Tarski Lectures

Per Martin-Löf

The 24th annual Alfred Tarski lectures
Three lectures given at the University of California, Berkeley
21, 22 and 24 February 2012

Transcriber's note

I am indebted to Paolo Mancosu for tracking down and sharing with me the video recording on which this transcription is based. He promptly responded to my inquiry of late September 2025 and, moreover, obtained confirmation from UC Berkeley that Per Martin-Löf owns the copyright to the lectures. The video was prepared by the organizers.

Ansten Klev

Contents

Assertion and inference

21 February 2012

First of all I would like to thank you for giving me this opportunity of speaking in honour of Alfred Tarski. There was no hesitation on my part in accepting this, because I have actually done some work recently that has directly to do with Tarski's treatment of logical consequence in particular, based on his previous work on the notion of truth. That gives a natural filling of these three lectures, because I want to reach Tarski in the last lecture, Tarski's notion of logical consequence. One of the things that I hope to make you understand is that before Tarski's metamathematical notion of logical consequence there is the ordinary notion of logical, or formal, consequence—pre-mathematical, if you want, but I think I will call it the ordinary notion of logical consequence. That notion Tarski simply took for granted. In the second lecture, or maybe towards the end of this lecture, I will give a precise explanation of it.

A further notion is needed in explaining the ordinary notion of logical consequence, namely the notion of inference. Since inference is needed in explaining consequence, inference has to be treated first, contrary to the standard way in ordinary textbooks of the metamathematical style. With the notion of inference, we are so close to the beginnings of logic that we might as well start from the very beginning, and ask, What is logic? How is the subject to be delimited?

The word logic comes from the Greek ἡ λογική—that is the adjective λογικός substantivized in the feminine, meaning that it was routinely taken to be followed by τέχνη or ἐπιστήμη, art or science, so logical art or logical science. This does not go back to Aristotle, but it is from the Ancient time, not too long before the birth of Christ. We know how it is understood, because already at the time of Cicero, it was clear that by logic one understood *ars logica*, so the art of reasoning—whether we say the art of reasoning or the art of inference does not matter much, but the standard formulation is, no doubt, the art of reasoning.

I think neither art nor science is good here. Art is fine for skilled activity, as is taught in elementary logic courses, both in philosophy and mathematics:

these are, basically, training courses helping the students to reason correctly. That is, however, not what we primarily have in mind when we speak about logic. Science, on the other hand, is no good at all, because science is ordinarily associated with demonstrative science, and logic is definitely not demonstrative primarily—rather, logic is the theory of demonstrative science: it brings out the structure of demonstrative science.

I think a less committing term like, simply, the study of reasoning or, best of all, the theory of reasoning, is a better choice than either art or science. A piece of reasoning is a gapless chain of immediate inferences. The notion of reasoning thus presupposes the notion of inference, hence it is an improvement in the formulation to say that logic is the study, or the theory, of inference—but that is such a small improvement that it does not matter much whether you say the one or the other.

What is an inference then? An inference occurs when we pass from certain assertions, or judgements, that we have previously made to a new assertion, or judgement, which becomes justified precisely by the inference that we are performing. From this explanation it is clear that the notion of inference is not the very first notion, because we need to speak first about assertion, or judgement. An inference is a passage from premisses, which are all assertions, or judgements, to a conclusion, which is a further assertion, or judgement. We should therefore say something like, Logic is the study of—whether we say assertion, or judgement—and inference: assertion, or judgement, has to come first, before inference.

You have already noted a hesitation here on my part, namely whether to use the term assertion or the term judgement. On this point I am really ambivalent. The term judgement was the standard term in the whole era of modern philosophy, beginning with Descartes via Kant up to Frege and Husserl. How come that we lost it? That is, essentially, because of the revolt against British idealism about a hundred years ago. Russell started to use the excellent term assertion already in the *Principles of Mathematics* from 1903. The term judgement was really a hot potato at the time in British philosophy. Moore's famous paper "The nature of judgement", which was his revolt, was from 1899. The term judgement thus got out of use, but now a hundred years have passed, and it is not so hot any longer, so we are free to use which one of these terms we prefer.

There is a well-established convention, from several hundred years ago, of using judgement for the mental act, as well as for the object of that act, and on the other hand, assertion for the verbal expression of the mental. One might thus say, Fine, we have both these terms, so let us use judgement for the mental and assertion for the linguistic, the verbal. The trouble with that solution is that it does not fit the needs of logic. As logicians we are equally interested in the laws of thought as we are in the laws of reasoning, whether in speech, or in writing, or even, nowadays, by the aid of a computer system, like the present computer implementations of type theory. We are using different media, and the important insight of logic is that the rules are the same: we do not care if it is mental or expressed in the verbal medium.

Hence, if we were to stick to this standard usage, then we would somehow lose both terms, because we want a term that is neutral. I think that is the reason for my own oscillation here. I am used to using these terms interchangeably, and if we need in some situation to qualify ourselves by saying that we are particularly interested in how this is performed mentally, then we can speak of the mental judgement, or assertion, and conversely we can speak of the verbal one. I will therefore not stick to this practice of using judgement for the mental and assertion for the verbal, for the reason that I just said. Nor am I going to take any stand on the order of priority between language and thought, because I need not. The stand that I have already taken is that logic deals with both, and the laws are the same—just as in elementary arithmetic, we do not care if the calculations are performed in our minds, so we are doing mental arithmetic, or if we are doing them in the standard way on the blackboard or whatever medium we use.

We have arrived at the definition of logic as the study, or theory, of assertion and inference, and the question is, To whom is that due? As the study of reasoning—that is ancient, as I have already said, but this formulation cannot be traced as far back as that. Unexpectedly, the earliest place I know of is in Bradley and Bosanquet, precisely the British idealists that Russell and Moore revolted against. They both took judgement and inference to be the main headlines for logic. Still in Britain, when Cook Wilson's posthumous papers were edited in 1926, one chose as title for those two volumes *Statement and Inference*, which amounts to the same in this case, but just reflects Cook Wilson's realism, I mean, Oxford realism—it is again the revolt against the previous idealism, but whether we say assertion and inference or statement and inference does not matter in this connection.

I said a moment ago that the term judgement is tied to the era of modern philosophy from Descartes up to Husserl. During the whole of that era, epistemology was at the centre, so the question is, What is the relation now between assertion, or judgement, and the concept of knowledge, more precisely demonstrative knowledge? To my mind it is this, that demonstrative knowledge is to be equated with reasoned judgement, or assertion:

$$\text{demonstrative knowledge} = \left.\begin{array}{r}\text{reasoned}\\ \text{grounded}\\ \text{justified}\\ \text{demonstrated}\end{array}\right\}\begin{array}{l}\text{judgement}\\ \text{assertion}\end{array}$$

Whether we say reasoned or grounded or justified or, as in mathematics, demonstrated does not matter: the important thing is that we have one term here for what justifies or grounds an assertion, or judgement. In mathematics, of course, there is only one word for what justifies an assertion, namely proof, or demonstration. A demonstration is a chain of immediate inferences, and an immediate inference is one in which, given the premisses, the conclusion neither needs nor is capable of any further demonstrative justification, but rather is justified because of how the concepts involved in the formulation of the premisses and the conclusion have been defined. In the more modern formulation, the conclusion of an immediate inference is justified in virtue of the meanings of the terms in terms of which the inference is formulated.

If this is the relation between demonstrative knowledge and the concept of assertion, or judgement, then you see that we have one term, demonstrative knowledge, with two components in it. These two components, assertion and demonstration, were the ones that made up my previous formulation of what logic is, so we can actually, if we want, change it into the more compact formulation that logic is the theory, or study, of demonstrative knowledge, of the system of demonstrative knowledge. This, to my mind, is the reason why Bolzano was right in calling his logic *Wissenschaftslehre*, correctly translated into English as *Theory of Science*. One would have expected it to be called his *Logic*, because it was his logic, but it is actually called *Theory of Science*. One explanation for the choice of title, I take it, is that after Fichte had published his *Wissenschaftslehre* Bolzano wanted to show what a decent Wissenschaftslehre ought to look like, hence he chose his title for the big work.

You see that the definition of logic that I have arrived at is given in epistemological terms. It is very instructive to see what happens if we try instead to define logic in objective or ontological terms. What would you say? What objects is logic about? Well, you would start by saying that it is about propositions, truth and consequence, which I have chosen for the title of my second lecture. That is a good first approximation, but on second thought you realize, That is just propositional logic, much too limited—we must include at least predicate logic also, so logic deals then also with individual domains, and functions from an individual domain into itself, and with propositional functions of any number of variables over an individual domain. The quantifiers presuppose these latter items. Now I have mentioned some other objectivities: could this serve as a definition of logic, that it is about these? No, the same problem arises immediately: what about arithmetic? Zero, successor, induction and recursion are not included in the previous list, but we need that also. Now we have perhaps something that is adequate for arithmetic, but then that is not all in logic: we also deal with transfinite or general inductive definitions, so then comes all the concepts connected with that. This is still not everything, because there are other ideas, like Grothendieck universes, whether in the sense of ZF set theory or in the sense of type theory: a universe in that sense is a new idea. These considerations are enough to make clear that, however we try to delimit logic in objective, or ontological, terms, we do not succeed: we merely succeed in delimiting a particular logical system which is not all-encompassing.

If one were contacted by Webster Dictionary to write the lemma on logic, it could perhaps look something like this:

> *logic*, n.
>
> 1 (uncountable) the study of judgement, or assertion, and inference in general
>
> 2 (countable) a specific system of forms of judgement and inference, like propositional logic, predicate logic, etc.

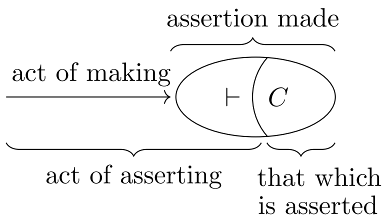
We have two senses of logic. In the first sense, logic is not plural-forming, but then there is the more modern second sense, where it is countable: a logic refers to a system of forms of judgement and inference. The difference between these two senses now is that the first is epistemological, whereas the second is given in ontological terms. It is clear that we have an order of priority here, which we can do nothing about, namely that the first has to come before, because it is presupposed by, the second. If we take the term logical system, then the adjective logical that I use there is precisely the adjective formed from logic in sense one, because it is a system of judgement and inference. Sense one is thus, no doubt, prior to sense two.

This gives already an indication of something which I want to get back to later, namely that we have an order of priority between the epistemological and the ontological, namely that the epistemological notions have to come before the ontological notions. This will come out more clearly in the sequel in the form that when explaining the ontological notions, the epistemological notions of assertion and inference have already to be in place. In particular, as I said already, in the case of the notion of logical consequence, which is an ontological notion—we need inference in order to explain it.

Now let me go over to a more systematic treatment of this. Where to start from then? The only viable starting point that I can see here is to go back to the act/object dichotomy, of Scholastic origin, so the notion of object here is the Scholastic notion of object: when I wish something, there is that which I wish and my wishing it, and when I fear something, there is the object of my fear, which may be the same as someone else's wish. A closely related dichotomy is the process/product dichotomy, which is, above all, associated with Twardowski—it is the centenary this year of Twardowski's well-known paper on the subject.

I am starting from this in order to narrow it down to the notion of speech act. A speech act is a clear example of a productive act, and its product is the complete sentence that we utter. Only by uttering a complete sentence can we say something meaningfully: it is, so to say, the unit by means of which we perform a speech act. We have an enormous variety of different acts, whether we think of them as speech acts or as acts of thinking: acts of thinking was the old way—Descartes described the cogito, after all, not the dico—but with Austin it was speech acts that came into focus instead. In agreement with what I said previously, however, it does not matter for logic whether we are dealing with acts of thinking or speaking as long as the structure is the same.

Among this variety of speech acts, or acts of thinking, perhaps the most basic one of all—although I have no good argument for that—is the act of asserting, or judging, with the correlated assertion, or judgement. Let us look at the act/object distinction and the process/product distinction in the case of assertion.



From the process/product point of view, there is an act of making, and the product is that which is made, so in the case of assertion, it is the assertion made. The assertion made consists of two parts, namely what Frege called the force and the content. In the case of assertion, it is standard to use the assertion sign, $\vdash$, for the force, and then we have the content, C. The process is the act of making, and the product is the assertion made—assertion because that is the force which is present here. This is the process/product dichotomy. If we instead start from the act/object dichotomy, we get the division line after the assertion sign instead, so we get the act of asserting, and what remains then is that which is asserted, which is the content.

In general, if we perform a speech act, then there is the speech product, which is a complete sentence. If we look at the structure of the complete sentence, the first that we meet is the force/content structure, which it is natural to write as

$$\text{force(content)}$$

since the force is unsaturated and needs a content in order to form a complete sentence—Austin may have been the first to write it in this functional form, $F(C)$. That is the first structure that we meet. There are many forces—I have given examples of some of them—and for each force, there will have to be a semantical explanation of that force, laying down under what condition we have the right to utter a sentence of the form in question.

I am not going to treat any other forces here than the assertoric force. That corresponds to a distinction between logic in the wide and the narrow sense. I mean by logic in the narrow sense the logic of assertion, which deals only with assertions, whereas it is quite natural to conceive of logic in the wider sense as dealing with all kinds of forces, in particular, questions and commands and conjectures and doubts and so on. Of course, assertion is the main object and is, at the present, infinitely better treated than these other forces, so I will limit myself to the assertoric force.

We are faced with the semantical question, How are the two components, the assertoric force, $\vdash$, and the content, C, to be explained? To my mind, the meaning of the assertoric force, that is, of the assertion sign, is given by stipulating that the condition under which you have the right to make a judgement, or an assertion, is that you have justified, or grounded, or demonstrated, it. When I make a judgement, even if it happens to be correct in the sense that someone else might have correctly made the same judgement—if I do not have grounds for it, then I am violating this rule which determines the meaning of the assertion sign. A more captious way of formulating it is negatively: do not make any ungrounded assertions. For an intuitionist, the negative formulation is not so good, and that is why I put it in the positive form: in order to have the right to make an assertion, you must have good grounds for it. That is the semantic explanation of the assertoric force.

What about the other component, the content? How is the content defined? It has to be the complementary thing, namely by laying down what it is that you must know, which is to say, have justified, grounded or demonstrated, in order to have the right to make it. What knowledge is it that you must have in order to have the right to make it? What that is in a particular case will depend on the form of the content in turn, so then we are beginning to go into the inner structure of the content. Before we go into the inner structure of the content, the only thing we can say in general is what I just said, namely that the content is defined by laying down what it is that you must know in order to have the right to make an assertion with that content.

Now we have, so to say, peeled off the outermost meaning component of an assertion, and what remains is the content. If we are going to say anything more than what I just said about the content, we cannot do that in general, but we have to go into the inner structure of the content. That means that we are now precisely at the borderline between the epistemological part of logic, with which I had to begin, and the ontological part of logic. The step from the epistemological to the ontological occurs precisely when we pass from the force/content structure to the inner structure of the content: that is when the ontological notions begin to appear.

I said already that I am going to limit myself to the assertoric force, but maybe I should say that if one considers logic in the wide sense, where one has, not only the assertoric force, but also other forces, then epistemology is not sufficiently wide to include the other forces, because epistemology is tied to the assertoric force. If you allow other forces, like conjecture and doubt and wish and so on, then you will have to use another term than epistemology, and the best that I can find is to say pragmatics: pragmatics would be defined as the theory of the various forces, whereas ontology is the theory of content.

Now let us start going into the content. That has to be done by starting to display at least one form, but presumably several forms, of content, and they have to come in a particular order, some of them before others. I will take a rather big step now and display directly the logical-consequence form of judgement, so that we can see what we have to do before actually coming to it. Since the lecture as a whole is centred around the notion of logical consequence, let us display the form—I forgot to make a remark regarding the term logical consequence, so let me do that first.

I distinguished the ordinary notion of logical consequence from Tarski's metamathematical reconstruction of it. What was the ordinary notion of logical consequence, and what was the term used for it? As far as I know, the term logical consequence is from Tarski: I have not seen any use of logical consequence before Tarski. The term that was used before, and Tarski was well aware of that, because he says in at least one place in the consequence paper from 1936, "logical, i.e. formal consequence"—Tarski was well aware that these terms, formal versus material consequence, were Scholastic terms, which had been there for a very long time. What I just quoted shows that Tarksi equated logical consequence with formal consequence, so he might just as well have chosen the term formal consequence.

Formal consequence is implicit in Aristotle's syllogistics. The very term consequence is from the Latin translation of the Greek ἀκολούθησις, which is in Aristotle's *Organon*, in particular in the *Prior Analytics*. It is there in maybe only one place, so it is not a common term, but it is the term that was translated into *consequentia* in Latin by Boethius. That the notion of formal consequence is implicit in Aristotle is quite clear if you look at his definition of the syllogism and his whole treatment, showing which syllogisms are valid and which syllogisms are not valid by giving counterexamples: the syllogisms that are valid are precisely those that are expressed by formal consequences. I do not think it is correct to say that it is more than implicit in the *Prior Analytics*. It became a chapter of its own in logic only in the 1400s, a chapter standardly called *De consequentiis*, where we get the distinction between formal versus material consequence that Tarski used. Such chapters begin to appear with Ockham, and Buridan is particularly important here. What is more well known in connection with Tarski is Bolzano's notion of Ableitbarkeit. It is astonishing that Bolzano, with his very thorough historical references, finds this new term and does not use the Scholastic terminology. I think the only explanation is that, on this point, he simply did not have access to the Scholastic treatises on consequences, so Bolzano seems to be independent from the Scholastic sources. In all these cases, it is what I called the ordinary notion of consequence, formal or material, that is at stake, not the metamathematical notion.

A formal consequence says that one proposition, A, is true provided finitely many other propositions are true, however certain parts are varied:

$$A \text{ true } \ (x_1 : \alpha_1, \ldots, x_m : \alpha_m, A_{m+1} \text{ true}, \ldots, A_n \text{ true}) \qquad \text{(Con)}$$

We vary finitely many parts, $x_1, \ldots, x_n$, and they have to be typed, and then we have the finitely many other propositions, $A_{m+1}, \ldots, A_n$, whose truth guarantees the truth of A. The variables have to be typed, because we have to vary the individual domains and the predicates and the function constants. We see then the need for dependent types: the predicate constants will get a type that will depend on an individual domain and similarly with the function constants. I will therefore use dependent type theory—whether you want to call it constructive or intuitionistic or dependent type theory, that is all the same—in order to analyze the notion of formal, or logical, consequence, because it has got the conceptual machinery needed for this variation.

When the German translation of Tarski's *Wahrheitsbegriff*, originally published in Polish in 1933, had been made in 1935, Tarski added a rather long Nachwort. How I interpret this is that up to that time he had taken Leśniewski's doctrine of semantical categories quite seriously and had more or less identified that with simple type theory. Whether that is correct or not, I cannot judge, because I do not know Leśniewski enough. I would rather think that there are differences between them, but at least Tarski opted for simple type theory in the late 1920s when it got formulated and presented, in particular, in Carnap's *Abriss der Logistik* from 1929. In 1935, when he wrote this postscript, it was clear that he was already disenchanted with simple type theory, and the reason, I take it, was that it is unable to deal with the variation over all structures that you need when you define logical consequence. The simple theory of types does not have dependent types, so you cannot do that in any simple way. Tarski therefore switched, at this time, from simple type theory as his basic framework to ZF set theory: in set theory we do not have problems speaking about all structures, because we have the set-theoretic universe. This is at least my interpretation of what Tarski says in this postscript. Now, with dependent type theory, the situation is different, because we can make this variation.

As you can see, I am using the colon for the copula. That is now the standard notation, although it was not my first choice. I thought originally that one ought to honour Peano by using his ε, for ἐστί(ν), for the copula, but when this new form of type theory came into existence through De Bruijn's and my work, it was apparently felt that the epsilon was so strongly tied to ZF-style set theory that it was no good to use it. De Bruijn therefore introduced the colon instead.

Since I have displayed the form of judgement (Con), let me also explain what it means before we finish today. What do we expect as the meaning explanation of this form of content? We have to lay down what it is that we must know in order to have the right to make an assertion, or judgement, of this form. The answer is that we must know a proof of A from the assumptions $A_{m+1}, \ldots, A_n$, a proof which, moreover, is a free-variable proof with respect to the variables $x_1, \ldots, x_m$. In the informal but familiar notation of natural deduction, the proof could be written perhaps like this:

$$\begin{array}{ccc} \overset{x_{m+1}}{A_{m+1}(x_1, \ldots, x_m)} & \ldots\ldots\ldots\ldots & \overset{x_n}{A_n(x_1, \ldots, x_m)} \\ & \vdots & \\ & A(x_1, \ldots, x_m) & \end{array} \qquad \text{(Ded)}$$

The proof starts from the assumptions $A_{m+1}, \ldots, A_n$, and not only A, but also $A_{m+1}, \ldots, A_n$ may depend on the variables $x_1, \ldots, x_n$, as indicated in the diagram. In natural deduction the assumptions are labelled, and I use the labels $x_{m+1}, \ldots, x_n$. This notation is still defective, since it has no indication of the types of the variables. One could of course write those as superscripts, as one usually does in type theory, but it is difficult to work with that notation, so I am writing it in this way, which is more familiar.

My meaning explanation was that in order to have the right to make a judgement of the form (Con) we should have a deduction of A from the assumptions $A_{m+1}, \ldots, A_n$ that is, moreover, a free-variable deduction with respect to the variables $x_1, \ldots, x_m$. In type theory the whole deduction (Ded) becomes instead denoted by a single symbol, a, which is a proof of A depending on the variables x_1, of type α_1, up to x_m, of type α_m, and depending, moreover, on assumed proofs x_{m+1} of A_{m+1} up to x_n of type A_n,

$$a : A \ \ (x_1 : \alpha_1, \ldots, x_m : \alpha_m, x_{m+1} : \mathrm{pr}(A_{m+1}), \ldots, x_n : \mathrm{pr}(A_n))$$

So in type-theoretical notation, the whole figure (Ded) is written by means of a single letter, a, or if we wish to indicate the variable dependencies,

$$a(x_1, \ldots, x_m, x_{m+1}, \ldots, x_n)$$

The meaning explanation of (Con) is that in order to have the right to make an assertion of that form, you need to have in your possession such a proof a.

Proposition, truth and consequence

22 February 2012

Remember where we were at the end of yesterday's lecture. I introduced the form of judgement that expresses that a proposition, A, is a formal, or logical, consequence of some other propositions, $A_{m+1}, \ldots, A_n$.

$$A \text{ true } \ (x_1 : \alpha_1, \ldots, x_m : \alpha_m, A_{m+1} \text{ true}, \ldots, A_n \text{ true}) \qquad \text{(Con)}$$

That is simply a form of judgement introduced into constructive type theory. Moreover, I not only introduced the form, I also gave the meaning explanation for it. According to the general explanation of what a judgemental content is, one has to lay down what you must know in order to have the right to make a judgement of this form. The answer was that in order to have that right, or be allowed to do that, you need to possess a proof of A which depends on the quantities, $x_1, \ldots, x_m$, that are varied, and also on arbitrarily given proofs of the propositions $A_{m+1}, \ldots, A_n$ which I called $x_{m+1}, \ldots, x_n$.

$$a : \text{pr}(A) \ \ (x_1 : \alpha_1, \ldots, x_m : \alpha_m, x_{m+1} : \text{pr}(A_{m+1}), \ldots, x_n : \text{pr}(A_n)) \qquad \text{(P)}$$

If that is the meaning explanation, then it is clear that the inference rule with (P) as premiss and (Con) as conclusion is valid:

$$\frac{a : \text{pr}(A) \ \ (x_1 : \alpha_1, \ldots, x_m : \alpha_m, x_{m+1} : \text{pr}(A_{m+1}), \ldots, x_n : \text{pr}(A_n))}{A \text{ true } \ (x_1 : \alpha_1, \ldots, x_m : \alpha_m, A_{m+1} \text{ true}, \ldots, A_n \text{ true})}$$

If you know the premiss, you have the right to make the judgement in the conclusion. This can just as well be turned around: we can just as well say that the meaning of the judgement that occurs in the conclusion is determined by this rule. In the first way, one gives a meaning explanation which is phrased in such a way that it sounds as if it is independent of the rules, but then, once you have written this rule down, you realize that you may as well say that the meaning of the form of judgement in the conclusion is determined by this rule. This is a pattern that is followed by all other meaning explanations, so we shall have many examples of this.

I have not said what the presuppositions are of the judgements (Con) and (P). They are as follows:

$$\begin{aligned}
\alpha_1 &: \text{type} \\
&\vdots \\
\alpha_m &: \text{type} \;\; (x_1 : \alpha_1, \ldots, x_{m-1} : \alpha_{m-1}) \\
A_{m+1} &: \text{prop} \;\; (x_1 : \alpha_1, \ldots, x_m : \alpha_m) \\
&\vdots \\
A_n &: \text{prop} \;\; (x_1 : \alpha_1, \ldots, x_m : \alpha_m) \\
A &: \text{prop} \;\; (x_1 : \alpha_1, \ldots, x_m : \alpha_m)
\end{aligned}$$

That the types and propositions are allowed to depend on variables is essential if we want to have generality with respect to all first-order structures.

What do I mean by presuppositions? That really belongs in the first part of yesterday's lecture. Immediately after you have introduced the notion of assertion, or judgement, you should say that, in general, an assertion, or judgement, has presuppositions—presupposition means that certain other judgements have to have been established in order for the judgement in question even to make sense. The most typical example is if you say that a is an element of a set A—this presupposes that A is a set. If you say that the proposition A is true, then it is presupposed that A is a proposition, and if you say that f is a function from a set A into a set B, then it is presupposed that both A and B are sets.

The numbers m and n in (Con) are arbitrary, except that m is less than or equal to n. There are several cases we may distinguish. When both m and n are zero, then we have the rule

$$\frac{a : \text{pr}(A)}{A \text{ true}}$$

This rule is meaning determining for the categorical form of judgement A true, in which you hold a proposition, A, to be true. If m is zero, but n is bigger, then we have no variables, but we have hypothses, or assumptions, so we get the ordinary notion of consequence. If m is greater than zero and equal to n, then we have no assumptions, only the variables, so we get logical, or formal, truth. Finally, if we have both variation and assumptions, then we get logical consequence.

$0 = m = n$	truth
$0 = m < n$	consequence
$0 < m = n$	logical truth
$0 < m < n$	logical consequence

It should be observed here that it is not that you first have the definition of categorical truth and then use that to define consequence. It is rather the other way around, that you have the consequence judgement that you explain in general, as I did it, and it specializes to the case where both m and n are zero.

In these meaning explanations you see an example of what I tend to call insertion, respectively deletion, or suppression, of proof objects. Giovanni Sambin at Padua calls this the forget-restore principle—forget corresponds to delete and restore to insert. In the logical-consequence form (Con) there are no proof objects of the proposition A. We just have the locution that a proposition is true, both in the thesis and in the hypotheses. But when I explained it, what I did was to insert proof objects, both in the thesis and in the hypotheses, following the rule that in the hypotheses you have to insert variables, whereas in the thesis, the positive part, you have to insert an in general complex proof object. That is an example of insertion of proof objects. In the other direction you have suppression of proof objects. What is important whenever you suppress proof objects is that you must be able to insert them again, since it is the full judgement, with the proof objects inserted, that forms the rock bottom of the semantics.

A categorical judgement of the form A true is the most important case of suppression of proof object. We have a proof of A, but we do not give it explicitly, we just say that A is true. In order to have the right to do that, we must have access to a proof, in one way or another. Let us say that A is a disjunction, $P \vee Q$. If I assert that A is true, then someone else might come and ask me whether it is P or Q that is true. I shall not be able to tell him unless I have access to the proof object, which is a program that allows me to do that, because of how disjunction is defined.

By insertion of the proof objects, we reduced the form of judgement (Con) to the form (P). It remains, of course, to explain (P), but now we are much closer to the bottom, because the basic forms of judgement of constructive type theory are these:

$$\alpha : \text{type} \;\; (x_1 : \alpha_1, \ldots, x_n : \alpha_n) \tag{1}$$

$$\alpha = \beta : \text{type} \;\; (x_1 : \alpha_1, \ldots, x_n : \alpha_n) \tag{2}$$

$$a : \alpha \;\; (x_1 : \alpha_1, \ldots, x_n : \alpha_n) \tag{3}$$

$$a = b : \alpha \;\; (x_1 : \alpha_1, \ldots, x_n : \alpha_n) \tag{4}$$

The form (P) is an example of (3), because of the rule

$$\frac{A : \text{prop}}{\text{pr}(A) : \text{type}}$$

If A is a proposition, then $\mathrm{pr}(A)$ is a type, namely the type of proofs of A—I am not writing out possible assumptions here, but use the more informal natural-deduction style. Once we have this rule and remember the presuppositions, then it is clear that (P) is an instance of (3).

Let me at least sketch the semantical explanations that you have to give for the basic forms. That begins with the first two forms, (1) and (2), which concern types and type equalities. I should say that these equalities are what Gödel called in his *Dialectica* paper intensional, or definitional, equalities. I will pay as little as possible attention to them in these lectures, so you could put a parentheses around (2) and (4) and concentrate on (1) and (3).

For the first form of judgement,

$$\alpha : \mathrm{type} \;\; (x_1 : \alpha_1, \ldots, x_n : \alpha_n) \tag{1}$$

I have to explain what a type is, a type that may, in general, depend on variables. What the types are is, in the first instance, explained by giving the rules of type formation for this language, and they are a direct generalization of Church's rules for the simple theory of types.

Here are the three rules of type formation, written again in the informal natural-deduction style,

$$\mathrm{prop} : \mathrm{type} \qquad \frac{A : \mathrm{prop}}{\mathrm{pr}(A) : \mathrm{type}} \qquad \frac{\alpha : \mathrm{type} \quad \begin{matrix}(x : \alpha)\\ \beta : \mathrm{type}\end{matrix}}{(x : \alpha)\beta : \mathrm{type}}$$

Following the Curry–Howard correspondence, the two first rules can instead be written

$$\mathrm{set} : \mathrm{type} \qquad \frac{A : \mathrm{set}}{\mathrm{el}(A) : \mathrm{type}}$$

Type theory is based on the Curry–Howard idea of identifying propositions and sets—as I have decided to call them. One might object to this terminological choice, since set is so tied to Zermelo's conception of set, the cumulative hierarchy conception of set. In predicate logic we speak of individual domain and quantificational domain, so domain sounds perfect, but then the problem is with Dana Scott's domain theory—especially in computer science, the term domain is so tied to that, so it is not good either. Many people therefore use type here, instead of my set, but for me type is so tied to type in the sense of simple type theory, where you have the simple type structure, which is not what you have here. Some compromise therefore has to be made. My choice has been to use set, thinking that, Why should we pay so big attention to Zermelo's use of it? After all, the notion of set comes from Cantor, and his use of it is much closer to what appears in the Curry–Howard correspondence. That is my reason for the choice here.

I should add that what I have to say in these lectures does not depend at all on the identification of sets and propositions, so we need not rely on it. The only thing that happens if you do not identify sets and propositions is that you would have two distinct types, a type of sets and a type of propositions. I leave that possibility open.

Let us make the comparison now with the type structure of the simple theory of types,

$$o : \text{type} \qquad \iota : \text{type} \qquad \frac{\alpha : \text{type} \qquad \beta : \text{type}}{(\alpha)\beta : \text{type}}$$

Church had a ground type o, the type of propositions, so that is only notationally different from the type prop above. Then he had the type ι of individuals. Here we have a vast generalization, because we have not only one type of individuals, but any proposition, or set, gives rise to a type, namely, in the case of a set, the type of elements of that set, and in the case of a proposition, the type of proofs of that proposition. Moreover, the propositions, or sets, are built up within the theory itself. To interpret Church's type structure, we can let ι be some fixed type $\text{el}(A)$, where A is a set,

$$\iota = \text{el}(A) : \text{type}$$

In the final rule, $(\alpha)\beta$ is the function type from α to β written in Schütte's notation. When β does not depend on any variable, then we can pick a variable, x, and define

$$(\alpha)\beta = (x : \alpha)\beta : \text{type}$$

We thus see that Church's type structure is included in this dependent type structure.

These are the formal rules for the type structure. It is, however, not enough just to give a notation for types: we must also explain what an object of any given type is. In particular, we have to answer the questions, What is a proposition? and, What is a proof of the proposition A? And then we have to answer the question, What is an object of the dependent function type, $(x : \alpha)\beta$?

I will take the dependent function type first. Assume that the premisses,

$$\alpha : \text{type} \quad \text{and} \quad \begin{array}{c} x : \alpha \\ \beta : \text{type} \end{array}$$

have been established. What is an object of the type $(x : \alpha)\beta$? It is a function, f, such that if you take an object a of type α, then you get an object, $f(a)$, of type $\beta(a/x)$, and if you take equal objects, a and a', of type α, then you get equal objects, $f(a)$ and $f(a')$, of type $\beta(a/x)$. This is simply taking seriously what is said about functions in elementary textbooks. Again we are in the situation that I could either say what I just said and then continue by saying that, in virtue of those explanations, these rules are valid:

$$\frac{f : (x : \alpha)\beta \qquad a : \alpha}{f(a) : \beta(a/x)} \qquad\qquad \frac{f : (x : \alpha)\beta \qquad a = a' : \alpha}{f(a) = f(a') : \beta(a/x)}$$

We could, however, just as well say that the meaning of the form of judgement $f : (x : \alpha)\beta$ can be read off from these rules.

It is natural to extend Gentzen's use of the term introduction and elimination rule and say that when a judgement whose meaning is being defined occurs in the conclusion, it is defined by its introduction rule, whereas when it occurs as premiss, it is defined by the elimination rule. In this terminology we can say that the form of judgment (Con) is defined by its introduction rule, whereas the form of judgement $f : (x : \alpha)\beta$ is defined by its elimination rules.

When a form of judgement is explained by an elimination rule, we can use a very good term introduced by Ryle, namely inference licence, or inference ticket. A judgement of the form in question is then an inference licence which licenses you to infer the conclusion from the additional premisses. This works only for those forms of judgement that occur as premisses, which is to say, that you infer from. It does not work for the forms of judgement that are defined by introduction rules, like all the logical operations, for instance.

That was the semantical explanation that goes with the rule of dependent function type formation, the third rule above. The rule claims that $(x : \alpha)\beta$ is a type, which means that we have to explain what an object of that type is—which I have just done. It remains now the crucial cases of the base types: the type of propositions and the type of proofs of a proposition.

The question we first have to answer is, What is a proposition? Many proposals have been given, but I shall discuss only the two that are viable from my point of view. Of the traditional definitions, the one that I will consider, because I think it is the best, is saying that a proposition is defined by its truth conditions. I want to show in particular that that answer, properly interpreted, is fine also from a constructive point of view: you just have to make the truth conditions more detailed by turning them into proof conditions, by insertion of proof objects. Truth conditions are obtained from proof conditions by deletion of proof objects, and conversely, proof conditions, which is to say the clauses in the Brouwer–Heyting–Kolmogorov interpretation, are obtained from truth conditions by insertion of proof objects. Let me exemplify that in the case of conjunction.

The truth condition for conjunction is that $A \mathbin{\&} B$ is true if both A and B are true. Now the question is, How do you interpret that? There are many ways in which that has been interpreted in the literature. There is something of that sort even in Tarski's truth definition, which is completely different from what I am going to say now. I am going to write the truth condition for conjunction in this form:

$$\frac{A \text{ true} \qquad B \text{ true}}{A \mathbin{\&} B \text{ true}} \tag{\&I-tc}$$

The horizontal line, as always, indicates inference, so the truth condition for conjunction is this inference rule, which is taken to define conjunction. This rule therefore serves as a justification of the formation rule for conjunction,

$$\frac{A : \text{prop} \qquad B : \text{prop}}{A \mathbin{\&} B : \text{prop}}$$

To justify this rule semantically, we have to explain what it means for A & B to be true under the assumption that we have been given the truth conditions of both A and B. That is precisely what the rule (&I-tc) above tells us.

In the Brouwer–Heyting–Kolmogorov interpretation, the proposition A & B is defined by laying down what its canonical proofs look like,

$$\frac{a : \text{pr}(A) \qquad b : \text{pr}(B)}{(a, b) : \text{pr}(A \mathbin{\&} B)} \tag{\&I-pc}$$

This is a different explanation from the truth-condition explanation, but you see that they are related to each other in a precise way: the step from (&I-tc) to (&I-pc) is the insertion of proof objects, and the step from (&I-pc) to (&I-tc) is the suppression of proof objects. According to the explanation of truth that I gave earlier, in order to have the right to assert A true, you must be in possession of a proof, a, of A. When we take the premisses in (&I-tc) to be given, there are, therefore, implicitly, proofs a of A and b of B. Those two proofs can be combined by pairing to form the proof (a, b) of A & B.

I used the term canonical a moment ago, and that is because not every proof of a conjunction has the form of a pair. There are also proofs obtained by elimination rules in Gentzen's sense, whereas the pair corresponds to a proof obtained by the introduction rule for conjunction. A proposition is defined, not by telling what an arbitrary proof of it is, but by telling how the canonical proofs of the proposition are formed. An arbitrary proof of a proposition, A, is then defined to be a method, or a program, which when executed yields a canonical proof of A as result—what a canonical proof of A is is just what the definition of A lays down. The term canonical has the alternative normal, which was used by Prawitz, for instance, in his treatment of natural deduction. It is just the standard mathematical terminology that we use, for instance, in speaking of the Jordan normal, or canonical, form of matrices. The one who introduced canonical in this connection is Brouwer, in his first paper where he introduced proof objects, in the proof of the so-called Bar Theorem—he spoke there about kanonisierte Beweise.

This suffices as an explanation of the type of propositions: a proposition is defined by laying down its truth conditions or, in more detail, its proof conditions. Suppose now that we have a proposition A, and we have to explain what an object of type $\mathrm{pr}(A)$ is, which is to say what a proof of A is—but I have just done that by saying what a canonical proof of A is and then what an arbitrary proof of A is.

I have yet to explain the third form of judgement,

$$a : \alpha \;\; (x_1 : \alpha_1, \ldots, x_n : \alpha_n) \tag{3}$$

of which the form of judgement (P) is an instance. The explanation is that a judgement of this form means that if we take objects of the argument types,

$$a_1 : \alpha_1, \ldots, a_n : \alpha_n(a_1/x_1, \ldots, a_{n-1}/x_{n-1})$$

then

$$a(a_1/x_1, \ldots, a_n/x_n) : \alpha(a_1/x_1, \ldots, a_n/x_n),$$

and if we take equal objects of the argument types,

$$a_1 = a_1' : \alpha_1, \ldots, a_n = a_n' : \alpha_n(a_1/x_1, \ldots, a_{n-1}/x_{n-1})$$

then

$$a(a_1/x_1, \ldots, a_n/x_n) = a(a_1'/x_1, \ldots, a_n'/x_n) : \alpha(a_1/x_1, \ldots, a_n/x_n)$$

From this explanation the validity of these two rules is clear:

$$\frac{\begin{array}{c} a : \alpha \;\; (x_1 : \alpha_1, \ldots, x_n : \alpha_n) \\ a_1 : \alpha_1 \;\; \ldots \;\; a_n : \alpha_n(a_1/x_1, \ldots, a_{n-1}/x_{n-1}) \end{array}}{a(a_1/x_1, \ldots, a_n/x_n) : \alpha(a_1/x_1, \ldots, a_n/x_n)}$$

$$\frac{\begin{array}{c} a : \alpha \; (x_1 : \alpha_1, \ldots, x_n : \alpha_n) \\ a_1 = a_1' : \alpha_1 \;\; \ldots \;\; a_n = a_n' : \alpha_n(a_1/x_1, \ldots, a_{n-1}/x_{n-1}) \end{array}}{a(a_1/x_1, \ldots, a_n/x_n) = a(a_1'/x_1, \ldots, a_n'/x_n) : \alpha(a_1/x_1, \ldots, a_n/x_n)}$$

Again we can just as well say that these rules are meaning determining for the form of judgement which occurs as their major premiss. These are then elimination rules for that form of judgement, since the judgement in question occurs as premiss and not as conclusion.

Now we have finally got down to the bottom, because this was the meaning explanation for the third basic form of judgement of type theory. Building up successively on that, we reach the consequence form of judgement. This consequence form is thus not just taken for granted, as it was by Tarski, but I have provided a detailed semantical explanation of what it means—a semantical explanation, moreover, which justifies the rules that hold for it.

The most obvious rule that holds for judgements of the consequence form is the following:

$$\frac{\begin{array}{c} A \text{ true } (x_1 : \alpha_1, \ldots, x_m : \alpha_m, A_{m+1} \text{ true}, \ldots, A_n \text{ true}) \\ a_1 : \alpha_1 \; \ldots \; a_m : \alpha_m(a_1, \ldots, a_{m-1}/x_{m-1}) \\ A_{m+1}(a_1/x_1, \ldots, a_m/x_m) \text{ true} \\ \vdots \\ A_n(a_1/x_1, \ldots, a_m/x_m) \text{ true} \end{array}}{A(a_1/x_1, \ldots, a_m/x_m) \text{ true}} \tag{R}$$

This serves as an elimination rule for the consequence form of judgement. Why is it valid under the interpretation that I have given? We assume that the premiss judgements have been established, and I have to explain why we have the right to make the conclusion judgement. According to the explanation of the consequence form of judgement, what permits us to make the judgement that is the first premiss of (R) is that we possess a proof object,

$$a : \text{pr}(A) \;\; (x_1 : \alpha_1, \ldots, x_m : \alpha_m, x_{m+1} : \text{pr}(A_{m+1}), \ldots, x_n : \text{pr}(A_n))$$

The premisses on the second line provides us with arguments, $a_1, \ldots, a_m$. What permits us to make the judgement that is the third premiss is that we possess a proof object,

$$a_{m+1} : A_{m+1}(a_1/x_m, \ldots, a_m/x_m)$$

and similarly for the rest of the premisses. From these premisses, we can thus restore proof objects $a_{m+1}, \ldots, a_n$. Now we have, not only the object a, but also $a_1, \ldots, a_m, a_{m+1}, \ldots, a_n$. Taken together, this gives us a proof object for the proposition $A(a_1/x_1, \ldots, a_n/x_n)$, namely

$$a(a_1/x_1, \ldots, a_n/x_n)$$

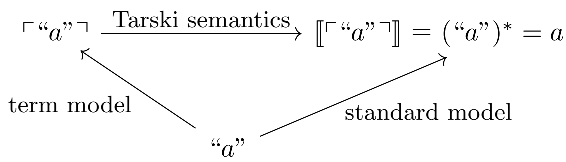
by the meaning of the third form of judgement. We therefore have the right to make the conclusion judgement. The rule (R) is thus justified by restoring the proof objects: that is how this obvious rule is validated according to the meaning explanations that I have given here.

Suppose you do not have any proof objects. You still want to explain what the consequence form of judgement, (Con), means. The obvious explanation is, of course, that it just means what it has to mean in order for the rule (R) to be valid. That is to say, without proof objects, the rule (R) becomes meaning determining for the consequence form of judgement. That is clearly how the ordinary notion of logical consequence was interpreted before proof objects were born, which was only in the 1920s. Bolzano, in his treatment of Ableitbarkeit, for instance, does not say anything explicitly on this point, but it is clear that that is how he interprets it. That must also have been the way in which the notion of formal consequence was interpreted in the 14th century, and it is what Etchemendy called the substitutional account of logical consequence.

Tarski's metamathematical reconstruction of the notions of truth and consequence

24 February 2012

Today it will be explicitly about Tarski's work, and that makes it much easier, in a way, than the previous lectures, although I was then speaking about my own things. Concerning Tarski semantics I think it is enough, since we all know it, to summarize it in the following diagram:



We have an expression, “a”, of the object language—I will use the usual quotation marks to indicate expressions rather than what they mean. An expression of the object language can be interpreted in many models, and the basic ones for Tarski semantics are the term model, on the one hand, and the standard model, on the other hand. For the term model I use the corners, so $\ulcorner$“a”$\urcorner$ is the interpretation of “a” in the term model, or syntactic model. For the intended interpretation, the interpretation in the standard model, I will use the star, so (“a”)* is the meaning of the expression “a”. This presupposes, as Tarski always does, that the object language is an interpreted language, so that the expressions in the object language have a meaning, an intended meaning. In general it is enough if (“a”)* can be expressed in the metalanguage, but if the object language is part of the metalanguage, then (“a”)* is equal to a.

Tarski semantics is a homomorphism from the term model to the standard model: it takes $\ulcorner$“a”$\urcorner$, which Tarski called the structural-descriptive name of the expression “a” of the object language, but which we nowadays rather think of as the interpretation of “a” in the term model, and yields $[\![\ulcorner$“a”$\urcorner]\!]$. Tarski's material adequacy condition says just that $[\![\ulcorner$“a”$\urcorner]\!] =$ (“a”)*. Tarski, of course, did not describe what he was doing in this way. This is a contemporary way of describing it, though I am not completely sure to whom it should be attributed—

maybe Lawvere in the 1970s had something like it.

This pattern is independent of the language we are considering and can, in particular, be applied to constructive type theory. Recall the form of judgement

$$\alpha : \text{type} \;\; (x_1 : \alpha_1, \ldots, x_n : \alpha_n)$$

What becomes of this in the term model? Corresponding to it we have the form of judgement

$$\tau : \ulcorner\text{type}\urcorner \;\; (\ulcorner x_1 \urcorner : \tau_1, \ldots, \ulcorner x_n \urcorner : \tau_n)$$

I use τ for the entities in the term model that interpret types and call these entities type symbols, naturally written $\ulcorner\text{type}\urcorner$. The context $(x_1 : \alpha_1, \ldots, x_n : \alpha_n)$ also has to be interpreted in the term model, and its interpretation is precisely what we call a signature. The signature $(\ulcorner x_1 \urcorner : \tau_1, \ldots, \ulcorner x_n \urcorner : \tau_n)$ says that the metamathematical variable $\ulcorner x_1 \urcorner$, which is not a variable any longer, but a constant, has type symbol τ_1, and so on up to $\ulcorner x_n \urcorner$, which has type symbol τ_n.

Suppose we have objects of the right types to associate with the metamathematical variables, $\ulcorner x_1 \urcorner, \ldots, \ulcorner x_n \urcorner$. So we have objects

$$a_1 : [\![\tau_1]\!], \ldots, a_n : [\![\tau_n]\!](a_1/\ulcorner x_1 \urcorner, \ldots, a_{n-1}/\ulcorner x_{n-1} \urcorner)$$

If we have such objects—and I am now formalizing a bit on the metalevel—then we may form the semantic value of τ, $[\![\tau]\!]$, which is a type under the metamathematical assignment of $a_1, \ldots, a_n$ to the metamathematical variables $\ulcorner x_1 \urcorner, \ldots, \ulcorner x_n \urcorner$—it is not really an assignment, but an association list that correlates $\ulcorner x_1 \urcorner$ with a_1, and so on.

$$\frac{\tau : \ulcorner\text{type}\urcorner \;\; (\ulcorner x_1 \urcorner : \tau_1, \ldots, \ulcorner x_n \urcorner : \tau_n) \qquad a_1 : [\![\tau_1]\!] \;\ldots\; a_n : [\![\tau_n]\!](a_1/\ulcorner x_1 \urcorner, \ldots, a_{n-1}/\ulcorner x_{n-1} \urcorner)}{[\![\tau]\!](a_1/\ulcorner x_1 \urcorner, \ldots, a_n/\ulcorner x_n \urcorner) : \text{type}}$$

Corresponding to the form of judgement

$$a : \alpha \;\; (x_1 : \alpha_1, \ldots, x_n : \alpha_n)$$

we have, in the term model, the form of judgement

$$t : \tau \;\; (\ulcorner x_1 \urcorner : \tau_1, \ldots, \ulcorner x_n \urcorner : \tau_n)$$

The corresponding rule is this:

$$\frac{t : \tau \;\; (\ulcorner x_1 \urcorner : \tau_1, \ldots, \ulcorner x_n \urcorner : \tau_n) \qquad a_1 : [\![\tau_1]\!] \;\ldots\; a_n : [\![\tau_n]\!](a_1/\ulcorner x_1 \urcorner, \ldots, a_{n-1}/\ulcorner x_{n-1} \urcorner)}{[\![t]\!](a_1/\ulcorner x_1 \urcorner, \ldots, a_n/\ulcorner x_n \urcorner) : [\![\tau]\!](a_1/\ulcorner x_1 \urcorner, \ldots, a_n/\ulcorner x_n \urcorner)}$$

Tarski's great achievement is to have properly defined these semantic values by recursion over the buildup—in this case, it is over the structure of the type symbols, respectively over the structure of the terms. This I take to be given now: doing Tarski semantics for type theory. That was the subject of the talk I gave at the LMPS conference in Florence in 1995, and I just take it for granted now, because we all know how to do Tarski semantics for any language that we are given, in particular for type theory.

When the metamathematical variables are assigned the variables themselves, $(x_1/\ulcorner x_1 \urcorner, \ldots, x_n/\ulcorner x_n \urcorner)$, supposing, of course, that we are in a surrounding context where we have introduced the variables $x_1, \ldots, x_n$ of the correct types, then I will write

$$[\![\tau]\!](x_1/\ulcorner x_1 \urcorner, \ldots, x_n/\ulcorner x_n \urcorner)$$

simply as $[\![\tau]\!]$, and

$$[\![t]\!](x_1/\ulcorner x_1 \urcorner, \ldots, x_n/\ulcorner x_n \urcorner)$$

as $[\![t]\!]$.

What is Tarski's definition of formal, or logical, consequence? Recall the logical-consequence form of judgement, the form of assertion that expresses an ordinary, as I called it, logical, or formal, consequence:

$$A \text{ true } \ (x_1 : \alpha_1, \ldots, x_m : \alpha_m, A_{m+1} \text{ true}, \ldots, A_n \text{ true}) \qquad \text{(Con)}$$

This is the ordinary notion of logical consequence. What is Tarski's notion? Instead of the initial context $(x_1 : \alpha_1, \ldots, x_m : \alpha_m)$ we introduce instead a signature $(\ulcorner x_1 \urcorner : \tau, \ldots, \ulcorner x_m \urcorner : \tau_m)$. We go from a context to its metamathematical counterpart, which is a signature. And whereas $A_{m+1}, \ldots, A_n, A$ were propositions in the context $(x_1 : \alpha_1, \ldots, x_m : \alpha_m)$, so they were propositional functions, we now have formulas in the signature $(\ulcorner x_1 \urcorner : \tau, \ldots, \ulcorner x_m \urcorner : \tau_m)$,

$$F_{m+1}, \ldots, F_n, F : \ulcorner \text{prop} \urcorner \ \ (\ulcorner x_1 \urcorner : \tau, \ldots, \ulcorner x_m \urcorner : \tau_m)$$

Formulas are what interprets propositions in the term model,

$$\ulcorner \text{prop} \urcorner = \text{formula}$$

We can now state the definition. The formula F is a logical consequence in the sense of Tarski of the formulas $F_{m+1}, \ldots, F_n$ provided

$$[\![F]\!] \text{ true } \ (x_1 : [\![\tau_1]\!], \ldots, x_m : [\![\tau_m]\!], [\![F_{m+1}]\!] \text{ true}, \ldots, [\![F_n]\!] \text{ true}) \qquad \text{(Tar)}$$

I am relying here on the abbreviations explained above for when metamathematical variables are assigned the variables themselves, so that I can write, for instance, $[\![F]\!]$ instead of $[\![F]\!](x_1/\ulcorner x_1 \urcorner, \ldots, x_m/\ulcorner x_m \urcorner)$ and $[\![\tau_m]\!]$ instead of

$[\![\tau_m]\!](x_1/\ulcorner x_1\urcorner, \dots, x_{m-1}/\ulcorner x_{m-1}\urcorner)$.

In (Tar) we have a well-formed assertion, an instance of the form (Con) obtained by inserting the semantical values of the arbitrary formulas $F_{m+1}, \dots,$ F_n, F for the arbitrary propositions $A_{m+1}, \dots, A_n, A$. I am in the good position now that last time I explained what an assertion of the form (Con) means, so it is something that we have access to, hence I can use it to define logical consequence in the sense of Tarski. Observe that the ordinary, or pre-metamathematical, notion of logical consequence, as expressed by (Con), is taken for granted here. Tarski simply presupposed that we all understand what it means for the proposition $[\![F]\!]$ to be true provided the propositions $[\![F_{m+1}]\!], \dots, [\![F_n]\!]$ are true, whatever we assign as values to the variables $x_1, \dots, x_m$. He was right in this, in a sense: we have an ordinary notion of logical, or formal, consequence, which, as I said, is implicit in Aristotle and was quite explicit in the 14th century and in Bolzano. Tarski relies on that without any further ado here. It is only I who have been much more careful in providing a precise meaning explanation of the form of judgement (Con).

We may say, very simply, that Tarski's notion of logical consequence, which is a relation between formulas rather than propositions, is obtained by composing the ordinary notion of logical consequence with Tarski's semantic evaluation procedure, indicated by $[\![\cdot]\!]$ here. Tarski's achievement is really the semantic evaluation procedure. Once the semantic evaluation procedure is in place, then we can, of course, retrace the ordinary notion of logical consequence, expressed by (Con), from propositions to formulas.

This is the basic thing that I had to say about Tarski's notion of logical consequence in relation to the ordinary notion of logical consequence. Maybe it is illuminating to see what happens when we use it in an example that is as simple as possible. Let us take the example with two propositional constants, P and Q, and conjunction, &. From the uninterpreted propositional constants P and Q we obtain formulas $\ulcorner P\urcorner$ and $\ulcorner Q\urcorner$, and from the proposition $P \,\&\, Q$ we obtain the formula $\ulcorner P \,\&\, Q\urcorner$, which is equal to $\ulcorner P\urcorner\ulcorner\&\urcorner\ulcorner Q\urcorner$. These are all formulas in the signature $(\ulcorner P\urcorner : \ulcorner \mathrm{prop}\urcorner, \ulcorner Q\urcorner : \ulcorner \mathrm{prop}\urcorner)$. In particular,

$$\ulcorner P \,\&\, Q\urcorner : \ulcorner \mathrm{prop}\urcorner \ (\ulcorner P\urcorner : \ulcorner \mathrm{prop}\urcorner, \ulcorner Q\urcorner : \ulcorner \mathrm{prop}\urcorner)$$

Suppose that X and Y are two arbitrary propositions,

$$X, Y : \mathrm{prop}$$

This is an ordinary assumption: let X and Y be two arbitrary propositions. Then we can form the semantic value in Tarski's sense of the formula $\ulcorner P \,\&$ $Q\urcorner$ under the assignment $(X/\ulcorner P\urcorner, Y/\ulcorner Q\urcorner)$, which, if we grind out the Tarski algorithm, we find to be $X \,\&\, Y$, which is a proposition in the context X :

prop, Y : prop.

$$[\![\ulcorner P \mathbin{\&} Q \urcorner]\!](X/\ulcorner P\urcorner, \ulcorner Q\urcorner/Y) = X \mathbin{\&} Y : \text{prop } (X : \text{prop}, Y : \text{prop})$$

When we do it for P and Q separately we get X, respectively Y.

$$[\![\ulcorner P \urcorner]\!](X/\ulcorner P\urcorner, Y/\ulcorner Q\urcorner) = X : \text{prop } (X : \text{prop}, Y : \text{prop})$$

$$[\![\ulcorner Q \urcorner]\!](X/\ulcorner P\urcorner, Y/\ulcorner Q\urcorner) = Y : \text{prop } (X : \text{prop}, Y : \text{prop})$$

Inserting all of this into (Tar) we get

$$X \mathbin{\&} Y \text{ true } \ (X : \text{prop}, Y : \text{prop}, X \text{ true}, Y \text{ true})$$

So here $m = 2$ and $n = 4$. The upshot is that the formula $\ulcorner P \mathbin{\&} Q\urcorner$ is a logical consequence in the sense of Tarski of the formulas $\ulcorner P\urcorner$ and $\ulcorner Q\urcorner$ precisely if $X \mathbin{\&} Y$ is an ordinary logical consequence of X and Y—which is, of course, what we expect.

Now let me consider the special case when both m and n are 0. Then we get simply that the closed formula F is true in the sense of Tarski precisely if the proposition $[\![F]\!]$ is true. Here I have a comment about how wise or unwise it is to call the semantic value of F, that is $[\![F]\!]$, $\text{Tr}(F)$, as Tarski does in the English translation. If we call the semantic value $\text{Tr}(F)$, then we have two true's here. From one point of view, that is disturbing, and from another point of view, natural. I want to explain in what way it is disturbing and in what way it is natural.

It is disturbing, of course, to use the same word for two different notions at the same time. To avoid that difficulty, the simplest way out is not to call $[\![F]\!]$ $\text{Tr}(F)$, but to call it $\text{Val}(F)$, the semantic value of F, which is what it is.

There is, however, a justification for calling it $\text{Tr}(F)$, as Tarski did, and the justification is this. Let me take an analogous example. Instead of formulas, we consider numbers, and instead of the property of being true in Tarski's sense, we consider the property of being odd. The assertion

$$n \text{ is odd}$$

is formalized as

$$\text{Odd}(n) \text{ is true}$$

where $\text{Odd}(n)$ is defined as $(\exists x)(n = 2x + 1)$. Two notions of odd are involved here: one is the predicate in the first assertion and the other is the propositional function that I just defined, whose value for an arbitrary n is true precisely when n is odd. This is, of course, the standard way in which we proceed, and that is exactly what Tarski did: the propositional function $\text{Tr}(F)$ reflects the property of being true in the sense of Tarski into a proposition that is true if, and only if, that property holds of the formula in question. This is the logic behind calling $[\![F]\!]$ Tr(F), as Tarski did.

With this I have finished what I had to say about the semantic evaluation procedure and its use to get Tarski's notion of logical consequence. In the final part of these lectures, I would like to compare certain ontological or logical notions with their formalistic counterparts. In the left column below I write certain logical, or if you prefer, ontological, concepts and in the right column their formalistic counterparts:

individual	term
proposition	formula
proof	derivation
truth = provability = existence of a proof	derivability = existence of a derivation
hypothetical proof	derivation from assumption formulas
consequence = hypothetical provability = existence of a hypothetical proof	derivability from assumption formulas = existence of a derivation from assumption formulas

Let this suffice as a list of the most basic concepts that I have considered and their formalistic counterparts.

The situation in the 1930s was that everybody was heavily syntactically prejudiced as a result of Hilbert's step of making the logical symbolism itself an object of mathematical study. One became preoccupied with these new combinatorial objects. Nobody then had any problems with the notions on the right hand side here: on the contrary, that was what everybody was interested in, because these were new mathematical objects that required investigation. This preoccupation with the syntactic had a dark side, however, namely that one felt very shaky about the semantic side, or the ontological side, if you want. This was precisely the time when Ryle wrote a paper with the famous title "Are there propositions?". To a mathematician that is like asking, Are there complex numbers?, or something like that—a completely absurd question. Of course, there are numbers. The question, What are they?, that is one thing—but we could not have the mathematics that we have unless they were. Similarly in logic, we have, after all, propositional logic, and how could we have that if there were no propositions? These worries were so strong that the propositional calculus changed name to the sentential calculus, and one started to speak about sentences instead of propositions—which shows how troublesome even the second line in the table above,

proposition	formula

was at the time.

Tarski very cleverly avoided this problem. He defined satisfaction—and of course, that a certain sequence satisfies a formula, that is expressed in a proposition. Tarski did not say that, however, but managed without saying it. He simply defined satisfaction directly in ordinary mathematical terms. Hence he did use propositions, but he did so without offending anyone by speaking about them.

When we come to the third line of the table,

proof	derivation

the situation is much worse than it was with the second line. In the third line there was a gap, a lacuna, in the logical ontology at the time: there were no proof objects. There were no mathematical objects that we could sensibly assign as semantic values to derivations. Well, actually there were, because of Brouwer and Heyting, but nobody knew about that at the time, because it had not been sufficiently systematically developed. Moreover, neither in the Vienna Circle nor in the Polish school, where they were occupied by the relation between the two sides in the table, did they take intuitionism seriously. The Vienna Circle was basically determined by Schlick and Carnap, and both followed Hilbert, with nothing in depth about intuitionism, and as far as I know, it was the same in the Polish school.

Maybe this is one of the most important points of this series of lectures, namely the introduction of a proof ontology. The lack of it before is very conspicuous. If you think of Bolzano's logic, which has a very clear architectural structure and which takes an ontological viewpoint from the start—the main chapters of his logic are

Vorstellungen an sich
Sätze an sich
Wahrheiten an sich

In the *Organon*, the first two correspond to the *Categories* and the *De interpretatione*. The third is a novelty, where we get the objective notion of truth, which applies to Sätze and sich. These are the three main chapters of Bolzano's logic. The upshot of my discussion is that this is too poor an ontology, because it lacks something that we should be able to assign as semantic values to derivations. There would have been needed a chapter

Beweise an sich

Although Bolzano calls the notion of formal consequence, or logical consequence, Ableitbarkeit, which is to say deducibility, there is no such chapter in his logic: he has deducibility without deductions. That is the same situation as we have had in the 1930s: we had a semantic notion of consequence, but no ontological notion of hypothetical proof.

If we look back now on these three lectures, the main points, in my eyes, are the following.

- The priority of epistemology over ontology, more specifically of inference over consequence.

Remember why: when we explain the notion of consequence semantically, the notion of inference had to be in place already.

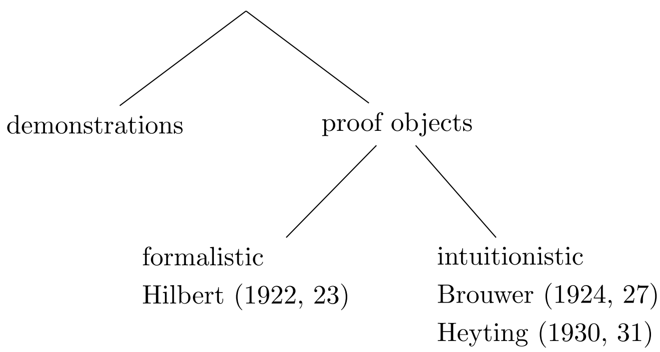
- The priority of the ordinary notion of logical consequence over Tarski's metamathematical reconstruction of it.

That is because the ordinary notion occurs in the definition of Tarski's notion of logical consequence.

- The introduction of the notion of proof object.

Proof objects are appealed to in the explanation of the ordinary notion of logical consequence.

Let me end by elaborating this last point a little bit. The notion of proof object is distinct both from the epistemological notion of demonstration and from the metamathematical notion of formal derivation, or proof term, as we often call them nowadays. You may look upon it as follows.



You have the notion of demonstration—a demonstration is a chain of immediate inferences, if you remember from the first lecture—and distinct from that, the notion of proof object. Proof objects, to my mind, were introduced first by Hilbert, but he understood them formalistically. This was in his papers that inaugurated proof theory: the "Neubegründung" paper from 1922 and the more sober *Mathematische Annalen* paper from 1923. I take it that Brouwer read everything that Hilbert wrote of a foundational character at the time, so I think we can take it for granted that these papers by Hilbert were read by Brouwer as soon as they appeared. Only a year later, Brouwer had, with his creative mind, fruitfully rethought that idea and put it to use in his own work. This is the first appearance of intuitionistic proof objects, namely in Brouwer's first proof of the Bar Theorem from 1924 and then in the more well-known proof from 1927—later, they were very fruitfully used by Heyting in his explanations of the meanings of the logical operations from 1930 and 1931. I cannot find references further back to anything that should be reminiscent of proof objects, so it is in 1920s that this proof ontology, proofs as objects, came into being.

The trouble with the formalistic conception of proof objects is what Brouwer immediately noticed. He was a relentless critic of Hilbertian formalism, because he thought that it is not the linguistic forms, but the meaning that is the important thing, so we should have the semantic counterpart to the formalistic notions. That fits, of course, with the discovery in the 1930s of the discrepancy between truth and derivability, the left and the right column in the fourth line of the table above,

truth	derivability

It is astonishing that it has taken such a long time to realize that once we introduce the intuitionistic proof objects rather than the formalistic ones, the formal derivations, then we reestablish the equation between truth and provability: truth is distinct from derivability, but once we introduce proof objects, the equality, now between truth and provability, gets reestablished,

$$\text{truth} = \text{provability}$$

In Heyting's work, the proof objects enter only in the interpretation of the logic. The step taken in intuitionistic type theory was to introduce—if these proof objects are so important for the interpretation, why should they not be there in the formal system itself? Once you put them in, you arrive at type theory.