1 INTRODUCTION

A new computational method is presented to implement the system of deductive logic described by Aristotle in *Prior Analytics* [1]. Each Aristotelian problem is interpreted as a parametric probability network in which the premises give constraints on probabilities relating the problem’s categorical terms (major, minor, and middle). Each probability expression evaluates to a linear function of the parameters in the probability model. By this approach the constraints specified as premises translate into linear equalities and inequalities involving a few real-valued variables. Using linear optimization methods, the minimum and maximum feasible values of certain queried probabilities are computed, subject to the constraints given as premises. These computed solutions determine precisely which conclusions are necessary consequences of the premises; thus logical deduction is accomplished by numerical computation. The problem’s figure (schema) describes which specific probabilities are constrained, relative to those that are queried.

This work is a synthesis of several existing methods, with the addition of a few new ideas. The most relevant prior work is that of Boole, who presented several innovations in his 1854 treatise on the *Laws of Thought* [4]. Boole demonstrated that logical propositions can be represented as algebraic formulas; more specifically, that statements of what we now call propositional calculus can be expressed as polynomials with real-number coefficients. Boole showed how to compute interesting results about logical propositions by solving systems of polynomial equations. Boole also showed useful relationships between statements of logic and statements of probability. In the late 20th century Pearl and others developed techniques for graphical probability models (Bayesian networks) which offer several benefits regarding representation and inference [20, 9]. Several investigators have described methods for symbolic inference in probability networks, which can be used to calculate polynomial expressions for queried probabilities [7, 5, 14, 22]. Boole already formulated optimization problems with polynomial objectives and constraints derived from logical formulas and from probability expressions; he solved these problems by *ad hoc* algebraic manipulations [4]. Now there are general methods for solving linear and nonlinear polynomial optimization problems. For the linear case, efficient computational methods were developed in the middle of the 20th century [6, 12]. It remains a challenge to compute exact global solutions to unrestricted nonlinear polynomial optimization problems; there are many promising methods which use various kinds of approximation [3, 21, 13, 17].

This work introduces two new ideas to complement these existing methods. First, a taxonomy of Aristotelian categorical statements is developed, with a distinction drawn between ‘primary’ and ‘composite’ relations. In this taxonomy, primary relations are mutually exclusive, whereas composite relations may overlap; also for both kinds of relations the case of an impossible subject (antecedent) is handled explicitly. This taxonomy results in more types of categorical statements than are usually considered; one advantage is that existential fallacies are prevented. Second, the concept of ‘complementary’ syllogism is introduced, to contrast with ‘classical’ syllogism. In a complementary syllogism the subject of the deduced categorical statement is held to be false instead of true. Complementary syllogism lets us extract additional information from categorical premises that would otherwise be lost to analysis. These new ideas are combined with the existing methods mentioned above (Boole’s mathematical logic; probability networks and symbolic probability inference; and linear optimization) to provide the method of analysis presented here. Let us call this synthesis the ‘probability-optimization paradigm’ for framing Aristotle’s logic.
Table 1  Four figures of Aristotelian problems, using major term $A$, middle term $B$, and minor term $C$. The major premise has type $m$, the minor premise has type $n$, and the goal is to find the implied types $s$ for the query statement $AsC$. The possible relation types are defined in Table 4. In the diagrams each arrow points from the subject to the predicate of a conditional statement; solid arrows indicate premises and dotted arrows indicate queries. The ‘flat’ diagrams follow Aristotle’s original (textual) descriptions; the ‘triangular’ diagrams provide a different view.

1.1 Review of Aristotle’s Logical Problems

The problems in Aristotle’s Prior Analytics involve three categorical terms, called ‘major’, ‘middle’, and ‘minor’, each of which can be either true or false. Let us use $A$ for the major term, $B$ for the middle term, and $C$ for the minor term. We abbreviate truth as $T$ and falsity as $F$. The major and minor terms are also called the ‘extreme’ terms, in contrast to the middle. Each Aristotelian problem consists of two premises and a query. Each premise has a subject and a predicate, each of which is one of the three categorical terms. Each premise also has a type that quantifies the relationship between its subject and predicate terms. Five main types of relationships are recognized here: universal-affirmative, universal-negative, particular-affirmative, particular-negative, and particular-intermediate. The universal relationships are further specialized into material and existential subtypes (as described in detail below); this expands the repertoire of categorical relationships to seven types. Note that these seven types of relationships are not mutually exclusive; certain combinations can hold simultaneously.

Aristotle imposed a few restrictions on how the various categorical terms may be used in a problem’s premises and query. Both of the problem’s premises must use its middle term $B$. One premise must use the problem’s major term $A$, and the other premise must use the problem’s minor term $C$ (hence these premises are called ‘major’ and ‘minor’ themselves). Within each premise either term may occupy the position of subject or of predicate. These restrictions allow four possible figures for Aristotelian problems, as shown in Table 1 below. Regardless of which figure is used in a problem, the main query is always the same: to find out what relationship between the major and minor terms is required by the given premises. In this query the major term $A$ is used as predicate and the minor term $C$ used as subject; this choice of positions is precisely what distinguishes major from minor. The inferred solution is a subset of the seven types of categorical relationships introduced above. We say that a ‘syllogism’ (deduction) has occurred when at least one of these types of relationships must hold. On the other hand there is no syllogism when the premises do not require any particular relationship between the major and minor terms.

Note that mathematical relationships may be asserted as constraints or derived as solutions; these are two different roles. For example, it is one thing to assert $x = 2$ as a constraint that should be satisfied, and a different thing to derive $x = 2$ as a solution to some other system of constraints (for example, as one of
the two real solutions to the equation $x^2 = 4$. We shall view Aristotle’s problems as systems of equations, both philosophically and practically. Philosophically, we shall regard categorical statements like ‘$A$ belongs to some $B’$ as relations like ‘$x = 2’$, to be used in either of the two roles just mentioned: sometimes asserted as constraints, and sometimes derived as solutions to other constraints. Practically, we shall translate logical statements about true/false terms into algebraic statements about real-valued variables (through the intermediate device of a probability model), and then use the tools of ordinary algebra (with real numbers) to compute solutions to the original logical problems.

1.2 A Basic Probability Model

The basic probability model described here represents the joint probabilities of the three categorical terms $A$, $B$, and $C$. With two possible truth values for each of the three terms, there are $2^3 = 8$ possible combinations of truth values. To each combination $i$ of truth values we assign a symbolic parameter $x_i$ that represents its probability. These parameters and their associated combinations of truth values are shown as the input probability table $Pr_0(A, B, C)$ in Table 2. To respect the laws of probability, these parameters $x_1, \ldots, x_8$ are constrained by $0 \leq x_i \leq 1$ and $\sum_i x_i = 1$.

| $A$ | $B$ | $C$ | $Pr_0(A, B, C)$ |
|-----|-----|-----|----------------|
| T   | T   | T   | $x_1$          |
| T   | T   | F   | $x_2$          |
| T   | F   | T   | $x_3$          |
| T   | F   | F   | $x_4$          |
| F   | T   | T   | $x_5$          |
| F   | T   | F   | $x_6$          |
| F   | F   | T   | $x_7$          |
| F   | F   | F   | $x_8$          |

Table 2: The input probability table $Pr_0(A, B, C)$ for the basic model. The $x_i \in \mathbb{R}$ are subject to $0 \leq x_i \leq 1$ and $\sum_i x_i = 1$.

Parametric probability networks such as this basic model are used like databases to answer queries. Each query requests an unconditioned probability or a conditional probability. Each response is a polynomial or a quotient of polynomials in the model’s parameters, computed from sums of the appropriate full-joint probabilities. Each full-joint probability is the product of one value from each input probability table. For example, starting from the inputs shown in Table 2, the respective probabilities that $B$ is true, that $B$ and $A$ are both true, that $B$ is true and $A$ is false, and that $A$ is true given that $B$ is true can be computed as the following expressions:

\[
\begin{align*}
Pr(B) & \Rightarrow x_1 + x_2 + x_5 + x_6 & (1) \\
Pr(B, A) & \Rightarrow x_1 + x_2 & (2) \\
Pr(B, \overline{A}) & \Rightarrow x_5 + x_6 & (3) \\
Pr(A|B) & \Rightarrow (x_1 + x_2) / (x_1 + x_2 + x_5 + x_6) & (4)
\end{align*}
\]

As you can see, each of these calculated values is either the sum of several input probabilities from Table 2 or the quotient of two such sums. It happens with this basic probability model that all computed probabilities are linear functions of the $x_i$ parameters (or quotients of such linear functions). Other probability models can yield nonlinear polynomials and quotients (when the full-joint probability has been factored into multiple input tables). Table 3 shows several output probabilities computed from the inputs in Table 2 which will be useful for the analysis that follows.

The essential methods of symbolic probability inference were described well enough several centuries ago [15]. There have since been developed rigorous mathematical formulations, powerful graphical models, and efficient inference algorithms [10, 8, 9]. The author has developed some computational methods for
parametric probability networks [16][18]. These methods include some idiosyncratic notation that is touched upon presently. 

Input and output probabilities are distinguished from one another. Input probabilities, used to specify the probability model, are written with the subscript 0, as in \( \Pr_0(A, B, C) \). Output probabilities, computed from the inputs, are written without a subscript, as in \( \Pr(B = T) \). The double right arrow \( \Rightarrow \) is used to indicate computation, such as the evaluation of a symbolic probability expression or the simplification of an arithmetical formula. This is distinct from the test or assertion of equality denoted with the usual equal sign \( = \). Finally, probability tables and their elements share similar notation. A probability expression such as \( \Pr(A, B) \) may refer to a table containing several values, such as the four elements shown as Table 3 (a).

But we can also use for example \( A \) to abbreviate the event \( A = T \) and \( \overline{A} \) to abbreviate the event \( A = F \), and hence use \( \Pr(A, B) \) to mean the individual element \( \Pr(A = T, B = T) \). The default used here is that probability expressions refer to individual elements; it will be announced in the neighboring text when a probability expression refers instead to an entire table containing several elements.

### 2 FROM CATEGORICAL STATEMENTS TO LINEAR EQUALITIES AND INEQUALITIES

Let us now translate Aristotelian categorical statements into linear equalities and inequalities involving the parameters of the basic probability model presented in Section 1.2. To begin, we regard a categorical statement with predicate \( P \) and subject \( Q \) as a relation involving \( \Pr(P | Q) \), the conditional probability that \( P \) is true given that \( Q \) is true. Here \( P \) and \( Q \) can be any of the three categorical terms \( A, B, \) or \( C \). At first glance, Aristotle’s Prior Analytics describes four types of relations between categorical terms, which suggest the following conditional-probability statements:

- **Universal-affirmative**: ‘\( P \) belongs to all \( Q \)’; \( \Pr(P | Q) = 1 \)
- **Universal-negative**: ‘\( P \) belongs to no \( Q \)’; \( \Pr(P | Q) = 0 \)
- **Particular-affirmative**: ‘\( P \) belongs to some \( Q \)’; \( \Pr(P | Q) > 0 \)
- **Particular-negative**: ‘\( P \) does not belong to some \( Q \)’; \( \Pr(P | Q) < 1 \)

Recall that conditional probabilities are defined as quotients of unconditioned probabilities:

\[
\Pr(P | Q) = \frac{\Pr(P, Q)}{\Pr(Q)} \tag{5}
\]

There are two troublesome issues with the four types of relations listed above. First, there is no prescription for what to do in case \( \Pr(Q) = 0 \) (i.e. it is impossible a priori for the subject term \( Q \) to be true). Since the laws of probability require that \( \Pr(P, Q) = 0 \) when \( \Pr(Q) = 0 \), this exceptional case would force the quotient shown in Equation 5 to have the indefinite value 0/0. It is unclear whether equations such as 0/0 = 0 and 0/0 = 1 should be considered satisfied or not (there is a reasonable argument to be made either way). Second, the listed relations are not mutually exclusive. For example \( \Pr(P | Q) = 1 \) requires also \( \Pr(P | Q) > 0 \), and conversely \( \Pr(P | Q) > 0 \) leaves it possible but not certain that \( \Pr(P | Q) = 1 \).

Note some potential confusion regarding the negation of a particular-affirmative premise. The English phrase ‘\( P \) does not belong to some \( Q \)’ leaves some ambiguity about what is negated. This phrase
could be interpreted to mean, ‘It is not the case that $P$ belongs to some $Q$’—suggesting the conditional-probability constraint $\Pr (P \mid Q) = 0$. Or this phrase could be interpreted to mean, ‘The negation of $P$ belongs to some $Q$’—suggesting the constraint $\Pr (\neg P \mid Q) > 0$ (or its equivalent $\Pr (P \mid Q) < 1$ which is listed above). Here the second of these interpretations is assumed.

2.1 Primary and Composite Relations

In order to address the issues with the naive categorical relations presented above, let us develop a different initial set of relation types—making direct use of the numerator $\Pr (P, Q)$ and denominator $\Pr (Q)$ from Equation 5 instead of their quotient $\Pr (P \mid Q)$. Let us call the following four items the primary relations concerning a predicate term $P$ and a subject term $Q$:

I. IMPOSSIBLE-SUBJECT: $\Pr (Q) = 0$
   ‘There are no $Q$’

II. UNIVERSAL-AFFIRMATIVE-EXISTENTIAL: $\Pr (P, Q) = \Pr (Q)$ and $\Pr (Q) > 0$
   ‘$P$ belongs to all $Q$, and there are some $Q$’

III. UNIVERSAL-NEGATIVE-EXISTENTIAL: $\Pr (P, Q) = 0$ and $\Pr (Q) > 0$
   ‘$P$ belongs to no $Q$, and there are some $Q$’

IV. PARTICULAR-INTERMEDIATE: $\Pr (P, Q) > 0$ and $\Pr (P, Q) < \Pr (Q)$
   ‘$P$ belongs to some but not all $Q$’

Both troublesome issues identified above have been addressed: these four primary relations are mutually exclusive, and they include explicit consideration of the case $\Pr (Q) = 0$.

Now, we can use various subsets of these four primary relations to define composite relations which include the seven types of categorical relationships promised in the introduction. These compositions should be understood as logical disjunctions. For example, combining primary relations II and III yields the composite relation that either $\Pr (Q) = 0$, or $\Pr (Q) > 0$ and $\Pr (P, Q) = \Pr (Q)$: in other words, the combined statement that ‘Either there are no $Q$, or there are some $Q$ and $P$ belongs to all of them’. This composite relation can be specified as the single constraint $\Pr (P, Q) = \Pr (Q)$. Recalling Equation 5, this single constraint requires either that the conditional probability $\Pr (P \mid Q) = 1$, or that $\Pr (Q) = 0$ and consequently $\Pr (P \mid Q)$ has the indefinite value $0/0$. This composite relation parallels the statement of material implication $Q \rightarrow P$ from the propositional calculus; thus we call it the UNIVERSAL-AFFIRMATIVE-MATERIAL relation.

Table 4 shows 7 composite relations defined as disjunctive combinations of the primary relations introduced above. This set of composite relations is not exhaustive: given 4 primary relations there are $2^4$ or 16 possible sets of such. Table 5 gives conditional-probability and natural-language descriptions of the selected 7 composite relations. In both tables, the composite relations are assigned codes to abbreviate them, based on the letters introduced in medieval times to designate different types of Aristotelian premises. The traditional codes $a$, $e$, $i$, and $o$ are supplemented with accented characters $\acute{a}$ and $\acute{e}$ that distinguish existential subtypes of universal statements from their material counterparts. Also, the letter $u$ has been added to designate the particular relation meaning ‘some but not all’. Within this document, for abbreviations $PaQ$, $PiQ$, and so on, the predicate $P$ is displayed before the relation code, and the subject $Q$ after it. Be aware that some authors use the opposite convention. Aristotle’s original texts did not use such abbreviations at all.

2.2 Instantiation as Linear Equalities and Inequalities

It remains to instantiate the composite relations defined in Table 4 into specific equalities and inequalities involving the parameters $x_1, \ldots, x_8$ of the basic probability model from Section 1.2 when particular categorical terms ($A$, $B$, or $C$) have been chosen as the predicate $P$ and as the subject $Q$. This instantiation is accomplished by using the results of symbolic probability inference shown in Table 3 to supply algebraic formulas for the relevant probability expressions.
| Composite Relation                        | Code   | Primary Relations | Probability Definition                                                                 |
|------------------------------------------|--------|-------------------|----------------------------------------------------------------------------------------|
| Universal-affirmative-material           | PaQ    | •                 | Pr (P, Q) = Pr (Q)                                                                     |
| Universal-affirmative-existential        | PaQ not 0 | •              | Pr (P, Q) = Pr (Q) and Pr (Q) > 0                                                      |
| Universal-negative-material              | PeQ not 0 | •                 | Pr (P, Q) = 0                                                                         |
| Universal-negative-existential           | PeQ not 0 | •                 | Pr (P, Q) = 0 and Pr (Q) > 0                                                           |
| Particular-affirmative                   | PiQ not 0 | •                 | Pr (P, Q) > 0                                                                         |
| Particular-negative                      | PoQ not 0 | •                 | Pr (P, Q) < Pr (Q)                                                                     |
| Particular-intermediate                  | PuQ not 0 | •                 | Pr (P, Q) > 0 and Pr (P, Q) < Pr (Q)                                                   |

Table 4 Composite relations between a categorical predicate P and a subject Q, using disjunctions of the primary relations I, II, III, and IV from Section 2.1. Each • indicates that the given primary relation is included in the given composite: for example universal-affirmative-material holds if either primary relation I or II holds. The probability relations that define the composite categorical relations are shown.

| Composite Relation                  | Code   | Conditional Prob. | Natural-Language Description                               |
|-------------------------------------|--------|-------------------|------------------------------------------------------------|
| Universal-affirmative-material      | PaQ    | Pr (P | Q) = 1 or 0/0 | P belongs to all Q, or there are no Q                      |
| Universal-affirmative-existential   | PaQ not 0 | Pr (P | Q) = 1 | P belongs to all Q, and there are some Q                   |
| Universal-negative-material         | PeQ not 0 | Pr (P | Q) = 0 or 0/0 | P belongs to no Q, or there are no Q                      |
| Universal-negative-existential      | PeQ not 0 | Pr (P | Q) = 0 | P belongs to no Q, and there are some Q                   |
| Particular-affirmative              | PiQ not 0 | Pr (P | Q) = 0 | P belongs to some Q                                       |
| Particular-negative                 | PoQ not 0 | Pr (P | Q) < 1 | The negation of P belongs to some Q                       |
| Particular-intermediate             | PuQ not 0 | 0 < Pr (P | Q) < 1 | P belongs to some but not all Q                           |

Table 5 Composite relations between a categorical predicate P and a subject Q, described in terms of conditional probability and natural language. These relations are defined by the unconditional probability statements shown in Table 4.
For example, let us consider the **UNIVERSAL-AFFIRMATIVE-EXISTENTIAL** statement with predicate term A and subject term B. As Table 3 shows, this categorical statement \( A\alpha B \) says that ‘A belongs to all B’, and there are some B*. Using Table 4 and substituting A for the predicate term P and B for the subject term Q, this statement \( A\alpha B \) is defined by the probability relations \( \Pr (B, A) = \Pr (B) \) and \( \Pr (B) > 0 \). Table 5 gives the algebraic formulas for the relevant probability expressions. The probability \( \Pr (B, A) \), meaning \( \Pr (B = T, A = T) \), is given in the first row of the computed table \( \Pr (A, B) \) which appears as part (a) of Table 3. The probability \( \Pr (B) \), meaning \( \Pr (B = T) \), is given in the first row of the computed table \( \Pr (B) \) which appears as part (d) of Table 3. Substituting these values into the probability relations \( \Pr (B, A) = \Pr (B) \) and \( \Pr (B) > 0 \) derived from Table 3, the categorical statement \( A\alpha B \) is therefore defined by the following algebraic relations:

\[
\begin{align*}
x_1 + x_2 &= x_1 + x_2 + x_5 + x_6 \quad (6) \\
x_1 + x_2 + x_5 + x_6 &> 0 \quad (7)
\end{align*}
\]

Equation (6) simplifies to \( x_5 + x_6 = 0 \) by rearranging its terms. As Table 5 reminds us, the constraints \( \Pr (B, A) = \Pr (B) \) and \( \Pr (B) > 0 \) that define the categorical statement \( A\alpha B \) require that the conditional probability \( \Pr (A \mid B) \) must have the definite value 1. This is evident from inspecting Equations (6) and (7) along with the algebraic formula computed for the conditional probability:

\[
\Pr (A = T \mid B = T) \Rightarrow (x_1 + x_2) / (x_1 + x_2 + x_5 + x_6)
\]

3 **COMPUTING CLASSICAL AND COMPLEMENTARY SYLLOGISMS**

The methods of Section 2 enable the translation of Aristotelian categorical statements into linear equalities and inequalities involving the real-valued parameters \( x_1, \ldots, x_8 \) of the basic probability model from Section 1.2. Such equalities and inequalities can be asserted as constraints themselves, or inferred as solutions to other systems of constraints. We now turn to the task of performing such inference: using linear optimization to compute some linear equalities and inequalities from others. These computed algebraic results can be used to deduce logical conclusions from Aristotelian premises. We shall define two kinds of deductions: **classical** syllogisms, which follow Aristotle’s original practice; and **complementary** syllogisms, which offer a useful variation on the theme. As explained further in Section 3.3, the difference is that for classical syllogisms the minor term is held to be true, and in complementary syllogisms the minor term is held to be false (in the queries evaluated to determine whether syllogisms are present).

### 3.1 Linear Programming to Bound Queried Probabilities

Consider an objective function \( f \) and some constraint functions \( g_1, g_2, \ldots, g_m \), each of which is a linear function of a list \( x = (x_1, \ldots, x_n) \) of variables. Standard linear optimization methods can compute the minimum and maximum feasible values of the objective \( f(x) \) subject to linear equality and inequality constraints such as \( g_1(x) = 0, g_2(x) \geq 0 \), and so on [12]. Standard linear optimization methods can also detect inconsistent constraints, a potential exception that is important to recognize.

For probabilistic translations of Aristotelian problems it is necessary to reason with strict inequalities (\( > \) and \( < \)) as well as weak ones (\( \leq \) and \( \geq \)). However common linear optimization methods treat all inequalities as weak. The following ‘epsilon-inequality reformulation’ works around this limitation. We choose a constant value \( \epsilon \) such as 0.01 that is small relative to 1, and replace each strict-inequality constraint \( g(x) > h(x) \) with a weak-inequality constraint \( g(x) \geq h(x) + \epsilon \). The optimization problem thus reformulated is solved using standard linear programming methods to find the minimum and maximum feasible values of the objective function \( f(x) \) subject to the given constraints. Any computed minimum solution greater than or equal to

---

*It happens that following Aristotle’s restrictions on terms and premises, there are no infeasible problems. However with more than two premises, or with terms that do not follow the typical major/minor/middle arrangement, it is quite possible to give inconsistent categorical statements. One simple example is the premise \( A \alpha A \) with the **PARTICULAR-NEGATIVE** relation type and the same term used as both subject and predicate. This premise translates to the unsatisfiable probability constraint \( \Pr (A, A) < \Pr (A) \) which says that the probability that A is true is strictly less than itself.*
\( \varepsilon \) is interpreted to mean that the objective \( f(x) \) must be strictly greater than zero; likewise any computed maximum solution less than or equal to \( 1 - \varepsilon \) is interpreted to mean that the objective \( f(x) \) must be strictly less than one. These qualitative solutions are sufficient for Aristotelian deduction; for this application the precise value of \( \varepsilon \) is not important.

To illustrate, consider the problem defined by the premises \( BeA \) and \( BiC \). Referring to Tables 3 and 5, the first premise \( BeA \) says ‘\( B \) belongs to no \( A \), or there are no \( A \)’ and translates to the probability constraint \( \Pr(A, B) = 0 \) (requiring that the conditional probability \( \Pr(B \mid A) \) has the value 0 unless it is undefined). As shown in Table 3(a) the probability expression \( \Pr(A = T, B = T) \) has the algebraic value \( x_1 + x_2 \). Hence this first premise \( BeA \) gives the linear constraint:

\[ x_1 + x_2 = 0 \tag{9} \]

The second premise \( BiC \) says ‘\( B \) belongs to some \( C \)’ and translates to the probability constraint \( \Pr(B, C) > 0 \) (requiring that the conditional probability \( \Pr(B \mid C) > 0 \) as well). As Table 3(b) shows, \( \Pr(B = T, C = T) \Rightarrow x_1 + x_5 \). Thus this second premise \( BiC \) asserts the linear constraint:

\[ x_1 + x_5 > 0 \tag{10} \]

Applying epsilon-inequality reformulation gives the modified constraint \( x_1 + x_5 \geq \varepsilon \). Let us choose as our objective function the probability \( \Pr(C, \overline{A}) \) that \( C \) is true and \( A \) is false. As shown in Table 3(c), the probability expression \( \Pr(C = T, A = F) \) evaluates to the algebraic formula \( x_5 + x_7 \).

Combining these results, including the general constraints \( 0 \leq x_i \leq 1 \) and \( \sum_i x_i = 1 \) from the basic probability model in Section 1.2 and choosing \( \varepsilon = 0.01 \) to encode strict inequality leads to the following linear optimization problem to find the minimum feasible value of the chosen objective probability \( \Pr(C, \overline{A}) \) subject to constraints that translate the categorical premises \( BeA \) and \( BiC \):

\begin{align*}
\text{Minimize :} & \quad x_5 + x_7 \\
\text{subject to :} & \quad x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 = 1 \\
& \quad x_1 + x_2 = 0 \\
& \quad x_1 + x_5 \geq \varepsilon \\
& \quad \varepsilon = 0.01 \\
& \quad 0 \leq x_1 \leq 1 \\
& \quad 0 \leq x_2 \leq 1 \\
& \quad 0 \leq x_3 \leq 1 \\
& \quad 0 \leq x_4 \leq 1 \\
& \quad 0 \leq x_5 \leq 1 \\
& \quad 0 \leq x_6 \leq 1 \\
& \quad 0 \leq x_7 \leq 1 \\
& \quad 0 \leq x_8 \leq 1 \\
\end{align*} \tag{11}

Linear optimization yields the numerical solution 0.01. This computed minimum value says directly that the given constraints require \( x_5 + x_7 \geq 0.01 \). Reversing epsilon-inequality encoding, we interpret this numerical solution as the strict algebraic inequality \( x_5 + x_7 > 0 \) which says that the value of the objective function is constrained to be greater than zero at all feasible points. The complementary problem to find the maximum value of the same objective subject to the same constraints yields the solution 1. In other words the given constraints require \( x_5 + x_7 \leq 1 \) (which we already knew from the general constraints included in the basic probability model).

### 3.2 Problems in Four Figures

Before discussing the details of how probability and optimization results indicate logical deductions, let us return to the four figures of Aristotelian problems introduced in Section 1.1. Table 1 describes these four figures in symbolic and graphical notation. The symbolic notation indicates which term is the predicate
and which is the subject of each categorical statement (the two statements asserted as premises, and the one statement used as a query). Italic letters A, B, and C stand for the major, middle, and minor terms. Gothic letters m, n, and s stand for types of categorical relations drawn from the set \{a, â, e, ë, i, o, u\} of types defined in Table 4. m for the relation type of the major premise, n for the relation type of the minor premise, and s for the relation type of the queried statement. For example, a problem in the second figure has a major premise $BmA$ with predicate $B$, subject $A$, and relation type $m$; it has a minor premise $BnC$ with predicate $B$, subject $C$, and relation type $n$; and it has query $AsC$ with predicate $A$, subject $C$, and relation type $s$. The example problem from Section 3.1 follows the second figure. Its major premise $BeA$ uses the \textsc{universal-negative-material} relation (denoted e) as m, and the minor premise $BiC$ uses the \textsc{particular-affirmative} relation (denoted i) as n. The meaning of the query $AsC$ is discussed in the next section.

It is interesting that Aristotle already described his figures in graphical language, indicating the positions of the various categorical terms on the page. For example, regarding his second figure, Aristotle wrote:

...by middle term in it I mean that which is predicated by both subjects, by extremes the terms of which this is said, by major extreme that which lies near the middle, by minor that which is further away from the middle. The middle term stands outside the extremes, and is first in position. (\[1\] 26b35)

Aristotle described his third figure in this way:

...by extremes I mean the predicates, by the major extreme that which is further from the middle, by the minor that which is nearer to it. The middle term stands outside the extremes, and is last in position. (\[1\] 28a15)

The graphical diagrams included in Table 1 realize Aristotle’s original textual descriptions in one view, and use a different graph layout in an alternative view. As has become customary, a fourth figure has been added. In graphical terms the fourth figure is the mirror image of the first, with the positions of the major and minor terms reversed.

3.3 Criteria for Classical and Complementary Syllogism

All four Aristotelian figures use the same query, in which the major term $A$ is the predicate and the minor term $C$ is the subject. The meaning of this common query $AsC$, which is shown in Table 1, is as follows. We seek every relation type s from the set \{a, â, e, ë, i, o, u\} of composite types in Table 4 for which the categorical statement $AsC$ is a necessary consequence of the two categorical statements which have been asserted as premises. There may be zero, one, or many such satisfactory relation types s. If there is at least one satisfactory relation type, we say that a ‘syllogism’ (deduction) is present; but if there are no satisfactory relation types then there is no syllogism. Let us say that a ‘classical’ syllogism concerns a deduced statement in which the minor term $C$ is held to be true. We shall also consider ‘complementary’ syllogisms in which the minor term $C$ is held to be false. Hence to find a complementary syllogism, the query uses the form $As\overline{C}$ whose subject is the negation of the minor term. It is possible to have complementary syllogism with or without classical syllogism (and likewise classical syllogism with or without complementary syllogism).

The results of certain probability-optimization problems indicate precisely which relation types are necessary consequences of the premises provided. To wit, there are four objective functions whose minimum and maximum feasible values must be computed, subject to the constraints translated from the provided premises (and subject also to the general constraints in the basic probability model that reflect the laws of probability). These four objective functions are the joint probabilities of the various combinations of truth and falsity of the extreme terms $A$ and $C$. As shown in Table 6 the symbols $\alpha_j$ and $\beta_j$ are used to represent the computed minimum and maximum values for the four objectives. Table 3 part (c) gives the algebraic values of the relevant probability expressions:

\[ Pr (C, \overline{A}) \Rightarrow x_1 + x_3 \quad Pr (\overline{C}, \overline{A}) \Rightarrow x_5 + x_7 \quad Pr (\overline{C}, A) \Rightarrow x_2 + x_4 \quad Pr (\overline{C}, A) \Rightarrow x_6 + x_8 \]
### Table 6

| Objective Probability | Computed Minimum | Computed Maximum |
|-----------------------|------------------|------------------|
| \( \Pr (C, A) \)      | \( \alpha_1 \)   | \( \beta_1 \)   |
| \( \Pr (C, \overline{A}) \) | \( \alpha_2 \) | \( \beta_2 \) |
| \( \Pr (C, A) \)      | \( \alpha_3 \)   | \( \beta_3 \)   |
| \( \Pr (C, \overline{A}) \) | \( \alpha_4 \) | \( \beta_4 \) |

| \( \Pr (C, A) \) | \( \Pr (C, \overline{A}) \) | Inferred Probability Relations | Categorical Deduction |
|-----------------|-----------------------------|-------------------------------|-----------------------|
| \*              | \( \beta_2 = 0 \)          | \( \Pr (C, A) = \Pr (C) \)  | \( AaC \)             |
| \( \alpha_1 > 0 \) | \( \beta_2 = 0 \) | \( \Pr (C, A) = \Pr (C) \) and \( \Pr (C) > 0 \) | \( A\hat{a}C \) |
| \( \beta_1 = 0 \) | \* | \( \Pr (C, A) = 0 \) | \( AeC \) |
| \( \beta_1 = 0 \) | \( \alpha_2 > 0 \) | \( \Pr (C, A) = 0 \) and \( \Pr (C) > 0 \) | \( A\hat{e}C \) |
| \( \alpha_1 > 0 \) | \* | \( \Pr (C, A) > 0 \) | \( A\hat{i}C \) |
| \*               | \( \alpha_2 > 0 \) | \( \Pr (C, A) < \Pr (C) \) | \( AoC \) |
| \( \alpha_1 > 0 \) | \( \alpha_2 > 0 \) | \( \Pr (C, A) > 0 \) and \( \Pr (C, A) < \Pr (C) \) | \( AuC \) |

| | | | |
|---|---|---|

Table 7 Criteria for classical syllogism, describing how the major term \( A \) must be predicated upon the minor term \( C \) when \( C \) is true, using the given premises as constraints. Here \( \alpha_1 \) is the minimum and \( \beta_1 \) the maximum feasible value of \( \Pr (C, A) \); likewise \( \alpha_2 \) is the minimum and \( \beta_2 \) the maximum feasible value of \( \Pr (C, \overline{A}) \). Multiple criteria may apply.

These algebraic values are used as objective functions during the formulation of optimization problems. Because these objectives represent probabilities, each pair of computed minimum and maximum values is already constrained by \( 0 \leq \alpha_j \leq \beta_j \leq 1 \). Therefore it would be trivial to compute a minimum value \( \alpha_j = 0 \) or a maximum value \( \beta_j = 1 \) for any queried probability, as the laws of probability already require these bounds. A proper deduction requires the computation of a nontrivial upper or lower bound on at least one of the queried probabilities. Note that the precise numerical solutions computed depend on the value \( \varepsilon \) chosen to encode strict inequality (following the epsilon-inequality encoding scheme discussed in Section 3.1).

The computed minimum and maximum values \( \alpha_j \) and \( \beta_j \) defined in Table 6 determine the presence or absence of syllogism, according to the rules discussed presently. Table 7 shows the criteria for classical syllogism, describing the necessary relationships between the major term \( A \) as predicate and the (affirmative) minor term \( C \) as subject. Table 8 shows the criteria for complementary syllogism, describing the necessary relationships between the major term \( A \) as predicate and the negation \( \overline{C} \) of the minor term as subject. As may be evident already, the criteria presented in Tables 7 and 8 for deducing categorical relationships based on relations involving probabilities are none other than the criteria presented in Table 4 for defining categorical relationships based on relations involving probabilities, instantiated for the predicate term \( A \) and the subject term \( C \) (or its negation \( \overline{C} \)).

To illustrate, let us consider how the computed optimization result \( \alpha_2 > 0 \) gives the categorical deduction \( AoC \), as shown in the penultimate row of Table 7. As shown in Table 8, the symbol \( \alpha_2 \) designates the minimum feasible value of the probability \( \Pr (C, \overline{A}) \) subject to the provided constraints. A solution \( \alpha_2 \) which is strictly greater than zero means that the constraints require \( \Pr (C, \overline{A}) > 0 \). The laws of probability provide that:

\[
\Pr (C) = \Pr (C, A) + \Pr (C, \overline{A}) > 0
\]  

(12)
As all probabilities are nonnegative, the inequality $\Pr (C, \overline{A}) > 0$ joined with Equation\textsuperscript{12} requires in turn that $\Pr (C, A) < \Pr (C)$. Returning to Table\textsuperscript{4} this derived inequality $\Pr (C, A) < \Pr (C)$ is precisely the definition of the PARTICULAR-NEGATIVE categorical relationship with predicate $A$ and subject $C$, abbreviated $AoC$. Also, as Table\textsuperscript{5} shows, this unconditioned-probability inequality $\Pr (C, A) < \Pr (C)$ implies the conditional-probability inequality $\Pr (A \mid C) < 1$.

Applying this criterion to the example problem from Section 5.1 produces the deduction that $AoC$ is a necessary consequence of the premises $BeA$ and $BiC$. The solutions calculated for the example in Section 5.1 are the bounds $\alpha_2 = 0.01$ and $\beta_2 = 1$ on the objective probability $\Pr (C, \overline{A})$ (whose algebraic value is $x_5 + x_7$), when this objective is subjected to the constraints $x_1 + x_2 = 0$ and $x_1 + x_5 > 0$ (translated from the categorical premises $BeA$ and $BiC$) and the constraints $0 \leq x_i \leq 1$ and $\sum x_i = 1$ (from the basic probability model). The computed solution $\alpha_2 = 0.01$ says directly that the given constraints require that the objective $\Pr (C, \overline{A}) \geq 0.01$ at all feasible points. Following the chain of reasoning outlined above, we interpret this solution to mean $\Pr (C, \overline{A}) > 0$ and in turn $\Pr (C, A) < \Pr (C)$. This last inequality is the definition of the categorical relationship $AoC$ (from Table\textsuperscript{4}). Thus is it derived using probability and optimization that the premises $BeA$ and $BiC$ require the conclusion $AoC$. In other words, the premises ‘$B$ belongs to no $A$, or there are no $A$’ and ‘$B$ belongs to some $C$’ require the conclusion that ‘The negation of $A$ belongs to some $C$’ (which might alternatively be stated as ‘$A$ does-not-belong to some $C$’).

Following convention we abbreviate each classical syllogism as mns-$k$ where $m$ is the code of the major premise’s relation type, $n$ is the code for the minor premise, $s$ is the code for the deduced relation type, and $k$ is the number of the problem’s figure. Hence ‘aaa-1’, ‘eio-2’, and so on. By analogy each complementary syllogism is abbreviated as mns-$\overline{k}$, using the bar over the figure number to indicate the negation of the minor term in the deduced statement. Syllogistic deductions can also be displayed in alternative notation using the turnstile symbol or a tabular arrangement of formulas, as in:

$$BeA, BiC \vdash AoC$$ \hspace{1cm} (13)

or:

\[
\begin{align*}
BeA \\
BiC \\
\therefore \: AoC
\end{align*}
\]

for the pattern eio-2.
4 EXHAUSTIVE ANALYSIS OF ARISTOTELIAN PROBLEMS

Using the probability and optimization methods presented above, let us now analyze all possible Aristotelian problems with the structure given in Section 1.1 and the types of categorical relationships enumerated in Table 3. For this task we shall consider all 4 figures, and within each figure all 7 possible relationship types for each of the 2 premises. This gives $4 \times 7^2$ or 196 distinct Aristotelian problems. For each Aristotelian problem we set up 8 optimization problems: one problem to find the minimum feasible value $a_j$ and one to find the maximum feasible value $\beta_j$ of each of the 4 objective probabilities from Table 6 (describing the possible combinations of truth and falsity of the major and minor terms $A$ and $C$). This gives 8 optimization problems for each of the 196 Aristotelian problems, hence 1,568 optimization problems altogether. For each Aristotelian problem, we compare the results of its 8 optimization problems with the deductive criteria listed in Tables 7 and 8 for classical and complementary syllogism. Using each table of deductive criteria, we record a yes/no answer for whether the optimization results satisfy the criterion for each of the 7 types of categorical relationships; thus there are 14 yes/no answers for each of the 196 Aristotelian problems.

The results of this exhaustive analysis are displayed in Tables 9 through 12 in the following format. Each results table has two parts: part (a) which shows classical syllogisms, and part (b) which shows complementary syllogisms. Tabulated within each cell of an inner table are the codes for all of the valid deductions from the premises indexed by that row and column, using the figure indicated. These valid deductions are syllogisms.

For concreteness let us focus on Table 9 part (a) which describes the classical syllogisms in Aristotle’s first figure, whose major premise $AmB$ and minor premise $BnC$ have types $m$ and $n$. The third row (labeled ‘$e$’) indicates the major premise $AeB$, the first column (labeled ‘$a$’) indicates the minor premise $BaC$, and the solitary cell entry ‘$e$’ at this row and column says that the only valid deduction from these premises is $AeC$. That is, from the premises ‘$A$ belongs to no $B$’ or there are no $B$’ and ‘$B$ belongs to all $C$’ or there are no $C$’ there follows the conclusion ‘$A$ belongs to no $C$’, or there are no $C$’. This syllogistic pattern ‘$eae$’ is also known by the medieval name ‘Celarent’. (The vowels in the medieval names give codes for categorical relations in the same sequence as the $mn$-$k$ abbreviation. The consonants in the Medieval names also convey information, which is not essential to the analysis presented here.)

Remaining in the third row but moving over to the second column (labeled ‘$u$’) gives slightly different result. Here there are three cell entries $\ddot{e}$, $e$, and $o$. These indicate that all three statements $A\ddot{e}C$, $AeC$, and $AoC$ are valid deductions from the premises $AeB$ and $B\ddot{a}C$. That is, from the premises ‘$A$ belongs to no $B$’, or there are no $B$’ and ‘$B$ belongs to all $C$’ there follow three necessary conclusions:

$A\ddot{e}C$: ‘$A$ belongs to no $C$, and there are some $C$’

$AeC$: ‘Either $A$ belongs to no $C$, or there are no $C$’

$AoC$: ‘Either $A$ belongs to all $C$, or there are some $C$’
Table 10  Analysis of Aristotelian problems in the Second Figure. Tabulated is every type s for which the categorical statement AsC (classical) or AsC (complementary) is a valid deduction from the indexing premises BmA and BnC.

(a) Classical syllogism (AsC)  (b) Complementary syllogism (AsC)

| Minor premise (BnC) | Minor premise (BnC) |
|---------------------|---------------------|
| a  b  e  é  i  o  u | a  b  e  é  o  o  o |
| a  á  e  é  o  o  o | a  á  e  é  o  o  o |
| e  e  é  e  o  o  o | e  e  é  e  o  o  o |
| e  é  é  é  o  o  o | e  é  é  é  o  o  o |

Table 11  Analysis of Aristotelian problems in the Third Figure. Tabulated is every type s for which the categorical statement AsC (classical) or AsC (complementary) is a valid deduction from the indexing premises AmB and CnB.

(a) Classical syllogism (AsC)  (b) Complementary syllogism (AsC)

| Minor premise (CnB) | Minor premise (CnB) |
|---------------------|---------------------|
| a  b  e  é  i  o  u | a  b  e  é  i  i  i |
| a  á  e  é  i  i  i | a  á  e  é  i  i  i |
| e  o  o  o  o  o  o | e  o  o  o  o  o  o |
| é  o  o  o  o  o  o | é  o  o  o  o  o  o |
| i  i  i  i  i  i  i | i  i  i  i  i  i  i |
| o  o  o  o  o  o  o | o  o  o  o  o  o  o |
| u  u  u  u  u  u  u | u  u  u  u  u  u  u |

4.1 Classical Modes of Syllogism Reproduced

The results of probability-optimization analysis shown in Tables 9 through 12 reproduce the standard modes of Aristotelian syllogism, as described for example in [2] and [11]. Let us focus for a moment on the four relation types \{a, e, i, o\}, excluding the PARTICULAR-INTERMEDIATE type u and also excluding the EXISTENTIAL subtypes á and é of the UNIVERSAL-AFFIRMATIVE and UNIVERSAL-NEGATIVE types. Table 9(a) indicates the following four patterns of syllogism for Aristotle’s first figure (displayed here with their medieval names):

aaa-1 (Barbara)  aii-1 (Darii)  eae-1 (Celarent)  eio-1 (Ferio)

Remember, for each syllogism denoted mns-k, the type m of the major premise gives the row, the type n of the minor premise gives the column, and the type s of the deduced statement appears in the cell at that row and column in the specified results table; k is the number of the figure. Table 11(a) shows the following four modes of syllogism for the second figure:
(a) Classical syllogism ($AsC$)  

(b) Complementary syllogism ($As\bar{C}$)

Table 12  Analysis of Aristotelian problems in the Fourth Figure. Tabulated is every type $s$ for which the categorical statement $AsC$ (classical) or $As\bar{C}$ (complementary) is a valid deduction from the indexing premises $BmA$ and $CnB$.

Table 11 (a) shows the following four modes of syllogism for the third figure:

- aii-3 (Datisi)
- eio-3 (Ferison)
- iai-3 (Disamis)
- oao-3 (Bocardo)

Table 12 (a) shows the following three modes of syllogism for the fourth figure:

- aee-4 (Camenes)
- eio-4 (Fresison)
- iai-4 (Dimaris)

There are few notable absences from the modes of syllogism listed here: aai-3 (Darapti); eao-3 (Felapton); aai-4 (Bramantip); and eao-4 (Fesapo). These cases are addressed in the next section.

4.2 Existential Fallacies Revealed

Discipline about the different subtypes of universal statements prevents existential fallacies from contaminating our analysis of Aristotelian problems. For example, Table 11 (a) shows that there are no valid syllogisms ‘aai-3’ nor ‘eao-3’. In other words the patterns ‘Darapti’ and ‘Felapton’ are invalid using the material interpretation of their universal premises. However the patterns ‘Darapti’, ‘Darapti’, and ‘Darapti’ are all valid, as are ‘Felapton’, ‘Felapton’, and ‘Felapton’. Valid syllogism in these cases requires that the truth of the middle term $B$ is not impossible 

Likewise, Table 12 (a) shows no valid syllogisms ‘aai-4’ nor ‘eao-4’. In other words the patterns ‘Bramantip’ and ‘Fesapo’ are invalid using the material interpretation of their universal premises. However the following patterns are valid: ‘Brámantip’, ‘Brámántip’, ‘Fésapo’, and ‘Fésapo’. For the ‘Bramantip’ patterns, syllogism requires that the truth of the major term $A$ is not impossible a priori (i.e. they require the existence of $A$, which is the predicate of the major premise). For the ‘Fesapo’ patterns, syllogism requires that the truth of the middle term $B$ is not impossible a priori (i.e. they require the existence of $B$, which is the predicate of the minor premise).

4.3 Complementary Syllogisms Added

As an additional benefit, analysis according to the probability-optimization paradigm finds instances of complementary syllogism which were not heretofore appreciated. Complementary syllogisms can recover information contained within the premises that would otherwise be lost. For example, consider the problem in the first figure with premises $AiB$ and $BeC$. There is no classical syllogism in this case: no particular relation is required when $A$ is predicated on $C$. However there is a complementary syllogism when the negation of the minor term $C$ is used as the subject of the query. As Table 12 (b) shows, the deduction
AiC is a necessary consequence of the premises AiB and BeC. That is, there follows from the premises ‘A belongs to some B’ and ‘B belongs to no C, or there are no C’ the necessary consequence that ‘A belongs to some non-C’. This instance of complementary syllogism is abbreviated ‘iei-1’ (note the bar over the figure number; there is no valid classical syllogism iei-1). As it happens, all four patterns iei-1, iei-2, iei-3, and iei-4 represent valid deductions:

\[
\begin{align*}
AiB, BeC & \vdash Ai\overline{C} \quad (15) \\
BiA, BeC & \vdash Ai\overline{C} \quad (16) \\
AiB, CeB & \vdash Ai\overline{C} \quad (17) \\
BiA, CeB & \vdash Ai\overline{C} \quad (18)
\end{align*}
\]

Despite their different figures, each pair of premises here shares the identical (and solitary) consequence that ‘A belongs to some non-C’.

5 CONCLUSION

Using probability and optimization, it is possible to compute solutions to the logic problems that Aristotle described in Prior Analytics. The requisite calculations take advantage of two kinds of mappings: first between categorical statements and relations involving probabilities; and second between probability expressions and algebraic expressions. These mappings allow categorical statements to be translated to and from linear equalities and inequalities involving a few real-valued variables. To begin the analysis, Aristotelian premises are translated into systems of linear constraints. Numerical bounds are then computed on the feasible values of certain objective functions, subject to these constraints. These computed bounds reveal precisely which categorical statements are necessary consequences of the premises that were asserted. Every valid syllogism from an Aristotelian problem can be computed in this way.

There are several benefits to this probability-optimization formulation of Aristotle’s logic. First, the inference uses quite ordinary mathematical methods: symbolic probability inference (which is essentially arithmetic with polynomial expressions) and linear programming. It is straightforward to write computer programs to automate these calculations. Second, the results of probability-optimization analysis not only reproduce the known modes of Aristotelian syllogism; they also add new results. The analysis clarifies the role of existential import in certain patterns of syllogism (such as the incorrect modes aai-3, eao-3, aai-4, and eao-4). The analysis adds new appreciation for ‘complementary’ syllogisms: deduced consequences in which the subject is held to be false instead of true. For example, there is no valid classical syllogism from a PARTICULAR-AFFIRMATIVE major premise and a UNIVERSAL-NEGATIVE minor premise (whether material or existential) in any figure. But in every figure this combination of premise types leads to a valid complementary syllogism: the major term A must have the PARTICULAR-AFFIRMATIVE relation to the negation \( \overline{C} \) of the minor term (the shared deduction \( Ai\overline{C} \) says that ‘A belongs to some non-C’).

Finally, there are many ways to extend this computational framework to provide even more capabilities. In addition to computing the categorical relationships that are necessary consequences of the given premises, it is possible to compute the relationships that are merely potential consequences of the given premises, as well as those relationships that are inconsistent with the premises (using criteria modified from those given in Tables 7 and 8). Also, it is possible to use any numbers of terms and premises, with the terms distributed among the premises and query in an arbitrary fashion. It is not necessary to stick to Aristotle’s original restrictions of using two premises with a common middle term that is not included in the query. Moreover, it is possible to use probability models other than the basic one introduced here. As discussed further in [19], different probability models (with the full-joint probability distribution over categorical terms factored into several input tables) allow the use of other semantic types of conditional statements (such as subjunctives), at the expense of using nonlinear polynomials instead of linear ones.
REFERENCES

[1] Aristotle. Prior analytics. In Jonathan Barnes, editor, The Complete Works of Aristotle: The Revised Oxford Translation, volume 1, pages 39–113. Princeton University Press, 1984. Translated by A. J. Jenkinson.
[2] Aristotle. Prior Analytics. Hackett, 1989. Translated, with introduction, notes, and commentary by Robin Smith.
[3] Dimitri P. Bertsekas. Nonlinear Programming. Athena Scientific, Belmont, Massachusetts, 1995.
[4] George Boole. An Investigation of the Laws of Thought, on Which Are Founded the Mathematical Theories of Logic and Probabilities. Walton and Maaberly, London, 1854.
[5] Enrique Castillo and José Manuel Gutiérrez. A new method for efficient symbolic propagation in discrete Bayesian networks. Networks, 28:31–43, 1996.
[6] Goerge B. Dantzig, Alex Orden, and Philip Wolfe. The generalized simplex method for minimizing a linear form under linear constraints. Pacific Journal of Mathematics, 5(2):161–319, 1955.
[7] Adnan Darwiche. Modeling and Reasoning with Bayesian Networks. Cambridge University Press, 2009.
[8] Finn V. Jensen. An Introduction to Bayesian Networks. Springer, New York, 1997.
[9] Daphne Koller and Nir Friedman. Probabilistic Graphical Models: Principles and Techniques. MIT Press, Cambridge, MA, 2009.
[10] Andréi Nikolaevich Kolmogorov. Foundations of the Theory of Probability. Chelsea, New York, second edition, 1956. Translated to English by Nathan Morrison; originally published in German as Grundbegriffe der Wahrscheinlichkeitsrechnung in 1933.
[11] Henrik Lagerlund. Medieval theories of the syllogism. In Edward N. Zalta, editor, The Stanford Encyclopedia of Philosophy. Winter edition, 2012. URL http://plato.stanford.edu/archives/win2012/entries/medieval-syllogism/.
[12] David G. Leuenberger. Introduction to Linear and Nonlinear Programming. Addison-Wesley, Reading, Massachusetts, 1973.
[13] Han-Lin Li and Ching-Ter Chang. An approximate approach of global optimization for polynomial programming problems. European Journal of Operational Research, 107:625–632, 1998.
[14] Zhaoyu Li and Bruce D’Ambrosio. Efficient inference in Bayes networks as a combinatorial optimization problem. International Journal of Approximate Reasoning, 11(1):55–81, 1994.
[15] Abraham de Moivre. The Doctrine of Chances: or, a Method for Calculating the Probabilities of Events in Play. Woodfall, London, second edition, 1738.
[16] Joseph W. Norman. Analysis of Imprecise Decision Models. PhD thesis, Stanford University, 2002.
[17] Joseph W. Norman. Bounded global optimization for polynomial programming problems. Technical report at arXiv:1205.6459 [math.OC], May 2012.
[18] Joseph W. Norman. A tutorial introduction to the logic of parametric probability. Technical report at arXiv:1201.3142 [math.LO], May 2012.
[19] Joseph W. Norman. Probability distinguishes different types of conditional statements. Technical report at arXiv:1307.3802 [math.LO], February 2014.
[20] Judea Pearl. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufman, San Francisco, revised second edition, 1988.
[21] Hanif D. Sherali and Cihan H. Tuncbilek. A global optimization algorithm for polynomial programming problems using a reformulation-linearization technique. Journal of Global Optimization, 2:101–112, 1992.
[22] Sampath Srinivas and John S. Breese. ideal: A software package for analysis of influence diagrams. In Proceedings of the Sixth Conference on Uncertainty in Artificial Intelligence, pages 212–219, 1990. Available as arXiv:1304.1107 [cs.AI].