Extending Syllogistic Reasoning

Robert van Rooij

ILLC

Abstract. In this paper syllogistic logic is extended first to propositional logic, and then an interesting fragment of predicate logic that includes relations.

1 Introduction

Traditional logic, also known as term logic, is a loose term for the logical tradition that originated with Aristotle and survived until the advent of modern predicate logic in the late nineteenth century. Modern logicians used quite a number of arguments as to why traditional logic should be abandonded. First and foremost, the complaint is that traditional logic is *not rich enough* to account for mathematical reasoning, or to give a serious semantics of natural language. It is only a small fragment of predicate logic, which doesn't say anything about propositional logic, or multiple quantification. Russell (1900) blamed the traditional logical idea that every sentence is of subject-predicate form for giving sentences misleading logical forms. Due to the development of Montague grammar and especially Generalized Quantifier Theory in the 1960s-1980s the misleading form thesis of early proponents of modern logic is not a mainstream position anymore, and analyzing sentences in subject-predicate form is completely accepted again.

In this paper I will first quickly discuss traditional Aristotelian syllogistics, and how to extend it (also semantically) with negative and singular terms. Afterwards I will discuss how propositional logic can be seen as an extension of Aristotelian syllogistics. Thus, in distinction with polish logicians like Lukasiewicz and others, I won't assume that to understand traditional logic we have to presuppose propositional logic, but instead formulate propositional logic by presupposing syllogistic reasoning. Afterwards I will follow (the main ideas, though not the details of) Sommers (1982) and his followers in showing how traditional logic can be extended so as to even account for inferences involving multiple quantification that almost all modern textbooks claim is beyond the reach of traditional logic: A woman is loved by every man, thus Every man loves a woman.

2 From syllogistics to propositional logic

Syllogisms are arguments in which a categorical sentence is derived as conclusion from two categorical sentences as premisses. As is well-known, a categorical sentence is always of one of four kinds: *a*-type ('All men are mortal'), *i*-type ('Some men are philosophers'), *e*-type ('No philosophers are rich'), or *o*-type ('Some men are not philosophers'). A rather standard proof theory **SYL** for syllogistic reasoning with negative terms (if P is a term, $\overline{P}$ is a term as well) which only makes use of a and i propositions can make use of the fact whether a term occurs distributively, or monotone decreasingly/negatively within a sentence, or not. Denoting a distributed term by $-$ and an undistributed term by $+$, the following follows at once: S^-aP^+, S^+iP^+, S^-eP^-, and S^+oP^-, which we might think of now as a *syntactic* characterisation. The proof system then consists of the following set of axioms and rules (with sentence-negation '¬' defined as follows: $\neg(SaP) \stackrel{def}{=} SoP$, $\neg(SiP) \stackrel{def}{=} SeP$, $\neg(SeP) \stackrel{def}{=} SiP$, and $\neg(SoP) \stackrel{def}{=} SaP$):

(1) $MaP, \Gamma(M)^+ \vdash \Gamma(P)$	Dictum de Omni,
where $\Gamma(M)^+$ is a sentence where M occurs undistributed.	
(2) $\vdash TaT$	Law of identity
(3) $\vdash T \equiv \overline{\overline{T}}$[1]	Double negation
(4) $Sa\overline{P} \vdash Pa\overline{S}$	Contraposition
(5) $\Gamma, \neg\phi \vdash \psi, \neg\psi \quad \Rightarrow \quad \Gamma \vdash \phi$	Reductio per impossible[2]
(f) $\vdash \neg(Ta\overline{T})$ (i.e. $\vdash TiT$)	Existential Import

We will now slightly extend syllogistic reasoning in some seemingly innocent ways. First, we add a distinguished 'transcendental' term '$\top$' to our language, standing for something like 'entity'. Obviously,

[1] By this I really mean $\vdash Ta\overline{\overline{T}}$ and $\vdash \overline{\overline{T}}aT$.
[2] I will always assume that $\phi_1, \phi_2 \vdash \psi$ iff $\phi_2, \phi_1 \vdash \psi$.

the sentence $Sa\top$ should always come out true for each term S. To reflect this, we will add this sentence as an axiom to **SYL**. But adding $\top$ as an arbitrary term to our language gives rise to a complication once we accept existential import for *all* terms, including negative ones: for negative term $\overline{\top}$ existential import is unacceptable. One way to get rid of this problem is to restrict existential import to *positive* categorical terms only. Next, we add *singular*, or *individual* terms to our language. In contrast to in standard predicate logic, we will not assume that there is a type-difference between individual terms and standard predicates. Following Leibniz (1966b) and Sommers (1982), we will assume, instead, that for singular propositions, a and i propositions coincide, just like e and o propositions. Thus, 'Plato sleeps' is represented by a sentence like 'PaS', which is equivalent with 'PiS'. Finally, we will add a rule (due to Sherperdson, 1956) saying what to do with empty terms. We will denote the system consisting of (1), (2), (3), (4), (5) together with the following four rules by $\mathbf{SYL}^+$.

(6) $\vdash \neg(Ta\overline{T})$, for all *positive* categorical terms T
(7) $\vdash Sa\top$
(8) for all singular terms I and terms P: $IiP \dashv\vdash IaP$.
(9) $\overline{S}aS \vdash \overline{S}aP$ (for any P)

To think of **propositional logic** in syllogistic terms, we will allow for 0-ary predicates as well. We will assume that if S and P are terms of the same arity, SaP, SiP etc. are formulas of arity 0. Moreover, if S and P are 1-ary predicates, and ϕ a 0-ary predicate, something like $(Si\overline{P})a\phi$ will be 0-ary predicates as well. Starting with a non-empty domain D, we will (extensionally) interpret n-ary terms as subsets of D^n (thus $D^0 = \{\langle\rangle\}$). If S and P are 0-ary or 1-ary terms, the categorical sentences are interpreted as follows: $V_M(SaP) = \{\langle\rangle : V_M(S) \subseteq V_M(P)\}$, $V_M(SiP) = \{\langle\rangle : V_M(S) \cap V_M(P) \neq \emptyset\}$, and the e and o-propositions as negations of them. It is easy to see that all types of complex propositional formulas can be expressed in categorical terms ($[\phi]a[\psi] \equiv$ '$\phi \rightarrow \psi$', $[\phi]i[\psi] \equiv$ '$\phi \wedge \psi$', $[\phi]e[\phi] \equiv$ '$\neg\phi$', and $[[\phi]e[\phi]]a[\psi] \equiv$ '$\phi \vee \psi$'), and receive the correct interpretation.

Let us see now how things work from a proof-theoretic point of view. To implement the above suggestions, we will add to $\mathbf{SYL}^+$ the following three ideas: (i) 0-ary terms don't allow for existential import, (ii) $\top^0$ is a singular term, and (iii) P^0 is equal to $\top^0 i P^0$. The first idea is implemented with the help of axiom (6) by stipulating that 0-ary terms are not categorical.

(10) 0-ary terms are no categorical and $\top^0$ is a singular term.
(11) $P^0 \dashv\vdash \top^0 i P^0$.

We will denote the system $\mathbf{SYL}^+$ together with (10) and (11) by $\mathbf{SYL}^+\mathbf{PL}$. The claim of this section of the paper is that this system can indeed account for all inferences in proposotional logic. It is almost immediately clear that Modus Ponens, Modus Tollens, the Hypothetical Syllogism, and the Disjunctive Syllogism can be thought of as ordinary valid syllogisms of the form Barbara, Camestres, Barbara, and Camestres, respectively. Also other 'monotonicity-inferences' follow immediately from the Dictum de Omni.

To show that $\mathbf{SYL}^+\mathbf{PL}$ is enough, we will show that also the following hold '$p \vdash p \vee p$', '$p \vee p \vdash p$', '$p \vee q \vdash q \vee p$', and '$p \rightarrow q \vdash (r \rightarrow p) \rightarrow (r \rightarrow q)$'. The reason is that we can axiomatize propositional logic by these four rules, together with modus ponens (cf. Goodstein, 1963, chapter 4). We can conclude that propositional logic *follows* from syllogistic logic if we (i) make the natural assumption that propositions are 0-ary terms, (ii) assume that $\top^0$ is a singular term, and (ii) treat singular terms as proposed by Leibniz.

It is important to realize that we represent $p \vee q$ by $\overline{p} a q$. '$p \vdash p \vee p$' immediately follows from the validity of $\overline{p} a \top$, the equivalence $p \equiv \top a p$ and the Dictum. As for disjunctive elimination, note that because $p \vee p \equiv \overline{p} a p$, we can conclude by (9) to $\overline{p} a \bot$. Via contraposition and double negation we derive $\top^0 a p$. Because $\top^0$ is a singular term (rule 10)) it follows by (8) that $\top^0 i p$, and thus via (11) that p. So we have validated $p \vee p \vdash p$. It is easier to validate '$p \vee q \vdash q \vee p$': it immediately follows by contraposition and double negation.

Notice that '$p \rightarrow q \vdash (r \rightarrow p) \rightarrow (r \rightarrow q)$' follows from the Dictum if we could make use of the *deduction theorem*: $\Gamma, P \vdash Q \Rightarrow \Gamma \vdash PaQ$ (for this to make sense, P and Q have to be 0-ary terms, obviously). But this deduction theorem follows from $\mathbf{SYL}^+\mathbf{PL}$: Assume $\Gamma, P \vdash Q$ and assume towards contradiction that $\neg(PaQ)$. This latter formula is equivalent to $Pi\overline{Q}$. We have seen above that $\top iP$ can be derived from $Pi\overline{Q}$ in $\mathbf{SYL}^+$. Because $\neg(PaQ) \vdash P$, it follows from the assumption $\Gamma, P \vdash Q$ that $\Gamma, \neg(PaQ) \vdash Q$. From this we derive $\top aQ$, and together with the validity of $Pa\top$ we derive via the Dictum that PaQ. Thus, from Γ and assuming $\neg(PaQ)$ we derive a contradiction: $\Gamma, \neg(PaQ) \vdash \neg(PaQ), PaQ$. By the reductio-rule we conclude that $\Gamma \vdash PaQ$.

3 Relations

Traditional logicians were well aware of an important limitation of syllogistic reasoning. In fact, already Aristotle recognized that the so-called 'oblique' terms (i.e. ones expressed in a grammatical case other than the nominative) gives rise to inferences that cannot be expressed in the ordinary categorical syllogistic. An example used by Aristotle is 'Wisdom is knowledge, Of the good there is wisdom, thus, Of the good there is knowledge'. This is intuitively a valid inference, but it, or its re-wording, is not syllogistically valid: 'All wisdom is knowledge, Every good thing is object of some wisdom, thus, Every good thing is object to some knowledge'. The re-wording shows that we are dealing with a binary relation here: 'is object of'. Aristotle didn't know how to deal with such inferences, but he noted that *if* there is a syllogism containing oblique terms, there must be a corresponding syllogism in which the term is put back into the nominative case.

It is generally assumed that in traditional formal logic there is no scope for relations. Thus — or so the Frege-Russell argument goes — it can be used neither to formalize natural language, nor to formalize mathematics. What we need, – or so Frege and Russell argued – is a whole new logic. But the Frege-Russell argument is only partly valid: instead of inventing a whole new logic, we might as well just extend the traditional fragment. As far as semantics is concerned, it is well-known how to work with relations. The main challenge, however, is to embed relations into the traditional theory, and to extend the inference rules such that also proofs can be handled that crucially involve relations. As it turns out, part of this work has already been done by medieval logicians, and also by people like Leibniz and de

Morgan when they were extending syllogistic reasoning such that it could account for inferences involving oblique terms, or relations.

We want to combine relations with monadic terms by means of the 'connectives' a, i, e, and o to generate new terms. This will just be a generalization of what we did before: When we combine a monatic term P with a monadic term S (and connective 'a', for instance), what results is a new 0-ary term like SaP. The generalization is now straightforward: if we combine an n-ary term/relation R with a monadic term S (and connective 'a', for instance), what results is a new $n-1$-ary term $(S^1aR^n)^{n-1}$. The semantics should now determine what such new terms denote. The $n-1$ ary term $(S^1aR^n)^{n-1}$, for instance, would denote $\{\langle d_1, ..., d_{n-1}\rangle : V_M(S) \subseteq \{d_n \in D : \langle d_1, ..., d_n\rangle \in V_M(R^n)\}$.[3]

Now we can represent the natural reading of a sentence like 'Every man loves a woman' as $Ma(WiL^2)$. The meaning of this formula is calculated as follows:

$$V_M(Ma(WiL^2)) = \{\langle\rangle : I(M) \subseteq \{d \in D : \langle d\rangle \in V_M(WiL^2)\},$$
$$\text{with } V_M(WiL^2) = \{d_1 : I(W) \cap \{d_2 \in D : \langle d_1, d_2\rangle \in I(L^2)\} \neq \emptyset\}.$$

To represent the sentence 'There is woman who is loved by every man' we will follow medieval practice and make use of the *passive form* of 'love': *being loved by*. For every binary relation R, we represent its passive form by $R^\cup$, interpreted as indicated above: $V_M(R^\cup) = \{\langle d_2, d_1\rangle : \langle d_1, d_2\rangle \in I_M(R)\}$.[4] Now we represent 'There is woman who is loved by every man' as follows: $Wi(MaL^\cup)$. This sentence is true iff:

$$V_M(Wi(MaL^\cup)) = \{\langle\rangle : I(W) \cap \{d \in D : \langle d\rangle \in V_M(MaL^\cup)\} \neq \emptyset\},$$
$$\text{with } V_M(MaL^\cup) = \{\langle d_1\rangle : I(W) \subseteq \{d_2 \in D : \langle d_1, d_2\rangle \in V_M(L^\cup)\}\}.$$

[3] This by itself is not general enough. To express the mathematical property of *density*, for instance, we need to be able to combine a binary relation with a ternary relation.

[4] Of course, the active-passive transformation only works for binary relations. For more-*ary* relations it fails. Fortunately, we can do something similar here, making use of some functions introduced by Quine in his proof that variables are not essential for first-order predicate logic. We won't go into this here.

Both truth conditions are intuitively correct, and correspond with those of the two first-order formulas $\forall x[M(x) \rightarrow \exists y[W(y) \wedge L(x,y)]]$ and $\exists y[W(y) \wedge \forall x[M(x) \rightarrow \wedge L(x,y)]]$, respectively.

What we want to know, however, is how we can reason with sentences that involve relations. Let us first look at the re-wording of Aristotle's example: 'All wisdom is knowledge, Every good thing is object of some wisdom, thus, Every good thing is object of some knowledge'. If we translate this into our language this becomes WaK, $Ga(WiR) \vdash Ga(KiR)$, with 'R' standing for 'is object of'. But now observe that we immediately predict that this inference is valid by means of the Dictum de Omni, if we can assume that 'W' occurs positively in '$Ga(WiR)$'! We can mechanically determine that this is indeed the case.[5] First, we say that if a sentence occurs out of context, the sentence occurs positively. From this, we determine the positive and negative occurrences of other terms as follows:

$\overline{P}$ occurs positively in Γ iff P occurs negatively in Γ.
If (SaR) occurs positively in Γ, then $S^{-}aR^{+}$, otherwise $S^{+}aR^{-}$.
If (SiR) occurs positively in Γ, then $S^{+}iR^{+}$, otherwise $S^{-}iR^{-}$.

Thus, first we assume that '$Ga(WiR)$' occurs positively. From this it follows that the 'WiR' occurs positively, from which it follows in turn that 'W' occurs positvely. Assuming that 'WaK' is true, the Dictum allows us to substitute K for W in $Ga(WiR)$, resulting in the desired conclusion: $Ga(KiR)$. Something very similar was done by medieval logicians (cf. Buridan, 1976).

As far as I know, this is how far medieval logicians went. But it is not far enough. Here is one classical example discussed by Leibniz (1966a): 'Every thing which is a painting is an art (or shorter, painting is an art), thus everyone who learns a thing which is a painting learns a thing which is an art' (or shorter: everyone who learns painting learns an art). Formally: $PaA \vdash (PiL^2)a(AiL^2)$. Semantically it is immediately clear that the conclusion follows from the premiss. But the challenge for traditional logic was to account for this inference in a proof-theoretic way. As Leibniz already observed, we can account for this inference in traditional logic if we add the extra

[5] Cf. Sommers (1982) and van Benthem (1983).

(and tautological) premiss 'Everybody who learns a thing which is a painting learns a thing which is a painting', i.e. $(PiL^2)a(PiL^2)$. Now $(PiL^2)a(AiL^2)$ follows from PaA and $(PiL^2)a(PiL^2)$ by means of the *Dictum the Omni*, because by our above rules the second occurrence of 'P' in $(PiL^2)a(PiL^2)$ occurs in a monotone increasing position.[6]

To account for other inferences we need to assume more than just a tautological premiss. For instance, we cannot yet account for the inference from 'There is a woman who is loved by every man' represented by $Wi(MaL^{\cup})$ to 'Every man loves a woman' represented by $Ma(WiL^2)$. In standard predicate logic one can easily prove the equivalence of $\forall x[M(x) \rightarrow \forall y[W(y) \rightarrow R(x,y)]]$ with $\forall y[W(y) \rightarrow \forall x[M(x) \rightarrow R(x,y)]]$. But in contrast to predicate logic, our system demands that the sequence of arguments of a relational term is in accordance with the scope order of the associated terms. Because of this, we have to use something like passive transformation to express 'reverse scope'. Thus, to reason with relations, we have to say which rules passive transformation obeys. To do so, we will follow medieval logicians such as Buridan and enrich our system **SYL+PL** with the rule of *oblique conversion* (12), and the passive rule (13) (for binary relations R, and predicates S and O), and the more general formulation of the Dictum in (1′):

(12) Oblique Conversion: $Sa(OaR) \equiv Oa(SaR^{\cup})$
from 'every man loves every woman' we infer that 'every woman is loved by every man' and the converse of this.[7]

(13) Double passive: $R^{\cup\cup} \equiv R$

(1′) **Dictum de Omni**: $\Gamma(MaR)^+, \Theta(M)^+ \vdash \Gamma(\Theta(R))$

Let us see how we can account for the inference from $Wi(MaL^{\cup})$ to $Ma(WiL)$:

[6] In terms of our framework, Leibniz assumed that all terms being part of the predicate within sentences of the form SaP and SiP occur positively. But this is not necessarily the case once we allow for all types of complex terms: P doesn't occur positively in $Sa(PaR)$, for instance. On Leibniz's assumption, some invalid inferences can be derived (cf. Sanches, 1991). These invalid inferences are blocked by our more fine-grained calculation of monotonicity marking.

[7] From this rule we can derive that $Si(OiR) \equiv Oi(SiR^{\cup})$ is also valid. And that is correct: the sentence 'A man loves a woman' is truth-conditionally equivalent to 'A woman is loved by a man'.

1.	$Wi(MaL^{\cup})$	premiss
2.	$(MaL^{\cup})a(MaL^{\cup})$	a tautology (everyone loved by every man is loved by every man)
3.	$Ma((MaL^{\cup})aL^{\cup\cup})$	from 2 and (12) ($S = (MaL^{\cup})$ and $S' = M$)
4.	$Ma((MaL^{\cup})aL)$	by 3 and (13), substitution of L for $L^{\cup\cup}$
5.	$Ma(WiL)$	by 1 and 4, by the Dictum de Omni $(1')$[8]

Many other examples can be accounted for in this way as well.

4 Decidable Fragment of Predicate Logic

In this paper I have argued with Sommers (1982) and others that it is possible to think of logic as an extension of traditional syllogistics. Singular propositions straightforwardly fit into the system, and the syllogistics can easily be extended to account for propositional reasoning, and even for reasoning with relational terms as well. Though we used neither a distinguished relation of identity, nor make use of variables to allow for binding, we have seen that we could nevertheless adequately express many types of sentences for which these tools are normally used in predicate logic. This doesn't mean that our extended syllogistics is as expressive as standard first-order logic. What we cannot (yet) represent are sentences which *crucially* involve variables/pronouns and/or identity. Some examples for which these tools are crucial are the following: 'Every/some man loves himself', 'All parents love their children', 'Everybody loves somebody else', 'There is a unique king of France', and 'At least 3 men are sick'. As it turns out, we can extend our language with *numerical quantifiers* (cf. Murphee, 1997) and Quinean predicate functors to solve these problems, but these extensions have their price. Pratt-Hartmann (2009) shows that syllogistic systems with numerical quantifiers cannot be axiomatized, and adding Quinean predicate functors brings us over the decidability border. In the formal system we have so far, the sequence of arguments of a relational term will always be in accordance with the scope order of the associated terms. Thinking of this system as a fragment of FOL means that this logic has a very interesting property. Following an earlier suggestion of Quine, Purdy

[8] With $\Gamma = Ma, \Theta = Wi, R = L$, and $M = MaL^{\cup}$.

(1998) shows that the limits of decidability are indeed close to the limits of what can be expressed in (our fragment of) traditional formal logic. One can argue that (our fragment of) traditional formal logic is thus the *natural* part of logic. Indeed, a small contingent of modern logicians (e.g. Suppes, Sommers, van Benthem, Sanchez, Purdy, Pratt-Hartmann, Moss) try to develop a system of *natural logic* which is very close to what we have done in this paper in that it crucialy makes use of monotonicity (or the Dictum de Omni) and is essentially variable-free.

References

1. Benthem, J. van (1983), 'A linguistic turn: New directions in logic', in R. Marcus et al. (eds.), *Logic, Methodology and Philosophy of Science*, Salzburg, pp. 205-240.
2. Burdidan, J. (1976), Tractatus de Consequentiis', in Hubien, H. (ed.) *Johannis Buridani tractatus de consequentiis*, critical edition, Publications universitaires, Louvain.
3. Goodstein, R.L. (1963), *Boolean Algebra*, Pergamon Press, London.
4. Leibniz, G. (1966a), 'A specimen of a demonstrated inference from the direct to the oblique', in Parkinson (ed.), *Leibniz. Logical Papers*, pp. 88-89.
5. Leibniz, G. (1966b), 'A paper on 'some logical difficulties'', in Parkinson (ed.), *Leibniz. Logical Papers*, pp. 115-121.
6. Lukasiewicz, J. (1957), *Aristotle's Sollogistic from the standpoint of Modern Formal Logic*, Garland Publishers, New York.
7. Lyndon, R.C. (1959), 'Properties preserved under homomorphism', *Pacific Journal of Mathematics*, **9**: 142-154.
8. McIntosh, C. (1982), 'Appendix F', in F. Sommers (1982), *The Logic of Natural Language*, Clarendon Press, Oxford, pp. 387-425.
9. Murphee, W.A. (1997), 'The numerical syllogism and existential presupposition', *Notre Dame Journal of Formal Logic*, **38**: 49-64.
10. Pratt-Hartmann, I. (2009), 'No syllogisms for the numerical syllogistic', in O. Grumberg et al. (eds.), *Francez Festschrifft, LNCS 5553*, Springer, Berlin, pp. 192-203.
11. Pratt-Hartmann, I and L. Moss (to appear), 'Logics for the Relational Syllogistics', *Review of Symbolic Logic*.
12. Purdy, W. (1996), 'Fluted formulas and the limits of decidability', *Journal of Symbolic Logic*, **61**: 608-620.
13. Quine, W.V.O. (1976), 'Algebraic logic and predicate functors', pp. 283-307 in
14. Russell, B. (1900), *A critical exposition of the Philosophy of Leibniz*, Cambridge University Press, Oxford.
15. Sanchez, V. (1991), *Studies on Natural Logic and Categorial Grammar*, PhD thesis, Universiteit van Amsterdam.
16. Sherpherdson, J. (1956), 'On the interpretation of Aristotelian syllogistic', *Journal of Symbolic Logic*, **21**: 131-147.
17. Sommers, F. (1982), *The Logic of Natural Language*, Clarendon Press, Oxford.