

The Early Wittgenstein on Logic and Metaphysics

Jinho Kang

(Department of Philosophy, Seoul National University)

It is fairly well known that Wittgenstein was deeply critical of metaphysics in the *Tractatus*. It is little known, however, that he did not begin with such a critical attitude. In the first substantial philosophical work he produced in October 1913, "Notes on Logic", Wittgenstein writes:

Philosophy consists of logic and metaphysics: logic is its basis.
(Wittgenstein 1979, p.106)

Far from being critical, Wittgenstein seems to suggest here that metaphysics does comprise one of the two legitimate parts of philosophy. On the other hand, he expresses in no uncertain terms in 6.53 of the *Tractatus* that all metaphysical remarks are nonsense and that the task of philosophy is to reveal their nonsensicality through the analysis of language:

The right method of philosophy would be this. To say nothing except what can be said, *i.e.* the sentences of natural science, *i.e.* something that has nothing to do with philosophy: and then always, when someone else wished to say something metaphysical, to demonstrate to him that he had given no meaning to certain signs in his sentences. (*Tractatus*, 6.531)

This passage originally occurs in an entry Wittgenstein made in a wartime notebook on December 1916 (Wittgenstein 1979, p.91). We can thus see that there was a fundamental change in Wittgenstein's attitude on metaphysics between October 1913 and December 1916. But why and how did this change take place? Given the importance of the issues concerning the status of metaphysics in Wittgenstein's early philosophy, it is surprising that no one has discussed the question I just raised, or even has paid attention to the change I pointed out.

The aim of this paper is to make sense of the change in Wittgenstein's attitude on metaphysics by investigating the relevant materials in the pre-*Tractatus* manuscripts, which chiefly include the 1913 "Notes on Logic", the 1914 "Notes Dictated to Moore in Norway" ("Moore Notes", for short), the three wartime *Notebooks* from August 1914 to January 1917, and the manuscript called the *Prototractatus* which contains an early version of the *Tractatus* (Wittgenstein 1979, 1996). I argue that the change in Wittgenstein's attitude on metaphysics resulted from important changes in his account of logic: Whereas his earlier account of logic in "Notes on Logic" and "Moore Notes" has a consequence that logic has metaphysical implications, the later account Wittgenstein develops in the wartime *Notebooks* and the *Prototractatus* leads to an opposite consequence. I then sketch a non-metaphysical reading of the statements in 1s and 2s of the *Tractatus*. Despite Wittgenstein's vehemence against metaphysics in 6.53, many commentators assumed that these statements did propose a complex and intricate metaphysical system. The main reason for this assumption is that otherwise we simply do not seem to have any idea of how to make

1) I follow Ogden's translation of the *Tractatus* (Wittgenstein 1922) rather than Pears and McGuinness's (Wittgenstein 1961).

sense of those statements, for instance “The world is the totality of facts, not of things” (1.1) and “An atomic fact [*Sachverhalt*] is a combination of objects” (2.01): If *these* are not metaphysical statements, what else could be? But I suggest that there are reasons to suspect the intelligibility of the statements in 1s and 2s of the *Tractatus*, even though they do initially seem to make sense.

§1. Logic and Metaphysics in “Notes on Logic” and “Moore Notes”

Let us begin with Wittgenstein's account of logic in “Notes on Logic” and “Moore Notes”. It primarily represents Wittgenstein's earlier attempt to solve the problems besetting Russell's logical theory.²⁾ Russell came to advocate a strong form of realism at the turn of the 20th century in opposition to the British Idealism, and his logical theory is crucially shaped by this realistic outlook. Russell follows the traditional view of logic as the study of inference, but he denies the popular idea of construing the notion of inference psychologically, as a certain sort of thought process (Russell 1996, §12 and §37). According to Russell, an inferential relation is a relation among *propositions*, and propositions for him are non-mental, non-linguistic entities existing independently of us. So logic for Russell is as much a study of the objective world as physics, chemistry, biology, etc. are. Logic is different from these sciences in that it is the *maximally general* science of the world, and this is because the inferential relations studied by logic are based on the most general features of the world. Therefore, under

²⁾ Wittgenstein also has Frege's logical theory in mind, but the main target is Russell's.

Russell's conception of logic, the traditional idea of logic as *the study of inference* is transformed to the idea of logic as *the study of the most general features of the world and the principles governing them*. Note that logic becomes virtually indistinguishable from metaphysics under this conception, for metaphysics was traditionally conceived as the study of the most general features of the world and the principles governing them. Indeed Russell hardly uses the term "metaphysics" in his philosophical writings from 1900 to 1913. I suspect that this is because Russell thinks that the new logic he has developed now *replaces* what is traditionally called metaphysics, and thus that he does not need metaphysics other than logic for his philosophy.

The most general features of the world on which logical inferences are based are what Russell calls "logical constants" in his 1903 *Principles of Mathematics*. Russell's list of logical constants, however, changed considerably after the *Principles of Mathematics*, as he had a great difficulty giving an account of exactly which properties should count as logical constants and what their natures are. By the time Wittgenstein was working with Russell from 1911 to 1913, the list chiefly included various types of ontological categories (individual, first-level monadic relation, first-level binary relation, etc.), forms of various types of atomic propositions (first-level subject-predicate form, first-level binary relation form, etc.), truth-functions of different levels (negation of the first-level, second-level, etc., conjunction of the first-level, second-level, etc.), and quantifiers of different levels. Note the difference between this list of logical constants and the usual contemporary one which chiefly include truth-functional connectives, quantifiers, and identity.

Now Wittgenstein does share Russell's basic assumption that logic is the maximally general science of the world. However, he comes to arrive at a radically different conception of logic with his

distinctive understanding of *logical* generality. According to Wittgenstein, the generality distinctive of logic must be understood as originating from the *formal* and *necessary* features of the world, the features that would remain whatever the world is like. Logical constants are none other than these formal and necessary features of the world. It is not yet clear exactly how this conception of logic is different from Russell's, but here is a crucial difference: According to Wittgenstein, logical constants understood as the formal and necessary features also give rise to the features necessarily possessed by any languages, given that languages aim to represent the world. Wittgenstein more or less accepts Russell's list of logical constants, except that truth-functions and quantifiers do not form hierarchies of levels for him. So Wittgenstein amounts to claiming that, however a language is set up, it will always possess the features corresponding to various types of ontological categories, logical forms of atomic propositions, truth-functions, and quantifiers. Now a language is set up based on conventions that govern the uses of the linguistic symbols it has. If there are features that are necessarily possessed by *any* languages, however, these features cannot be based on conventions, for nothing that is established by conventions can be necessary. Hence no linguistic symbols can represent logical constants.

It is the thought that no linguistic symbols can represent logical constants that Wittgenstein calls his "fundamental thought" [*Grundgedanke*] in the *Tractatus*:

The possibility of sentences is based upon the principle of the representation of objects by signs.

My fundamental thought is that the "logical constants" do not represent. That the *logic* of facts cannot be represented. (*Tractatus*, 4.0312)

But this sounds very odd: don't we have expressions that do represent logical constants, expressions such as "thing", "binary relation", "subject-predicate form", "not", "and", "some", and so on? Wittgenstein's answer is that ordinary language misleads us here. It is only an illusion created by ordinary language to think that those expressions represent logical constants; they are pseudo-expressions.

Wittgenstein accordingly proposes that logical constants cannot be represented by symbols but only be *shown* by the *ways symbols symbolize in language*, so that different logical constants will be shown by the different ways symbols of different logical types symbolize. For example, let us think of a (non-linguistic) atomic proposition *Dante loves Beatrice*, and of the logical constant of binary relationhood³⁾ possessed by a relation *love* in it. This atomic proposition is represented by a sentence "Dante loves Beatrice." In "Notes on Logic" and "Moore Notes", Wittgenstein accepts the natural view that names like "Dante" and "Beatrice" symbolize simply by being used as arbitrary labels for the things they refer to. But binary relation symbols, according to Wittgenstein, symbolize in a way entirely different from that of names. The symbol for the binary relation *love* in the sentence "Dante loves Beatrice" is not the sign "loves", but the relation holding between the sign "loves" and two names, i.e. the relation *that "loves" stands between two names*. Therefore, unlike a name that gets its meaning on its own, a binary relation symbol cannot symbolize without combining with two names; the mode of combination for a binary relation symbol is already built into the way it symbolizes. It is this different way of symbolization of the binary relation symbol, not a pseudo-predicate "ξ is a binary relation", which shows the logical feature of the

³⁾ Strictly speaking, I should say "*first-level* binary relationhood", in accordance with Russell's type theory.

binary relation that it has the possibility of combining with two things. For no expression that can combine with the predicate “ ξ is a binary relation” can have the feature essential to a binary relation symbol that it symbolizes always in combination with two names.

Accordingly, Wittgenstein proposes the ideas of the *theory of symbolism* and the *logically perspicuous notation*. The theory of symbolism is a specification of how symbols of different types symbolize in different ways. The logically perspicuous notation is a symbolism where symbols of different types can be recognized in a clear and non-misleading way. (For example, small letters for names; capital letters for relations of various types, with superscripts indicating their levels and subscripts indicating the number of their argument places). With this, the logically perspicuous notation is purified of pseudo-expressions such as “thing”, “binary relation”, “subject-predicate form”, “negation”, “all”, and so on, so that those who use this notation are no longer under the illusion that logical constants can be represented by symbols.

But it is not yet clear how truth-functional connectives and quantifiers can be shown by the ways symbols symbolize, and how the expressions for these logical constants can accordingly be eliminated. Wittgenstein provides an answer to this problem with an account of how *elementary sentences* of various types symbolize. According to Wittgenstein, an elementary sentence is a *fact* of a certain type: an elementary sentence “Dante loves Beatrice”, for example, is a fact that the sign “loves” stands between “Dante” and “Beatrice” in that order. Being a fact, an elementary sentence can be compared with various atomic facts, and accordingly it symbolizes by being true or false: if there is an atomic fact that corresponds to the elementary sentence, the sentence is true; otherwise it is false. So an elementary sentence thus understood always has a *specific truth-falsehood possibility*, depending on which atomic facts make it

true and which make it false. This specific truth-falsehood possibility of an elementary sentence is what Wittgenstein calls its *sense*.⁴⁾ So the possibility of being true or false is intrinsic to the way an elementary sentence symbolizes, which shows the logical feature of an atomic fact that it can exist or not exist. Again, this feature of the possibility of existence and non-existence of an atomic fact cannot be captured by a pseudo-predicate “ ξ is an atomic fact”, for no expressions that can fill the gap in this predicate can have the feature essential to elementary sentences that they are capable of being true or false.

Wittgenstein devises a symbolism which makes perspicuous the feature of an elementary sentence that it always has a specific possibility of being true or false. This is what he calls the “*ab*-notation” in “Notes on Logic” and “Moore Notes”, which is later presented in the apparently more familiar truth-table notation in the *Tractatus*. In Wittgenstein's truth-table notation, an elementary sentence “*p*” in Russell's notation is written as

$$\begin{array}{c} p \\ \top \\ \text{F} \end{array}$$

indicating that “*p*” is a symbol that must be capable of being true and capable of being false. It is with this feature of the “bi-polarity” of an elementary sentence that Wittgenstein eliminates symbols for truth-functional connectives and quantifiers in his symbolism, thereby substantiating further his “fundamental thought” that logical constants

⁴⁾ So Wittgenstein's notion of sense is quite different from Frege's. Frege admits of a sentence which can be neither true nor false but still has a sense, whereas this is simply impossible for Wittgenstein given that the sense of a sentence for him *is* its specific truth-falsehood possibility.

do not have representatives. For example, “ $p \vee q$ ” in Russell's notation is written in Wittgenstein's truth-table notation as

p	q	
T	T	T
T	F	T
F	T	T
F	F	F

where “ \vee ”, which looks like a symbol for a binary relation in Russell's notation, simply disappears. Symbols for other truth-functional connectives can similarly be eliminated. Wittgenstein does not seem to be entirely clear about how to eliminate symbols for quantifiers in “Notes on Logic” or “Moore Notes”, but in the *Tractatus* he construes general sentences as (potentially infinite) conjunctions or disjunctions of the sentences sharing common symbols. In the truth-table notational renderings of general sentences thus construed, quantifier symbols can also be eliminated.

Now it is crucial to bear in mind that, in spite of the superficial similarity in their appearance, Wittgenstein's truth-table must be sharply distinguished from the truth-table in contemporary logic. The truth-table in contemporary logic is a device for *semantics*, where each row or each line in the table for a sentence represents an *interpretation* of the sentence. But Wittgenstein's truth-table is a linguistic *notation*. A rendering of a sentence in truth-table notation for him is merely another way of writing the sentence, making it perspicuous that the way a sentence symbolizes is such that it always has a specific truth-falsehood possibility, the possibility which is determined by the truth-falsehood possibilities of elementary sentences it consists of. For Wittgenstein, therefore, the notions of truth and falsehood cannot be separated from that of a sentence; one

cannot understand the notion of sentence without understanding what it is to be capable of being true and capable of being false. This conception of sentence is quite different from the one in contemporary logic, according to which whether a string of signs can count as a sentence is specified purely syntactically, i.e. in terms of formation rules in a formal system.

Based on a consideration which is similar to Donald Davidson's observation on the creativity of language, Wittgenstein supposes that *every* (genuine) sentence in our ordinary language can ultimately be analyzed into a truth-functional compound of elementary sentences. Moreover, he supposes that every inference is truth-functional in its character. This enables Wittgenstein to maintain in "Moore Notes" that inferential relations among sentences can be read off directly from the sentences themselves once we rewrite each sentence in its completely analyzed form in the truth-table notation. For example, we can see directly from the truth-table renderings of " p " and " $p \vee q$ " that the latter follows from the former, for the truth-table renderings make it perspicuous that the latter cannot be false when the former is true. Unlike Frege and Russell, therefore, Wittgenstein does not need axioms or rules of inference in order to explain logical inference.

It thus turns out that, once we establish a proper theory of symbolism in Wittgenstein's sense, we will not need a further account of what logical constants are or what logical inferences are: logical constants will already be explained away by the specification of the ways symbols symbolize, and inferential relations among sentences can be read off directly from the sentences themselves. We have seen that Russell thought of logic as the study of logical constants and the principles governing them. Therefore, under Wittgenstein's account of logic in "Notes on Logic" and "Moore Notes", the whole of logic as Russell characterizes it is based on

the theory of symbolism. Indeed, logic *is* the theory of symbolism, according to this account.

Now, it is because of the treatment of logic as the theory of symbolism that Wittgenstein declares in "Notes on Logic" that logic is the basis of metaphysics. Wittgenstein's distinction between logic and metaphysics is not surprising to us, but it may have been surprising to Russell. For we have seen that Russell's conception of logic makes logic virtually indistinguishable from metaphysics. Wittgenstein can make a distinction between logic and metaphysics because the theory of symbolism, which he equates with logic, is not itself a metaphysical theory; as we have seen, it is a specification of how various types of *symbols* symbolize. On the other hand, however, different types of symbols in Wittgenstein's theory of symbolism correspond to different ontological categories such as things, first-level properties, first-level binary relations, second-level properties, and so on. His elimination of symbols for logical constants also has an important ontological consequence that logical constants are not entities. In other words, Wittgenstein's theory of symbolism has metaphysical *implications*, even though it is not itself metaphysics. This is why Wittgenstein takes logic, understood as the theory of symbolism, to be the basis of metaphysics.

§2. The Problems of the Elementary Sentence and Wittgenstein's "Class Theory"

Wittgenstein's account of logic as the theory of symbolism in "Notes on Logic" and "Moore Notes" solves many problems of Russell's logical theory, which I will not go into here. However, by

the time Wittgenstein begins his philosophical entries in the 1914 wartime *Notebook*, he comes to realize that the account has some serious problems.

We have seen that logical inference under the account of logic in “Notes on Logic” and “Moore Notes” is entirely truth-functional and also entirely intrinsic to the sense of a sentence. Therefore, for this account to work, it must be guaranteed that every sentence is analyzable into a truth-functional compound of elementary sentences and that every elementary sentence has a sense, i.e. its specific truth-falsehood possibility. But this does not seem to be guaranteed. A sentence like “Smith believes that Dante loves Beatrice and Abelard loves Eloise.” does not seem to be a truth-functional compound of elementary sentences. Also, there seem to be a well-formed elementary sentence that does not have a truth-falsehood possibility (e.g. “Excalibur is sharp,” which is neither true nor false), and a well-formed but nonsensical elementary sentence (e.g. “Mt. Everest loves Mississippi River.”) One might try to deal with these counter-examples along the lines of contemporary approaches, for example by developing semantics of empty names or developing lexical semantics which blocks nonsensical strings of symbols by appealing to the so-called “selection rules” which govern the so-called “selection restrictions” of lexical entries. But Wittgenstein does not take these strategies. The strategy he takes is to propose that genuine elementary sentences are radically different from apparent elementary sentences in ordinary language. (So my previous example of an elementary sentence, “Dante loves Beatrice”, is not really a genuine one.) Genuine elementary sentences can be discovered only as a result of logical analysis. Once an analysis is done, it will indeed turn out that every sentence in ordinary language is a truth-functional compound of elementary sentences, and that every elementary sentence is guaranteed to have a sense; or so

Wittgenstein thinks.

But this strategy raises two problems, which I will call the *problems of the elementary sentence*: First, what are these genuine elementary sentences like? If we do not know what they are like, we will not be able to know when our analysis should end, so the very idea of analysis will not get a foothold. Second, which types of symbols will occur in these genuine elementary sentences? Since the types of symbols occurring in elementary sentences also determine their logical forms, this is equivalent to asking what the possible logical forms of elementary sentences are. Will there be an elementary sentence with a 4-place relation form? What about an elementary sentence with a 27-place relation form? If ordinary elementary sentences were genuine ones, we could reply in each case by finding out whether there are elementary sentences having the logical form in question. But ordinary elementary sentences are supposedly not genuine ones, and thus their apparent logical forms cannot be any indicators of genuine logical forms; hence we cannot reply in this way. Note that we will not be able to establish the theory of symbolism if we fail to solve this second problem: The theory includes the specifications of the ways various types of symbols occurring in elementary sentences symbolize. But whence the guarantee that these types of symbols would really occur in genuine elementary sentences?

Therefore, Wittgenstein needs to solve the problems of the elementary sentence unless he is willing to give up the idea of analysis or the idea of the theory of symbolism. But he finds that satisfactory solutions are hard to come by. The main reason is due to a requirement which was already implicit in "Notes on Logic" and became explicit in Wittgenstein's remark in the 1914 *Notebook* that "logic must take care of itself." (Wittgenstein 1979, p.2) This is what I call the requirement of the *autonomy of logic*. The

requirement is that, as logic is the most fundamental theory, it cannot be based on any other theory, be it metaphysics, theory of knowledge, theory of meaning, psychology, or what not. Since logic is none other than the theory of symbolism for Wittgenstein, the autonomy of logic demands that the theory of symbolism should not be based on any further theory. For example, we have seen that Wittgenstein did not choose to develop semantics of empty names or lexical semantics in order to defend his assumption that every elementary sentence has a sense. Under the requirement that logic must be autonomous, however, Wittgenstein now *cannot* choose this option, for logic would then be based on semantics and this would directly undermine the autonomy of logic. Another strategy Wittgenstein cannot take is Russell's: Russell thought that elementary sentences were those whose terms only referred to what we were directly *acquainted* with. This strategy, however, bases logic on epistemology, for it takes the notion of acquaintance, which is an epistemological notion, to be most fundamental. Hence it also violates the requirement that logic must be autonomous.

So Wittgenstein's task is to give answers to the two problems of the elementary sentence purely by logic. I have argued elsewhere that it was in order to shed light on this task that Wittgenstein developed his famous analogy between sentence and picture (Kang 2007). I have also argued that there is ultimately no such thing as the "picture theory" of language in Wittgenstein's early philosophy, for his analogy between sentence and picture turns out to be a *mere* analogy and does not give any genuine answer to the problems of the elementary sentence. The genuine answer comes from what Wittgenstein calls the "class-theory" in the 1915 *Notebook* (Wittgenstein 1979, p.55). It is dubbed the "class theory" because it treats a sentence as a class of symbols sharing common features, and it turns out that these common features are the *inferential*

relations a sentence has to other sentences. Under the class theory, there are two most basic uses of a sentence, *asserting* and *negating*. That a sentence can be true or false means that it can be asserted or negated. Again, that a sentence can be asserted or negated means under the class theory that there are sentences that follow and those that do not follow from it. Two sentences assert the same thing, or have the same sense, if and only if the same set of sentences follow from each of them. So the sense of a sentence is entirely determined by its inferential relations to other sentences, and the meanings of sub-sentential parts by their roles in these inferential relations. Two sentences “*p*” and “*q*” are the same sentence if and only if they have the same logical consequences or, in other words, imply each other. It is evident that the class theory is an inferential account of language.

Two points should be immediately made, however, to forestall a misunderstanding: First, as in “Notes on Logic” and “Moore Notes”, every inference is truth-functional for Wittgenstein. So Wittgenstein's inferential account of language does not leave room for what Robert Brandom calls the “material inference”, an inference that does not solely depend upon the logical forms of sentences (Brandom 1994, pp.97-105). According to Wittgenstein, an inference from “A is red” to “A is not green”, for example, *is* a truth-functional inference in spite of its appearance. This inference does not look truth-functional simply because the genuine logical forms of the two sentences are hidden.

Secondly, Wittgenstein's conception of inference is not a *semantic* one, or at least not in the usual contemporary sense. It seems as if this was the case, for Wittgenstein seems to characterize inference in terms of the truth-table and this characterization seems to be a semantic one in that it explains inference in terms of the semantic notions of truth and falsehood. However, recall that Wittgenstein's

truth-table is not a semantic device; it is simply another notation, another way of writing a sentence. I have pointed out that this would not make sense unless the notion of sentence according to Wittgenstein already contains the notions of truth and falsehood. The class theory makes this point prominent: We have seen that, according to this theory, that a sentence can be true or false means that it can be asserted or negated, and that a sentence can be asserted or negated means that there are sentences that follow from it and those that do not follow from it. But a symbol cannot be a sentence under the class theory unless it has such inferential connections; hence a symbol cannot be a sentence unless it can be true or false.

Now, this does not mean that Wittgenstein attempts to explain truth and falsehood in terms of inference either. Rather, the relation between the notion of inference and those of truth and falsehood is reciprocal for Wittgenstein: that a sentence can be true or false *is* that there are sentences which follow from it and those that do not follow from it. Figuratively speaking, notions of truth, falsehood, assertion, negation, and inference all form a single whole, so that one cannot understand any of these notions independently of the others. But the assumption of explanatory priority of truth over inference occupies a central place in semantics in contemporary logic; hence Wittgenstein's conception of inference is not a semantic one.

Rather, the notion of inference is to be treated as *primitive* in Wittgenstein's class theory: If an inferential relation holds between two sentences, there is nothing further by which we can explain or justify why this is the case. Only under this view can we do a full justice to Wittgenstein's dictum that "logic must take care of itself," for otherwise inference must be further explained by some other theory.

We have so far seen what Wittgenstein's class theory is and what kind of conception of inference is operative behind this theory. The class theory provides answers to the two problems of the elementary sentence. We can think of a terminus in a chain of truth-functional inferences from all sentences in ordinary language to their logical consequences. If there is such a terminus, the logical consequences at the final stage will be truth-functionally independent of one another, for under this circumstance none of the consequences will imply nor contradict one another, and hence further inferences among them won't be possible. It is these sentences that are treated as elementary sentences in the class theory. That is, an elementary sentence under the class theory is a member of the maximally large set of the logical consequences of all sentences in ordinary language such that its members are truth-functionally independent of one another.

Note that this characterization of the elementary sentence is purely *logical*, for it is given in terms of inferential independence among elementary sentences and inference is precisely what logic is about. Also, the characterization provides a *substantial* guide for the analysis of ordinary sentences into elementary sentences. This is because we do have an *a priori* ability to trace inferential patterns among sentences. To use an example I introduced before, we know that neither "A is red" nor "A is green" is an elementary sentence as it is characterized in the class theory, even though we have no idea of what the logical forms of these two sentences are, for we know that the first sentence contradicts the second and hence that they are not truth-functionally independent of each other. According to Wittgenstein, it is this *a priori* ability of ours to trace inferential patterns without knowing the logical forms of sentences that enables us to find the set of genuine elementary sentences. In sum, the class theory gives a purely logical and substantial solution to the problem

of what elementary sentences are like.

What about the second problem of the elementary sentence, the problem of what the possible forms of elementary sentences are? We can simply think of the possible forms of elementary sentences as those that are possessed by the members of the set of the elementary sentences specified in the class theory. Hence the class theory also provides a solution to the second problem of the elementary sentence purely by logic alone.

§3. Wittgenstein's New Account of Logic and the Demise of Metaphysics

However, this solution to the second problem of the elementary sentence is not entirely satisfactory for Wittgenstein's purpose. Even though it gives us an account of how to find the possible forms of elementary sentences, it still does not tell us *what* they are. But unless we know what the possible forms of elementary sentences are, we cannot complete our list of logical constants, and we cannot establish the theory of symbolism either.

Along with other related problems, Wittgenstein struggled with the problem of what the possible forms of elementary sentences are until the end of the 1915 *Notebook*, but failed to come up with a definite answer (Wittgenstein 1979, p.61, p.65). Wittgenstein reaches the final answer only later, and it is a very radical one: Contrary to his view in "Notes on Logic" and "Moore Notes", Wittgenstein no longer thinks that we can establish the possible forms of elementary sentences *a priori*. Logical forms of elementary sentences are *not* necessary features of language:

It would be completely arbitrary to give any specific form.
(*Prototractatus*, 5.4012)

Clearly we have some concept of elementary sentences quite apart from their particular logical forms. (*Prototractatus*, 5.4221)

If I cannot say *a priori* what elementary sentences there are, then the attempt to do so must lead to obvious nonsense.
(*Prototractatus*, 5.4223)

These remarks in the *Prototractatus* occur almost verbatim in the *Tractatus* 5.554, 5.555, and 5.5571, respectively, so we can safely assume that they express Wittgenstein's final view. According to this view, we can no longer assume *a priori*, for example, that there will be an elementary sentence that has a subject-predicate form; there may be none. Indeed, this view seems to be a natural consequence of the class theory: Recall that, according to this theory, the possible logical forms of elementary sentences are whatever forms possessed by the members of the set of elementary sentences specified in the theory. Under the class theory, therefore, what the possible logical forms of elementary sentences are dependent on its specification of the set of elementary sentences. But this specification crucially depends on what kind of sentences we actually have in our ordinary language, given that the elementary sentences specified must be jointly sufficient to analyze sentences in ordinary language. Again, what kind of sentences we have in ordinary language depends on what the world is like, given that they must be jointly sufficient to represent the world. Therefore, the specification of elementary sentences in the class theory is ultimately dependent on what the world is like. But then it follows that the possible forms of elementary sentences cannot be established *a priori*. Hence they are not necessary features of the world.⁵⁾

I said that Wittgenstein's denial of the logical forms of elementary sentences as necessary features of the world was radical, for it has the following surprising consequence: It now turns out that what Wittgenstein calls the theory of symbolism is not an *a priori* theory, that it shows at best the *contingent* features of the world. Remember that the theory of symbolism consists of the specifications of the ways symbols of various types occurring in elementary sentences symbolize, whose types also determine the logical forms of elementary sentences. Therefore, if the logical forms of elementary sentences are the contingent features of the world, the specifications of the types of symbols in the theory of symbolism can no longer be treated as *a priori* specifications; the theory of symbolism becomes *a posteriori*.

But wasn't the theory of symbolism supposed to be logic? And wasn't logic supposed to be a study of the *a priori* features of reality? As Wittgenstein puts it, "logic *precedes* every experience - that something is *so*." (*Tractatus*, 5.552). So if the theory of symbolism is not an *a priori* theory, how could it be logic? Wittgenstein agrees. Logic can no longer be equated with the theory of symbolism, as the latter is not an *a priori* theory. The account of

5) The last transition from *a priori* to necessity will raise the eyebrows of the contemporary readers who are familiar with Saul Kripke's *Naming and Necessity*. However, given Wittgenstein's view that logic is purely *a priori* and that the only necessity is *logical* necessity (cf. *Tractatus*, 6.3), he would not accept Kripke's notion of a *posteriori* necessity.

There is also another complication. Since the set of Wittgenstein's elementary sentences provides the complete description of not only the actual world but also every possible world, the logical forms of elementary sentences will be the same in every possible world. Therefore, if we follow the possible world analysis of necessity, the logical forms of elementary sentences will indeed be the necessary features of the world. This consequence suggests that Wittgenstein's notion of necessity in the *Tractatus* should not be understood in terms of possible worlds.

logic in "Notes on Logic" and "Moore Notes" must be denied.

What accordingly emerges is a new account of logic. The account in "Notes on Logic" and "Moore Notes" took as basic various types of elementary sentences and the symbols occurring in them and then showed logical constants and inferential relations among sentences based on the specification of how these symbols symbolize. Now, Wittgenstein takes the notion of (truth-functional) inference as basic; there is no theory on which inference is based. Along with the notion of inference, those of truth, falsehood, assertion, negation, and other truth-functional connectives come altogether, as they are inseparable. So all of these still count as logical constants, as they did in "Notes on Logic" and "Moore Notes." What also count as logical constants are three cross-categorical notions, fact, atomic fact, and object, which correspond to three cross-type linguistic categories, sentence, elementary sentence, and name. Sentences are symbols of whatever types that can be true or false. Elementary sentences are symbols of whatever types that can be true or false and that are truth-functionally independent of other elementary sentences. Names are simple symbols of whatever types that occur in elementary sentences of whatever types. Under the new category of the name, therefore, now fall not only names but also relation symbols of various types in "Notes on Logic" and "Moore Notes".

On the other hand, what correspond to *logical forms of elementary sentences* and *types of symbols* occurring in them no longer count as logical constants. Wittgenstein now sharply distinguishes between logic *per se* and the *application* of logic, claiming that logical forms of elementary sentences and types of symbols can be discovered only as a result of applying logic to sentences in ordinary language (*Tractatus*, 5.557). As logical forms of elementary sentences and types of symbols can be discovered by the application of logic, Wittgenstein continues to call them formal properties or formal

concepts (e.g. *Tractatus*, 4.1274). Though they are formal, however, they are no longer treated as *a priori* and necessary features of language.

The upshot, then, is that we can now think of two layers of formal features: the lower one to which various logical forms of elementary sentences and types of symbols belong, and the higher one to which logical constants properly called belong. Only those formal features that belong to this higher layer belong to logic.

Now, what remain as logical constants under this new account of logic, i.e. truth, falsehood, assertion, negation, conjunction, disjunction, etc., are extremely general features. Indeed, Wittgenstein insists that they are not even parts of the logical forms of sentences: He calls truth-functional connectives *operations*, and claims in the *Tractatus* that operations are not the “marks” of the logical forms of sentences (*Tractatus*, 5.241). This claim surely strikes us as very bizarre. Shouldn't Wittgenstein treat the logical form of the sentence, say “ $p \vee q$ ”, differently from the logical form of the sentence “ $p \& q$ ”, whether they are written in this Russellian notation or in the truth-table notation where signs “ \vee ” and “ $\&$ ” are eliminated? But if truth-functional operations are not the marks of the logical form of a sentence, then “ $p \vee q$ ” and “ $p \& q$ ” will turn out to have the same logical form! In fact, Wittgenstein does not deny that operations can *affect* the logical form of a sentence to which they are applied, but only that they are *parts* of the logical form. So “ $p \vee q$ ” and “ $p \& q$ ” do have different logical forms, although truth-functional operations like disjunctive and conjunctive here do not belong to their logical forms. Now this may sound even more bizarre: how could an operation affect the logical form of a sentence without its being a part of the logical form? But Wittgenstein has a good reason to think so: Let us think of a sentence, “ p ”, and its double negation “ $\sim \sim p$ ”, for example. We have seen that the class theory treats any symbols

having the same inferential relations as the same sentence. So “ p ” and “ $\sim\sim p$ ” must be the same sentence under the class theory, and accordingly they must have the same logical form. But if one treats negation as a part of the logical form of a sentence, then “ p ” and “ $\sim\sim p$ ” would have different logical forms.⁶⁾ The same thing can also be said for the sentence pair “ $\sim p \vee q$ ” and “ $p \supset q$ ”, and in general for any pair of logically equivalent sentences. This is why Wittgenstein does not treat truth-functional operations as parts of the logical forms of sentences. As he puts it metaphorically in the *Tractatus*, “Operations can vanish” (5.254), whereas logical forms of elementary sentences cannot.

But if truth-functional operations are not even the marks of logical forms, and if they can “vanish”, it is no longer clear in what sense they can count as the features of the *world*. Again, think of the two sentences “ $\sim\sim p$ ” and “ p ”. They are exactly the same sentence, according to Wittgenstein, and hence they represent exactly the same fact. How could this be if the so-called feature of negation is a feature of the world, be it structural or formal or whatever?

So what still remain as logical constants in Wittgenstein's new account of logic are no longer treated as features of the world, although they are *a priori* and necessary. On the other hand, logical forms of elementary sentences and types of symbols occurring in them are no longer treated as *a priori* and necessary, although they are still formal features of the world. The upshot then is that there is no feature that is both a feature of the world and *a priori*. In the 1915 *Notebook*, Wittgenstein once asked: “The great problem round which everything that I write turns is: Is there an order in the world *a priori*, and if so what does it consist in?” (Wittgenstein 1979,

⁶⁾ I should add that these two sentences do count as having different logical forms in any notion of logical form available in contemporary logical theories, including the so-called LF in linguistics.

p.53). Now Wittgenstein's answer is: "There is no *a priori* order of things." (*Prototractatus*, 5.33546; *Tractatus*, 5.634).

Therefore, under Wittgenstein's new account of logic, logic no longer has any metaphysical implications. In fact, the very idea of metaphysics must be abolished, as there is no *a priori* structure of the world. Contrary to the view in "Notes on Logic" and "Moore Notes", the aim of logical analysis is no longer taken to reveal the *a priori* structure of the world. Rather, its aim now becomes a purely critical one, that of showing to someone who puts forwards metaphysical statements that he or she is deeply confused.

§4. The "Metaphysics" of the *Tractatus*?

An obvious question will arise at this point, however: Does not logic, under this new account, still have *some* metaphysical implications? It seems that not all ontological distinctions are gone, for we still have cross-type distinctions among sentences, elementary sentences, and names, which correspond to the distinctions among facts, atomic facts, and what Wittgenstein calls "objects" in the *Tractatus* which encompass things, properties, and relations of various types. Aren't these three ontological categories still perfectly legitimate metaphysical distinctions? And if so, shouldn't the features by which they are distinguished be the features of reality? Indeed, the metaphysical remarks in 1s and 2s of the *Tractatus* seem to be concerned precisely with the ontological categories of fact, atomic fact, and object, along with the distinctive features they have.

An adequate reply will require another paper, as it must include a detailed analysis of the remarks in 1s and 2s of the *Tractatus*. Let me just indicate how I will deal with this question. The examination

of these remarks shows that the features by which the ontological distinctions among fact, atomic fact, and object are distinguished correspond to the logical constants under Wittgenstein's new account of logic, which is just what is expected. But we have seen that no sense can be made of these logical constants being the "features of reality". Accordingly, the so-called ontological distinctions among fact, atomic fact, and object turn out to fall apart under closer scrutiny, as the "features" by which they are distinguished are not really the features of reality in any proper sense.⁷⁾

First, about facts: What are facts? Facts are what are represented by sentences, and sentences are distinguished from names in that they can be true or false. The ontological counterpart to this linguistic feature is that facts can exist or not exist. Therefore, the features by which facts are distinguished from objects are those of existence and non-existence. But are existence and non-existence really the features of a fact? It is hard to see how they could be: Firstly, when a fact exists, it is not as if the fact had a feature called "existence" in addition to whatever features it has; it is simply that there is the fact. Similarly, if a fact does not exist, it is not as if there was a fact that had a feature called "non-existence"; it is simply that there is no such a fact. Secondly, the same fact can exist or not exist. But if existence and non-existence were really the features of a fact, it couldn't be the case that a fact when it exists is identical with the fact when it does not exist, as the two facts possess different features.

The peculiarities I have just pointed out concerning the supposition of existence and non-existence as features of reality hold for an atomic fact as well, as it is also a fact. The supposedly distinctive

⁷⁾ I am indebted to Ricketts 1996 and Goldfarb 1997 for the general idea that the apparently intelligible ontological remarks in the *Tractatus* come to fall apart under closer scrutiny. The application of this idea in what follows is mine.

feature of an atomic fact turns out to be problematic, too: Recall that an atomic fact corresponds to an elementary sentence. An elementary sentence is what is truth-functionally independent of other elementary sentences. Hence the distinctive feature of an atomic fact is that it can exist or not exist independently of the existence and non-existence of other atomic facts. However, if it is doubtful to think of existence and non-existence as genuine features of reality, it is of course as much doubtful to think of this feature of “independence” as a genuine feature, for it is characterized precisely in terms of the features of existence and non-existence.

Because of these problems, it is very tempting to think of atomic facts, which Wittgenstein calls “*Sachverhalten*” in the *Tractatus*, as *possible* facts, and thus to think of its features of existence and non-existence as those of actuality and non-actuality. It seems that this is why Pears and McGuinness render “*Sachverhalten*” as “states of affairs” in their translation of the *Tractatus*, not as “atomic facts” as in Ogden's translation, and thus try to be neutral about whether Wittgenstein's *Sachverhalten* are facts or possible facts. But once we introduce the category of possible facts, we should of course assume that possible facts exist, only to maintain that there are possible facts. And the feature of existence in the sense of the existence of possible facts must be something more general than the feature of existence in the sense of the actuality of possible facts. If so, however, what Wittgenstein means by the existence and non-existence of *Sachverhalten*, understood as their actuality and non-actuality, cease to be the most general features. But this is an intolerable consequence, given that the features logic is concerned with are supposed to be the most general features at the highest level for Wittgenstein.

What about objects? Wittgenstein remarks that objects are “what is unalterable and existent.” (*Tractatus*, 2.0271; translation modified),

whereas facts are what is changing and variable. So he seems to claim that objects are the entities that necessarily exist, whereas facts are those that can exist or not exist. Again, we have seen that the feature of existence can hardly count as a feature of reality, and it is doubtful whether this feature of “necessary existence” would do any better. Apart from this problem, however, a further reflection reveals that it is very peculiar for Wittgenstein to talk about the existence of objects at all. Recall that the feature of the possibility of existence or non-existence of a fact is an ontological counterpart to the linguistic feature of a sentence that it can be true or false. So the ontological claim that an object “necessarily exists” corresponds to a linguistic claim that a name is necessarily true. But of course a name is not a linguistic item that can be true or false, let alone be necessarily true. So Wittgenstein's characterization of objects as what “is unalterable and existent”, while seemingly intelligible, leads to nonsense as we think through its implications.

Wittgenstein also remarks that objects are “simple” (*Tractatus*, 2.02). By this remark, he means that an object does not have parts that can further be decomposed by logical analysis. But this does not mean that an object does not have any features, which would directly undermine Wittgenstein's ascription of the feature of simplicity to an object. Wittgenstein distinguishes between material features and formal features, and claims that an object necessarily possesses only the latter. Therefore, an object is simple in the sense that the features it necessarily possesses do not include any material features but only formal features. Now, since the feature of simplicity is supposedly a necessary feature of an object, it must be one of the formal features of an object. It turns out, however, that what Wittgenstein means by the formal features of an object is the possibilities of its occurrence in atomic facts (*Tractatus*, 2.0141, 2.0231). Whatever the feature of simplicity is, it is certain that it is

not one of such possibilities of occurrence in atomic facts. Therefore the feature of simplicity does not belong to the formal features of an object. However, the only features necessarily possessed by an object are its formal features. The upshot, then, is that the feature of simplicity is the one that is necessarily possessed by an object but cannot be possessed by the object. Again, we find that an apparently intelligible claim about the simplicity of an object turns out to be a mere piece of nonsense.

§5. Conclusion

I have argued that, upon closer scrutiny, the features that supposedly distinguish the ontological categories of fact, atomic fact, and object cannot be taken to be the features of reality, and hence that the distinctions among these categories in 1s and 2s of the *Tractatus* become unintelligible. At *this* level of generality, it seems indeed nonsensical to think of the “features of reality” or “categorical distinctions” based upon them. My argument was rather sketchy, but it seems to me to be a promising start. Even if my argument does not turn out to be exactly right, the initial question I raised about how to make sense of the change in Wittgenstein's attitude on metaphysics still remains. The changes I have discerned concerning Wittgenstein's account of logic also remain, and I do not see how one would be able to make sense of the change in Wittgenstein's attitude on metaphysics without invoking the changes in his account of logic.

I will now conclude with brief remarks on the implication of my discussion for the recent debate about how to read the *Tractatus*. Over the last fifteen years or so, there has emerged an interpretation

of the *Tractatus* which takes its central message to be a radically anti-theoretical and anti-metaphysical view of philosophy. According to this “resolute” or “therapeutic” reading of the *Tractatus*, the statements in the *Tractatus* should not be thought of as advancing any substantial philosophical theses or theories, be they about language, the world, the self, or what not. Rather, these statements should really be taken as plain, genuine nonsense, as Wittgenstein notoriously urges in the penultimate remark of the *Tractatus*.

As one may guess from my “deflationary” interpretation of the ontological remarks in the *Tractatus*, I am quite sympathetic to this line of reading. I agree that the aim of the *Tractatus* is not to advance any philosophical theses that can answer philosophical questions. In my view, Wittgenstein's real aim in the *Tractatus* is to make readers recognize *via* the logically perspicuous notation he proposes toward its end that there can be no philosophical theses, including the ones in the *Tractatus*. What the readers are expected to realize from such recognition is that the very assumption that philosophical questions are genuine ones is an illusion created by our ordinary language, an illusion that reflects human reason's deeply-rooted tendency to transcend its limits while seeking after the most universal and absolute truths about reality and ourselves. This, I think, is the ultimate message Wittgenstein wants to deliver in the *Tractatus*.

Despite my general sympathy, however, I think that various versions of the resolute reading currently pursued by a number of commentators, most notably by Cora Diamond and James Conant (Diamond 1988, Conant 2001), share a common problem. While these commentators attempt to articulate a non-metaphysical reading of the *Tractatus*, none of them recognizes the fundamental changes in Wittgenstein's account of logic. Accordingly they attribute to the *Tractatus* an account of logic that is essentially the one I find in

“Notes on Logic” and “Moore Notes”. Most importantly, resolute readers assume that Russellian type distinctions are still retained in Wittgenstein's logic in the *Tractatus*. But then it is difficult to see how they could defend a non-metaphysical reading of the *Tractatus*, as Wittgenstein himself thought that the account of logic in “Notes on Logic” and “Moore Notes” did have metaphysical implications. Unless resolute readers give up the assumption that Russellian type distinctions belong to the logic of the *Tractatus*, I do not think that they will be able to even *begin* to develop a satisfactory non-metaphysical reading of the *Tractatus*.⁸⁾

⁸⁾ I am grateful to Warren Goldfarb, Richard Heck, Charles Parsons, and Thomas Ricketts for helpful questions and comments. This paper is dedicated to Professor Myung-Hyun Lee in honor of his retirement.

References

- Brandom, Robert(1994), *Making It Explicit*(Cambridge, MA: Harvard University Press).
- Conant, James(2001), "The Method of the *Tractatus*", in E. Reck, ed., *From Frege to Wittgenstein*(Oxford: Oxford University Press), pp.374-462.
- Diamond, Cora(1988), "Throwing Away the Ladder: How to Read the *Tractatus*", *Philosophy* 63, rep. in Diamond, *The Realistic Spirit*(Cambridge: MIT Press, 1991), pp.179-204.
- Goldfarb, Warren(1997), "Metaphysics and Nonsense: On Cora Diamond's *The Realistic Spirit*", *Journal of Philosophical Research* 22, pp.57-73.
- Kang, Jinho(2007), "'Picture Theory' of Language?", unpublished manuscript.
- Ricketts, Thomas(1996), "Pictures, Logic, and the Limits of Sense in the *Tractatus*." in H. Sluga and D. Stern, ed., *The Cambridge Companion to Wittgenstein*(Cambridge: Cambridge University Press, 1996), pp.59-99.
- Russell, Bertrand(1996), *Principles of Mathematics*(New York: W.W. Norton & Company).
- Wittgenstein, Ludwig(1922), *Tractatus Logico-Philosophicus*, trans. C.K. Ogden(London: Routledge).
- _____ (1961), *Tractatus Logico-Philosophicus*, trans. D. Pears and B. McGuinness(London : Routledge).
- _____ (1979), *Notebooks 1914-1916*, 2nd edition, ed. G.H von Wright and G.E.M. Anscombe, trans. G.E.M. Anscombe(Oxford: Blackwell).
- _____ (1996), *Prototractatus: An Early Version of Tractatus Logico-Philosophicus*, 2nd edition, ed. B.F. McGuinness, T. Nyberg and G.H. von Wright(London: Routledge).