

FULL-LENGTH REPRINT – differs from the original in layout but not in contents

Janowicz M., Ochnio L., Chmielewski L.J., Orłowski A.,
Application of automated theorem-proving to philosophical thought: Spinoza's *Ethics*.
In K. Kim and N. Joukov (Eds.), *Information Science and Applications 2017: Proc. ICISA 2017*,
volume 424 of *Lecture Notes in Electrical Engineering*, pages 512-518,
Macau, China, 20-23 Mar 2017. doi:10.1007/978-981-10-4154-9_59
The copyright owner of the paper is Springer.
The final publication is available at Springer via https://doi.org/10.1007/978-981-10-4154-9_59.

Application of automated theorem-proving to philosophical thought: Spinoza's *Ethics*

M. Janowicz, L. Ochnio, L.J. Chmielewski, and A. Orłowski

Faculty of Applied Informatics and Mathematics (WZIM),
Warsaw University of Life Sciences (SGGW), Poland
ul. Nowoursynowska 159, 02-775 Warsaw, Poland
{maciej_janowicz,luiza_ochnio,arkadiusz_orlowski}@sggw.pl
<http://www.wzim.sggw.pl>

Abstract. We have applied the automatic theorem-prover Prover9 to prove first eleven theorems of Part 1 of Benedict de Spinoza's *Ethics*. We have used a previous formalization of that segment developed by Blum and Malinovich. We have found Prover9 to be very efficient, providing proofs in tens of milliseconds. It appears that the only, but fundamental, limitation for testing philosophical reasoning is related to the difficulty of unique formalization.

Keywords: automated reasoning; theorem proving; Spinoza's philosophy

1 Introduction

As one might expect, automated reasoning has found its most important applications in the fields of mathematical logic, foundations of mathematics, and computer science. To convince oneself about that, it is sufficient to browse recent issues of the leading magazine – the Journal of Automated Reasoning. However, one can readily conceive other applications. In fact, automatic theorem-provers can also be used to any field of human thought which admits formalization. By “formalization” we mean here (to avoid lengthy discussion) simply a procedure which results with a set of terms, definitions, and axioms added to the classical first-order logic with the inference rules inherited from the latter.

In particular, it should be possible to apply automatic reasoning programs to prove, or find counterexamples to, statements made by philosophers. This is especially so in the context of what is known as “analytic philosophy” (as contrasted with “continental philosophy”, usually associated with (post-)Kantian, (post-)Hegelian, or phenomenological/existential schools of thought). This is not because the “analytic” way of philosophical considerations is “better” or “more rigorous”, but because it is more oriented towards solving specific problems and closer, it seems, to the way the mathematical logic works.

We would like to argue that the philosophical thought has always been and remains relevant in everyday life as it can generate consistent worldviews and help us to take more rational and more moral decisions about ourselves. We also

argue that automatic theorem proving can greatly help us to select a life philosophy which is logically consistent and morally viable. This is because automatic theorem proving can radically amend the time effectiveness and reduces costs in every situation where it can be applied.

As our first attempt to fulfill the program outlined above, we have decided to try and employ a well-known automated theorem-proving program Prover9 [1] written by William McCune to a part of *Ethics*, a fundamental work by Benedict de Spinoza [2]. This has been done for three reasons. Firstly, Spinoza belongs to the venerable grandsires of all the modern thought, invoked both by Hegel and the analytic school. Secondly, his *opus magnum*, the *Ethics*, has the structure very favorable from our point of view: it is made of definitions, axioms, and theorems. Thirdly, the hardest work, namely, the formalization of a segment of *Ethics*, has already been performed several times (see, e.g., [3,4]). Below, we shall use the formalization provided by Blum and Malinovich [4], who have also provided formal (human-made) derivations of the theorems. As far as we are aware, this is the first attempt to apply automated theorem proving to philosophy; in a sense, it constitutes the first real attempt to fulfill Leibniz's dream about such an application.

The main body of this work is organized as follows. In Section 2 we describe the Blum-Malinovich formalization of the relevant part of Spinoza's work. In Section 3 we list the theorems proved by both Blum and Malinovich and by Prover9 as well as an example of an output file from Prover9. Section 4 contains some concluding remarks.

2 Blum-Malinovich formalization of Spinoza's *Ethics*

In [4] the terms, necessary for the formalization of *Ethics* have been introduced (see Tab. 1). The list appears to be rather lengthy, but all the concepts there are necessary to characterize the Spinoza thought. Actually, additional terms, not listed here, would have been introduced were we to check the whole body of *Ethics*. For instance, we have found the relations "is conceived by" and "having more reality than" painfully lacking in the list.

What is more, let us provide the list of Spinoza's definitions in the Blum-Malinovich settings. We will use the notation of Prover9, that is, all the entries look almost exactly the same as in the Prover9 input file. "Almost" means that \rightarrow represents the sign " $-$ " followed by " $>$ " and \leftrightarrow by " $<$ " followed by " $-$ " and " $>$ " in the input file. We believe the notation is self-explanatory, except that all the formulae have to be ended with a dot.

1. $\text{all } x \ (C(x, x) \leftrightarrow N(x)).$
2. $\text{all } x \ (Kf(x) \leftrightarrow \exists y \ (K(x, y) \ \& \ L(y, x) \ \& \ \neg(x = y))).$
3. $\text{all } x \ (S(x) \leftrightarrow I(x, x)).$
4. $\text{all } x \ (S(x) \leftrightarrow D(x, x)).$
5. $\text{all } x \ (S(x) \rightarrow \neg(\exists y \ (D(x, y) \ \& \ \neg(y = x)))).$
6. $\text{all } x \ \text{all } y \ (A(x, y) \leftrightarrow (S(y) \ \& \ E(x, y))).$
7. $\text{all } x \ (\exists y \ (S(x) \rightarrow A(y, x))).$

Table 1. List of terms in the Blum-Malinovich formalization of *Ethics*.

Term	Meaning
$A(x, y)$	x is an attribute of y
$C(x, y)$	x causes y
$D(x, y)$	x depends on y
$Et(x)$	x is eternal
$E(x, y)$	x is essence of y
$F(x)$	x is finite
$H(x)$	x is absolutely infinite
$I(x, y)$	x is contained in y
$Kf(x)$	x is finite after its kind
$K(x, y)$	x and y are of the same kind
$L(x, y)$	x limits y
$M(x, y)$	x is a mode of y
$N(x)$	x has necessary existence
$P(x, y)$	x is prior to y
$Q(x)$	x is free
$S(x)$	x is a substance
$T(x, y)$	x is the effect of y
$U(x, y)$	x knows y
$W(x, y, z)$	x and y have z in common

8. all x all y ($M(x, y) \leftrightarrow (S(y) \& I(x, y) \& (\neg(x = y)) \& D(x, y))$).
9. all x ($G(x) \rightarrow (S(x) \& H(x))$).
10. all x ($Q(x) \leftrightarrow (N(x) \& C(x, x))$).
11. all x ($Et(x) \leftrightarrow N(x)$).

Spinoza's axioms have been written as follows:

1. all x ($I(x, x) \mid \exists y (\neg(x = y) \& I(x, y))$).
2. all x all y ($(\neg(x = y) \& \neg D(x, y)) \rightarrow D(x, x)$).
3. all x ($\exists y (C(x, y) \rightarrow \exists z (T(z, x))) \& \forall x (\exists y (T(x, y) \rightarrow \exists z (C(z, x))))$.
4. all x all y ($C(x, y) \rightarrow \forall z (U(z, x) \rightarrow U(z, y))$).
5. all x all y ($(\neg \exists z (W(x, y, z)) \rightarrow (\exists v (U(v, x) \& \neg(U(v, y))) \& (\exists v (U(v, y) \& \neg(U(v, x))) \& \neg(D(x, y)) \& \neg(D(y, x))))$).
6. all x ($N(x) \rightarrow \forall y (E(y, x) \rightarrow \forall z (A(z, y) \rightarrow \exists u (z = u)))$).

One additional axiom schema has been introduced by Spinoza ("A true idea must correspond with its ideate or object") which has not used here. There have been, however, several concepts which Spinoza most likely considered as too obvious to define or characterize in the axiomatic form. And yet, their explicit introduction is necessary in the formal proofs. That fact have compelled the authors of [4] to introduce the following additional "postulates":

1. all x all y ($(\neg(x = y) \& D(x, y)) \rightarrow P(y, x)$).
2. all x ($D(x, x) \leftrightarrow C(x, x)$).
3. all x all y ($(D(x, y) \mid D(y, x)) \leftrightarrow \exists w (W(x, y, w))$).

4. all x all y all u all v ((E(u, x) & E(v, y)) → ((x = y) ↔ (u = v))).
5. all x (Q(x) ↔ (¬(exists y (L(y, x))))).

We have had Prover9 found proofs for the first eight- and the eleventh theorem from *Ethics*:

1. all x all y ((S(x) & M(y, x)) → P(x, y)).
2. all x all y ((S(x) & S(y) & exists z exists v (A(z, x) & A(v, y) & (z != v))) → (¬(exists w (W(x, y, w)))).
3. all x all y ((¬(exists z (W(x, y, z)))) → ((¬C(x, y)) & (¬C(y, x)))).
4. all x all y ((S(x) & S(y) & (x != y)) → ((exists z (A(z, x) & ¬A(z, y))) | (exists v (M(v, x) & ¬M(v, y))))).
5. all x all y ((S(x) & S(y) & (¬(x=y))) → (¬(exists z (W(x, y, z))))).
6. all x all y ((S(x) & S(y) & (¬(x=y))) → (¬(C(x, y) | C(y, x)))).
7. all x all y (S(x) → N(x)).
8. all x (S(x) → ((¬Kf(x)) & (¬(exists y (L(y, x)))))).
9. all x (G(x) → N(x)).

As for the Theorem 9, “The more reality or being a thing has the greater the number of its attributes.”, we believe it is actually a definition of the relation $R(x, y)$, “ x have more reality or being than y ”. Prover9 has demonstrated, however, a proposition which remotely resembles Theorem 9 without using a new definition:

- all x all y ((all z (E(z, x) → E(z, y))) → (all v (A(v, x) → A(v, y)))).

Similarly, it is difficult to convincingly formalize Theorem 10: “Each particular attribute of the one substance must be conceived through itself.”

We have played with several formulae which, again, very remotely resemble the meaning of the above proposition. Among other things, Prover9 has proved the following:

- all x all y ((S(y) & A(x, y)) → (I(x, x))).

The difficulties in formalization of Theorems 9 and 10 have probably been the reason of their omission in [4].

3 Example of the output from Prover9

After some testing, Prover9 has been used by us as a black box. We knew in advance that all the propositions presented to it have already been proved in [4].

The output files from Prover9 are rather lengthy as they provide quite detailed information about the proof search. In this section we just show the text of final part of the output file containing proof of the famous Theorem 11.

We do this mainly to demonstrate beyond any doubt that the Prover9 has really performed the required action!

```
=====
% Proof 1 at 0.02 (+ 0.00) seconds.
% Length of proof is 15.
% Level of proof is 4.
% Maximum clause weight is 6.000.
% Given clauses 45.

1 (all x (C(x,x) ↔ N(x))) # label(non_clause). [assumption].
4 (all x (S(x) ↔ D(x,x))) # label(non_clause). [assumption].
9 (all x (G(x) → S(x) & H(x))) # label(non_clause). [assumption].
20 (all x (D(x,x) ↔ C(x,x))) # label(non_clause). [assumption].
24 (all x (G(x) → N(x))) # label(non_clause) # label(goal). [goal].
31 ~S(x) | D(x,x). [clausify(4)].
39 ~G(x) | S(x). [clausify(9)].
40 G(c1). [deny(24)].
43 ~G(x) | D(x,x). [resolve(39,b,31,a)].
64 ~C(x,x) | N(x). [clausify(1)].
77 ~D(x,x) | C(x,x). [clausify(20)].
81 ~N(c1). [deny(24)].
108 D(c1,c1). [resolve(43,a,40,a)].
128 ~C(c1,c1). [ur(64,b,81,a)].
135 $F. [resolve(108,a,77,a),unit_del(a,128)].

===== end of proof =====

===== STATISTICS =====

Given=45. Generated=95. Kept=70. proofs=1.
Usable=45. Sos=25. Demods=0. Limbo=0, Disabled=99. Hints=0.
Kept_by_rule=0, Deleted_by_rule=0.
Forward_subsumed=24. Back_subsumed=0.
Sos_limit_deleted=0. Sos_displaced=0. Sos_removed=0.
New_demodulators=0 (0 lex), Back_demodulated=0. Back_unit_deleted=0.
Demod_attempts=0. Demod_rewrites=0.
Res_instance_prunes=0. Para_instance_prunes=0. Basic_paramod_prunes=0.
Nonunit_fsub_feature_tests=24. Nonunit_bsub_feature_tests=92.
Megabytes=0.28.
User_CPU=0.02, System_CPU=0.00, Wall_clock=0.

===== end of statistics =====

===== end of search =====
```

THEOREM PROVED

4 Concluding remarks

To our best knowledge, this work describes the first attempt to employ an automated-reasoning computer program to the philosophical thought. We have not (or, better to say, not yet) attempted to write our own software dedicated to that purpose. This is because the major difficulty in such applications is the proper formalization of a philosophical theory. Indeed, every formalization is at the same time an interpretation. And even in our case, Spinoza could have had complained that Blum-Malinovich formalization does not fully reflect his thought, regardless of the differences of historical epoch and in the language. For instance, one could compare the above symbolic Definition 1 with its original verbal counterpart: “By that which is self-caused, I mean that of which the essence involves existence, or that which the nature is only conceivable as existent”. Clearly, the formalization here does not only interpret the original in some way, but also trivializes it. The last statement can, of course, be used against any attempt to seek for computer-generated proofs. But we would like to conclude with a more positive statement: there are many levels of possible formalizations, and, by iterating, we can eventually approach the philosopher’s intentions arbitrarily close. And the endeavor itself is valuable as some thinkers sometimes have some important things to say to us.

Let us finally remark that automated theorem proving has recently found application in the analysis of optical system [5]. That interesting development possibly heralds new ways in which automated provers will be applied in both theoretical thought and technology.

We would like to dedicate this work to the memory of late Professor William McCune.

References

1. Prover9 <http://www.cs.unm.edu/~mccune/prover9/manual/2009-11A/> [Accessed 2016.12.20]
2. B. Spinoza, *Ethics*, English translation from Latin in: *The Chief Works of Benedict De Spinoza* by R.H.M. Elwes, Dover Publication, New York (1951)
3. C. Jarrett, The logical structure of Spinoza’s Ethics, Part I *Synthese*, **37**, 15-65 (1978)
4. A. Blum and S. Malinovich, A formalization of a segment I of Spinoza’s *Ethics*, *Metalogicon*, **VI**, 1 (1993)
5. O. Hasan, S. Khan Afshar, S. Tahar, Formal Analysis of Optical Waveguides in HOL. In: Berghofer S., Nipkow T., Urban C., Wenzel M. (eds) *Theorem Proving in Higher Order Logics. TPHOLs 2009*. Lecture Notes in Computer Science, vol 5674, pp. 228-243. Springer, Heidelberg (2009). doi:[10.1007/978-3-642-03359-9_17](https://doi.org/10.1007/978-3-642-03359-9_17)