The conditional introduction rule and human reasoning: Findings from the mental models theory[1]

La regla de introducción del condicional y el razonamiento humano: Hallazgos de la teoría de los modelos mentales

Miguel López Astorga
Institute of Humanistic Studies Juan Ignacio Molina, Talca University, Talca, Chile
milopez@utalca.cl

Received: 15-08-2014. **Accepted**: 08-12-2014.

Abstract: Based on the mental models theory, Orenes and Johnson-Laird analyze certain inferences that, being valid in standard logic, individuals tend to reject. One type of such inferences refers to the logical rule that introduces conditionals and they propose a mechanism of modulation that makes these inferences acceptable. This brief paper shows that Orenes and Johnson-Laird's results reveal that, if the idea that human reasoning follows formal rules is accepted, a restriction on the use of the conditional introduction rule needs to be assumed. That restriction is that the denial of the antecedent must allow deriving the consequent.

Keywords: Conditional introduction, formal rules, logic, mental models, modulation.

Resumen: Basándose en la teoría de los modelos mentales, Orenes y Johnson-Laird analizan ciertas inferencias que, a pesar de que son válidas en la lógica estándar, tienden a ser rechazadas por los individuos. Un tipo de tales inferencias hace referencia a la regla lógica que introduce condicionales y ellos proponen un mecanismo de modulación que convierte a estas inferencias en aceptables. Este breve trabajo muestra

[1] This paper is a result of the Project N. I003011, "Algoritmos adaptativos e inferencias lógicas con enunciados condicionales", supported by the Directorate for Research of Talca University (Dirección de Investigación de la Universidad de Talca), Chile. The author, who is also the main researcher of that Project, would like to thank the mentioned institutions for their help in funding this paper, and Phil Johnson-Laird for a critical reading of it.

que los resultados de Orenes y Johnson-Laird revelan que, si se acepta la idea de que el razonamiento humano sigue reglas formales, es necesario asumir una restricción para la utilización de la regla de la introducción del condicional. Tal restricción es que la negación del antecedente tiene que permitir derivar el consecuente.

Palabras clave: Introducción del condicional, reglas formales, lógica, modelos mentales, modulación.

1. Introduction

There is no doubt that the mental models theory (from now on, MMT) has proven to be a very useful theory for describing human reasoning. Its predictive range is conspicuously wide, and it can easily explain most actual human inferential processes. The basic assumptions of this theory, which can be found in many texts and papers (e. g., Byrne & Johnson-Laird, 2009; Johnson-Laird, 1983, 2001, 2006, 2012; Johnson-Laird & Byrne, 2002; Johnson-Laird, Byrne, & Girotto, 2009; Orenes & Johnson-Laird, 2012), are that formal rules do not lead human reasoning and that human beings make inferences by means of mental models. Such mental models are representations of the situations described in inferences, and, according to MMT, individuals tend to prefer the models that are consistent with the sentences included in the inferences, i.e., with the premises and the conclusions.

As said, MMT is successful. However, the most important aspect for this paper is that its findings make it difficult to hold a theory on human reasoning based only on logical formal rules. Thus, the frameworks more or less similar to those of Henle (1962) or Rips (1994) can no longer ignore one of the most important theses of MMT: the role of semantics in human inference. Certainly, MMT shows us that syntax itself is not strong enough for explaining reasoning. The meaning of sentences has an influence on it and, if it is wished to continue to try to describe real inferential processes by means of formal rules, it is inevitably necessary to accept certain restrictions for such rules.

An example of this fact is analyzed in the present paper. The conditional introduction rule ($\alpha \vdash \beta \rightarrow \alpha$) is a valid logical rule. Nevertheless, the research carried out by Orenes and Johnson-Laird (2012) –from now on, OJL- demonstrates that individuals not always consider it right in practice. This means that, if we insist that human reasoning uses this rule, we need to adopt a strong restriction for it. Evidently, other options are also possible. An alternative solution to this difficulty can be, for example, to consider the material interpretation not to be the appropriate interpretation for conditionals and conditionals to need to be interpreted as defective (Adams, 1998). But the present paper only tries to show that, if the material interpretation of conditional is assumed, a constraint for the conditional introduction rule is required.

In order to explain this point, firstly, the paper describes the problems related to the conditional introduction rule detected by OJL and the solution based on MMT that they propose for such problems. It then outlines the restriction needed for that rule if it is related to human reasoning. Finally, the paper draws some consequences for cognitive science and logic.

2. A conditional paradox

OJL´s experiments are focused on paradoxical inferences that are logically valid and that, however, individuals tend to reject. In particular, they review three of these inferences, which are only paradoxical figuratively speaking, since, as said, they are logically valid. Nevertheless, only the first one, which is the inference referring to the conditional introduction rule, will be considered in this paper. The problem is the following:

OJL observed that most their participants responded negatively to a problem such as:

“Luisa didn´t play music. Does it follow that if Luisa played a game then she didn´t play music?” (OJL, p. 374).

This is a problem because, according to propositional calculus, the formalization of this inference is as follows:

$\neg q$

$p \rightarrow \neg q$

And, obviously, by the conditional introduction rule ($\neg q \vdash p \rightarrow \neg q$), it is correct. In fact, p → ¬q can also be drawn from ¬q by means of other rules. For example, by means of the disjunction introduction rule, ¬q can be transformed into ¬p ∨ ¬q, and, by the equivalence of disjunction to conditional, ¬p ∨ ¬q can be transformed into p → ¬q. Nonetheless, this paper will mainly consider the deduction by means of the conditional introduction rule. There is a very simple reason for doing so: the conditional introduction rule is a basic rule and the equivalence of disjunction to conditional is a derivate rule that involves other rules (the disjunction elimination rule, the *Reductio ad Absurdum* rule and the conditional introduction rule itself).

But what is interesting for this paper is that OJL think that MMT can easily explain the mentioned problem. Following MMT, individuals reason considering the models consistent with the sentences of inferences (in this case, the premise *Luisa didn´t play music* and de conclusion *If Luisa played a game then she didn´t play music*), and the difficulty in this example is that the conclusion is a conditional sentence and, according to the material interpretation of conditionals (and hence according to the conditional truth table), there are three possibilities in which it is true:

Luisa played a game	Luisa did not play music
Luisa did not play a game	Luisa did not play music
Luisa did not play a game	Luisa played music

These three possibilities correspond to fully explicit models of the conditional conclusion. OJL argue that their participants did not accept this inference because the third model (*Luisa did not play a game* & *Luisa played music*) is not compatible with the premise (*Luisa didn´t play music*). However, in their view, it is possible to make the inference acceptable. It is enough to block the third model, and it can be made by modulating the conditional conclusion. OJL propose different modulation mechanisms, but the relevant mechanism for this inference is to use a conditional sentence in which, when the antecedent is false, the consequent is necessarily true, that is, to use an inference such as this one:

“Andrés didn´t play soccer. Does it follow that if Andrés played a game then he didn´t played soccer?” (OJL, p. 374).

Given that soccer is a game, only two models are possible:

Andrés played a game Andrés did not play soccer
Andrés did not play a game Andrés did not play soccer

The model that would be problematic (*Andrés did not play a game* & *Andrés play soccer*) is not possible, since Andrés cannot play soccer if he does not play a game. As MMT predicts, most participants often accept modulated inferences such as this one. It is therefore clear that meaning and semantics are relevant in human inferential processes, and that prima facie an approach based only on formal rules cannot explain the present phenomenon. However, in my view, the conditional introduction rule could be consistent with these results, although it would be necessary adopt a strong restriction. In addition, that restriction would not be the only difficulty. It would be also necessary to use a calculus more complex than propositional calculus: first-order predicate calculus. In the next section, these points are described in details.

3. A restriction for the conditional introduction rule

In principle, it is not a problem to use the first-order predicate calculus. It is legitimate to adopt a logic framework and, at the same time, to consider that that framework is based on first-order predicate calculus. Nevertheless, what is a problem is that first-order predicate calculus is more complex than propositional calculus, which is an additional difficulty for any approach based on formal rules. Such an approach would also have to explain how the complicated logical forms of first-order predicate calculus are recovered from natural language.

In any case, the restriction on the conditional introduction rule that OJL´s results imply is the following:

(R) $\beta \rightarrow \neg\alpha$ can be drawn from $\neg\alpha$ only if $\neg\alpha$ can be drawn from $\neg\beta$

(R) comes from OJL´s modulation mechanism, which refers to the necessity of using a conditional in which, if its antecedent is false, its consequent is true. Its effectiveness can be noted if both the paradoxical version of the previous example and its modulated version are formalized following first-order predicate calculus. To do that, I will use the following abbreviations:

Pxy: x plays y
Gx: x is a game
l: Luisa
m: Music
a: Andrés
s: Soccer

Thus, the formalization of the paradoxical version would be as follows:

$\neg Plm$

$\exists x\,(Gx \wedge Plx) \rightarrow \neg Plm$

According to (R), this inference would be valid if $\neg\exists x\,(Gx \wedge Plx) \vdash \neg Plm$. But the case is that $\neg\exists x\,(Gx \wedge Plx) \nvdash \neg Plm$, and it is easy to note why. A possible derivation is:

(1) $\neg\exists x\,(Gx \wedge Plx)$	(premise)
(2) $\forall x\,\neg(Gx \wedge Plx)$	(by definition of $\exists$, 1)
(3) $\neg(Gm \wedge Plm)$	(by $\forall$ elimination, 2)
(4) $\neg Gm \vee \neg Plm$	(by De Morgan, 3)
(5) $\neg Gm$	(assumption for eliminating $\vee$)
(6) …	

As it is well-kown, the disjunction elimination rule establishes that $\alpha \vee \beta \vdash \gamma$ if $\alpha \vdash \gamma$ and $\beta \vdash \gamma$, i. e., if the same formula can be deduced from each disjunct. This means, in this case, that it is necessary that $\neg Gm \vdash \neg Plm$ and $\neg Plm \vdash \neg Plm$. It is obvious that $\neg Plm \vdash \neg Plm$ by identity, but it is not possible to accept that $\neg Gm \vdash \neg Plm$ in this scenario. Evidently, other derivations are also possible, e. g., $\forall x\,\neg(Gx \wedge Plx)$ can be transformed into $\forall x\,(Gx \rightarrow \neg Plx)$, but $\neg Plm$ cannot be drawn in any way.

The case of the modulated version is different. Its formalization would be as follows:

$\neg Pas$

$\exists x\,(Gx \land Pax) \rightarrow \neg Pas$

Following (R), the requirement that is needed now is $\neg \exists x\,(Gx \land Pax) \vdash \neg Pas$, and it is evident that that requirement can be fulfilled in this version:

(1) $\neg \exists x\,(Gx \land Pax)$	(premise)
(2) $\forall x\,\neg(Gx \land Pax)$	(by definition of $\exists$, 1)
(3) $\neg(Gs \land Pas)$	(by $\forall$ elimination, 2)
(4) $\neg Gs \lor \neg Pas$	(De Morgan, 3)
(5) $\neg Gs$	(assumption for eliminating $\lor$)
(6) Pas	(assumption)
(7) Gs	(by general knowledge)
(8) $Gs \land \neg Gs$	($\land$ introduction, 7, 5)
(9) $\neg Pas$	(Reductio ad absurdum, 6-8)
(10) $\neg Pas$	(assumption for eliminating $\lor$)
(11) $\neg Pas$	(identity, 10)
(12) $\neg Pas$	($\lor$ elimination, 4, 5-9, 10-11)

Therefore, the fact that the participants in experiments carried out by OJL tended to reject the paradoxical version of this kind of inference and to accept its modulated version can be explained from first-order predicate calculus (note that, if propositional calculus is used, it is very difficult to show that, in the case of the modulated version, $\neg\beta \vdash \neg\alpha$). OJL use more examples of this type of inference with different thematic content, but the structure of those other examples is very similar to that of the examples analyzed here. For this reason, it can be said that the two derivations of this section can be easily applied to such examples.

There are other real problems. Firstly, it is necessary to explain how the logical forms both of the premises and of the conclusions are recovered from natural language (they are forms corresponding to first-order predicate calculus). Secondly, (R) can be considered to be a very severe restriction that transforms logical calculus and limits its scope and its possibilities (this restriction does not allow deducting certain theorems). Thirdly, the deduction of the modulated version has twelve steps (it can be thought that they are many steps for one inferential process). And finally, step 7 in that deduction can be controversial and problematic (if the role of general knowledge is accepted, then the role of semantics, and not only that of syntax or that of the formal rules, is also accepted).

4. Conclusions

Even if the previous four problems are put aside, difficulties remain. The literature, including the paper of OJL, offers us many more examples of controversial inferences that are hard to explain from a framework only based on formal rules. In this way, as an illustration, I can mention that OJL analyze other two types of inferences that are logically valid and that, however, are not often accepted by individuals. In particular, these inferences are $\alpha \vdash \beta \vee \alpha$ and $\neg\alpha \vdash \alpha \rightarrow \beta$. OJL also present modulated versions of these inferences that individuals tend to admit and, perhaps, we can find restrictions for them that allow us to explain why they are not accepted when they are not modulated. Nevertheless, this could transform logical calculi into very different systems with excessive restrictions. In fact, probably, calculi would no longer be calculi. The restrictions would make them into systems with logical symbol that try to describe human reasoning (whether successful or not) and that do not have certain basic characteristics of logical calculi.

Undoubtedly, further research can be made in order to determine whether or not all the problems detected and explained by MMT can also be taken into account and understood from a syntactic approach exclusively based on formal rules. Nonetheless, it could not be forgotten (even in a scenario in which that activity was successful) that the resulting framework would be indebted to MMT and it would be an heir of it. Semantics is very relevant for MMT and step 7 of the second deduction presented in the previous part recalls us that all the frameworks trying to explain human reasoning must consider it. Furthermore, the problem of the recovery of logical forms, that of the considerable number of steps needed in logical deductions describing human inferences, and that of fact that the possible restrictions for formal rules can be very difficult to accept show us that MMT is ahead and has decisive advantages at the moment.

Works cited

Adams, Ernest. *A Primer Probability Logic*. Stanford: Center for Study of Language and Information Publications, 1998.

Byrne, Ruth M. & Philip Johnson-Laird. "'If' and the problems of conditional reasoning." *Trends in Cognitive Science* 13 (2009): 282-287.

Henlé, Mary. "On the relation between logic and thinking." *Psychological Review* 69 (1962): 366-378.

Johnson-Laird, Philip. *Mental Models. Towards a Cognitive Science on Language, Inference and Consciousness*. Cambridge, MA: Harvard University Press, 1983.

Johnson-Laird, Philip. "Mental models and deduction." *Trends in Cognitive Science* 5 (2001): 434-442.

Johnson-Laird, Philip. *How We Reason*. Oxford: Oxford University Press, 2006.

Johnson-Laird, Philip. "Inference with mental models." In Holyoak, K. J. and R. G. Morrison (eds.), *The Oxford Handbook of Thinking and Reasoning* (pp. 134-145). New York: Oxford University Press, 2012.

Johnson-Laird, Philip & Ruth M. Byrne. "Conditionals: A theory of meaning, pragmatics, and inference." *Psychological Review* 109 (2002): 646 - 678.

Johnson-Laird, Philip, Ruth M. Byrne & Vittorio Girotto. "The mental models theory of conditionals: A reply to Guy Politzer." *Topoi* 28 (2009): 75 - 80.

Orenes, Isabel & Philip Johnson-Laird. "Logic, models and paradoxical inferences." *Mind and Language* 27 (4) (2012): 357-377.

Rips, Lance. *The Psychology of Proof*. Cambridge: Massachusetts Institute of Technology (MIT) Press, 1994.