Φιλοσοφια
Volume 23, 2: 2022

INDUCTION IN HUMAN REASONING: GAUTAMA'S SYLLOGISM AND SYSTEM K

Miguel López-Astorga
University of Talca, Talca Campus - Chile

The literature has shown that the theory of mental models is able to describe human inductive processes. That theory has been related to the structure of inductive inferences, such as Gautama's Syllogism indicates. On the other hand, the theory of mental models has also been linked to modal system K. This paper argues that there can be a connection between Gautama's Syllogism and system K, not in rigorous logical deductions but in describing how the human mind can work. They can refer to two different moments of inductive reasoning; the rule of necessitation of K can be a key element in the second of those moments.

Keywords: Gautama's Syllogism; induction; rule of necessitation; system K; theory of mental models

INTRODUCTION

Several points have been provided in the recent literature on philosophy, cognitive science, and psychology. The first one is that there is a theory that can clearly explain the process of why individuals make inductions. That theory is the theory of mental models (e.g., Byrne & Johnson-Laird 2020, 760-780; Espino, Byrne, & Johnson-Laird 2020, 1263-1280; Johnson-Laird, Quelhas, & Rasga 2021, 951-973; Khemlani, Byrne, & Johnson-Laird 2018, 1887-1924; Khemlani & Johnson-Laird 2019, 219-228; Khemlani & Johnson-Laird 2022, 289-312). Its account of induction is to be found in works such as that of Johnson-Laird (2012, 134-145). The second one is that the explanation of induction given by the theory of mental models can be related to inferences such as that described by Gautama's Syllogism, which is the inductive inference in the book *Nyāyasūtra* Colebrooke named the 'Hindu Syllogism' (e.g., Ganeri 2004, 321). In this case, the idea has been that there is a correspondence between the inferential steps in Gautama's Syllogism and the manner the theory of mental models understands induction (e.g., López-Astorga 2016, 351-358). Finally, the theory of mental models has also been linked to a modal syntax as basic as the one offered by logical system K (e.g., López-Astorga 2020, 160-169).

Therefore, what remains to be done is to review whether there are connections between Gautama's Syllogism and system K. That is the aim of the present paper. It will try to show that the key can be the only rule of inference included in K, the rule of necessitation. This is because that rule seems consistent with what people do in the second moment of induction processes described by Gautama's Syllogism and the theory of mental models. However, two points are important here. The paper will propose only a possible way to connect Gautama's Syllogism and system K. They are very different frameworks. So, the main goal will be to look for a possible link to move from one of them to the other. On the other hand, system K and its rule of necessitation will be understood in a loose sense. The present paper will not try to show rigorous logical deductions based on that system. On the contrary, the idea is to offer a description of how the human mind can work. System K can help with this, even if its essential rule is taken figuratively and what is mainly assumed is its language. As indicated below, this means to apply the rule of necessitation to what is considered a fact by individuals but keeping the modal operators and the semantics of possible worlds.

The paper will have four main sections. The first section will be devoted to the theory of mental models. It will comment on some of its most important theses, specifically, how it accounts for induction. Then, the paper will present Gautama's Syllogism and the relations that have been established between it and the theory of mental models. The third section will address the links that have been proposed between that very theory and system K. The last section will try to indicate how Gautama's Syllogism and system K can collaborate to help understand mental inductive processes.

INDUCTION AND THE THEORY OF MENTAL MODELS

The theory of mental models considers, *inter alia*, all the connectives in classical propositional logic. Its purpose with that is to attempt to explain how people reason about sentences, including those connectives. The general idea of the theory is that individuals tend to relate different possibilities to the connectives (See also, e.g., Johnson-Laird & Ragni 2019). Thereby, given a conditional such as (1),

(1) If today is Tuesday, then you have English class.

According to the theory of mental models, these models can be assigned to it (the following way to express the models is akin to the one habitually used by the proponents of the theory and in papers such as López-Astorga 2020, 160-169):

(2) Possible (p & q) & Possible (¬p & q) & Possible (¬p & ¬q) & Impossible (p & ¬q)

Where 'p' stands for the antecedent, 'q' refers to the consequent, '¬' indicates negation, and '&' represents conjunction.

What (2) reveals is that there are three cases in which (1) can be true (its three first conjuncts, which are possibilities) and one case in which (1) is false (its fourth conjunct, which is an impossibility). Since the only scenario in which (1) is false is, hence, the circumstance in which today is Tuesday, and you do not have English class, one might think that this account is not very different from the one that standard logic offers by means of the truth table of the conditional. However, it is. The theory of mental models claims that semantics and pragmatics can modulate the possibilities of sentences (e.g., Johnson-Laird & Byrne 2002, 646-678; Orenes & Johnson-Laird 2012, 357-377). That is the case of, for instance, (3).

(3) If this writer is Portuguese, then he is José Saramago.

Because José Saramago is, in fact, a Portuguese writer, the possibilities corresponding to (3) are not (2) but (4).

(4) Possible (p & q) & Possible (p & ¬q) & Possible (¬p & ¬q) & Impossible (¬p & q)

Clearly, (4) violates the truth table of the conditional in classical logic. It allows cases of p and ¬q, and forbids cases of ¬p and q.

The situation is similar in the case of disjunction. Following the theory of mental models, the models of a sentence such as (5) are the ones in (6) (see, e.g., Johnson-Laird 2012, 134-145).

(5) Either you eat rice, or you eat potatoes.
(6) Possible (p & q) & Possible (p & ¬q) & Possible (¬p & q) & Impossible (¬p & ¬q)

In (6) 'p' corresponds to the left disjunct in (5) and 'q' denotes the right disjunct in (5). So, (5) is an inclusive disjunction. The theory considers cases of exclusive disjunction such as (7) as well (see also, e.g., Khemlani, Orenes, & Johnson-Laird 2014, 1-7).

(7) Either today is Tuesday or today is Wednesday.

The meaning of the words in (7) makes the first conjunct or possibility in (6) impossible. Thus, the models of (7) are (8).

(8) Possible (p & ¬q) & Possible (¬p & q) & Impossible (p & q) & Impossible (¬p & ¬q)

Nevertheless, modulation can also have an influence on disjunction (See also, e.g., Quelhas & Johnson-Laird 2017, 703-717). An example in this regard of an inclusive disjunction is (9).

(9) Either this writer is José Saramago, or he is Portuguese.

Its models are those in (10).

(10) Possible (p & q) & Possible (¬p & q) & Impossible (p & ¬q) & Impossible (¬p & ¬q)

This last set of models infringes a truth table, too (in this case, that of disjunction). It states that cases of p and ¬q are not valid.

As far as conjunction is concerned, if there is no modulation, it only indicates one possibility. The theory deems this possibility as a fact (e.g., Johnson-Laird & Ragni 2019). For example, (11) just refers to the only possibility in (12).

(11) There is a cat, and there is a dog.
(12) Possible (p & q) & Impossible (p & ¬q) & Impossible (¬p & q) & Impossible (¬p & ¬q)

In (12) 'p' stands for the left conjunct in (11) and 'q' represents the right conjunct in (11).

However, if (12) only includes one possibility and is, for that reason, a fact, it can be equivalent to the statement in (13).

(13) p & q

All of this is interesting to give a general view of the framework of the theory of mental models. Nonetheless, as pointed out, what is important to this paper is how the theory describes inductive processes. Following, for example, Johnson-Laird (2012, 146), that kind of processes happens in inferences such as (14).

(14) They are falling asleep. Therefore, they stay overnight.

According to the theory of mental models, (14) is an inference involving an induction. This is because, in these cases, people tend to act as if conditional (15) were correct.

(15) If they stay overnight, then they will fall asleep.

The possibilities of (15) are obvious. In principle, (2) captures them. However, the premise in (14) (they are falling asleep) modifies the third conjunct in (2), which is not a possibility anymore and becomes an impossibility. Thus, the set of models resulting is (10). However, that is not all. The induction occurs because people might ignore the second conjunct in (10) and think that the only possibility for (15) is actually that they stay overnight and they fall asleep (p & q). The reason for that is that individuals might consider the first conjunct in (10) to be more probable than the second one, that is, the pair p and q to be more likely than the pair ¬p and q (or, in other words, the fact that they stay overnight and they fall asleep to be more probable than the fact that they do not stay overnight and they fall asleep). When this happens, an induction is made (see, in addition to Johnson-Laird 2012, 146, e.g., López-Astorga 2016, 355).

THE THEORY OF MENTAL MODELS AND GAUTAMA'S SYLLOGISM

As said, this account of the theory of mental models about induction has been related to Gautama's Syllogism (e.g., López-Astorga 2016, 351-358). Descriptions of this last syllogism are to be found in several works (see, in addition to López-Astorga 2016, 352-353, e.g., Ganeri 2004, 321ff). Following those works, it seems that the proponent of the syllogism was Gautama Aksapāda, that it was presented in the *Nyāyasūtra*, and that Colebrooke gave it the name 'Hindu Syllogism.' There are different ways to express Gautama's Syllogism. Nevertheless, this section will basically resort to the expressions used in papers such as the one of López-Astorga (2016, 351-358).

Gautama's Syllogism tries to demonstrate that the reason why an element a has property Q is that a also has property P. In order to do that, the syllogism takes previous knowledge into account. In particular, that b is akin to a, that b has property Q too, and that b has property P as well. This can be formally expressed by means of the formula (16).

(16) $Pa \wedge Qa \wedge (a \Leftrightarrow b) \wedge Pb \wedge Qb$

With other symbols, (16) appears in López-Astorga (2016, 356). It is not really presented as a formula there but as a representation of possibilities corresponding to the theory of mental models. However, for the aims of this paper, it can be deemed as a formula. In it, 'Px' and 'Qx' indicates that x has property P and property Q, respectively. '$\wedge$' denotes logical conjunction. On the other hand, '$\Leftrightarrow$' stands for similarity relation.

According to works such as that of López-Astorga (2016, 351-358), what is relevant about this syllogism is that it is based on a process that is not very different from the one described in the previous section. There is an alternative possibility for P, Q, a, and b that Gautama's Syllogism also ignores. That possibility is that to which (17) refers.

(17) $Pa \wedge \neg Qa \wedge (a \Leftrightarrow b) \wedge Pb \wedge Qb$

This formula is in López-Astorga (2016, 356) too. As (16), (17) is not actually a formula in López-Astorga's paper and is expressed there by means of other symbols. Nonetheless, the important point (17) makes is that it presents other possible scenarios in which most of the facts keep being the same: a has property P, a is similar to b, b has property P, and b has property Q. However, a does not have property Q now. In López-Astorga's view, what happens in Gautama's Syllogism is that (17) is ignored, and only (16) is considered. That is because (16) is deemed as a more probable circumstance. So, the parallel to the account of the theory of mental models is evident.

THE THEORY OF MENTAL MODELS AND MODAL LOGIC

In general, the proponents of the theory of mental models claim that there is no link between the theory and any system of modal logic (e.g., Khemlani, Hinterecker,

& Johnson-Laird 2017, 663-668). However, relations between it and system K have been provided (E.g., López-Astorga 2018, 120-136; 2020, 160-169). K is a very simple system of modal logic. As it is well known, it receives its name from Kripke's work and, as an extension of classical logic, includes the operators of possibility and necessity, a rule, and an axiom. The rule es the rule of necessitation, and the axiom is the axiom of distribution (see, e.g., Garson, 2018). (18) is the rule of necessitation, and (19) is the axiom of distribution.

(18) $p \therefore Np$

Where '$\therefore$' represents logical consequence and 'N' is the operator of necessity.

(19) $N(p \rightarrow q) \rightarrow (Np \rightarrow Nq)$

Where '$\rightarrow$' denotes material implication.

The way the theory of mental models has been linked to this framework is easy to see. The theory considers the models in sets such as (2), (4), (6), (8), (10), or (12) to be conjunctions (e.g., Khemlani et al. 2017, 663-668). So, if, in them, '&' is replaced with '$\wedge$,' and, given that the models are possibilities, 'Possible' is replaced with the operator of possibility '$\Diamond$,' well-formed formulae in K such as (20), (21), (22), (23), (24), and (25) can be obtained.

(20) $\Diamond(p \wedge q) \wedge \Diamond(\neg p \wedge q) \wedge \Diamond(\neg p \wedge \neg q) \wedge \neg\Diamond(p \wedge \neg q)$
(21) $\Diamond(p \wedge q) \wedge \Diamond(p \wedge \neg q) \wedge \Diamond(\neg p \wedge \neg q) \wedge \neg\Diamond(\neg p \wedge q)$
(22) $\Diamond(p \wedge q) \wedge \Diamond(p \wedge \neg q) \wedge \Diamond(\neg p \wedge q) \wedge \neg\Diamond(\neg p \wedge \neg q)$
(23) $\Diamond(p \wedge \neg q) \wedge \Diamond(\neg p \wedge q) \wedge \neg\Diamond(p \wedge q) \wedge \neg\Diamond(\neg p \wedge \neg q)$
(24) $\Diamond(p \wedge q) \wedge \Diamond(\neg p \wedge q) \wedge \neg\Diamond(p \wedge \neg q) \wedge \neg\Diamond(\neg p \wedge \neg q)$
(25) $\Diamond(p \wedge q) \wedge \neg\Diamond(p \wedge \neg q) \wedge \neg\Diamond(\neg p \wedge q) \wedge \neg\Diamond(\neg p \wedge \neg q)$

Formulae (20) to (25) correspond, respectively, to sets of models (2), (4), (6), (8), (10), and (12). Although perhaps not exactly the same symbols, formulae akin to (20) to (25) can be found in several works relating the theory of mental models to K (e.g., López-Astorga 2018, 120-136; 2020, 160-169). However, before continuing, two points should be highlighted. First, as indicated, 'Possible' has been replaced with '$\Diamond$' in the previous formulae. Accordingly, 'Impossible' has been replaced with '$\neg\Diamond$.' Second, as also mentioned, for the theory of mental models, a sentence such as (11) is not really linked to a possibility such as (12), but to a fact such as (13). Therefore, the most suitable formula for (11) would not be (25), but (26).

(26) $p \wedge q$

Now, what remains to be done from this perspective is to check whether or not the way the theory of mental models understands induction can also be related to K. This point is not made in the literature. Nevertheless, it is not hard to make if the rule of necessitation is considered in a figurative sense, that is, if the rule is applied to facts such as that corresponding to the formula (26). From what has been argued, the formula that could be assigned to (15) would be (20). But the premise in (14) would transform (20) into (24). Thus, as, in this case, induction consists of ignoring the second possibility in (24), the result is just one possibility, that is, $\Diamond(p \wedge q)$. Nonetheless, as it has been said, under the theory of mental models, when there is only one possibility (as in the case of (11)), it is not correct to speak about a possibility, but about a fact. This means that $\Diamond(p \wedge q)$ is transformed into (26). And this explains inference (14). By (18), (26) is transformed into (27).

(27)$N(p \wedge q)$

Which leads individuals to think that what necessarily occurs is that they fall asleep and they overnighted, ignoring any other alternative possible world in which they fall asleep, and they did not stay overnight. Of course, as said, this is not a rigorous use of system K and its rule of necessitation. It is only an adaptation of them to the way the human mind seems to work following the theory of mental models.

Previous papers have already shown how the theory of mental models and K can be related to account for most of the conclusions to which people come when making deductive inferences (e.g., López-Astorga 2018, 120-136; 2020, 160-169). The previous explanation is just a suggestion to research the case of induction. However, a point still needs to be analyzed. If the theory of mental models can be related to both Gautama's Syllogism and K, can Gautama's Syllogism and K be related to each other? Of course, this question is asked taking into account that, actually, representative works in the literature about Gautama's Syllogism (e.g., Ganeri 2004, 321ff) do not attribute modal characteristics to it.

GAUTAMA'S SYLLOGISM AND K

It seems that the answer to that question is positive. There are several facts known for sure in Gautama's Syllogism:

(28)Pb
(29)Qb
(30)$a \Leftrightarrow b$
(31)Pa

So, (32) also expresses a fact.

(32)$Pb \wedge Qb \wedge (a \Leftrightarrow b) \wedge Pa$

And, by virtue of (18), (33) can be obtained.

(33)$N[Pb \wedge Qb \wedge (a \Leftrightarrow b) \wedge Pa]$

In principle, this does not suffice to explain why Qa should be derived given Pa, even if (30) holds. This is because there are two possibilities:

(34)$\Diamond(Pa \wedge Qa)$
(35)$\Diamond(Pa \wedge \neg Qa)$

Nevertheless, if, following both the theory of mental models and Gautama's Syllogism, (35) is considered less likely and, therefore, ignored, (36) must be accepted.

(36)$N(Pa \wedge Qa)$

This is true from two perspectives. On the one hand, (31) leads to (37).

(37)$N(Pa)$

And, if (35) is rejected, that means (38).

(38)$\neg\Diamond(Pa \wedge \neg Qa)$

Thereby, (37) and (38) lead to (36).

On the other hand, if (34) is the only possibility, it is not a possibility, but a fact, that is, (39).

(39)$Pa \wedge Qa$

But, by (18), (36) can be inferred from (39).

Of course, one might think about an alternative broader proposal. That proposal would not only apply to an element such as a, but to any element. In this way, given (28) and (29), it could be thought that the real induction process here consists of removing b and taking just P and Q into account. Thus, (40) would be a fact.

(40)$Px \wedge Qx$

And, by (18), (41) would have to be inferred.

(41)$N(Px \wedge Qx)$

Nonetheless, if the true nature of Gautama's Syllogism is considered, this alternative proposal does not appear to be better. In the abstract, it seems to suggest. However, it does not pay attention to a very important characteristic of that syllogism. The key in Gautama's Syllogism is the similarity between a and b. Without that similarity relation, the syllogism cannot be applied (E.g., López-Astorga 2016, 351-358). Hence, the link between P and Q is not valid for any x, but only for any x that is akin to b. So, (40) and (41) cannot be assumed as resources to express the process involved in Gautama's Syllogism by means of K. Only formulae such as (36), (37), (38), and (39) are admissible; they suppose (30).

CONCLUSIONS

Accordingly, mental inductive processes can be related not only to the theory of mental models and Gautama's Syllogism but also to the machinery offered by K. Obviously, it is necessary to assume that K works in the second moment, that is, in the service of a previous process. That previous process is explicitly indicated by the theory of mental models and is implicitly in Gautama's Syllogism (e.g., López-Astorga 2016, 351-358). That is the process by means of which when there are two possibilities, and one of them is more probable, the other one is eliminated. This, from the point of view of the theory of mental models, transforms the remaining possibility into a fact. In this way, the rule of necessitation can be applied, of course, in a lax sense, and one might come to the conclusion that the fact is true in all the possible worlds (note that, although a fact is not a theorem, the rule of necessitation provides that p implies that p is necessary). This seems to be the manner a bridge between a deductive system such as K (or, at least, between the language of a deductive system such as K) and the mental inductive activity can be built.

The necessity of a previous mental process reveals something obvious: K cannot explain induction by itself (even if it is deemed in a weak and figurative way, and what is considered of the system is essentially its linguistic characteristics). At most, it can account for what happens once a possible alternative is removed in a way that is not deductively justified. However, this is important. On the one hand, it shows that the human mental activity is much more complex than the inferential actions that can be captured by systems akin to K. However, on the other hand, it also shows that some human mental processes could be, to an extent, compatible with this system.

In fact, that is what has been raised in the literature (e.g., López-Astorga 2018, 120-136; 2020, 160-169). The relations of the theory of mental models to K do not imply that the human mind works just as indicated by K. The theory of mental models is intended to describe how that mind works and proposes a very comprehensive framework to do that. The framework also includes mechanisms of information processing. Thus, because K does not have those mechanisms, this last system (or its weak interpretation here) is only compatible with the mental activities the theory of mental models assigns to the human mind after processing the information it receives.

On the other hand, Gautama's Syllogism appears to be consistent with the actions of information processing described by the theory of mental models. In those actions, certain possibilities are sometimes ignored (at least, in the case of induction). Therefore, it could be stated that the theory of mental models has compatibilities with several philosophical proposals. First, its description of mental activities by means of which the relevant information is selected can be coherent with what Gautama's Syllogism indicates (at a minimum, in the case of induction). Second, the deductive inferences that occur after those activities seem to be closer to systems such as K.

There is no doubt this is relevant. It allows seeing the theory of mental models as a unifying approach to integrating different efforts made in the history of thought.

The theory of mental models is a cognitive framework. Nevertheless, the search for relations such as those presented above can enable us to discover manners and directions in which human thought and other scientific theories have been built coming to the present time. For this reason, perhaps it is interesting to keep looking for relations of this kind.

REFERENCES

Byrne, Ruth M. J. and Philip N. Johnson-Laird. 2020. *If* and *or*: Real and counterfactual possibilities in their truth and probability. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 46 (4): 760-780.

Espino, Orlando, Ruth M. J. Byrne, and Philip N. Johnson-Laird. 2020. Possibilities and the parallel meanings of factual and counterfactual conditionals. *Memory & Cognition* 48:1263-1280.

Ganeri, Jonardon. 2004. Indian logic. In *Handbook of the History of Logic, Volume 1. Greek, Indian and Arabic Logic*. Edited by Dov M. Gabbay and John Woods. Amsterdam: Elsevier: 309-395.

Garson, James. 2018. Modal logic. In *The Stanford Encyclopedia of Philosophy*. Edited by Edward N. Zalta. Retrieved from: https://plato.stanford.edu/entries/logic-modal/.

Johnson-Laird, Philip N. 2012. Inference with mental models. In *The Oxford Handbook of Thinking and Reasoning*. Edited by Keith J. Holyoak and Robert G. Morrison. New York: Oxford University Press: 134-145.

Johnson-Laird, Philip N. and Ruth M. J. Byrne. 2002. Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review* 109 (4): 646-678.

Johnson-Laird, Philip N., Ana Cristina Quelhas, and Célia Rasga. 2021. The mental model theory of free choice permissions and paradoxical disjunctive inferences. *Journal of Cognitive Psychology* 33 (8): 951-973.

Johnson-Laird, Philip N. and Marco Ragni. 2019. Possibilities as the foundation of reasoning. *Cognition* 193. DOI: 10.1016/j.cognition.2019.04.019.

Khemlani, Sangeet, Ruth M. J. Byrne, and Philip N. Johnson-Laird. 2018. Facts and possibilities: A model-based theory of sentential reasoning. *Cognitive Science* 42 (6): 1887-1924.

Khemlani, Sangeet, Thomas Hinterecker, and Philip N. Johnson-Laird. 2017. The provenance of modal inference. In *Computational Foundations of Cognition*. Edited by Glenn Gunzelmann, Andrew Howes, Thora Tenbrink, and Eddy J. Davelaar. Austin: Cognitive Science Society: 663-668.

Khemlani, Sangeet and Philip N. Johnson-Laird. 2019. Why machines don't (yet) reason like people. *Künstliche Intelligenz* 33: 219-228.

Khemlani, Sangeet and Philip N. Johnson-Laird. 2022. Reasoning about properties: A computational theory. *Psychological Review* 129 (2): 289-312.

Khemlani, Sangeet, Isabel Orenes, and Philip N. Johnson-Laird. 2014. The negation of conjunctions, conditionals, and disjunctions. *Acta Psychologica* 151: 1-7.

López-Astorga, Miguel. 2016. The Hindu Syllogism, iconic representations, and human thought. *Revue Roumaine de Philosophie* 60 (2): 351-358.

López-Astorga, Miguel. 2018. Iconic representations, possible worlds, and system K. *Analele Universitatii din Craiova, Seria Filosofie* 42 (2): 120-136.
López-Astorga, Miguel. 2020. The recovery of syntactic structures cannot ignore temporality. *Analele Universitatii din Craiova, Seria Filosofie* 45 (1): 160-169.
Orenes, Isabel and Philip N. Johnson-Laird. 2012. Logic, models, and paradoxical inferences. *Mind & Language* 27 (4): 357-377.
Quelhas, Ana Cristina and Philip N. Johnson-Laird. 2017. The modulation of disjunctive assertions. *The Quarterly Journal of Experimental Psychology* 70 (4): 703-717.

Acknowledgments

PIA Ciencias Cognitivas, Centro de Investigación en Ciencias Cognitivas, Instituto de Estudios Humanísticos, Universidad de Talca.

Fondo Fondecyt de Continuidad para Investigadores Senior, código FCSEN2102, Universidad de Talca.