EXCLUSIVITY AND KINDS OF MODELS IN MODULATED DISJUNCTIVE INFERENCES

Exclusividad y tipos de modelos en inferencias disyuntivas moduladas

Miguel López-Astorga

Fecha de recepción: 03 marzo de 2017
Fecha de aceptación: 24 mayo de 2017
Abstract

The disjunction introduction rule, that is, the schema in classical logic that, from a premise, leads to conclude a disjunction transforming the premise into a disjunct of that disjunction, is a cognitive problem. This is so because people only use this rule on a few occasions. Therefore, any theory trying to account for human reasoning must explain this phenomenon as well. Based on the mental models theory, Orenes and Johnson-Laird provide such an explanation. In this way, this paper is intended to show that, although it can seem that their account is in contradiction with some essential assumptions and important developments of the aforementioned theory, that is not actually so. The key points in this regard that will be analyzed are the way the mental models theory really understands disjunction and the distinction that this last approach presents between Mental Models and Fully Explicit Models.

Keywords: disjunction, inference, logic, models, reasoning.

Resumen

La regla de introducción de la disyunción, esto es, el esquema de la lógica clásica que, a partir de una premisa, lleva a concluir una disyunción transformando la premisa en uno de los términos de tal disyunción, es un problema cognitivo. Esto es así debido a que las personas solo utilizan esta regla en ciertas circunstancias. Por tanto, toda teoría que trate de describir el razonamiento humano tiene que explicar también este fenómeno. Basándose en la teoría de los modelos mentales, Orenes y Johnson-Laird proponen una explicación a este respecto, y este trabajo pretende mostrar que, aunque puede parecer que su explicación es contradictoria con algunos supuestos esenciales y ciertos desarrollos importantes de la teoría mencionada, tal no es verdaderamente el caso. Los puntos clave en este sentido que serán analizados son el modo en que la teoría de los modelos mentales realmente entiende la disyunción y la distinción que este mismo enfoque plantea entre Modelos Mentales y Modelos Completamente Explicitos.

Palabras clave: disyunción, inferencia, lógica, modelos, razonamiento.
Introduction

A habitual phenomenon in human cognition seems to prove that our reasoning is not absolutely logical, at least if the criterion assumed to check that is classical logic. In particular, the problem is a rule that not only is valid in this last logic, but that it can even considered as a basic schema in that logic. Indeed, Deaño (1999: 154) shows that, in a framework based on Gentzen’s (1934-1935) work, that is, a framework whose structures must be well-formed formulae of classical logic, the aforementioned rule has that status. The rule is the disjunction introduction rule (from now on, DIR), which is often expressed in this way:

\[
\begin{array}{c}
P \\
\hline
\text{So } p \lor q
\end{array}
\]

Or, if preferred,

\[
\begin{array}{c}
q \\
\hline
\text{So } p \lor q
\end{array}
\]

Where ‘\lor’ represents disjunction.

This schema causes difficulties because most of the time people do not apply it or think that it is incorrect. It is accepted only in certain specific cases. Of course, this fact can lead one to assume, as said, that the human inferential activity is not logical -or that that activity is not necessarily coherent with standard logic- and to think that, if that were so, the rule should be used. However, what is important for this paper in this way is that, obviously, any framework proposed with the intention to explain reasoning cannot ignore this phenomenon. Such a framework has to account for it and predicts the situations in which the rule will be applied and the circumstances in which it will not.

The mental models theory (from now on, MMT) appears to be able to do that. Indeed, Orenes and Johnson-Laird (2012) give an explanation following this last approach that proposes the reasons why people only use DIR on certain particular occasions and the usual behavior is to reject it. Nevertheless, it can be thought that Orenes and Johnson-Laird’s (2012) account of DIR is not completely coherent with the general framework of MMT, the reason being that this theory includes certain assumptions that can lead to predictions absolutely different from those raised by them. Thus, possible problems that can be indicated are the difference between Mental Models and Fully Explicit Models that MMT provides, and the fact that further developments of this very approach (e.g., Khemlani, Orenes, & Johnson-Laird, 2014) have made it evident that the habitual way people understand disjunction does not match that described by Orenes and Johnson-Laird (2012). However, my main aim here is to show that such problems are not real from Orenes and Johnson-Laird’s (2012) framework and that hence their arguments are absolutely compatible with the general approach of MMT.
In this way, to achieve that goal, firstly, I will comment on Orenes and Johnson-Laird’s (2012) proposal about DIR. Then I will address the points of MMT that are apparently in conflict with their explanation to argue that there are not really contradictions between such points and that explanation. Accordingly, I begin focusing on essentially Orenes and Johnson-Laird’s (2012) work.

1. DIR in Orenes and Johnson-Laird’s (2012) paper

Actually, the study carried out by Orenes and Johnson-laird (2012), which not only considers disjunction but also the conditional, has been analyzed and reviewed in several different papers (e.g., López-Astorga, 2015). Nonetheless, I will try to keep to just what is claimed in Orenes and Johnson-Laird’s original text about disjunction here. In this regard, it can be said that their basic idea is that, as indicated by MMT (see, e.g., Johnson-laird, 2012: 138, Table 9.2), a disjunction such as \( p \lor q \) refers to a model with three semantic possibilities or possible scenarios:

\[
\begin{align*}
[1]: & \ p \land q \\
[2]: & \ p \land \neg q \\
[3]: & \neg p \land q
\end{align*}
\]

where ‘\( \neg \)’ expresses negation and ‘\( \land \)’ can be understood as conjunction.

This means that \( p \lor q \) can describe [1], i.e., a situation in which both of the disjuncts are true, [2], i.e., a situation in which only the first disjunct (p) is true, or [3], i.e., a situation in which only the second disjunct (q) is true. So, if q is the premise, the second possibility [2] is precisely the problem with DIR, since, as it can be noticed, it indicates that q is false. One example can be helpful to explain this in a clearer way. Let us think about this inference with the formal structure of DIR:

"Viv is here.
Therefore, Pat is here or Viv is here, or both” (Orenes & Johnson-Laird, 2012: 362).

As pointed out and explicitly shown by Orenes and Johnson-Laird (2012: 362), the disjunction embedded into the conclusion can be true in three cases:

\[
\begin{align*}
[1]: & \ Pat \land Viv \\
[2]: & \ Pat \land \neg Viv \\
[3]: & \neg Pat \land Viv
\end{align*}
\]

Clearly, [2] is incompatible with the premise Viv is here, as it provides that Viv is not here. Thus, this is the cause that people tend not to consider DIR to be correct, since its conclusion includes a possibility that is absolutely inconsistent with the previous information, that is, the premise.

But, as said, there are also cases in which individuals accept the rule and, if a theory wants to exactly describe the way the human deductive activity really works, it should provide an account of them too. As also mentioned, Orenes and Johnson-Laird (2012) present such an account as well, which is supported
by an important thesis of MMT. According to this last theory, when people are identifying the models of a sentence, modulation processes can happen. Such processes are related to semantics and pragmatics and can modify the content of a particular model (by adding or eliminating possibilities) by virtue of the meaning of the words that the sentence includes and the circumstances in which it is expressed (in addition to Orenes & Johnson-Laird (2012), see, e.g., Johnson-Laird (2012), Johnson-Laird & Byrne (2002), or Johnson-Laird, Khemlani, & Goodwin (2015). This can be seen in an obvious way by means of another example with the same structure as the previous one:

"Lucia wore jewelry.

Therefore, Lucia wore the bracelet or she wore jewelry" (Orenes & Johnson-Laird, 2012: 363).

Now, a modulation process occurs and its result is the elimination of [2] as a possibility for the disjunction in the conclusion:

[1]: Bracelet & Jewelry
[3]: ¬Bracelet & Jewelry

And the reason is not hard to understand: if pragmatics and the exact meanings of the words are taken into account, according to MMT, it can be thought that it is not possible wearing a bracelet and not wearing jewelry, as the former pragmatically and semantically implies the latter. So Orenes and Johnson-Laird (2012) claim that, in cases such as this one, which are called by them cases of modulated inference, people should consider DIR to be valid, since there is no contradiction between the premise and the possibilities of the conclusion (neither [1] nor [3] are in conflict with the premise). And, certainly, their second experiment appeared to confirm their predictions. In that experiment they used tasks of the two types commented on and the results were clear: most of their participants thought that the non-modulated inferences were not correct and that the modulated ones were valid.

Of course, there are works that criticize this account and propose an alternative explanation for the results presented by Orenes and Johnson-Laird (2012), for example, that of López-Astorga (2015), which also gives a commentary on their proposal akin to the one just provided above. I will briefly refer to López-Astorga’s (2015) paper below but, as far as my aims here are concerned, what should be done now is to indicate the reasons why certain Orenes and Johnson-Laird’s (2012) arguments that can seem problematic assuming the basic framework of MMT are not truly so. I do that in the next section.

2. The exclusive character of disjunction and the difference between Mental Models and Fully Explicit Models

A first difficulty in Orenes and Johnson-Laird’s (2012) explanation can be that, if later papers supporting MMT are taken into account, it appears that their account needs to be reviewed. Indeed, for example, the results of one of the experiments included in Khemlani et al.’s (2014) research make it evident that people usually interpret disjunctions as exclusive. In a condition of that experiment, they presented to
their participants a disjunction such as \( p \lor q \) and asked them to respond with which of these possibilities that disjunction was compatible:

\[
[1]: p \land q \\
[2]: p \land \neg q \\
[3]: \neg p \land q \\
[4]: \neg p \land \neg q
\]

The majority response was that [2] and [3] were the possibilities linked to \( p \lor q \), [1] and [4] being ignored. [4] is not a problem, since it was not considered by Orenes and Johnson-Laird (2012). Nevertheless, [1] does be a difficulty because they included it in their comments on their examples. In this way, if we pay attention to the results of this experiment, which was carried out with thematic content, and not, as my description here, with just formal structures, it must be stated that the real possibilities corresponding to the example of non-modulated inference above that should have been mentioned by Orenes and Johnson-Laird (2012) are just:

\[
[2]: \text{Pat} \land \neg \text{Viv} \\
[3]: \neg \text{Pat} \land \text{Viv}
\]

Likewise, in the case of the modulated inference, only one scenario should have been taken into account:

\[
[3]: \neg \text{Bracelet} \land \text{Jewelry}
\]

And this is so because, as indicated, these are the actual possibilities that, following Khemlani et al.'s (2014) results, people could identify if faced to disjunctions such as those embedded into the conclusions of the examples cited. Of course, it must be acknowledged that the fact that, as pointed out, the expression \textit{or both} appears in the non-modulated version can lead individuals to notice that the disjunction is inclusive and that [1] is a possible scenario. Nevertheless, in the same way, it must also be said that, in the materials that were really used in their experiment, the non-modulated inferences did not have such an expression (see Orenes & Johnson-Laird, 2012: 375), that the example of modulated inference, as shown, did not include it either, and that the actual modulated inferences presented to their participants lacked it as well (see also Orenes & Johnson-Laird, 2012: 375).

Thus, it can be thought that Khemlani et al.'s (2014) experimental results are already enough to undermine Orenes and Johnson-Laird's (2012) explanation. However, this is not truly so. On the one hand, that its disjunction is exclusive does not have an influence on the non-modulated inference at all. [2] continues to be a possible scenario and, as said, it denies the premise. On the other hand, although one might think that the situation is different in the modulated inference, MMT has the necessary resources to argue in favor of Orenes and Johnson-Laird's (2012) account regardless of whether the disjunction is inclusive or exclusive. The case in which the disjunction is regarded as inclusive has already been commented on above. As far as the alternative in which it is considered as exclusive is concerned, modulation
can solve the problem as well and show why the inference must be accepted. It is true that firstly it can be thought that the prediction should also be that people would reject the inference. And this is so because the only possibility of the conclusion is [3], which provides that the bracelet is not worn, that is, an information that cannot be derived just from the premise, i.e., from the fact that jewelry is worn. However, this could only be actually thought after a superficial reading of the theses raised by MMT.

As explained, [2] is removed because modulation reveals that it is not possible a scenario with the bracelet and without jewelry. Nevertheless, that very modulation process can make the disjunction inclusive, since it shows that [2] is not possible because, as also commented on, whenever there is a bracelet, there is jewelry as well. And this clearly leads to [1] as another possibility. Certainly, modulation does not only eliminate possible scenarios. It can also modify the elements of a particular scenario (Johnson-laird et al., 2015: 202). Hence it can be admitted that, in the case of the modulated inference, modulation does not just removes [2], but that it transforms [2] into [1] (by virtue of general knowledge, it is known that it is not possible having a bracelet without having jewelry, but it is also known that, if a bracelet is had, jewelry is had too). Accordingly, [1] and [3] can be the possibilities again. Thus, it can be interpreted that what the disjunction in the conclusion really means is that, whether the bracelet is worn [1] or not [3], as indicated in the premise, jewelry is always worn (which is true in both [1] and [3]).

So, given that the possibilities can be the same if the disjunction is understood as exclusive, this is not a great problem for Orenes and Johnson-laird’s (2012) proposal, as their general arguments can continue to be raised. In this way, what Khemlani et al.’s (2014) study actually implies is not that the explanation given by Orenes and Johnson-laird (2012) does not hold, but only that that explanation needs to be updated by means of the information achieved with more current research.

Nonetheless, another point that appears to be a true difficulty for that explanation and allow questioning it is the fact that, from the beginning, MMT distinguishes between Mental Models and Fully Explicit Models (see, e.g., Johnson-Laird, 2004, 2006, 2012; Johnson-Laird et al., 2015; Oakhill & Garnham, 1996). The distinction is based on an essential principle of the theory: the principle of truth. According to it, the Mental Models “…represent only what is possible given the truth of assertions, not what is impossible…” (Johnson-Laird, 2012: 138). And this implies that only secondly individuals can identify the Fully Explicit Models, that is, the sets of possibilities with the complete scenarios standing for both what is true in them and what is false in them. Thus, although transforming Mental Models into Fully Explicit Models is possible, it is also an activity that can be hard and require deep reflection. For this reason, at first most people tend to only detect the Mental Models of each sentence.

For the case of an exclusive disjunction such as $p \oplus q$ (where $\oplus$ expresses exclusive disjunctive relationship), this means that the initial models, that is, the Mental Models, considered by most individuals are not even [2] and [3], which are the Fully Explicit Models, but simply:

\begin{align*}
[5]: & \quad p \\
[6]: & \quad q
\end{align*}

As it can be checked, [5] and [6] correspond to [2] and [3], the difference being that, certainly, in the former what is denied, that is, what is false ($\neg q$ in [5] and $\neg p$ in [6]) does not appear. But, if this is so, it must also be admitted that, given a non-modulated inference such as that described above, the scenarios
firstly considered are truly these ones:

[5]: Pat
[6]: Viv

This, clearly, is not a problem for Orenes and Johnson-Laird’s (2012) predictions, since, while [5] is not absolutely inconsistent with the premise Viv is here, it is obvious that the fact that Pat is here cannot be deduced from that premise alone. So the expected response continues to be that the inference is not correct.

However, although it can seem to be the contrary, the situation does not radically change in the case of the modulated inference either. From what has been said, it is clear that the Mental Models of the example indicated are:

[5]: Bracelet
[6]: Jewelry

Obviously, it can be thought that, at first, modulation cannot have an influence on [5]. If the information that the jewelry is not worn does not appear, it is impossible to note that the scenario expresses a contradiction (the bracelet is worn and jewelry is not) and has to be removed or modified. So, it could be stated that, as in the previous case of the non-modulated inference, although it is not incoherent with the premise Lucia wore jewelry, [5] cannot be inferred from it (that Lucia wore jewelry does not lead to that she wore the bracelet). In this way, it could also be claimed that really the inference should be considered as incorrect, and that Orenes and Johnson-Laird’s (2012) account can only be accepted if it is supposed that individuals always (or at least almost always) identify the Fully Explicit Models in inferences with the structure of the tasks used by them. Nevertheless, in their defense, Orenes and Johnson-Laird (2012) could also argue that, given that individuals know the meanings of both bracelet and jewelry, it is possible to complete [5] (or, if preferred, to transform it into [1]) with the information that, because the bracelet is worn, jewelry is worn as well. This is clearly possible in the theory because it provides that modulation can help identify the Fully Explicit Models as well (see, e.g., Johnson-Laird & Byrne, 2002: 658). Thus, the possibilities would be in this case [1] and [6], that is, that Lucia wore both the bracelet and jewelry and that she only wore jewelry. Such scenarios would be absolutely compatible with the premise, since what the conclusion would provide would be that jewelry is always worn ([1] and [6]), with the possibility that, in addition, the bracelet, which is, as said, jewelry too, is worn as well [1]. Hence it can be held that, even when individuals only identify the Mental Models, the prediction of MMT continues to be that they will consider the modulated inference to be correct. After all, the premise is not denied in any of the possible scenarios, and, as explained, the only additional element (the bracelet) may be [1] or may not be [6].

Therefore, neither the fact that people tend to interpret disjunction as exclusive nor the differentiation between Mental Models and Fully Explicit Models are a real problem for Orenes and Johnson-Laird’s (2012) account. Given that, based on the basic ideas of MMT, these are the two most immediate objections that can be presented against their explanation, it can be stated that their proposal is clearly coherent with such ideas.
Conclusion

Indeed, Orenes and Johnson-Laird's (2012) approach is incompatible neither essential theses of MMT (e.g., the distinction between Mental Models and Fully Explicit Models) nor further developments of this last theory (e.g., the finding that disjunction is mainly exclusive in the human thought). The first point, the one related to the differentiation between types of models, is especially important, since it has been said that a great weakness of MMT is that it does not explain in detail neither the exact situations in which it is more probable that the Fully Explicit Models are identified nor the exact circumstances in which only the Mental Models are often taken into account (e.g. López-Astorga, 2016). Nonetheless, even supposing that this weakness is real, as seen, it does not have an influence on the arguments provided by Orenes and Johnson-Laird (2012).

On the other hand, as said above, other theories present alternative explanations to the one of Orenes and Johnson-Laird (2012). However, such theories also have weaknesses. An example in this regard can be López-Astorga's (2015) argumentation. In his view, it is possible to continue to argue that a logic exists in the human mind. It is only necessary to acknowledge, as, for instance, the mental logic theory (e.g., Braine & O'Brien, 1998; Gouveia, Roazzi, O'Brien et al. 2003; O'Brien, 2014) does, that that logic is not exactly classical logic. Thus, it can be proposed, as the mental logic theory does too, that, while many rules of standard logic and the natural deduction calculus are easily applied by human beings, not all of them are so, DIR being precisely one of the schemata valid in the mentioned logic and the mentioned calculus that are difficult for people. In this way, the results of the non-modulated inferences are explained: if DIR is hard to individuals, it makes sense that that kind of inference is considered as incorrect.

As far as the results achieved by Orenes and Johnson-Laird (2012) with the modulated inferences are concerned, the solution given by López-Astorga (2015) is simple. If a disjunction such as *Lucia wore the bracelet or she wore jewelry* only refers to the scenarios [1] and [3], it is not a true disjunction. If a formula such as p ∨ q is only true in the cases of p and q, and ¬p and q, this fact reveals that its truth table in classical logic is not the one of disjunction, but that of a formula such as (p ∨ ¬p) → q, where ‘→’ represents conditional relationship, or just q. Thus, the real structure of the modulated inference is not

\[
\begin{align*}
&\text{q} \\
&\text{-------------} \\
&\text{So p ∨ q} \\
&\text{But} \\
&\text{q} \\
&\text{----------------} \\
&\text{So (p ∨ ¬p) → q} \\
&\text{or, still more clearly,} \\
&\text{q} \\
&\text{-----} \\
&\text{So q}
\end{align*}
\]
It is evident that, regardless of the content of \( q \), if this last formula is taken as a premise and it is asked whether or not it follows from it that \( q \) is true, most people will respond positively. And this shows that López-Astorga’s (2015) thesis that the actual logical form of the conclusion in the modulated inference is not that of disjunction can be a good explanation of why this last kind of inference is often considered as valid by individuals.

However, a problem of López-Astorga’s (2015) argumentation is that it must face results such as those presented by Khemlani et al. (2014) as well, according to which individuals generally interpret disjunctions as exclusive. This is an important point because it reveals that the only real scenario of the conclusion in the modulated inference can be, as indicated, just \([3]\), and that what has to be done, from this last account, is to think about a formula whose truth table in standard logic shows that it is only true in the case of the combination \( \neg p \) and \( q \). Nevertheless, addressing this issue, which, in principle, does not seem very difficult keeping López-Astorga’s (2015) basic approach, would take us off the main aims of this paper, which are focused on Orenes and Johnson-Laird’s (2012) arguments.

In this way, what is truly interesting here of López-Astorga’s (2015) proposal is that, as it can be noticed, while it can be an alternative explanation to that of Orenes and Johnson-Laird (2012), it is also clear that the former depends to certain extent on the latter. Indeed, as López-Astorga explicitly indicates (López-Astorga, 2015: 148), to recover the real logical form of the conclusion in the modulated inference is absolutely necessary to take into account the possibilities assigned to it by Orenes and Johnson-Laird (2012), and to assume, as MMT does, that modulation can remove some of these possibilities. And this means that, ultimately, it is necessary to accept that the possibilities of disjunction are usually \([1]\), \([2]\), and \([3]\), and that, in the modulated inferences, \([2]\) is not.

So, in short, it appears that we can speak about an important strength of MMT. Although an approach such as that of López-Astorga (2015) is assumed and it is thought that, in the modulated versions of the inference, \( p \lor q \) is not the real logical form of the conclusion, the methodology of MMT must be assumed too. And this is so because the possibilities of the disjunctive sentences have to be considered to detect the true logical forms (as said, López-Astorga, 2015, already comments on this fact).

Furthermore, although the criticism raised by López-Astorga (2016) is correct and MMT needs to clarify the factors that cause only the Mental Models to be identified in some cases and the Fully Explicit Models in entirety to be detected in other instances, there is no doubt that the distinction between these two kinds of models can be fruitful. In fact, as shown by the literature on cognition, the distinction is very useful to account for many intellectual phenomena. Hence, if MMT actually has problems in this way, it can be expected that its proponents will work on a clear description of the exact mechanisms that lead to one or other of these types of models in next studies.

References


