the models for the conclusion in (3) are those in (12). This causes difficulties again. The second model in (12) refers to a scenario in which the premise (\(q\)) does not occur. Accordingly, the conclusion is unacceptable. If it were accepted, a scenario in which the premise is not the case would be possible. That would lead to an inconsistency (this section is literature review; so, an explanation similar to this one with different examples is to be found in Orenes & Johnson-Laird, 2012). Nonetheless, in (6) modulation acts again (see also, e.g., Quelhas & Johnson-Laird, 2017). If ‘p’ denotes that ‘these people eat apples’ and ‘q’ keeps having the same meaning, the conclusion in (6) is not compatible with (12). Its models are those that (15) reveals.

\[
\Box(p \& q) \& \Box(\neg p \& q)
\]

As in the previous case, given that nobody can eat apples without eating fruits, the controversial model has disappeared. If the premise in (6) is true, its conclusion can be supported. The conclusion implies no model denying the premise. It is obvious why people rarely consider inferences such as (6) to be incorrect (this section is literature review; so, an explanation akin to this one with other examples is to be found in Orenes & Johnson-Laird, 2012).

**Introducing a conditional from first-order predicate logic**

This section and the next one develop the aim of the paper. As pointed out, and it can be noted below, the methodology is linked to the search for resources within classical logic. Those resources should be able to help account for the phenomena described above. In particular, the methodology is focused on the review of the logical forms of the inferences considered.

The differences between (1) and (5) can be explained from first-order predicate calculus if it is possible to show that their logical forms are not the same. This seems difficult, since both (1) and (5) consist of a premise and a conclusion with a conditional sentence. In both of them ‘if’ appears in the conclusion, which leads to the logical connective ‘\(\rightarrow\)’. This connective is understood in first-order predicate logic as a material implication.

However, the use of a word such as ‘if’ in a natural language (in this case, English) does not always mean that the logical form of the sentence is related to a conditional formula in logic. It is possible that the relation provided is biconditional. That is what happens, for example, in sentence (16).

\[
\text{Ex 16 If they are human beings, they have human rights}
\]

Two conditional relations exist in (16). One of them is from the first clause (‘they are human beings’) to the second one (‘they have human rights’). The other one is from the second clause (‘they have human rights’) to the first one (‘they are human beings’). If they are human beings, they have human rights. But, if they have human rights, they are human beings.
There is a lot of literature, which can be reviewed, about the reasons why a conditional can be deemed as a biconditional (see, e.g., López-Astorga, 2022b, where different works in that way are cited). The works in that literature often differentiate between sufficient and necessary conditions (e.g., Moldovan, 2009). The idea is that, whenever the antecedent of a sentence with ‘if’ can be interpreted not only as a sufficient condition for the second clause (which is what provides the conditional relation) but also as a necessary condition for that very clause, a biconditional logical form should be assigned to the sentence. In (7), the sufficient condition for the consequent \( q \) is the antecedent \( p \). Likewise, the necessary condition for the antecedent \( p \) is the consequent \( q \). Nevertheless, if the consequent \( q \) also had a necessary condition, and that necessary condition were the antecedent \( p \), the antecedent would have in turn a sufficient condition. That sufficient condition would be the consequent \( q \). This is what would lead to a biconditional interpretation of (7) (e.g., Moldovan, 2009).

This applies to (16). Being a human being is a sufficient condition for having human rights. In the same way, it is also a necessary condition for that. As indicated, if an individual has human rights, that implies that the individual is a human being, and there are no further possibilities. The reason is that to have human rights is a sufficient condition for being a human being, too (this is the case at least at present. In the future, rights such as the human rights might be extended beyond human beings. But nowadays understanding (16) as a biconditional sentence seems correct).

The problem is that, when context is limited, it is not easy to determine whether a particular sentence is a conditional or a biconditional (see also, e.g., López-Astorga, 2022b). That is what can occur with a sentence such as the conclusion in (1). It is known nothing about the scenario in which, when these people study geography, they do not eat apples. So, the conditional interpretation is possible in that case.

One of the points of this paper is that this can account for why people tend to deny inferences such as (1). If their conclusions are not conditional sentences, but biconditional sentences, their logical form is not (2), but (17).

\[
\exists x \lnot Ax \therefore \exists x (Gx \leftrightarrow \lnot Ax)
\]

Where ‘\( \leftrightarrow \)’ refers to biconditional relation.

But (18) cannot be deduced from (19) in first-order predicate calculus.

\[
\exists x (Gx \leftrightarrow \lnot Ax) \\
\exists x \lnot Ax
\]

Hence, (17) cannot be admitted in that very calculus. That can explain the rejection of inferences such as (1).

The circumstances are different for (5). The biconditional interpretation is harder for its conclusion. There is a semantic relation between the words ‘fruits’ and ‘apples’. People know that the fact of eating fruits is a necessary condition for the fact of eating apples, as well as that the fact of eating apples is a sufficient condition for the fact of eating fruits. For this reason, although the context is not detailed here either, it is difficult to establish that the consequent of the conclusion in (5), that is, ‘these people do not eat apples’ is a sufficient condition for its antecedent, that is, ‘these people eat fruits’. This is because what is known as correct is that the negation of the consequent of the conclusion in (5), that is, ‘these people eat apples’, is a sufficient condition for its antecedent, that is, ‘these people eat fruits’. If these people eat apples, obviously, these people eat fruits. However, if these people do not eat apples, it is not possible to state that these people eat fruits for sure.

On the other hand, it is also hard to assume that the antecedent of the conclusion in (5), that is, ‘these people eat fruits’, is a necessary condition for the consequent of that very conclusion, that is, ‘these people do not eat apples’. As mentioned, general knowledge shows that what is the case is that the antecedent of the conclusion in (5), that is, ‘these people eat fruits’, is a necessary condition for the denial of the consequent of the conclusion in (5), that is, ‘these people eat apples’.

A skeptic might argue that all of this does not make a biconditional interpretation of the conclusion in (5) impossible. That is true. Nevertheless, it reveals, at a minimum, that the biconditional interpretation of that conclusion is not easy. This circumstance can explain, from first-order predicate logic, why people many times embrace...
inferences such as (5). Its biconditional interpretation is not impossible, but it is less likely. Accordingly, it can be thought that most individuals understand that the logical form in these cases is (2), and not (17). (2) is correct in first-order predicate calculus, (17) is not. Of course, this account could be further supported by means of experimental results in this way, which can be obtained in future studies. However, the literature by itself already gives evidence in favor of the account.

Another possible objection against the account in this section could be that it only considers two examples, (1) and (5), which, in addition to being insufficient, do not directly come from the literature. Nonetheless, there is a response to this objection as well. Most of the inferences used in the literature to study these phenomena share the characteristics for (1) and (5) indicated in the present paper. The inferences people often deem as incorrect include, as (1), conclusions whose clauses do not have semantic relations. Besides, they are presented to participants without describing a context (they are just shown and the participants are asked for answering whether or not they are correct). On the other hand, the inferences usually admitted contain, as (5), conclusions in which there are semantic relations between their clauses such as that dealt with here. The negations of their consequents imply relations of necessary and sufficient conditions such as those explained above (many examples of these two kinds of inferences are to be found in, e.g., Orenes & Johnson-Laird, 2012. It would not be hard to apply the arguments this section provides to them).

**Introducing a disjunction from first-order predicate logic**

In principle, the situation is similar in this case. It could be stated that (3) and (6) are different only if their logical forms did not match. However, it seems that the logical form of both inferences is (4), that is, \( \exists x \ Fx : \exists x (Cx \lor Fx) \). Both of them have a premise and a conclusion consisting of a disjunctive sentence.

But, as known and indicated above, disjunctions can be exclusive or inclusive. Sometimes, it is necessary to mention ‘but not both of them’ to clarify that a disjunction is exclusive, and ‘or both of them’ to make it clear that it is inclusive. This is what has been done here up to now. Nevertheless, the semantic content of a disjunction can reveal whether it is exclusive or inclusive, too. For instance, (20) is an exclusive disjunction.

**Ex 20**  
*Either it is summer or it is winter*

If (20) is true, one of its two disjuncts requires to be true. But it is not possible for those disjuncts to be the case at the same time when the sentence is true.

The literature has experimentally revealed that, when a disjunctive sentence is ambiguous and the context does not reflect whether it is exclusive or inclusive, people typically deem it as an exclusive disjunction (e.g., Khemlani, Orenes, & Johnson-Laird, 2014). Disjunction is normally inclusive in classical logic, but logic has ways to capture exclusive disjunctive relations. In propositional logic, the logical form of (11) is (21).

**Ex 21**  
\( p \lor q \)

And that of (9) can be, for example, (22) (see, e.g., López-Astorga, 2015).

**Ex 22**  
\( (p \lor q) \land \neg (p \land q) \)

Where ‘\( \land \)’ represents conjunction.

Based on this, it can be thought that, given (3), whose conclusion is either these people eat cake or these people eat fruits, people will tend to understand its conclusion as an exclusive disjunctive sentence. Hence, its real logical form would not be (4), but (23).

**Ex 23**  
\( \exists x Fx : \exists x [(Cx \lor Fx) \land \neg (Cx \land Fx)] \)

The problem is that (23) does not hold in first-order predicate calculus. This can explain why people do not usually agree with inferences such as (3).
The case of (6) is not the same. Its logical form cannot be (23). Semantics shows that the disjunction in its conclusion is inclusive. If these people eat apples, then they eat fruits. This reveals not only that the second disjunct can be true when the first is true, but also that it is impossible that the first disjunct is true and the second one is false. So, the logical form of (6) is (4), which is admissible in first-order predicate calculus. This accounts for why people often adopt this kind of inference.

Again, an objection is possible. This section has focused only on two examples, (3) and (6). They do not directly come from the literature either. Nevertheless, the answer to this objection is similar to that given for the last objection in the previous section. The inferences used in the literature to show that introducing a disjunction is problematic has the same characteristics as (3). They are presented to participants without context. Accordingly, following experimental results (e.g., Khemlani et al., 2014), their disjunctions should lead to exclusive interpretations. On the other hand, the inferences people almost always admit are akin to (6). This means that they include disjunctive conclusions whose content indicates that they are inclusive. Hence, interpreting them in an exclusive way is not possible (many examples of these two types of inferences are to be found in Orenes and Johnson-Laird (2012). It would not be hard to apply the arguments this section provides to them).

Conclusions

The theory of mental models explains why some inferences that are valid in logic are, at the same time, difficult to endorse for people. This is the case of the logical rules to introduce the conditional and disjunction. The problem, in both cases, is that the conclusion is linked to three models, one of them being incoherent with the premise.

Nevertheless, the theory can also account for why those rules are sometimes validated. That happens in cases where the semantic content of the conclusions modulates. This makes the models inconsistent with the premises impossible, which in turn allows deeming the conclusions as correct conclusions.

The present paper has not tried to offer an account alternative to that of the theory of mental models. Without querying its explanation of these phenomena and human cognition in general, the paper has attempted to express these facts from a framework based on first-order predicate logic. The idea has not been to propose that human beings reason following that logic. It has been only to present the results achieved by the theory of mental models from the resources first-order predicate logic gives.

Thus, it can be said that, in the case of the conditional, the introduction rule is not embraced when the conditional conclusion can be, with a high probability, interpreted as a biconditional sentence. This occurs when, because of the lack of context, the clauses can be understood as both sufficient and necessary conditions. But, when the content refers to clearer necessary and sufficient conditions, the possibilities to interpret the conditional conclusion as a biconditional conclusion decrease. This is what causes the inference to be supported to a greater extent. As far as disjunction is concerned, the inference tends to be deemed as inadmissible because people often understand disjunctions as providing exclusive relations. However, the rule to introduce disjunctions in first-order predicate calculus requires disjunctions to be inclusive. For this reason, when the content reveals that a disjunction is inclusive with no doubt, the inference is rarely denied.

As stated, this does not mean that the human mind works as first-order predicate logic determines. It only indicates that there are cognitive phenomena that can be expressed resorting to formulae and rules from that logic. That can help distinguish, within the framework of first-order predicate logic, kinds of inferences that appear to be similar.

Conflict of Interest

The authors declare no conflict of interest regarding the publication of this article.
References


Santrauka

Miguel López-Astorga

Išvadų paradoksai, dvilypė interpretacija ir nesuderinamumas

MIGUEL LÓPEZ-ASTORGA
Professor, Institute of Humanistic Studies, University of Talca, Talca Campus, Chile

Research Interests
Philosophy of Language, relations between syntax, semantics, and pragmatics

Address
Instituto de Estudios Humaníticos, Universidad de Talca. Av. Lircay s/n, 3460000 Talca, Chile

E-mail milopez@utalca.cl

Orcid ID 0000-0002-6004-0587