Some arguments are probably valid

M.H. Tessler, Noah D. Goodman
{mhtessler, ngoodman}@stanford.edu
Department of Psychology, Stanford University

Abstract

This is not an abstract.
Keywords: Reasoning, probabilistic model

Syllogistic reasoning has been a topic of considerable interest in cognitive psychology for over one hundred years, and before that in philosophy, dating back to Aristotle. A syllogism is a two premise argument, which relates two terms by a middle term. The relation is a quantifier, and the quantifiers used in the classical syllogism are all, some, none, and not all. For example,

> All artists are bakers
> Some bakers are chemists

is a syllogistic argument. There are 64 possible syllogism which one can derive by replacing the quantifiers with others from the allowable set and reordering the terms (All As are Bs −¿ All Bs are As). Most syllogisms have no valid conclusion, like the one presented above.

People with no training in formal logic, and even some with training, find reasoning with syllogisms difficult. A recent meta-analysis showed that over the population, accuracy on producing valid conclusions ranges from 90 % to 1% (?, ?). For syllogisms with no valid conclusion, accuracy ranges from 76% to 12%.

This evidence would seem to suggest that people are not reasoning using deductive mechanisms in the mind. Yet, many theories take deduction as a given and try to explain errors as a matter of noise in the system. We take a different path. We propose, as has been proposed before, that people are reasoning according to their *e*veryday mode of reasoning. *E*veryday reasoning is understood as the type of reasoning that is refined for dealing with a world of uncertainty, a world in which you don't know how many people are in the hallway outside your door or whether or not the lion is going to start charging. This type of reasoning is most succinctly described in the language of probability theory. In this formalism, deduction emerges as that which is most probable.

In the rest of the paper, we review three theories of syllogistic reasoning – Mental Models, Mental Logics and Probability Heuristics – which we take to exemplify three quadrants of a two-dimensional theoretical space. We develop a computational level theory in the unexplored quadrant of the space, instantiated in a probabilistic model. We compare our model predictions with the predictions of the other theories as well to behavioral data from a recent meta-analysis. We explore the flexibility of the model to account for reasoning behavior in a study using the generalized quantifiers most and few. We conclude by discussing further predictions of the model and future directions.

Extant theories

A recent meta-analysis carved the space of reasoning theories into three partitions: those based on model or diagrammatic reasoning, those based on formal logical rules, and those based on heuristics (Khemlani & Johnson-Laird, 2012). We see the space differently. In one dimension, theories may be based on applying rules – be they heuristic or logical – or they are based on constructing concrete representations or models. In another dimension, theories may be considered fundamentally probabilistic or deterministic.

Mental Models

The Mental Models Theory (MMT) describes a psychological process by which people reason by constructing *iconic* mental representations or models, which represent the terms of a proposition as a collection of individuals. In syllogistic reasoning, a model is constructed for each premise, and premise models are consolidated so that the conclusion may be "read off" the joint model. For example, a model for the premise *All artists are bakers* could be represented as the following situation.

artist	baker
artist	baker
	baker

This shows 2 individuals who are both artists and bakers, and one individual who is a baker but not an artist. Thus, each row is a representation of the properties of an individual. The authors emphasize the need to search for counter-examples to check for logical validity. Errors arise in this search process.

The MMT captures the intuition that people are able to reason about sets of things explicitly and with respect to context. The a priori believability of propositions has been shown to have an important effect on reasoning (Oakhill, Johnson-Laird, & Garnham, 1989). Mental models are flexible in that content can motivate one to carry on the search process longer. The search process allows for individual differences insofar as some individuals may test many models and some may test just one. At the same time, the theory is not well defined insofar as it does not specify how various models come into existence, only that various models can come into existence.

Mental Logics

Rips (1994) proposed that people reason according to rules of *natural deduction*. The theory of Mental Logics, instantiated in the PSYCOP model, posits individuals construct *logical sentences* in a language of first-order predicate calculus

which are linked when the "individual recognizes [the link or inference] as intuitively sound". The model explains errors as a failure to recognize the applicability of a given formal rule, a failure to retrieve the rule, or a failure to carry out the necessary steps for that rule. People are especially prone to such errors when complex rules are needed so there is a predicted effect of rule complexity on difficulty. In this instance, the theory does not specify *which* conclusions will be drawn fallaciously, only that some will. Mental Logics posits that people reason according to logical, deterministic rules.

Probability Heuristics

Heuristic accounts be understood in a similar way: that people reason according to heuristic rules, which may be determined by probabilities. This the case with Chater and Oaksford's Probability Heuristic Model (PHM). Like our approach, the PHM is inspired by the notion that people are not fundamentally deductive reasoners, but instead are trying to gage degrees of plausibility for the conclusion. This amounts to computing the probability of a particular conclusion being true given that the premises are true. [equation] To accomplish this, the PHM relies on a number of *generation* and *test* heuristics which produce and quantify confidence in conclusions, and which they claim are justified by their computational level theory. The computational level theory includes a notion of informativeness, on which all their heuristics rely. We do not believe their heuristics are necessary for deriving a probabilistic model of reasoning, as we discuss below. Further, as is the case for theories based on formal rules, the very nature of their heuristics suggests reasoners are not engaging with the syllogisms at a semantic level. We also do not believe this to be the case.

The Conditional Semantics model

The motivation for our model stems from two intuitions: (1) people are using *everyday* reasoning in syllogistic reasoning tasks and (2) they do this by constructing situations [1] and reasoning over these situations. This formulation places the theory in the unexplored quadrant of the two-dimensional theoretical space described above: we consider reasoning over situations and situations to be constructed probabilistically. We use a naive binomial prior to sample situations. Each situation is composed of some number of objects, each with 1, 2 or 3 of the properties under discussion.

We draw on work in formal semantics by assuming that sentences are truth-functional operators. Thus, for each situation, some sentences are true while others are false. By sampling over many situations, we construct a distribution over sentences, shown in Figure 1. [Here we can show either a bar graph for n=6, p=0.25; or we can show a line graph with either n held constant or p held constant, or we can show a surface graph.]

For reasoning over syllogisms, the prior distribution is conditioned on the truth of the premise sentences. This is the distribution over sentences conditioned on the fact that the premises are true. In this way, we evaluate the Pr(conclusion | premises). This is essentially a metric of the plausibility of the argument. In this formulation, Pr(conclusion | premises) = 1 if and only if the syllogism is logically valid.

Model predictions

Our model assumes that properties are relatively rare of objects. A principle of rarity is often assumed in line with the intuition that properties are relative rare [2], and we assume a principle of rarity here. We tested a number of rarity factors and found that p=0.25 provided the best fit. This is also consistent with the Probability Heuristic's rarity assumption.

For each model, we report the total number of syllogisms the model gets correct as well as the overall fit. A syllogism is deemed correct if the model's maximum predicted response is the same as the meta-analysis data. This is a qualitative assessment of the fit, while correlation is a quantitative assessment.

The Prior   It might be the case that individuals are drawing conclusions simply from the prior probability of those conclusions being true. We show that this is not the case. From the prior, however, we can see that all conclusions are not equally likely. In particular, we observe an ordering which has previously been cited as an ordering of Informativity. The Probability Heuristics Model derives this ordering by positing categories as hyperspheres in a high-dimensional concept space (Chater & Oaksford, 1999). This ordering comes out of the Conditional Semantics Model with no further assumptions.

The prior only model gets only the syllogisms with a "Not all" conclusion correct. For the Chater & Oaksford meta-analysis, this amounts to 29 out of the 64 syllogisms.

Conditional Semantics   The Conditional Semantics model conditions on the truth of the premises. The model generates situations in which the premises are true, and determines which conclusions are true in the variety of situations. In the limit, this converges to the Pr(conclusion | premises).

The model matches people's maximum judgments on 37 of the 64 syllogisms. This includes all the syllogisms for which "not all" is the maximum, as well as 8 syllogisms for which "some" and "none" are favored. Crucially, it doesn't match any of the syllogisms for which people like the all conclusion. One such syllogism is the canonical "All/all" syllogism:

All Xs are Ys
All Ys are Zs

This illustrates one of the issues with the conditional semantics model: it is too literal. The model computes the prob-

---

[1]These situations are in no way different from mental models. We prefer the term situation so as to not confound the word model, which we take to refer to a computational model.

[2]This article is an article and it's about reasoning, but it's not a cat, and it's not a car, nor an elephant nor the color red. In fact, there's a very large number of things which this article is not.

ability of the conclusion conditioned on the premises being true. For logically valid syllogistic conclusion, this probability is equal to 1. In particular, for any syllogism for which one of the universal quantifiers is valid (i.e. *n*one and *a*ll), the particular quantifier (i.e. *n*ot all and *s*ome, respectively) will also be valid. Data from a number of experiments suggests that people do not treat *a*ll and *s*ome as equally correct or good conclusions. Instead, participants show a preference for conclusions the universal conclusions (all and none) when they are warranted.

To address this we draw on recent advances in computational models of pragmatic reasoning. These models address the problem of language understanding as that of a speaker trying to convey information about a world or situation that the speaker has observed. The speaker draws on common-ground of communicative goals to maximize information content of a given utterance. In particuar, choosing an utterance in proportion to the likelihood of a listener inferred the world about which the speaker wishes to convey information. In turn, a listener considers worlds in which the given utterance heard is

Mental Models Theory explains difficulty in syllogisms by demonstrating that multiple, distinct situations can be consistent with a pair of premises, giving rise to different possible conclusions. Valid syllogistic conclusions arise in every possible model. (This is the same notion of Pr(conclusion — premises) = 1).

This is always the case the logically invalid syllogisms (see discussion below), and is sometimes the case with logically valid syllogisms.

Logical invalidity For syllogisms that have no valid conclusion, the probabilistic reasoner computes the probability of the conclusion conditioned on the premises being true by sampling. For all invalid syllogisms, this yield a gradeds response across the different possible conclusion types.

Logical validity A conclusion is logically valid if and if only it is true in all possible situations. Recall that possible situations are restricted to those that are consistent with the premises. The universal quantifiers (all, none) entail their particular counterparts (some, not-all), respectively. This was noticed by Aristotle and commonly visualized in the *Square of Opposition*. For our probabilistic reasoner, confidence for *a*ll and *s*ome (to take the affirmative case) will be both maximal. *A*ll is always true and *s*ome is true whenever *a*ll is true.

Pragmatics

One thing to notice about the probabilistic reasoner is that it is agnostic among equally valid conclusions. However, human reasoners endorse some valid conclusions more so than others. What can account for this apparent illogicality?

Harness recent advances in quantitative models of conversational pragmatics (implicature paper).

Results

Meta-analysis data

Scatterplot, things that we get wrong. Perhaps a more detailed view of the ones we get wrong... would need a good diagram for this, the most simple which comes to mind is a series of bargraphs, each graph has 4 x 2 (=8) bars for the 4 conclusions x (model, data). Could also do this for the ones we get right... 64 syllogisms in total so we'll have to choose wisely.

Model comparison

Something creative here.

Most and few

A grand virtue of the model.

Model predictions

There are two experiments. As of today, the model does quite well with Experiment 1 (quantifiers: all, most, few, not-all) and less well with Experiment 2 (quantifiers: most, few, some, none). Interestingly or not, Chater and Oaksford have the same pattern (better with Exp 1) and I think our fit is better than theirs.

Chater & Oaksford Data vs. Our Model

Results here.

Model comparison

Something creative here.

Future directions

Belief bias as a difference of prior.

References

Chater, N., & Oaksford, M. (1999). The Probability Heuristics Model of Syllogistic Reasoning. *Cognitive psychology*, *258*, 191–258.

Khemlani, S., & Johnson-Laird, P. N. (2012, May). Theories of the syllogism: A meta-analysis. *Psychological bulletin*, *138*(3), 427–57. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22289108 doi: 10.1037/a0026841

Oakhill, J., Johnson-Laird, P., & Garnham, A. (1989). Believability and syllogistic reasoning. *Cognition*, *31*, 117–140.