Paul Bernays' Independence Proofs
Reconsidered

Cyril Welch

Abstract:
There are exactly 24 three-value models for proving the independence of the Tautology, Addition, and Permutation in Russell's sentential calculus, and I present 4 four-value models for proving the independence of Summation. A careful review of the way these models work reveals a variety of "properties" and thereby a nuanced understanding of the method. This nuanced understanding in turns allows us to consider the historical shift from the focus on form to the focus on structure. The significance of this shift of focus has become a central, if not bothersome question.

In his habilitation work of 1918, Paul Bernays introduces a technique for demonstrating the independence of four of the primitive sentential "propositions" of Principia Mathematica:¹

\begin{align*}
\sim(p \lor p) \lor p & \quad \text{Tautology} \\
\sim q \lor (p \lor q) & \quad \text{Addition} \\
\sim(p \lor q) \lor (q \lor p) & \quad \text{Permutation} \\
\sim(\sim q \lor r) \lor \left[\sim(p \lor q) \lor (p \lor r)\right] & \quad \text{Summation}
\end{align*}

The technique consists of re-interpreting the primitive symbols for negation and disjunction with three or four rather than the two values of truth and falsehood. Interpreting the two operators as follows:

\[
\begin{array}{c|cccc}
\lor & \alpha & \beta & \gamma & \sim \\
\hline
1 & \alpha & \alpha & \alpha & \alpha \times \\
2 & \beta & \alpha & \beta & \beta \times \\
3 & \gamma & \alpha & \gamma & \gamma \times \\
0 & \delta & \delta & \delta & \delta \times \\
\end{array}
\]

Bernays showed, first, that each and every "world" of Addition, Permutation, and Summation has the value \(\alpha\); second, that every application of Substitution and modus ponens on the latter three yields a formula every world of which has the value \(\alpha\); and, third, that not every world of Tautology has the value \(\alpha\). Employing strong induction, Bernays then concludes the independence of Tautology.

To demonstrate the independence of the other primitives, Bernays interprets the operators with four values:

\[
\begin{array}{c|cccc}
\lor & \alpha & \beta & \gamma & \delta \\
\hline
1 & \alpha & \alpha & \alpha & \alpha \times \\
2 & \beta & \alpha & \beta & \beta \times \\
3 & \gamma & \alpha & \gamma & \gamma \times \\
4 & \delta & \delta & \delta & \delta \times \\
\end{array}
\]

\[
\begin{array}{c|cccc}
\lor & \alpha & \beta & \gamma & \delta \\
\hline
1 & \alpha & \alpha & \alpha & \alpha \times \\
2 & \beta & \alpha & \beta & \beta \times \\
3 & \gamma & \alpha & \gamma & \gamma \times \\
4 & \delta & \delta & \delta & \delta \times \\
\end{array}
\]

\[
\begin{array}{c|cccc}
\lor & \alpha & \beta & \gamma & \delta \\
\hline
1 & \alpha & \alpha & \alpha & \alpha \times \\
2 & \beta & \alpha & \beta & \beta \times \\
3 & \gamma & \alpha & \gamma & \gamma \times \\
4 & \delta & \delta & \delta & \delta \times \\
\end{array}
\]

⁠[This exposé I wrote during the summer of 2003 with a view to a seminar on Heidegger and Logic that I conducted in the fall. The only explicit indication of this intent I inserted into the last endnote.]
For the demonstration of the independence of Addition, however, Bernays allows for two “distinguished” values: the worlds of each of the other three primitives have a mix of \( \alpha \) and \( \gamma \) as values, whereas Addition has a mix of \( \alpha \), \( \beta \) and \( \gamma \) as values.

The results are indisputable. The technique has become commonplace even beyond David Hilbert’s formalist program (which Bernays intended to illustrate). But the philosophical significance of the technique? It has become so commonplace that precisely those who become expert in its application need no longer recall its peculiar power, and therewith its inherent limitations.

In order better to contemplate the significance of the accomplishment, let us first review its workings in detail.

First of all, we may prove three of the four sentential primitives independent by employing only three-value interpretations. There are in fact six such interpretations demonstrating the independence of Tautology in the fashion already illustrated:

1. \[ \lor \begin{array}{ccc} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \end{array} \]
2. \[ \lor \begin{array}{ccc} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \end{array} \]
3. \[ \lor \begin{array}{ccc} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \end{array} \]
4. \[ \lor \begin{array}{ccc} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \end{array} \]
5. \[ \lor \begin{array}{ccc} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \end{array} \]
6. \[ \lor \begin{array}{ccc} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \end{array} \]

There are eight demonstrating the independence of Addition:

7. \[ \lor \begin{array}{ccc} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \end{array} \]
8. \[ \lor \begin{array}{ccc} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \end{array} \]
9. \[ \lor \begin{array}{ccc} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \end{array} \]
10. \[ \lor \begin{array}{ccc} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \end{array} \]

There are four demonstrating the independence of Permutation:

11. \[ \lor \begin{array}{ccc} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \end{array} \]
12. \[ \lor \begin{array}{ccc} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \end{array} \]
13. \[ \lor \begin{array}{ccc} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \end{array} \]
14. \[ \lor \begin{array}{ccc} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \end{array} \]

And to demonstrate the independence of Summation? An exhaustive review of the possibilities reveals none for three values. So we must resort to four values. There is Bernay’s, and I have discovered, by trial and error, three more:

15. \[ \lor \begin{array}{cccc} 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \end{array} \]
16. \[ \lor \begin{array}{cccc} 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \end{array} \]
17. \[ \lor \begin{array}{cccc} 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \end{array} \]
18. \[ \lor \begin{array}{cccc} 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \end{array} \]

The dominant property at issue in these twenty-two demonstrations reads “every value in its possible worlds is 0,” or: \[ \forall v \[ v \in W \rightarrow v = 0 \]. \]
whereas, the property of the maverick includes \( v = 1 \) (or 2), so: \[ \exists v \[ v \in W \& v \neq 0 \]. \]

However, Bernays demonstrates the one, Addition, by employing two preferred properties (distinguished values). And an exhaustive review of possible three-value interpretations reveals six that demonstrate the independence of Permutation in this fashion. Three that perform on the basis of 0 and 1 as the preferred properties:
And three with 0 and 2:

<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

In these six cases, as in Bernays’ for Addition, we may formulate the property enjoyed by the three other primitives to read:

\[ \forall v [v \in W \to (v = 0 \lor v = x)] \]

(where \( x = \) the other preferred value, i.e. 1 or 2), while in each case the possible worlds of Permutation (or of Addition, in Bernays’ example) include another value as well:

\[ \exists v [v \in W \land \neg (v = 0 \lor v = x)] \]

i.e. the property of the one primitive contradicts the dominant property.\(^3\)

For three-value models (interpretations), the proof of the Induction Step is simple:

<table>
<thead>
<tr>
<th>A</th>
<th>\neg A</th>
<th>\neg A \lor C</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td></td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

where we can complete the table according to the definitions of the operators and assure ourselves that whenever \( A \) and \( \neg A \lor C \) have the preferred value of 0, so too does \( C \).

The Induction Step for proofs allowing two preferred values requires us to look at six possible conditions for any three-value interpretation (eight for four-value models with two preferred values):

<table>
<thead>
<tr>
<th>A</th>
<th>\neg A</th>
<th>\neg A \lor C</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td></td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>x</td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>x</td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>x</td>
<td></td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

where, once again, we must assure ourselves that whenever *modus ponens* is employed, the Consequent will have one of the two preferred values.\(^4\)

Of course, the desired result requires only one suitable interpretation of the curl and the wedge for each primitive. Yet the very plethora of suitable interpretations allows a more careful consideration of the technique. Essential to this technique is the introduction of a distinction between logical and material discourse unknown to classical thinkers from Aristotle onwards, including Russell himself.

Classical logicians culled their logical principles from discourse already prevailing (geometry, oratory, . . .). Such discourse is significant in its own right, and its signification remains the touchstone not only for the original culling of inferential principles but also for various other logical decisions, most notably in the use of counter-examples.

Employing Bernays’ technique for demonstrating independence, we invent our own material discourse, one having no prior significance at all — twenty-four for three-value interpretations alone. Such invention takes place as we review all the permutations formally possible when stipulating three (or more) values for the two logical operators, and then discover those not only assuring a homogeneity of “worlds” for all but one of the primitives but also meeting the requirements for the bequeathal of the homogeneity.

Meanwhile, the logical discourse remains quintessentially classical. Both the homogeneity and the heterogeneity must be formulated according the classical principles of universality,
particularity, and contradiction, and the proofs of bequeathal follow the classical principles of Substitution (the form retains its force no matter what the instances) and modus ponens (implicit in Aristotle, explicit from the Stoics onwards). The only difference between classical logical discourse and the logical discourse introduced by Bernays is the inclusion in the latter of the distinctively mathematical principle of induction.⁵

The power of the technique lies in the relation between Bernays’ invented and Russell’s original material discourse: what holds for the invented holds for the original. How exactly does this relation work?

Classical logic (whether Aristotelian, Stoic, medieval, or modern) contemplates form: the variables take arguments while the operators remain constant. Here, the material discourse remains within the semantic sphere; in modern terms, it remains truth-functional.

Even three-or-more-value models retain a faint semblence of truth-functional discourse so long as there is only one preferred value (“true” then, with two or more ways of being “false”). This last vestige disappears when we resort to two or more preferred values: here we are constructing systems in which each decidable formula (“line”) has its own dynamism (“life”): its value may vary from “world” to “world”—a variation we otherwise associate with contingent formulas, yet one that in no way militates against the necessity essential to our constructions.

In short, with the operators themselves interpreted variously, the formulas no longer retain their semantic sense: they no longer await “instances” of p, q, . . . (either true or false, whether empirically or logically)—no longer mean anything in the traditional logical sense of deferred meaning. Thus we can formulate Russell’s primitives to reflect the neutrality of the logical operators, e.g.:

\[
\begin{align*}
\circ(p\circ p) & \circ p & \text{“Tautology”} \\
\circ q & \circ(p \circ q) & \text{“Addition”} \\
\circ(p \circ q) & \circ(q \circ p) & \text{“Permutation”} \\
\circ(q \circ r) & \circ[(\circ p \circ q) \circ(p \circ r)] & \text{“Summation”}
\end{align*}
\]

We can now understand the neutral \(\otimes\) and \(\oslash\) as taking substitution-instances from “domains” of permutations. In the case of three-value interpretations, there are \(3^3\) possibilities for \(\oslash\), only ten of which (on inspection) are fruitful for independence proofs; for \(\otimes\) there are \(3^3\) possibilities, only thirteen of which are fruitful. Thus we would have to devise twenty-three new symbols if we were to construct our 24 three-value systems separately.

But we can also revert to two-value interpretations for both \(\otimes\) and \(\oslash\)—among which we will find two (namely, 0-1 and 1-0) appropriate for the original truth-functional operators \(\neg\) and \(\vee\). Here, the resulting system is precisely that of the Theory of Deduction (sentential calculus) presented in Principia Mathematica. This system, the original, now appears as one among the unlimited number possible for the neutral formulations.

Given the difference between the logical and the material discourse, the object of the latter (illustrated by the neutral formulations) deserves a name of its own: structure—in contrast to form. From the standpoint of the material discourse, form is merely a “substitution instance” of structure. From the standpoint of the logical discourse, structure is merely a fabrication: something we make according to our own rules.

Reverting to structure enables us to decide questions regarding the transfinite possibilities of form (just as, in the classical manner, reverting to form enables us to decide questions regarding all possible empirical instantiations). Bernays’ technique for proving independence is itself only one example, very elementary, of this reversion to structure. The same reversion lies at the heart of proofs of consistency, completeness, and the decidability of (some) quantified propositional forms.⁶ Perhaps the most celebrated instance of such reversion is Gödel’s proof of the incompleteness of any arithmetic axiomatized on the basis of any axiomatized sentential and predicate logic.

However, just as Aristotle’s account of syllogistic unfolds in constant, even if often only implicit reference to the classical ontology of form, so Bernays’ various demonstrations unfold against the background of “formalist” program as spearheaded by David Hilbert. According to his biographer Constance Reid, Hilbert told his students already in his 1898-99 lectures on the “elements of geometry” that “Euclid’s definitions of point, straight line, and plane were really mathematically insignificant”: 
In other words, whether they were called points, straight lines, planes or were called tables, chairs, beer mugs, they would be those objects for which the relationships expressed by the axioms were true. And “true” here means crafted according to the rules — not pointing to how things really are, to anything in the “physical universe.” This memorable image of “tables, chairs, beer mugs” receives more careful formulation in Hilbert’s 1925 lecture “On the Infinite.” In contrast, he says, to the youthful (later abandoned) efforts of Frege and Dedekind to base their contemplation of the infinite on “pure logic” (its semantics) “independent of all intuition and experience”:

as a precondition for the employment of logical inference and for the activation of logical operations, something is already given in representation — namely, certain extra-logical concrete objects that are intuitively there as immediately experienced. For logical inference to be certain, these objects must permit us to survey them completely, in all their parts — and their exhibition, their distinction, their sequences and contiguities must be given as something that does not permit any reduction to anything else, or stand in need of such reduction. This is the basic philosophical disposition that I hold to be requisite for mathematics as well as for any scientific thinking, understanding and communicating. And, especially in mathematics, the signs themselves are the object of our contemplation, the configuration of this object being immediately clear and ever-again recognizable.

What’s “already given in representation” are the “marks on paper” — extra-logical and immediately experienced. Only these can be surveyed completely — unlike anything infinite, whether the number of points on a segment or number of possible lines in an axiomatic proof. This willingness to pull back into the signs — these as the objects of our contemplation, not what they signify (if anything) in the classical sense—is the requisite disposition, even talent, for “any scientific thinking, understanding and communicating,” dominantly so for mathematics (and for mathematical logic of the sort now expanded by Bernays and Gödel).
... In earlier epochs ... nature ... was a realm that lived according to its own laws, a realm into which man had somehow to fit himself and his life. In our age, though, we live in a world that has been so completely transformed by human agency that everywhere we constantly come up against humanly devised structures and in this sense only encounter ourselves.10

That is, no longer acknowledging the full force of the metaphorical expression “laws of, i.e. in nature,” modern natural science devises, works with, and presents to us laws about nature — better called structures. These structures pertain to what we might still call “nature”: they result from a definite manner of dealing with circumstances and they serve to strengthen our dealings. Indeed, proposed formulations must prove and improve themselves in such dealings. However, compared with its ancient homonym, science now provides not so much contact with or insight into the inner and restful nature of circumstances as rather strategies for dealing with the actual course of events in all their restlessness. But then we really only encounter laws, structures, and strategies—not nature herself. And these laws, structures, and strategies represent ourselves, i.e. inherited and projected achievements.

Second, the testimony of J.-Claude Piguet on the development of music in the 20th century. For instance, Piguet contrasts the accomplishments of Debussy and Stravinsky: while the first “begins with expression (objectivity in an image) in order to renew form,” the second “begins with form to achieve a new kind of expression”:

In effect, Stravinsky inaugurates a new manner of saying things in music, one that is not tied directly, but rather indirectly to the things being said ... a manner of saying just about anything, not according to the rules that govern the things being said but only according to the laws of saying itself.

Piguet also reviews a number of examples contrasting Stravinsky with Haydn and Mozart, both of whom composed instances within musical forms—rather than taking the form itself as the “subject.” One effect of retreating into “form,” i.e. into structure, is that new work can consist of modifying languages that already prevail.

At the same time, Stravinsky gets music to express what it never before was able to express, namely the fact that music is a language. His music is thus a formalism, but a formalism that is expressive because it is consistent with itself. ... Simply put, Stravinsky expresses nothing by means of form alone, he rather gets music to express the fact, in itself abstract, that music is always form.11

Bernays’ technique, too, engages us in a style of thinking “where we constantly come up against humanly devised structures and in this sense only encounter ourselves.” It, too, “inaugurates a new manner of saying things ... one that is not tied directly, but rather indirectly to the things being said.” Yet, unlike work in physics or music, in the social sciences or literary studies, work in logic does not itself intrude on our circumstances—does not reshape what we encounter. In this sense, it is entirely innocent. Thus the careful distinctions embedded in the application of this one very simple technique—especially that between logical and material discourse—provide a neutral ground for studying the momentous shift of “intentionality” from form to structure, the “subsumption” of the one under the other.12

Notes
2. I shall henceforth employ the now more familiar numerical designations, in keeping with James. B. Gerrie’s software program The Logical Theorist, with which I discovered the following interpretations. This program, designed to facilitate the construction of systems of sentential and predicate logic like Principia Mathematica, may be obtained, along with the manual, at http://www3.ns.sympatico.ca/jimgerrie/LOGIC53.ZIP
3. It is important to retain the logical formulation of the crucial properties. In their Principles of Mathematical Logic (1950; Preface 1928), Hilbert and Ackermann construe the properties arithmetically: \(\triangleright\) represents “ordinary
multiplication.” But this is very misleading, since the “arithmetic” here is itself arbitrarily construed; e.g., \(2 \times 2 = 0\), which hardly represents “ordinary multiplication.” Acknowledging the arbitrariness, Bernays’ employment of Greek letters accentuates the possibility of defining the properties logically.

4. Hopes can be dashed, as I learned. Having discovered that there are exactly four three-value models providing the Base Cases for proving the independence of Summation, I discovered that none passed the test of the Induction Step:

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0 1 2</td>
<td>0 1 2</td>
<td>0 1 2</td>
</tr>
<tr>
<td>0 1 2</td>
<td>0 1 2</td>
<td>0 1 2</td>
</tr>
<tr>
<td>0 1 2</td>
<td>0 1 2</td>
<td>0 1 2</td>
</tr>
</tbody>
</table>

Similarly, there are six more “dashed hopes” for Tautology and eight more for Permutation (none at all for Addition).

5. It has been fashionable to distinguish between “meta-language” and “object-language” (cf. Irving M. Copi’s *Symbolic Logic*, 5th ed. 1979, pp. 213 ff.). While it draws attention to the difference between discussing something and facing what is discussed, as well as that between rules and primitive formulas, the distinction obscures the focus of logical study itself. Thus, for instance, in their very admirable *Development of Logic* (1962), William and Martha Kneale speak apologetically about “the use of English in the formulation of [Huntington’s] postulates,” which “naturally carries with it the assumption of ordinary logical principles in the derivation of consequences…” (p. 696) — as though we would like to dispense with this assumption as much as possible. But what is logical study if not the careful detection and proper celebration of logical principles?


   Otto Blumenthal reports that in a discussion in a Berlin train station in 1891, Hilbert said that in the proper axiomatization of geometry, ‘one must always be able to say, instead of “points, straight lines, and planes,” “tables, chairs, and beer mugs.”’


8. “Über das Unendliche,” *Mathematische Annalen*, Vol. 95 (1925), pp. 167 and 171. An English translation is available in *Philosophy of Mathematics*, edited by Benacerraf and Putnam (1983, pp. 188 & 192). However, this translation considerably weakens Hilbert’s insistence on the “sensory” character of the “concrete objects,” i.e. symbols (e.g., by translating *Vorstellungen as “conceptions” rather than as “representations”*).


12. Martin Heidegger’s *Being and Time* (1927) addresses the question of structure under the rubric of “world.” In his later works, Heidegger argues that “world” has historically evolved into a “framework” (*Gestell*) whose only domain of instantiations is that of “inventory” (*Bestand*). As the heart of technology, this evolving primacy of structure is no longer innocent, as is Bernays’ technique. But those intent upon contemplating this historical evolution had best understand how it works “on site” — before issuing proclamations about its possibly pernicious results.