Proceedings of the KI 2017 Workshop on Formal and Cognitive Reasoning

Syllogistic Reasoning in Seven Spaces

Frieder Stolzenburg and Raimund Lüderitz

Department of Automation and Computer Sciences, Harz University of Applied Sciences, Friedrichstr. 57-59, 38855 Wernigerode, Germany
{fstolzenburg,u27732}@hs-harz.de

Abstract Syllogisms and syllogistic reasoning has been the subject of research and scientific discourse for more than two millennia. Syllogisms sum quantified assertions into an overall statement, usually consisting of two premises and one conclusion. While syllogistic reasoning can be modeled by classical first-order logic in a straightforward manner, it is an open question which of the possible syllogisms are accepted as valid by human reasoners. In this paper, we present an approach that models the reasoning process with seven spaces of a set diagram. It can easily be implemented by constraint logic programming. We distinguish several assumptions that humans may make during their reasoning process, in particular that all used categories are non-empty. In contrast to pure logic-based approaches, the proposed procedure allows to represent diverse models human reasoners may follow. The results show good correlation and coincidence with psychological investigations.

Keywords: syllogistic reasoning, set diagrams, constraint logic programming.

1 Introduction

Syllogistic argumentation can be traced back to Aristotle [11]. A classical syllogism at first consists of two quantified assertions about categories, e.g. A, B, and C, like 'Some A are B' and 'No B are C', connecting exactly two categories each time. Other so-called generalized quantifiers like 'Most' or 'Few' are possible among others (see e.g. [10]). The task is to derive logical consequences from these statements. For instance, 'Some A are not C' is a logical consequence of the given two statements in classical first-order logic.

Example 1. Another example with the concrete categories 'bakers', 'artists', and 'chemists' (cf. [5]) is:

Some artists are bakers.
All bakers are chemists.
$\therefore$ Some artists are chemists.

For this example, most people and also classical logic accept the last assertion as a consequence of the two premises.

It is easy to express a syllogistic statement in classical logic, namely by the logic of monadic assertions which is the subset of first-order logic where predicates, which correspond to categories in this context, always have exactly one argument. For instance, the assertion 'All A are B' can be expressed in monadic logic as:

$$\forall x \left(A(x) \rightarrow B(x) \right)$$

The overall syllogistic statement then consists of three such assertions φ_1, φ_2, and φ_3. It corresponds to the following logical implication:

$$\varphi_1 \wedge \varphi_2 \rightarrow \varphi_3$$

By means of a theorem prover or other kind of logical deduction system, valid syllogisms – in the sense of classical logic – may be determined automatically.

However, human reasoning does not strictly follow the rules of classical logic. Explanations for this may be incomplete knowledge, incorrect beliefs, or inconsistent norms. From the very beginning of artificial intelligence research, there has been a strong emphasis on incorporating mechanisms for such kind of rationality into reasoning systems. Rationality may be bounded, simply because it may be difficult for humans to have all possible cases in mind.

In addition, at latest since the famous Wason selection task [12], it is well-known that it makes a significant difference whether people have to solve an abstract reasoning task or a concrete task with concrete categories (like in Example 1). Hence for psychological investigations surely this has to be taken into account: It is very likely that the purely logical reasoning of humans is interfered in the presence of concrete categories, in particular if the assertions are not true in the real world. To see this consider again the assertion 'All bakers are chemists' (Example 1) which might not be true in general. Therefore every psychological experiment on syllogistic reasoning has to be carefully designed accordingly and should not be too complex, i.e. involve too many categories.

In the following, we first make the terminology more precise and define among others the meaning of the notions category, mood, and syllogism (Section 2). Then, we briefly discuss related approaches, including works from cognitive psychology (Section 3). After that, we state our approach that models the reasoning process with seven spaces of a simple set diagram (Section 4), which can easily be implemented by constraint logic programming. We present some results (Section 5) which show good correlation with other suited theories of syllogistic reasoning and also with psychological investigations and end up with some conclusions (Section 6).

2 Syllogisms – Notions and Terminology

Let us now define more precisely what a syllogism is. For this, we refer to mathematical notions like sets and probabilities. Clearly, we cannot assume that people without much training in mathematics and logic have knowledge on this. Even trained people probably do not apply strict formal rules while deriving logical

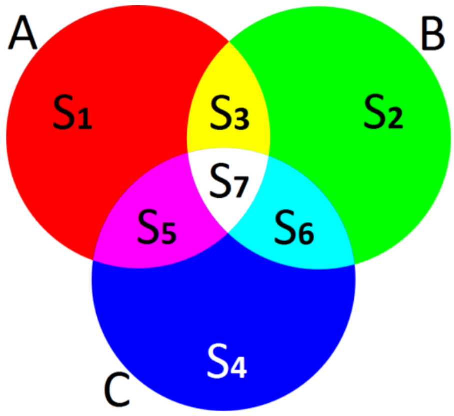

Figure 1. Venn diagram for three categorical sets A, B, and C. The spaces $S_1, \ldots, S_7$ correspond to the respective marked subsets. For example, $S_1 = A \cap \overline{B} \cap \overline{C}$ and $S_7 = A \cap B \cap C$.

consequences. Nevertheless in the following we try to capture the required notions and terminology reasonably formally. On the one hand, this helps us to analyze and model underlying assumptions made by human reasoners more precisely. On the other hand, the definitions lay a clear basis for the implementation of the proposed set-based approach with seven spaces presented in this paper.

We start with the notion category: A *category* C stands for a set of individuals or objects. It is associated with a monadic predicate $C(x)$, saying that x is an individual of the respective category. A classical syllogism combines in total three categories: A, B, and C. A *space* S is one of the seven possible subsets of $A \cup B \cup C$ in the set diagram for the three involved categories (see Figure 1). The set of all such spaces is:

$$M = \{A\cap\overline{B}\cap\overline{C}, \overline{A}\cap B\cap\overline{C}, A\cap B\cap\overline{C}, \overline{A}\cap\overline{B}\cap C, A\cap\overline{B}\cap C, \overline{A}\cap B\cap C, A\cap B\cap C\}$$

A *categorical assertion* or just *assertion* for short for two categories A and B is a statement of the form 'A certain quantity of A is (not) B'. Classically, four different *moods* of assertions are distinguished. They correspond to the quantifiers 'All' or 'Some' and are optionally combined with the negation of the second category involved in the assertion. In addition to the four classical moods, which are abbreviated **A**, **E**, **I**, and **O**, we also consider a fifth one, which we call **U**. Every mood can be directly interpreted by means of conditional probabilities $P(B|A)$ as follows:

A: 'All A are B'
This means $P(B|A) = 1$. Since $P(B|A) = P(A \cap B)/P(A)$ and A is the disjoint union of $A \cap B$ and $A \cap \overline{B}$, this is equivalent to $P(A \cap B)/\big(P(A \cap B) + P(A \cap \overline{B})\big) = 1$ and hence $P(A \cap \overline{B}) = 0$.

E: 'All A are not B'
This means $P(\overline{B}|A) = 1$. Since $P(\overline{B}|A) = 1 - P(B|A)$, this is equivalent to $P(A \cap B) = 0$.

I: 'Some A are B'
This means $P(B|A) > 0$, which is equivalent to $P(A \cap B) > 0$.

O: 'Some A are not B'
This means $P(\overline{B}|A) > 0$, which is equivalent to $P(A \cap \overline{B}) > 0$.

U: 'Some but not all A are B'
This means $0 < P(B|A) < 1$. It corresponds to the conjunction of mood **I** and the negation of mood **A** which is equivalent to mood **O**. Therefore we have $P(A \cap B) > 0$ and $P(A \cap \overline{B}) > 0$.

We consider the mood **U** here, because it allows to state more specific categorical assertions than with the mood **I**: Human reasoners or the experimental setting may assume that at most one assertion should be accepted as consequence of the premises, but from a logical point of view, there can be more than one valid conclusion, in particular if the classical mood **I** is used. For instance, 'All A are B' (mood **A**) implies 'Some A are B' (mood **I**) provided that the categories A and B are non-empty, but not 'Some but not all A are B' (mood **U**). In addition, mood **I** is a symmetric operator, i.e. 'Some A are B' and 'Some B are A' are logically equivalent. Thus for mood **I** the order of the categories in the conclusion does not matter, whereas in general it does for mood **U**. Interestingly, the variation of mood **U** 'Some but not all A are not B' is logically equivalent with mood **U**. To see this, just replace B by $\overline{B}$ and vice versa.

Now, three categorical assertions φ_1, φ_2, and φ_3 are combined to an overall statement, called *syllogistic statement* or just *syllogism* for short. In this context, φ_1 and φ_2 are the *premises* and φ_3 is called *conclusion* of the syllogism. If φ_1 and φ_2 together logically imply φ_3, then we speak of a *valid* syllogism. The categorical assertion φ_1 relates the categories A and B with each other, φ_2 the categories B and C, and φ_3 the categories A and C.

In each of the assertions, the two involved categories can occur in two orders, e.g. A–B or B–A for φ_1. Therefore the premises φ_1 and φ_2 can have four different orders in total, called *figures* (cf. Figure 2). Hence, a problem can be completely specified by the moods of the first and second statement and the figure. The conclusion φ_3 is also a quantified statement where the categories A and C can appear in either order.

Assertion	Figure 1	Figure 2	Figure 3	Figure 4
φ_1	A–B	B–A	A–B	B–A
φ_2	B–C	C–B	C–B	B–C

Figure 2. Four different figures of a syllogism. Note that in the literature – as done in this table, too – only the order of the categories in the premises are considered.

How many syllogisms are there to worry about? Actually, there is only a small, finite number of cases: Each categorical assertion may be in one of the classical $n = 4$ or extended $n = 5$ moods. The categories in each assertion may be in one of two orders. Hence, for the three categorical assertions φ_1, φ_2, and φ_3, we have altogether

$$N = (n \cdot 2)^3$$

different syllogisms to consider. This yields us $N = 512$ or $N = 1000$ for $n = 4$ or $n = 5$ moods, respectively. Although conjunction is a commutative operator in classical logic, the presentation of the order of premises may play a role for human reasoners. However, since we already consider both possible orders of the categories A and C in the conclusion, this aspect is already covered and thus needs no special treatment.

3 Related Works

Syllogistic reasoning has been extensively investigated by cognitive psychologists (cf. the meta-study in [5]). A common result is that the conclusions humans draw differ from those of classical first-order logical reasoning. Therefore, many theories have been developed to explain the human behavior, among them are: heuristic theories that capture principles that could underlie intuitive responses, e.g. probabilistically valid conclusions [10], theories based on formal rules which may include non-monotonic logic, and theories based on diagrams, models, or sets, e.g. Venn diagrams [9].

Analyses of syllogisms based on first-order logic formalizations cause several problems (cf. [5]): The meaning of quantifiers often cannot be expressed in first-order logic adequately, because it is not powerful enough to represent many determiners in ordinary language, such as ‘more than half’ or ‘most’. One way out could be second-order predicate calculus, which allows quantification over sets (and hence categories) as well as over individuals. Unfortunately, from a psychological standpoint, these generalized quantifiers are infeasible: The computation of, say, the set of all sets containing all individuals of category C is intractable and is likely to be too large to fit inside anyone’s brain [2].

Heuristic theories of syllogistic reasoning assume more or less that human reasoners do not apply any formal rules: According to the so-called atmosphere theory [15], people may be predisposed to accept a conclusion that fits the mood of the premises. This means, if a premise contains ‘Some’, use it in the conclusion; if a premise is negative, use a negative conclusion. Other theories claim the preference of conclusions in the same mood as the most informative or conservative premise [10,13]. In this context the moods are sorted in some preference ordering. For more details, the interested reader may consult [5] which provides a thorough review of several theories about syllogistic reasoning.

According to [5], at least some of the disadvantages of first-order logic can be overcome, if categorical assertions are stated as relations between sets. This approach accommodates the quantifiers that cannot be expressed in first-order logic. Cognitive scientists have argued that mental representations of quantified assertions are set-theoretic (e.g. [3]). It is also consistent with various diagrammatic systems of syllogistic reasoning. The mental model theory [4] also follows this line. It postulates that human reasoners can represent a set iconically and build a mental model of its members.

From a set-theoretic point of view, [6,16] introduce a syllogistic system that is transformed into fuzzy rules for syllogistic reasoning. Fuzzy existential quantifiers are introduced covering the range from 'Some' to 'All' in several steps. From that, relative truth ratios are calculated based on the cardinalities of the syllogistic cases. The approach is similar to the one presented in this paper. It is also based on set diagrams. However, we do not consider fuzzy or multi-valued logics here. We explain our approach in more detail in the next section (Section 4).

In summary, there are many different theories trying to explain the experimental findings. However, probably no single theory provides an adequate account (cf. [5]). It may be the case, that different persons apply very distinct reasoning strategies (cf. [13]). This would mean that only a collection of different theories can explain the whole picture.

4 Implementing Syllogistic Reasoning

In the literature and in the corresponding psychological experiments, often several prerequisites are assumed for human reasoners, e.g. that all categories are non-empty. This is done implicitly or explicitly, namely as part of the instructions to the test persons in psychological experiments. These assumptions have also to be modeled. We do this using probabilistic notions:

Let us consider the probability space where an event E may be a set in M, i.e. one of the seven spaces (cf. Section 2), or a finite union thereof. The cardinality of the whole probability space is thus $2^7 = 128$. Let P be a probability measure, assigning probabilities to the events. The following *assumptions* may be considered:

1. If a space S is non-empty, it has a non-zero probability, i.e. $P(S) > 0$ for $S \neq \emptyset$. In the following, we always adopt this assumption.
2. Quantification with 'Some' (moods **I** and **O**) induces the existence of individuals or objects of the respective category. Furthermore it can be understood non-inclusive, i.e. as 'Some but not all' (mood **U**).
3. All categories are non-empty, i.e. $A, B, C \neq \emptyset$ and thus $P(A) > 0$, $P(B) > 0$, and $P(C) > 0$.
4. All categories are non-equivalent, i.e. $A \neq B \neq C \neq A$. Each of these three inequalities can be interpreted probabilistically, e.g. $A \neq B$ by $P(A \backslash B \cup B \backslash A) > 0$, because $A \neq B$ is equivalent to $A \nsubseteq B$ or $B \nsubseteq A$ and hence to $A \backslash B \ \cup \ B \backslash A \neq \emptyset$.

Since assumption 1 is of more or less purely technical nature, we always adopt it hereinafter. The other assumptions may hold or not and can therefore be switched on or off explicitly in our implementation, although human reasoners may not even be aware whether they apply them. Assumptions 2 and 3 can be traced back to Aristotle's original works on analytics (*Analytica priora* and *Analytica posteriora*, cf. [1]). In contrast to this, assumption 4 seems to be considered explicitly only recently (e.g. in [6,16]).

Syllogistic reasoning (including these assumptions) can now be implemented in a more or less straightforward manner by *constraint logic programming* [8]. Constraint logic programming emerged from the field of logic programming. It is a form of constraint programming, in which logic programming is extended to include concepts from constraint satisfaction. A constraint logic program is a logic program that contains constraints that are specific conditions in the body of clauses. In the case of finite-domain constraints, we have variables that take their values out of a finite domain of integer numbers. Domains can be described as enumerations of possible values. For efficient reasoning, special comparison operators that are prefixed by the symbol `#` are used as constraint predicates.

As one can see, every assumption from above and every mood of an assertion can be expressed by one or more conditions of the form $P(E) = 0$ or $P(E) > 0$ for some event (set) E. By assumption 1 this eventually means, for a specific case, we simply have to distinguish which spaces are empty and which are not, where a *case* is defined by its corresponding set of non-empty spaces $S \in M$. If a case satisfies the assumptions 3 and 4, we call it *admissible*. There are $2^7 - 3 \cdot 2^3 + 3 + 2 = 109$ cases satisfying assumption 3 and $109 - (1 + 3 \cdot 4) = 96$ admissible cases, i.e. where in addition assumption 4 holds.

In [7] (thesis supervised by first author), the syllogistic reasoning in seven spaces has been implemented by means of constraint logic programming in the programming language SWI Prolog [14]. How, then, can a syllogism expressed by means of constraint logic programming? For this, we only need the following ingredients:

- That non-empty spaces have a non-zero probability (assumption 1), can be modeled easily by constraint variables for every space $S_1, \ldots, S_7$. Each one may take one of the integer values 0 or 1 with the meaning that the respective space is empty or non-empty, respectively. In Prolog syntax, this can be written as follows:

```
[S1, S2, S3, S4, S5, S6, S7] ins 0..1
```

- As said above, every mood can be interpreted by one or more conditions of the form $P(E) = 0$ or $P(E) > 0$ for some event (set) E. For example, the assertion 'All A are B' (mood **A**) is equivalent to $P(A \cap \overline{B}) = 0$. It holds $A \cap \overline{B} = S_1 \cup S_5$ (cf. Figure 1). Thus this allows us – together with assumption 1 – to implement the condition simply by the constraint $S_1 + S_5 = 0$, in Prolog syntax:

```
S1 + S5 #= 0
```

- Similarly, constraints of the form $P(E) > 0$ can be implemented: For example, 'Some A are B' (mood **I**) leads to the condition $P(A \cap B) > 0$. Because of $A \cap B = S_3 \cup S_7$ we obtain the constraint $S_3 + S_7 > 0$, in Prolog syntax:

```
S3 + S7 #> 0
```

– Assumption 3 can also be expressed easily by means of constraint logic programming. To see this, we consider the example that category C is non-empty, i.e. $P(C) > 0$. Because of $C = S_4 \cup S_5 \cup S_6 \cup S_7$, we arrive at the constraint $S_4 + S_5 + S_6 + S_7 > 0$, in Prolog syntax:

```
S4 + S5 + S6 + S7 #> 0
```

– Last but not least, assumption 4 is also feasible. For example, $A \neq B$ is equivalent to $A \backslash B \cup B \backslash A \neq \emptyset$. Because of $A \backslash B = S_1 \cup S_5$ and $B \backslash A = S_2 \cup S_6$, we obtain the constraint $S_1 + S_5 + S_2 + S_6 > 0$, in Prolog syntax:

```
S1 + S2 + S5 + S6 #> 0
```

Our constraint logic program for syllogistic reasoning provides implementations of all assumptions in the above-mentioned manner and also of all 30 categorical assertions (not listed here in detail) that are possible with $n = 5$ moods involving two of the three categories in any order. By means of our implementation, we can now find all solutions for every syllogism, simply by means of Prolog built-in set predicates like `bagof` [14]. Since syllogisms have the form $\varphi_1 \wedge \varphi_2 \to \varphi_3$ (cf. Section 1), we compute the number of cases that satisfy (a) $\varphi_1 \wedge \varphi_2$ and (b) $\varphi_1 \wedge \varphi_2 \wedge \varphi_3$ and compute their ratio. Our implementation [7] allows several settings, namely with or without:

– the additional mood **U** (assumption 2),
– possibly empty categories (assumption 3), or
– possibly identical sets (assumption 4).

Example 2. Let us illustrate this by the following example with abstract categories:

All A are B.
All B are C.
$\therefore$ All C are A.

Here, all assertions are in mood **A**, the premises have the form of the syllogistic figure 1 (cf. Figure 2), and in the conclusion, C is related to A, in this order. The premises induce that S_1, S_2, S_3, and S_5 must be empty (cf. Figure 1). The remaining three spaces S_4, S_6, and S_7 may be empty or not. Thus $2^3 = 8$ cases satisfy both premises, at least if we allow empty and equivalent categories. The conclusion enforces that S_4 and S_6 must be empty, too. This means all three categorical assertions hold only if all spaces are empty except possibly S_7. Therefore $2^1 = 2$ cases remain. Hence the ratio, i.e. the fraction of valid cases is only $2/8 = 25\%$ for this example.

The situation changes if we adopt one or more of the assumptions from above: For instance, if all categories are non-empty, then space S_7 must be non-empty. Hence the syllogism of Example 2 holds in one of four cases, if assumption 3 holds, which again leads to a ratio of $1/4 = 25\%$. If in addition assumption 4 holds, then spaces S_4 and S_6 must be non-empty, because otherwise some of the sets are equivalent. The ratio is then $0/1 = 0\%$.

5 Results

Table 1 shows the complete numbers and ratios for all possible syllogisms with the classical four moods, i.e. without mood **U**. Each three consecutive columns show the number of cases that satisfy all three assertions or only the premises of the respective syllogism and their ratio in percent. Since in the literature and in many psychological experiments it is normally given that all categories are non-empty, we adopt assumption 3 here. Nevertheless we allow that some of the category sets may be identical, i.e. assumption 4 does not necessarily hold. The syllogisms that are valid in 100% of the cases are the logically valid syllogisms. The tables for the other settings can be found in the technical report [7].

Since the ratio may vary and is not just 0 or 1, we can build a fine-grained model for modeling syllogistic reasoning and try to predict human responses. In psychological experiments, the setting often is that that two premises are given and then the user has to draw the valid conclusions or find out that there are no valid conclusions. This procedure is reported e.g. in [5]. Nonetheless, from a logical point of view, there can be more than one valid conclusion (cf. Section 2). Thus one could allow multiple answers. We start now with:

Hypothesis 1. People draw a specific conclusion the higher the percentage of its computed ratio is according to our model.

First evaluation results for this hypothesis are encouraging: A simple correlation analysis between the percentage values computed by the constraint logic implementation and the empirical findings reported in [5, Table 6] yields, as desired, moderate positive correlation for all possible settings (cf. Table 2). Here the case that there are no valid conclusions is ignored. Interestingly, the maximum is attained if none of the assumptions 2, 3, and 4 is adopted – although the test persons might be instructed that they should hold. For instance, with assumption 3 the assertions in mood **I** with A and C in any order are also valid conclusions for Example 2 but are not without them. Only the latter coincides with the empirical finding that most people conclude (only) the assertion 'All A are C' (mood **A**). Taking also the case of no valid conclusions into account, we modify Hypothesis 1 as follows:

Hypothesis 2. People draw a specific conclusion only if the percentage of its computed ratio is 100% and no valid conclusions otherwise.

If we repeat our comparison under Hypothesis 2, then we gain a coincidence of almost 90% over all $64 \cdot 9 = 576$ possible syllogisms including the case that two premises have no valid conclusions (see also Table 2). This is rather good. Here, a conclusion is adopted if the implemented model yields a percentage of 100% or the majority of people in the psychological experiments draw this conclusion, respectively. We thus map the real percentage values of the model and the psychological studies to just two values: yes or no. The coincidence with respect to no valid conclusions alone is about 70% (see again Table 2). This is significantly above chance. Clearly, this point needs further investigation and is still work in progress.

Table 1. Table for all possible syllogisms disregarding mood **U** and adopting assumption 3 but not assumption 4. The moods and figures of the premises are indicated in the leftmost column. Here, for example, AA1 means that the first and the second premise have mood **A** and the form of figure 1. The possible conclusions are given in the further columns. The moods and, by small letters, the order of the categories is given here. For example, Aca stands for the categorical assertion 'All C are A'. Each three consecutive columns show the number of cases that satisfy all three assertions or only the premises of the respective syllogism and their ratio in percent.

Syllogism	Premises	Aac			Aca			Iac			Ica			Oac			Oca			Eac			Eca		
AA1	Aab,Abc	4	4	100%	1	4	25%	4	4	100%	4	4	100%	0	4	0%	3	4	75%	0	4	0%	0	4	0%
AA2	Aba,Acb	1	4	25%	4	4	100%	4	4	100%	4	4	100%	3	4	75%	0	4	0%	0	4	0%	0	4	0%
AA3	Aab,Acb	4	10	40%	4	10	40%	8	10	80%	8	10	80%	6	10	60%	6	10	60%	2	10	20%	2	10	20%
AA4	Aba,Abc	4	8	50%	4	8	50%	8	8	100%	8	8	100%	4	8	50%	4	8	50%	0	8	0%	0	8	0%
AI1	Aab,Ibc	8	20	40%	4	20	20%	16	20	80%	16	20	80%	12	20	60%	16	20	80%	4	20	20%	4	20	20%
AI2	Aba,Icb	4	16	25%	8	16	50%	16	16	100%	16	16	100%	12	16	75%	8	16	50%	0	16	0%	0	16	0%
AI3	Aab,Icb	8	20	40%	4	20	20%	16	20	80%	16	20	80%	12	20	60%	16	20	80%	4	20	20%	4	20	20%
AI4	Aba,Ibc	4	16	25%	8	16	50%	16	16	100%	16	16	100%	12	16	75%	8	16	50%	0	16	0%	0	16	0%
AO1	Aab,Obc	4	18	22.22%	3	18	16.67%	12	18	66.67%	12	18	66.67%	14	18	77.78%	15	18	83.33%	6	18	33.33%	6	18	33.33%
AO2	Aba,Ocb	3	18	16.67%	6	18	33.33%	16	18	88.89%	16	18	88.89%	15	18	83.33%	12	18	66.67%	2	18	11.11%	2	18	11.11%
AO3	Aab,Ocb	4	12	33.33%	0	12	0%	8	12	66.67%	8	12	66.67%	8	12	66.67%	12	12	100%	4	12	33.33%	4	12	33.33%
AO4	Aba,Obc	0	14	0%	6	14	42.86%	12	14	85.71%	12	14	85.71%	14	14	100%	8	14	57.14%	2	14	14.29%	2	14	14.29%
AE1	Aab,Ebc	0	2	0%	0	2	0%	0	2	0%	0	2	0%	2	2	100%	2	2	100%	2	2	100%	2	2	100%
AE2	Aba,Ecb	0	6	0%	2	6	33.33%	4	6	66.67%	4	6	66.67%	6	6	100%	4	6	66.67%	2	6	33.33%	2	6	33.33%
AE3	Aab,Ecb	0	2	0%	0	2	0%	0	2	0%	0	2	0%	2	2	100%	2	2	100%	2	2	100%	2	2	100%
AE4	Aba,Ebc	0	6	0%	2	6	33.33%	4	6	66.67%	4	6	66.67%	6	6	100%	4	6	66.67%	2	6	33.33%	2	6	33.33%
IA1	Iab,Abc	8	16	50%	4	16	25%	16	16	100%	16	16	100%	8	16	50%	12	16	75%	0	16	0%	0	16	0%
IA2	Iba,Acb	4	20	20%	8	20	40%	16	20	80%	16	20	80%	16	20	80%	12	20	60%	4	20	20%	4	20	20%
IA3	Iab,Acb	4	20	20%	8	20	40%	16	20	80%	16	20	80%	16	20	80%	12	20	60%	4	20	20%	4	20	20%
IA4	Iba,Abc	8	16	50%	4	16	25%	16	16	100%	16	16	100%	8	16	50%	12	16	75%	0	16	0%	0	16	0%
II1	Iab,Ibc	16	80	20%	16	80	20%	72	80	90%	72	80	90%	64	80	80%	64	80	80%	8	80	10%	8	80	10%
II2	Iba,Icb	16	80	20%	16	80	20%	72	80	90%	72	80	90%	64	80	80%	64	80	80%	8	80	10%	8	80	10%
II3	Iab,Icb	16	80	20%	16	80	20%	72	80	90%	72	80	90%	64	80	80%	64	80	80%	8	80	10%	8	80	10%
II4	Iba,Ibc	16	80	20%	16	80	20%	72	80	90%	72	80	90%	64	80	80%	64	80	80%	8	80	10%	8	80	10%
IO1	Iab,Obc	8	76	10.53%	16	76	21.05%	64	76	84.21%	64	76	84.21%	68	76	89.47%	60	76	78.95%	12	76	15.79%	12	76	15.79%
IO2	Iba,Ocb	12	72	16.67%	12	72	16.67%	64	72	88.89%	64	72	88.89%	60	72	83.33%	60	72	83.33%	8	72	11.11%	8	72	11.11%
IO3	Iab,Ocb	12	72	16.67%	12	72	16.67%	64	72	88.89%	64	72	88.89%	60	72	83.33%	60	72	83.33%	8	72	11.11%	8	72	11.11%
IO4	Iba,Obc	8	76	10.53%	16	76	21.05%	64	76	84.21%	64	76	84.21%	68	76	89.47%	60	76	78.95%	12	76	15.79%	12	76	15.79%
IE1	Iab,Ebc	0	12	0%	4	12	33.33%	8	12	66.67%	8	12	66.67%	12	12	100%	8	12	66.67%	4	12	33.33%	4	12	33.33%
IE2	Iba,Ecb	0	12	0%	4	12	33.33%	8	12	66.67%	8	12	66.67%	12	12	100%	8	12	66.67%	4	12	33.33%	4	12	33.33%
IE3	Iab,Ecb	0	12	0%	4	12	33.33%	8	12	66.67%	8	12	66.67%	12	12	100%	8	12	66.67%	4	12	33.33%	4	12	33.33%
IE4	Iba,Ebc	0	12	0%	4	12	33.33%	8	12	66.67%	8	12	66.67%	12	12	100%	8	12	66.67%	4	12	33.33%	4	12	33.33%
OA1	Oab,Abc	6	18	33.33%	3	18	16.67%	16	18	88.89%	16	18	88.89%	12	18	66.67%	15	18	83.33%	2	18	11.11%	2	18	11.11%
OA2	Oba,Acb	3	18	16.67%	4	18	22.22%	12	18	66.67%	12	18	66.67%	15	18	83.33%	14	18	77.78%	6	18	33.33%	6	18	33.33%
OA3	Oab,Acb	0	12	0%	4	12	33.33%	8	12	66.67%	8	12	66.67%	12	12	100%	8	12	66.67%	4	12	33.33%	4	12	33.33%
OA4	Oba,Abc	6	14	42.86%	0	14	0%	12	14	85.71%	12	14	85.71%	8	14	57.14%	14	14	100%	2	14	14.29%	2	14	14.29%
OI1	Oab,Ibc	12	72	16.67%	12	72	16.67%	64	72	88.89%	64	72	88.89%	60	72	83.33%	60	72	83.33%	8	72	11.11%	8	72	11.11%
OI2	Oba,Icb	16	76	21.05%	8	76	10.53%	64	76	84.21%	64	76	84.21%	60	76	78.95%	68	76	89.47%	12	76	15.79%	12	76	15.79%
OI3	Oab,Icb	12	72	16.67%	12	72	16.67%	64	72	88.89%	64	72	88.89%	60	72	83.33%	60	72	83.33%	8	72	11.11%	8	72	11.11%
OI4	Oba,Ibc	16	76	21.05%	8	76	10.53%	64	76	84.21%	64	76	84.21%	60	76	78.95%	68	76	89.47%	12	76	15.79%	12	76	15.79%
OO1	Oab,Obc	8	69	11.59%	15	69	21.74%	60	69	86.96%	60	69	86.96%	61	69	88.41%	54	69	78.26%	9	69	13.04%	9	69	13.04%
OO2	Oba,Ocb	15	69	21.74%	8	69	11.59%	60	69	86.96%	60	69	86.96%	54	69	78.26%	61	69	88.41%	9	69	13.04%	9	69	13.04%
OO3	Oab,Ocb	14	75	18.67%	14	75	18.67%	68	75	90.67%	68	75	90.67%	61	75	81.33%	61	75	81.33%	7	75	9.33%	7	75	9.33%
OO4	Oba,Obc	12	73	16.44%	12	73	16.44%	60	73	82.19%	60	73	82.19%	61	73	83.56%	61	73	83.56%	13	73	17.81%	13	73	17.81%
OE1	Oab,Ebc	2	15	13.33%	6	15	40%	12	15	80%	12	15	80%	13	15	86.67%	9	15	60%	3	15	20%	3	15	20%
OE2	Oba,Ecb	2	11	18.18%	4	11	36.36%	8	11	72.73%	8	11	72.73%	9	11	81.82%	7	11	63.64%	3	11	27.27%	3	11	27.27%
OE3	Oab,Ecb	2	15	13.33%	6	15	40%	12	15	80%	12	15	80%	13	15	86.67%	9	15	60%	3	15	20%	3	15	20%
OE4	Oba,Ebc	2	11	18.18%	4	11	36.36%	8	11	72.73%	8	11	72.73%	9	11	81.82%	7	11	63.64%	3	11	27.27%	3	11	27.27%
EA1	Eab,Abc	2	6	33.33%	0	6	0%	4	6	66.67%	4	6	66.67%	4	6	66.67%	6	6	100%	2	6	33.33%	2	6	33.33%
EA2	Eba,Acb	0	2	0%	0	2	0%	0	2	0%	0	2	0%	2	2	100%	2	2	100%	2	2	100%	2	2	100%
EA3	Eab,Acb	0	2	0%	0	2	0%	0	2	0%	0	2	0%	2	2	100%	2	2	100%	2	2	100%	2	2	100%
EA4	Eba,Abc	2	6	33.33%	0	6	0%	4	6	66.67%	4	6	66.67%	4	6	66.67%	6	6	100%	2	6	33.33%	2	6	33.33%
EI1	Eab,Ibc	4	12	33.33%	0	12	0%	8	12	66.67%	8	12	66.67%	8	12	66.67%	12	12	100%	4	12	33.33%	4	12	33.33%
EI2	Eba,Icb	4	12	33.33%	0	12	0%	8	12	66.67%	8	12	66.67%	8	12	66.67%	12	12	100%	4	12	33.33%	4	12	33.33%
EI3	Eab,Icb	4	12	33.33%	0	12	0%	8	12	66.67%	8	12	66.67%	8	12	66.67%	12	12	100%	4	12	33.33%	4	12	33.33%
EI4	Eba,Ibc	4	12	33.33%	0	12	0%	8	12	66.67%	8	12	66.67%	8	12	66.67%	12	12	100%	4	12	33.33%	4	12	33.33%
EO1	Eab,Obc	4	11	36.36%	2	11	18.18%	8	11	72.73%	8	11	72.73%	7	11	63.64%	9	11	81.82%	3	11	27.27%	3	11	27.27%
EO2	Eba,Ocb	6	15	40%	2	15	13.33%	12	15	80%	12	15	80%	9	15	60%	13	15	86.67%	3	15	20%	3	15	20%
EO3	Eab,Ocb	6	15	40%	2	15	13.33%	12	15	80%	12	15	80%	9	15	60%	13	15	86.67%	3	15	20%	3	15	20%
EO4	Eba,Obc	4	11	36.36%	2	11	18.18%	8	11	72.73%	8	11	72.73%	7	11	63.64%	9	11	81.82%	3	11	27.27%	3	11	27.27%
EE1	Eab,Ebc	2	5	40%	2	5	40%	4	5	80%	4	5	80%	3	5	60%	3	5	60%	1	5	20%	1	5	20%
EE2	Eba,Ecb	2	5	40%	2	5	40%	4	5	80%	4	5	80%	3	5	60%	3	5	60%	1	5	20%	1	5	20%
EE3	Eab,Ecb	2	5	40%	2	5	40%	4	5	80%	4	5	80%	3	5	60%	3	5	60%	1	5	20%	1	5	20%
EE4	Eba,Ebc	2	5	40%	2	5	40%	4	5	80%	4	5	80%	3	5	60%	3	5	60%	1	5	20%	1	5	20%

Table 2. Comparison between the implemented model in different settings with respect to the assumptions from Section 4 and the empirical findings reported in [5, Table 6]. The table shows the respective (Pearson) correlation between the percentages and their coincidence. For the latter, a conclusion is adopted if the implemented model yields a percentage of 100% or the majority of people in the psychological experiments draw this conclusion, respectively. The last column shows the coincidence with respect to no valid conclusions alone.

assumptions					no valid
2	3	4	correlation	coincidence	conclusions
no	no	no	0.482	89.6%	65.6%
no	no	yes	0.377	89.2%	65.6%
no	yes	no	0.370	87.3%	73.4%
no	yes	yes	0.314	87.0%	73.4%
yes	no	no	0.450	88.9%	65.6%
yes	no	yes	0.367	88.5%	65.6%
yes	yes	no	0.371	87.5%	71.9%
yes	yes	yes	0.317	86.8%	71.9%

As already been noted (in Section 2), from a logical point of view, mood **I** is a symmetric operator. The same holds for mood **E**, i.e. 'All A are not C' and 'All C are not A' are logically equivalent. It follows that the approach for syllogistic reasoning in seven spaces presented here cannot distinguish the two categorical assertions for the moods **I** and **E**. Nonetheless, in the psychological experiments [5, Table 6], human reasoners definitely prefer the order A–C for the conclusion in 76,6% (mood **I**) and 70.3% (mood **E**) of the possible 64 cases (corresponding to the rows in Table 1). Furthermore, in all 16 cases where the premises have the form of figure 1, 100% of the human reasoners prefer this order or, at least, do not prefer the reverse order.

6 Conclusions

In this paper, we have introduced a set-based approach to syllogistic reasoning that allows a fine-grained analysis of syllogistic reasoning. It is cognitively simple, because it makes use of only seven spaces. In addition, there is a straightforward implementation by means of constraint logic programming, that is presented here for the first time. Last but not least, several assumptions humans implicitly may have can be tested, e.g. that all categories are non-empty.

Further work will concentrate on a more thorough analysis of syllogistic reasoning and comparison with the empirical findings. Possibly no single, monolithic theory can explain the whole picture. In consequence, the individual behavior of a test person may be derived from sample conclusions taken by that person. In addition, techniques from machine learning like case-based reasoning, clustering, decision tree learning, or neural networks may be employed.

References

1. Crivelli, P.: Aristotle's logic. In: Shields, C. (ed.) The Oxford Handbook of Aristotle, pp. 113–149. Oxford Handbooks, Oxford University Press (2012)
2. Hall-Partee, B.: Semantics – mathematics or psychology? In: Bäuerle, R., Egli, U., von Stechow, A. (eds.) Semantics from different points of view, pp. 1–14. Springer (1979)
3. Johnson-Laird, P.N.: Models of deduction. In: Falmagne, R.J. (ed.) Reasoning: Representation and Process: Children and Adults, pp. 7–54. Erlbaum, Springdale, NJ (1975)
4. Johnson-Laird, P.N.: Mental models: Towards a cognitive science of language, inference, and consciousness. Harvard University Press (1983)
5. Khemlani, S., Johnson-Laird, P.N.: Theories of the syllogism: A meta-analysis. Psychological Bulletin 138(3), 427–457 (2012)
6. Kumova, B.I.: Approximate reasoning with fuzzy-syllogistic systems. In: Beierle, C., Kern-Isberner, G., Ragni, M., Stolzenburg, F. (eds.) Proceedings of the KI 2015 Workshop on Formal and Cognitive Reasoning – 5th Workshop on Dynamics of Knowledge and Belief (DKB-2015) and 4th Workshop KI & Kognition (KIK-2015). pp. 55–62. CEUR Workshop Proceedings 1444, Dresden (2015), `http://ceur-ws.org/Vol-1444/paper5.pdf`
7. Lüderitz, R.: Analyse des syllogistischen Schließens. Projektarbeit, Fachbereich Automatisierung und Informatik, Hochschule Harz (2016)
8. Marriott, K., Stuckey, P.J.: Programming with Constraints. MIT Press, Cambridge, MA, London (1998)
9. Newell, A., Nickerson, R.: Reasoning, Problem Solving, and Decision Processes: The Problem Space as a Fundamental Category, pp. 693–718. Harvard, Cambridge, MA (1980)
10. Oaksford, M., Chater, N.: Bayesian Rationality: The Probabilistic Approach to Human Reasoning. Oxford Cognitive Science Series, Oxford University Press (2007)
11. Smith, R.: Aristotle's logic. In: Zalta, E.N. (ed.) Stanford Encyclopedia of Philosophy. Stanford University (2000), `http://plato.stanford.edu/entries/aristotle-logic/`, online encyclopedia, updated continuously, substantive revision 2017
12. Wason, P.C.: Reasoning about a rule. Quarterly Journal of Experimental Psychology 20(3), 273–281 (1968)
13. Wetherick, N.E., Gilhooly, K.J.: 'Atmosphere', matching, and logic in syllogistic reasoning. Current Psychology 14(3), 169–178 (1995)
14. Wielemaker, J.: SWI-Prolog Reference Manual. University of Amsterdam, The Netherlands (2015), `http://www.swi-prolog.org/download/stable/doc/SWI-Prolog-7.2.3.pdf`, updated for version 7.2.3
15. Woodworth, R.S., Sells, S.B.: An atmosphere effect in formal syllogistic reasoning. Journal of Experimental Psychology 18(4), 451–460 (1935)
16. Zarechnev, M., Kumova, B.I.: Truth ratios of syllogistic moods. In: Yazici, A., Pal, N.R., Kaymak, U., Martin, T., Ishibuchi, H., Lin, C., Sousa, J.M.C., Tütmez, B. (eds.) FUZZ-IEEE 2015 – IEEE International Conference on Fuzzy Systems. pp. 1–8. IEEE, Istanbul, Turkey (2015)