October 1, 2025

Week 5 Notes

1. **Introduction to Week 5**:

Begin by picking up the main theme from last time: **raising sortal consciousness**.

a) To understand *counting* (so, "How many?" questions, hence numbers) we must recognize a hidden relativity of counting to sortal kind concepts.
   That is, singular terms, to pick out *countables*, must fall under *sortal kind-concepts*, which supply their criteria of identity and individuation.
b) One large consequence should be realizing the need to **distinguish sharply between**
   i. Possible worlds, and
   ii. Tarskian models.

The difference is that Tarskian models require specification of a *domain*, or all the *individuals* or *objects* in the domain. Since we can't count individuals or objects as such, we must understand the domain to have been specified in some antecedent semantic metavocabulary, which *does* specify individuating sortals governing the elements of the domain. It is true that specifications of domains usually just say: a set of *things*. But sets have to be individuated by their elements—they are abstracted from those elements, since two *sets* of things A and B are the *same set* just in case their elements are the *same things*: iff [might mention witticism "When I say 'if' here, I mean 'if', and *only* 'if' (not if *and* only if)] that is iff $\forall x[x \in A \Leftrightarrow x \in B]$. Note how this goes beyond just a 1-1 correspondence. There is a definite answer to the question "How many *elements* (things) are there in the domain of this model?"

Possible worlds, by contrast, **don't have *domains***.

In *specifying* a possible world, one must *say* what 'things' are in it by using sortal kind terms in one's use-language or semantic metavocabulary.

The question: "How many *things* are there in this possible world?" precisely has the missing relativity to the sortals by which one counts them.

John Etchemendy's fine book *The Concept of Logical Consequence* is a good starting-point for articulating this difference between the possible-worlds framework and the model-theoretic framework.

This means that **one cannot, Pavel Tichy-wise, be a *finitist* about possible worlds, *tout court***.

That is, one cannot intelligibly stipulate that one wants, for instance, to restrict one's attention to *finite* worlds, worlds that have only a *finite* number of *things* in them. Or at least, finitism in *this* sense is only ultimately intelligible in the context of a fuller story that *does* bring in the sortals by which things are individuated, and so become countable.

Finitism is only ultimately intelligible as a thesis about *specifications* or *descriptions* of possible worlds.

Often philosophers try to move back and forth between the model-theoretic and the possible worlds frameworks, for instance by using applying model-theoretic results directly to possible worlds. Stalnaker and Jagwon Kim both do this in talking about *supervenience* of one vocabulary on another. Putnam does it as well, in arguing against certain semantic programs.

c) Rant about **the empty** set:

Notes on *Empty-Set Skepticism* as a Foundational Concern for Pure Set Theory

'Nothing', read as *no thing*, is in the same boat as 'everything', read as *every thing*.

Both suffer from the difficulty that one cannot *count things*, as things.

a) 'Thing', and its cognates like 'object', 'particular', 'item', 'entity', and so on have the grammatical form of sortal kind terms, but they do not have the criteria of identity and individuation that are essential to genuine sortal kind terms.

These *pseudo*sortals are merely *schematic placeholders* for genuine sortals.

They are often *prosortals*, pointing toward anaphoric antecedents that are genuinely individuating sortal kind terms. As with other anaphoric dependents or proforms, this includes cases where it is not clear what the antecedent is.

These pseudosortals should be understood as having a *parameter*, the genuine sortal or sortals that make what falls under them countable. Unless and until a value is supplied for that parameter, a real, individuating sortal kind term, they remains merely schematic.

b) Treating schematic pseudosortals as genuine sortals is a mistake.

It requires ignoring the triangular relationship of mutual presupposition of

i. Singular terms,
ii. Sortal kind terms that govern the use of those terms, and
iii. Identity claims, which license the intersubstitutions that both determine the use of singular terms and permit counting of what falls under the governing sortal kind term.

c) When quantifiers are used without sortal restrictions, when '$\forall$x' is read as meaning "for any (thing) *x*", or just, "for anything", the relevant parameter is not supplied, and the conceptual content (in the sense of inferential role, here determined substitutionally) of the singular terms that can be put in for 'x' is accordingly left indeterminate. The needed sortal parameter *can* be supplied by interpreting the quantifier model-theoretically, in terms of a *domain of quantification*. But this strategy will only be successful if the sortal appealed to in identifying and individuating elements of the domain ('element' itself being a schematic pseudo-sortal) is specified in the semantic metavocabulary used to say what the models are. If it is not, the problem is merely put off—the bump in the rug is just moved around.

These remarks express the lesson I learn from what Geach makes of the lessons of Frege's discussion of the triangle of singular terms, sortal kind terms, and identity locutions ("recognition judgments" as intersubstitution licenses as determining the conceptual content or inferential role of singular terms) in the *Grundlagen*.

So much is just rehearsing the story I told last time.

Here is an extension of that line of thought that I gestured at, but left implicit.

d) **'Nothing' has the same defects as 'everything.'** Or, more charitably, it incurs the same responsibility to redeem the promissory note implicit in it, namely to supply the missing parameter in the form of a sortal governing the 'things' one is quantifying over.

   'No tigers', 'no photons', 'no books', 'no holes', are all OK. The substitutional obligations one incurs by using these expressions to make claims are settled by the explicit sortal kind terms. Using them is to be understood in terms of the number of Ks, perhaps further specified ('…in the room', 'emitted by the reaction,' 'on the table', 'in the curtain'…), having number 0: they can be put in 1-to-1 correspondence with the number of natural satellites of Mercury, or the number of present Kings of France. But if we just say 'no thing', we are given no individuating sortal to make the relevant commitments substitutionally, and so inferentially, determinate.

e) A potentially important special case is the **empty set** in set theory. I understand it as follows. To specify a **set,** one specifies its *elements* or *members*. If a specification S includes some term *e* as an element, we say $e \in S$. An equivalence relation is then defined among such specifications: if two specifications of the elements specify just the *same* elements: every element specified in specification S_1 is also an element specified in specification S_2, and every element specified in specification S_2 is also an element specified in specification S_1, which we could write as $S_1 \approx S_2$, then the *sets* so specified are identical: $\text{set}(S_1) = \text{set}(S_2)$. This is a definition by *abstraction*, in Frege's sense. Notice that we can write this in quantificational terms as

   $\forall x[(x \in S_1 \Leftrightarrow x \in S_2) \Leftrightarrow (\text{set}(S_1) = \text{set}(S_2))]$.

   But abstraction is a process by which new sortals (here, 'set') are introduced on the basis of previous sortals, the ones individuating the 'elements' so that we can tell when one of the 'elements' of S_1 is identical to one of the elements of S_2. But 'element' can't do this job of individuation. It just stands in for, marks the place where a genuine individuating sortal would go. If we ask the set-theorists how the *elements* are to be identified and individuated, they will say that that is not their business. Give them the elements, which must be re-identifiable and distinguishable, and they will tell us about *sets* of those elements. But we must not forget that *some* sortal that does this job is being presupposed. The use of the schematic pseudosortal 'element' just tells us i) that there *is* some genuine sortal individuating them, and ii) that it doesn't matter (to the set theorist) *which* sortal it is.

f) **The empty set ∅ is defined as the set that has *no* members, the set in which *nothing* is a member.** That is, $\forall x[x \notin \varnothing]$. Put otherwise, ∅ is the set abstracted from the equivalence relation same-elements-as by concepts whose number is 0. But, 'element' is not an individuating sortal. It is a schematic pseudo-sortal, like 'thing'. Just as with the universal quantifier in terms of which we can specify the empty set, there is a missing sortal parameter in the claim. You say that the empty set has '*no* elements': the number of its elements is 0. As Frege insists, this is a perfectly determinate and intelligible claim to make about a concept. 0 is the number of natural satellites of Mercury, of divisors of 17 other than itself and 1, and of tigers more than a hundred feet tall. So, given a sortal K, it is determinate what it is for the elements of a set, say the set ∅, to be zero. But the fact remains, there must *be* a sortal kind concept in order for it to be determined what *is* the number associated with *that* sortal—by abstraction according to the equivalence relation of 1-to1 correspondence.

g) The set theorist says that ∅ has *no* elements. We can ask: Does that mean that there are no *tigers* in the empty set? And we will be told that is true. Also none of the statues Goliath and David, the lump of clay Lumpl, no passengers, no persons, no photons, no fridgeons…? And the response will be: Right. There is *nothing* in the set. It has no elements of *any* kind. (And it is not the set theorist's business as such to be able to say which identities involving these terms hold. They take for granted the identity and individuation of the 'elements' from which they then abstract sets, introducing the new sortal 'set' and the singular terms falling under it, based on the sortals that govern the 'things' appealed to as elements of sets.)
This is exactly parallel to the *wide-open* reading of the universal quantifier as meaning 'everything'. If $\forall x Fx$, then *everything* is an F: every tiger, every statue, every fridgeon…everything, every *thing*.

h) So what are the sortals that identify and individuate the elements, of which there might be exactly 0? The set theorist wants to say: it doesn't matter. In both cases (wide-open, unsortalized universal quantifiers and the empty set) there is an implicit quantification over *all* sortals. And there, I think, lies the rub, the source of tension, the discursive obligation shirked.

i) ***The* empty set is *the* set whose members number 0, *the* set that has *no* members of *any* kind. That is a definite description**. Frege properly analyzes the use of singular terms of this kind as essentially involving two commitments that must be redeemed in order to be entitled to use the definite description:

existence and uniqueness. There must *be* some set that has no members, and there must be exactly one of them. There *are* concepts that have the number 0. And we can use them to define a set. The set of tigers in the room with me now is empty. It has no tigers in it, and all the elements of *that* set are specified to be tigers. So it is *an* empty set. So the requirement of existence is satisfied. What about uniqueness? Is the set that is empty of tigers in the room with me now the *same* set as the set that is empty of fridgeons (because Fodor's fridge is not on)? The two sets can be put in 1-to-1 correspondence, since both have number 0. But to be the *same* set (and both of them to be ∅, *the* empty set) they are to have the *same* elements. We must be able to identify the absence of tigers from the set with the absence of fridgeons from the set. We are not saying the same thing when we say it is empty of tigers and empty of fridgeons, but have we specified the same *set*? **To stipulate that it *is* the same set, and so to be entitled to use the definite description curled up in the symbol '∅', namely *the* empty set, we must quantify over *all* the *kinds* of things, in order to specify that this is *the* set that has *none* of those things in it, has *no* things of *any* kind.** Did we really settle that *that* is the *same* set as the set of tigers in the room with me now? We didn't say anything about fridgeons. We were just talking about tigers.

j) The reason I have focused on the empty set—moving beyond considering just what would be required to entitle oneself to wide-open, sortally unrestricted universal quantification—is that ***pure* set theory** is distinguished from *applied* set theory just in that *all* the 'elements' of pure set theory are elaborated from the empty set. In this way, it is thought, the set-theorists can licitly cut all ties to the sortals that identify and individuate the elements, and so the sets that contain tigers, fridgeons, statues, and lumps of clay, freeing themselves from the obligation to worry about any principles of identity and individuation of the things from which sets are abtracted, in favor of the purely set-theoretic principles that appeal to the prior identification and individuation of *elements* of sets. Everything that matters for set theory can be elaborated from the empty set. Since it has *no* members, there need be no concern with the sortals that identity and individuate those members. But can that be right? Counting to zero is still *counting*, and that requires sortals to identify and individuate what is counted, (moons of Mercury, 100-foot-tall tigers), even if the result of counting is 0.

k) The thorny issue, I think, concerns the *suppressed* commitment (more charitably, the *unacknowledged* or merely *implicit* commitment) to **the intelligibility of quantifying over *all* sortal kind concepts** that is involved in introducing the sortal <u>set</u>, and the singular terms that it governs (such the definite description 'the empty set'). Abstraction should be thought of as a process that defines *new* sortal kind concepts (determining the conceptual content—the determinate intersubstitution licenses—expressed by identities relating their associated singular terms) from *old* ones. Set theory derives its generality from its indifference to what sortals one starts with. One can form sets of tigers, fridgeons, or what you like, and can mix and match sortals in one set, so long as the elements of each set can be identified and re-identified—since that is all that matters for the identity and individuation of the sets formed from those 'elements'.

But sortally unrestricted universal quantification and entitlement to use the definite description '*the* empty set' (which is essential to the ultimate intelligibility of *pure* set theory) both require, in effect, the intelligibility of the metaconcept of <u>all sortal kind concepts</u>.

l) Why might one think that the set of all sortal kind concepts is ill-defined?

   i. It can't do the work it is being called on to do if it is read as restricted to the sortal kind concepts expressible in some fixed vocabulary. For the sortally wide-open universal quantifier is meant to apply to *everything*, *not* just everything specifiable in a particular vocabulary. And the empty set is to contain *no* elements, not just no elements of kinds specifiable in a particular vocabulary.

   ii. Further, **given any set (I am happy to use the concept here) of sortals, we have well-established techniques of producing *new* ones from them.** Think of how <u>fridgeon</u>, <u>passenger</u>, or <u>undetached rabbit-part</u> are formed.

   One of those methods is *abstraction*. Indeed, the Russell paradox arises precisely because among the *kinds* of things sortally unrestricted quantifiers quantify over is the *sets* formed by abstraction from *other* sortals. The set of all sets can be specified by abstraction from any vocabulary that forms sets by abstraction from things falling under other sortals—and there is no exception for the empty set, which must not have that set of all sets as an element, along with all the other things it does not have as elements. It must be an element of itself. That is not true of every set, so we can form the set of all sets that are *not* members of themselves, and then that set can neither be nor not be (or must both be *and* not be) a member of itself.

   iii. **Is there a definite totality of *possible* sortal concepts—those expressible in *any* vocabulary at all? I am skeptical.** Given any collection of vocabularies, it is possible in various ways to elaborate new vocabularies, not included in the original collection. There might be room for some sort of limit operation forming a kind of limit ordinal comprising all of such an infinite sequence, but I have no clue how that might go.

> At any rate *if* there is sense to the concept <u>all possible sortal kind-concepts</u>, that sense must be *given* to it, specified by whoever wants to make use of it. It does not just *come with* a determinate sense that can be appealed to without further comment.

d) The same complaint applies to the basic parts-without-parts, **the atomic parts, of mereology**.

One cannot count 'parts', except of mereological wholes formed from them. The *basic* 'parts' must be specified using some sortal other than 'part'. It is all of *those*, together with all the mereological *wholes* formable from them, that are *not* parts of the ur-parts or atomic parts.

e) After *Grundlagen,* Frege never talks again, I think, about the special *kind* of concept that is sortal kind-concepts—though he has a *lot* to say about *concepts* in general.

This, it turns out, is precisely what gets him in trouble in the *Grundgesetze*, leading to the Russell paradox.

And the 'PM-ese' logics we get downstream, from Carnap and Quine and Tarski, don't distinguish sortal predicates within the class of attributive predicates generally.

Kripke raised interest again in the topic of *natural* kinds, as candidates for a special kind of modal rigidity.

But Lewis-Stalnaker possible worlds theories don't, as such, involve distinguishing sortals, even in their sophisticated Montague versions.

Nor do Fine's truth-maker semantics and their logics, not even (I blush to confess) the expressivist logics and the implication-space semantics for them Ulf and I introduce in our *Reasons for Logic, Logic for Reasons* book.

f) <u>**Type Theory is the formal theory of sortal kind concepts**</u>**:**

There *is*, however, a formal theory of *kinds* of things, namely ***type theory***.

It was developed in response to the fact that Frege's now-infamous Axiom V of *GG* leads to paradoxical results, and so the inconsistency of his system. Russell and Whitehead's *ramified type theory*, which lets one avoid commitment to impredicative and ungrounded sets, was the first formally adequate response to the troubles caused by Frege's louche attitude towards defining things of the kind <u>course of values</u> (extensions).

Alonzo Church made the next big step, with his ***typed lambda calculi***.

g) Modern type theory is introduced by **Per Martin-Löf**, in the form of his ***intensional type theory***. It is a theory of the classification of things into kinds or types, and in particular of ways of defining or constructing *new* types from *old* ones. At its core is a theory of *kind* or ***type*-constructors**. A paradigm is a constructor that forms a new type of *ordered pairs*, consisting of one object of type A and another object of type B.

h) The latest version of this is **HoTT, Homotopy Type Theory**, to which the group at CMU around Steve Awody is a central contributor.

Computer scientists are introduced to type theory in their introductory courses.

Philosophers, even technically minded ones, are still pretty much left to pick it up on the street.

2. Here are the topic headings from the handout:

1) Functions are ‘complete’ or ‘unsaturated’.
2) Values, arguments, functions.
3) Ranges or courses of values.
4) Ranges of values formed by *abstraction*.
5) Introducing truth values as objects, by abstraction.
6) Sense and Reference as functions and their values.
7) Concepts are functions whose values are truth-values.
8) Definition of ‘extension of a concept’ (as course of values).
9) Statements are species of equation
10) Discerning functions in statements
11) Next generalization: allow objects in general as arguments and values of functions.
12) Objects vs. Functions, Truth-values as objects
13) Courses of values and extensions are objects.
14) Concepts must have sharp boundaries, functions must have values for every argument
15) Truth-functions: both arguments and values of functions are truth-values.
16) Generality
17) Functions whose arguments are functions
18) Grasping functions requires generalizing

Plan (after recap on raising sortal consciousness, including speculations about the empty set).

3. Frege comes out of the *GL* with a picture relating:
Singular terms, sortal kind terms, identity statements, and intersubstitution.
“No entity without identity.” Quine.

The project Frege embarks on after *GL*, in the 7 years (1884-1891) on the way to this essay, is to see what he can do with the metaconcepts of identity claims, singular terms, and functions that he had put in play in *GL*. “Function and Concept” articulates a whole, intricate metaphysical and logical picture on this very spare basis—one that should be put in a box with that of Wittgenstein’s *Tractatus*.

a) This means that there is *one* kind of sentence of which he can fully specify the conceptual content—the inferential role (including premissory role, or consequences)—namely *identities*. For “in intersubstitution, all the laws of identity are contained.” Less Germanically:
All the laws of identity are contained in intersubstitution.
And as a consequence, he has a clear metaconcept of singular term (each purporting to pick out just one *object*), as what can flank the identity sign in some *true* identity claims.

b) **<u>Leibniz's Law of Identity, Subdivided:</u>**

This claim is a way of expressing what has come to be known as ***Leibniz's Law*****,** as expressing the essence of the concept of <u>identity</u>.

It can be decomposed into two parts:

i. The **indiscernibility of identicals**: if a=b then there is no property that *a* has that *b* does not, and *vice versa*.

ii. The **identity of indiscernibles**: If there is no property that *a* has that *b* does not, and *vice versa*, then a=b.

We appealed to the indiscernibility of identicals, the claim that if *a* and *b* are *identical*, then they must have all the same properties, to argue that Goliath the statue cannot be *identical* to Lumpl, the lump of clay, even if as a matter of fact they are spatiotemporally coincident—the lump only exists in the form of the statue, and both are destroyed together—because of the different *subjunctive* properties they have that are essential elements of the criteria of identity and individuation that go with their sortals: if the shape of Goliath *were* to be radically changed, Goliath would cease to exist, but Lumpl would not. (I dismissed the attempt to get around this by restricting to the range of the indiscernibility of identicals to "non-modal" properties, on the grounds that there aren't any.)

It has turned out to be **useful to distinguish these, the 'if' from the 'only if'** of the equivalence of identity and indiscernibility in the sense of having all the same properties, because the identity of indiscernibles is potentially more controversial. It presupposes that there are *enough* properties to distinguish any two items that are in fact not identical. In an expressively impoverished scheme, this might not be so. Early jewelers could not tell jadeite from nephrite, calling them both 'jade', but they turned out *not* to be identical, once we could determine chemical mineralogical composition.

4. Suppressing (leaving in the background) the sortals, the key schema for Frege is:

Frege's Master Formula: $\mathbf{a=b \Leftrightarrow \underline{f}(a)=\underline{f}(b)}$.

'⇔' here is "if and only if". (Cf. Gerry Cohen's *bon mot*.) That is endorsing *both* sides of Leibniz's Law.

5. In addition to the fundamental identity sign, there are expressions of two fundamentally different kinds in this formula:

i. the expressions that flank the identity signs, and

ii. expressions that can *only* do so if accompanied by expressions that *can* flank the identity sign.

- What can flank the identity sign is ***singular terms*****.** They stand for *objects*, since identity claims are *recognition judgments*, that is, judgments that express the recognition of an object as the same again.

- The other expressions, the '<u>f</u>'s, are ***function*-expressions**. They are 'incomplete' in the sense that in order to get an expression that can flank the identity sign, they need to be supplemented by some singular term, some expression that can appear on one side of an identity claim.

Note that one might worry that the identity sign appears on both sides of key formula.
We are to understand the one on the right in just the same way we do the one on the left:
<u>f</u>(a)=<u>f</u>(b) ⇔ <u>g</u>(<u>f</u>(a))=<u>g</u>(<u>f</u>(b)), for all <u>f</u> and all <u>g</u>.

6. At this point, Frege makes the first of **a series of fateful and decisive moves**:
He extends the notion of <u>function</u> from relating *numbers* (of various kinds) as *arguments* and *values*, to allowing arbitrary *objects*, as arguments and values.
This is a conceptually revolutionary extension of the 19^{th} century mathematical concept <u>function.</u>
Objects, for him, are anything that can stand in identity relations.
(I am following him in not being fussy, in this motivational discussion, about the distinction between linguistic expressions and worldly items.)
That is, he will take the Master Formula to allow arbitrary *functions* on the right, in a sense that is exactly co-ordinate with the arbitrary *objects* on the left.

7. The **second bold and decisive move** in the wake of endorsing Frege's Master Formula is to formulate the basis for **the *ontology* of "Function and Concept"**:

i. What there is can be divided, ***exclusively*** and ***exhaustively*** into objects and functions.

Both of these are bold, daring claims:

- **that nothing is *both* object *and* function**, and
- **that everything is *either* object *or* function**.

This is the essence of Frege's mature metaphysics.
It is every bit as bold, powerful, and strange as Wittgenstein's metaphysics in the *Tractatus*, even though it is much less well known and understood.

a) Dividing the world exclusively and exhaustively into objects and functions (with concepts being functions whose values are truth-values).

…An object is anything that is not a function, so that an expression for it does not contain any empty place.
A statement contains no empty place, and therefore we must regard what it stands for as an object. But what a statement stands for is a truth-value. Thus the two truth-values are objects. [32]

This is meant to be a division of ***what there is, or could be***—in a sense of <u>possibility</u> that is more than merely *nomological*, since it goes deeper than laws. It is *metaphysical*.
Further, he now *starts* with a notion of "self-standing" or "complete" items, which are *objects*, and the "unsaturated" or "incomplete" *functions*, which we are to

understand in terms of *substitution* of objects "in" other objects. (An important point here is that the *arguments* of functions are *not parts* of the *values* of the functions: Stockholm is the capital of Sweden, but Sweden is not *part* of Stockholm.)

b) Here **the *name-bearer model*** comes to the fore, being prior in the order of explanation.
   We saw it in play, defined substitutionally from *conceptual contents* as *begriffliche Inhalten*, by turning the crank one more time on the machinery of observing invariance (of the goodness of implications) under substitution (in *BgS* and *GL*).
   But now a new and very different line of thought is put in play—one that *starts* with the idea of *complete* or *saturated* items, understanding *functions* in terms of them, and understanding *truth-values* in terms of them (as kinds of them), and in effect understanding *sentences* in terms of *singular terms*—terms for those very special *objects the True* and *the False*.
   In fact things are more complicated than that, since Frege in fact thinks that the *only* way to understand these *objects*—truth values—is by knowing how to use sentences to make assertions. In effect, 'true' and 'false' are to be understood in terms of practical doxastic attitudes of *taking-true* (accepting-asserting) and *taking-false* (rejecting-denying).
   But the *kind* of thing (sortally) that truth-values are is to be understood in terms of a *prior* (in some sense) notion of <u>object</u>. One big issue is how to understand the metaconcept <u>object</u>. Where have all the sortals gone?
   One reply is: they are still absolutely in the picture. For our grip on the metaconcept <u>object</u> is going to come from what it is for a *concept* (a kind of *function*) to "divide what falls under it in a definite way", that is, to be a *sortal kind term*.
c) The idea of "functions as objects with holes in them" is the theme of "Concept and Object" (CO). Here my understanding of them in terms of equivalence classes of substitutional variants offers a reading of the metaphor.
d) What is the super-sortal, the genus, of which <u>object</u> and <u>function</u> are species? This is a topic of CO.

8. <u>Functions can take all and only *objects* as arguments and values</u>

Along the way, Frege floats a view that he does not end up addressing. That is the view that functions can take *all and only* ***objects*** as *arguments* and *values*.
This is *one* of his motivations for the important (and fraught) move to assimilating *sentences* semantically to *singular terms*. That move is applying the *name-bearer model* to *sentences*.
Here there is a tension with his (good Kantian) view in *GL* that we should only understand *sub*sentential expressions in terms of the contribution those expressions make to the meanings of

sentences. This is his "context principle", laid down as one of the three most important methodological principles of the *Grundlagen*:

In the enquiry that follows, I have kept to three fundamental principles:

1. Always to separate sharply the psychological from the logical, the subjective from the objective;
2. **Never to ask for the meaning of a word in isolation, but only in the context of a proposition** [im Satzzusammenhang];
3. Never to lose sight of the distinction between concept and object.

[Introduction, p. x]

[I have not mentioned this passage before, so should do so in Week 5.]

We can think of the two 1891 essays as expanding on the third principle.

Frege can still technically satisfy the second principle in "Function and Object."

But he violates the *spirit* of it by understanding sentences on the model of singular terms, and so introducing the concept of <u>truth-values</u> as a kind of object.

Truth values are introduced by abstraction: assimilating sentences by the effects of intersubstituting them. This is clearest for what show up as *truth functions*.

But it is also where the **slingshot argument** provides a rationale.

9. <u>**Sentences are construed as singular terms**</u>:

The **third fateful and decisive move** in Frege's mature metaphysics is to understand sentences as singular terms.

He does that on the basis of the exclusive and exhaustive division of what there is into what is *complete*, in the sense that it does not need supplementation, and what is *incomplete*, in that it does. Sentences must be *either* object-expressions, that is *singular terms*, or *function*-expressions, that is, one's that are *semantically self-standing* in that they do not require *adding singular terms* as *arguments* in order to be semantically interpretable.

If sentences are *not* functions, they *must be* singular terms.

For sentences *to be* singular terms, they must 'stand for' *objects*, that is, for *identity* claims involving them to express the *recognition* of *as the same again*.

This is where Frege uses his extension of the model of identity claims as intersubstitution licenses to write expressions that no-one had ever contemplated before:

'$(2+2=4) = (2^2=4)$'

These are identity signs flanked by *sentences* rather than *singular terms*.

How are we to understand such identities?

10. <u>So, Sentences are *all* to be understood in terms of *identities*:</u>

So *all* judgments (in the sense of judgeables, what is expressible by declarative sentences) end up being understood in terms of *identity* judgments. He understands and can specify the inferential role of identity claims, as licensing intersubstitution. We can, to begin with, understand this as the endorsement of an identity claim as being the endorsement of all the inferences sharing a *pattern*: endorsing all implications of the form f(a) so f(b) and f(b) so f(a). That is using the *BgS* metaconcept of conceptual content as *begrifflich Inhalt* as what is preserved or invariant under

substitution of the equated terms. Later on in this line of thought, **Frege proposes to do the work previously done by noting substitutional invariance *salva consequentia* by noting substitutional invariance *salva veritate*.**
(Later, *GG,* he will understand *asserting* a sentence as asserting an identity holding between the sentence and the True. But that is a further step.)

11. Abstracting *Truth Values*

Construing the *reciprocal implication* of sentences of the form f(a) and sentences of the form f(b) as an *identity* claim is a kind of *abstraction*. It is how *truth-values* are abstracted.
What flanks an identity sign must be singular terms. The identity claim both expresses the recognition of an *object* as the same again and does so by licensing intersubstitutions.
We can use the conditional of *Begriffsschrift* (written as '→') to parse the inferential significance of an identity of sentences, '$(2+2=4) = (2^2=4)$' ⇔ $((2+2=4)\rightarrow(2^2=4))$ & $((2^2=4)\rightarrow(2+2=4))$.
The first excludes the case where '$2+2=4$' is true and $2^2=4$ is false, and the second excludes the case where '$2^2=4$' is true and '$2+2=4$' is false.
This biconditional relation is an equivalence relation.
So, by abstraction, we can turn it into an identity, of items of a new sortal introduced by abstraction.
If '$2+2=4$' ↔ '$2^2=4$', then TruthValue('$2+2=4$') = TruthValue('$2^2=4$').
Frege introduces the sortal *truth-value* of a sentence (compare: direction of a line) by abstraction.
He accordingly understands *sentences* as *singular terms*, whose identities express the *recognition* of an *object* of a certain kind as the *same again*.
That sortally individuating kind of object is *truth values*.

12. This move amounts to turning the *BgS* and *GL* order of explanation on its head.

Instead of starting with *implications*, divided into good and bad, and deriving *conceptual contents* of *sentences* by assimilating sentences occurring in those implications as premises (and conclusions) as having the *same* conceptual content iff they are intersubstitutable *salva consequentia*, and then assimilating *singular terms* occurring in those sentences as having the same indirect conceptual content if they are intersubstitutable *salva consequentia*, so
Implications→Sentential Contents→Singular Term Contents,
the order is going to be:
Singular Terms/Objects→Identity Sentences→Implications
(where preservation of truth-value the True is not only *necessary* for a good implication, but also *sufficient*).
From **implications→sentences→terms**
to
terms→sentences→implications.

a) **Shift from a top-down**, consequence-based approach, drilling down from goodness of implications to conceptual contents substitutionally, and then again to conceptual contents of singular terms, accordingly as their intersubstitution (explicitly licensed by identities) preserves the conceptual content of sentences (from *BgS* and *GL*), **to a bottom-up, truth-based approach**.

b) Functions, still understood substitutionally, now are understood as taking *objects* as arguments, and having *objects* as values.
c) This object-based strategy can be extended to assimilate *sentences* substitutionally by noting invariance under substitution, according to the extended notion of <u>function</u>, by introducing truth-values as *objects*.
d) Then identities license intersubstitution *salva veritate,* rather than *salva consequentia.*
e) Why treating *sentences* as a subspecies of *names* (singular terms) is strange and bold.
f) **Eventually, Frege will in effect use *both* orders of explanation, top-down *and* bottom-up, the first for *senses* and the second for *referents* (*Bedeutungen*)**. That is a story for next week.
g) Punchline 1: Can see a rationale for it in the bottom-up order of explanation, starting with the idea of names and bearers, in ***the slingshot argument.***

13. If we take seriously that order of explanation—the one developed and articulated in detail by Quine—as the triumph of an extensional semantic approach based on the *name-bearer* model of interpretation, rather than the earlier (*BgS*) *inferential-role* plus substitutional analysis model, that is, going *bottom-up*, rather than *top-down*, in terms of the inclusions of terms in sentences and sentences in implications, then we can consider the argument rationalizing this move that is provided by Quine's student Donald Davidson: **the slingshot argument**.

The Slingshot argument:

In connection with Frege's introduction of *truth values* as objects:

a) Why this is momentous.
b) Why it deserves to be controversial.
c) Davidson's "slingshot argument".

Davidson makes this argument in "Truth and Meaning", and repeats it later.

He uses it to argue against the idea of <u>facts</u>. If there were facts, there would only be *one*.

d) The argument is by substitution.

The essence of the argument is to construct a definite description, so, a singular term formed from a concept, that contains a sentence as a component. Then we look at which substitutions for that sentence preserve the object that is the referent, the value of the function.

e) The argumentative strategy is Quinean, and bottom-up, rather than Fregean and top-down.

That is, it assumes a notion of <u>extensionality</u> according to which a context (predicate) is *extensional* in case substitution of *coreferential* terms is *salva veritate*.

So it assumes that we know what it is for terms to be coreferential.

We use that notion to understand extensional contexts.
Then argue that *any* expressions that are intersubstitutable *salva veritate* in *those* contexts must be coreferential.
Then look at definite descriptions of the form ιx[Fx & p], and consider sentences such as:
ιx[Fx & p] = Tom Doniphone.
For instance, "the man such that
i) he shot Liberty Valance, and
ii) The town of Shinbone is in Wyoming."
If Shinbone is in Wyoming, then that term refers to Tom Doniphone, since he actually shot Liberty Valance.
If not, then there is no-one such that *both* (i) and (ii) hold.
Suppose it is true.
The question then is: which substitutions for *p* preserve the truth of:
The man who (shot Liberty Valance and Shinbone is in the Western Territories) is (=) Tom Doniphone.
The answer is: *any* other true claim.
Just as we can substitute any coreferential term for 'Liberty Valance', such as "the character played by Lee Marvin the movie 'Who Shot Liberty Valance?'" *salva veritate,* so we can substitute any other true sentence for "Shinbone is in the Western Territories."
The conclusion is that all true sentences have the *same* referent.
So we can call that referent the truth-value *True*.

Worry: definite descriptions that include declarative sentences as conjuncts are weird. We can form them in *PM*-ese, but do they correspond to any natural language sentences?
If not, does that matter to the cogency of the argument?

Compare: x is a fridgeon iff x is a photon and Fodor's fridge is on.
The second conjunct is a sentence, and the definition lets us form a definite description: The fridgeon that was sent toward the screen in the two-slit experiment.
(Note: The human eye can detect a single photon—the light that would reach it from a candle a mile away. Can it detect fridgeons?)

14. **From truth *values* to truth *functions***:
Once we have abstracted truth-values, we can consider functions that have truth-values as *arguments* and *values*: these are truth *functions*.

15. Defining the Metaconcept Concept:

We can also consider functions that take *singular terms* as *arguments* and *truth-values* as *values*. These functions are Fregean *concepts*.
Treating truth-values as objects—by *abstraction*—is the crucial move in **understanding *concepts* as *functions***. For Frege, concepts are just functions that have truth-values as values, whatever the arguments. Along the way, he invents the concept <u>truth function</u>, as functions that take truth-values *both* as values *and* as arguments.

16. **<u>Courses of values (Wertverlaüfe)</u>**:
We have already looked at the Fregean Master Formula:
a=b ⇔ <u>f</u>(a)=<u>f</u>(b), where the function-expression '<u>f</u>' is schematic, allowing any function-expression to go in its place, we can also consider a different schema:
f(<u>x</u>) = g(<u>x</u>), for schematic singular term <u>x</u> and two specific function-expressions 'f' and 'g'. That generalized identity is an equivalence relation, and we can use *abstraction* to turn it into an identity among new singular terms of a new sortal kind: *courses of values*.
'εf(ε) = 'αg(α) iff ∀x(f(x)=g(x)).
Courses of values, as the (extensional) *graphs* of functions, are conceptually and explanatorily *less basic* than the functions from which they are abstracted.
This is done by *abstraction*: we generalize the notion of one-to-one correspondence, to say that two functions have the same course of values (compare: two lines have the same direction) in case ∀x[F(x) ⇔ G(x)].
This is the move that causes the Russell-paradox inconsistency in *Grundgesetze*.
Even if functions are not objects, and so not as such *countable*, their courses of values are objects, and are countable.
Forming courses of values, or extensions as singular terms. This is by abstraction: "***the*** course of values of the sortal concept K" (will be able to extend this to attributive concepts, so long as the terms they apply to are properly sortalized)—and the definition is *really* applicable to *all* functions, not just *concepts* (whether sortal or attributive, including relational ones of both kinds [note that there *are relational sortals* such as <u>parent</u>, or—in the recent English coinage modeled on <u>sibling</u>—<u>niblings</u> and <u>piblings</u>, [which sound like something Dorothy ran across in Oz])—should be seen as analogous to "the direction of line *l*," in that it *is* a definite description formed by abstraction from a prior sortal and the terms it governs. So must show existence and uniqueness of *the* course of values of any function. Again, one wants to say that what is abstracted *from*, the function, must be grasped *first*, and the objects defined from the functions by abstraction should be conceptually downstream from these.

17. <u>Courses of values of functions whose values are truth-values are the *extensions* of concepts.</u>
This redeems the promissory note issued (on a charitable reading) in the infamous footnote late in *GL* in which he says "**I assume it is known what the extension of a concept is.**"

Introducing *courses of values*. Doing this is making good on the jaw-dropping footnote in *GL*: "I assume it is known what the extension of a concept is." Here he provides the cash for that promissory note. This is a way of defining *objects* from *concepts* (a kind of *function*). It can be laid alongside forming *objects* from *concepts* by *definite descriptions* (by showing *existence* and *uniqueness*).

18. <u>Understanding the sense and reference of concepts in terms of functions and their values:</u>

Just as the singular terms '2+2' and '(2^2)' involve the application of different functions, but have the same value, so the expressions '2+2=4' and '2^2=4' express different senses, because they involve different functions. But they have the same reference, since (2+2=4) = (2^2=4).
This identifies *senses*, in effect, with *computations*, individuated by the arguments, and the functions applied to those arguments, and *referents*, the *values* of those functions for particular arguments, as the *results* of the computation.
Cf. the Curry-Howard version of *computational trinitarianism*.
Senses and referents of terms (including sentences).

a) Within the confines of the bottom-up, singular-term-centered order of explanation, distinguish
   i. referents of expressions as objects that are *values* of applying *functions* to objects that are *arguments*
   ii. senses as the function-argument pairs that yield the referents as values.
b) What makes sense about this. Its virtues.
   Specific sense in which *senses determine referents*.
c) Looking ahead: This is *not*, or at least not *obviously* (I claim), the view that we get in USB. For with functions, too, we distinguish senses from referents. The issue of *sense* functions.
d) It amounts to treating senses as means of *computing* referents as values from arguments.
   Senses are computations: programs.
   A series of *steps*, starting from *argument*, to arrive at *value* of function.
e) Frege sharply distinguishes:
   i. Functions, from
   ii. *Graphs* of functions: the pairs of (in the simplest case) arguments and values.

   He insists that the functions are *conceptually* or *explanatorily* prior to the *courses of values* that they determine. The courses of values of two functions can be identical, even though the functions whose courses of values they are are *not* identical.
   Further, courses of values are introduced *from* functions, by *abstraction*—that is, by assimilating functions by the values they assign to arguments.
   There is no backward road: from the course of values one cannot determine which of the many functions that could have produced it.
f) Functions can be understood as means of computing courses of values.

The function is a program that computes the course of values, assigning to each argument a value. That function = program is the *sense* of a functional expression.

As such, it is "incomplete" or "unsaturated".

When "completed", by giving it an *argument*, it determines an object (which might be at truth-value), the *value* of the *function* for the *argument*. That object is the *Bedeutung*, the referent, of the resulting "complete" expression.

g) Punchline: **Computational Trinitarianism.** This is showing an *isomorphism* between
   i. Functions
   ii. Programs
   iii. Proofs (specifically, intuitionist proofs)

h) This is *not* the full Curry-Howard-Lambek correspondence (isomorphism), which adds *category theory*.

Wikipedia:

The **Curry-Howard-Lambek correspondence** is a profound and elegant connection between three seemingly distinct domains:

- **Logic** (proofs and propositions)
- **Computation** (programs and types)
- **Category theory** (morphisms and objects)

It reveals that these domains are structurally analogous, meaning that concepts in one domain can be translated into the others in a principled way.

This means:

A **proof** of a proposition is like a **program** of a type.

A **type** in programming corresponds to a **logical formula**.

A **morphism** in a cartesian closed category (CCC) corresponds to a **proof** or **program**

Historical Development

- **Curry–Howard**: Initially discovered by Haskell Curry and William Howard, this part of the correspondence links **intuitionistic logic** with **typed lambda calculus**.
- **Lambek**: Joachim Lambek extended the correspondence to **category theory**, showing that cartesian closed categories can model both logic and computation.

This correspondence is foundational in:

Functional programming languages (like Haskell, OCaml, Scala)

Proof assistants (like Coq, Agda)

Type theory and formal verification

Category-theoretic semantics of computation

It allows us to reason about programs as if they were proofs, and vice versa—bridging logic, computer science, and mathematics in a unified framework.

The Curry–Howard correspondence is the observation that there is an isomorphism between the proof systems, and the models of computation. It is the statement that these two families of formalisms can be considered as identical.
If one abstracts on the peculiarities of either formalism, the following generalization arises: *a proof is a program, and the formula it proves is the type for the program*. More informally, this can be seen as an analogy that states that the return type of a function (i.e., the type of values returned by a function) is analogous to a logical theorem, subject to hypotheses corresponding to the types of the argument values passed to the function; and that the program to compute that function is analogous to a proof of that theorem. This sets a form of logic programming on a rigorous foundation: *proofs can be represented as programs, and especially as lambda terms*, or *proofs can be* ***run***.

Note that the account of senses (*Sinne*) in *USB* is different, although it builds on this one.
My idea is that the key to that account is that an expression can be variously analyzed into functions and arguments. For sentences, there are multiple functional decompositions or analyses of '2^4=16'.
All of these, the entire constellation of function-argument analyses, together make up the sense of the expression, the thought expressed by, '2^4=16'.
But this is the topic for next time.

19. **Higher-order functions**:
Q: How can we understand functions that take *functions* as arguments?
A1: In terms of their *courses of values*, which are *objects*.
But even doing this requires great care.
For what needs to be arguments and values of these, and so substitutable, in the case of metaconcepts applying to concpets, is *complex* predicates, not *simple* predicates, in Dummett's sense.

a) Frege floats an idea as a ladder he climbs up and goes beyond: that functions can relate *any* objects as arguments and values.
b) But he does not restrict the arguments or values to objects. He allows *functions* to be arguments and values of functions as well.
c) This is what gets him into trouble with Axiom V of *GG*.
d) But it is delicate in any case.
e) What makes it delicate is that functions are not substitutable-for in the way singular terms are. This is because of the cross-references between their argument-places.
f) The punchline is **Dummett's distinction between *simple* and *complex* predicates**.
Complex predicates are not *parts of* sentences, but *patterns in* sentences.
That is why they cannot straightforwardly be substituted *for*.
Need to discern complex predicates to codify certain important kinds of implication.

Since such is the essence of the function, we can explain why, on the one hand, we recognize the same function in $2*1^3 +1$ and $2*2^3 +2$ even though these expressions stand for different numbers, whereas, on the other hand, we do not find one and the same function in $2*1^3 +1$ and 4-1, in spite of their equal numerical values.
This passage shows that the cross-refencing is crucial. Only *some* of the '2's in the original formula are replaced, if we go back from the second to the first. $2*2^3+2$ is *also* an instance of the function $x*x^3+2$, which $2*1^3+1$ is *not*.

Relations of concepts (functions) and objects as *exclusive* categories of what there is, is the topic of "Concept and Object," which we'll talk about next time—as well as USB.

20. Six periods in the development of Frege's views about conceptual content—and in particular, the relations between inference-based and truth-based conceptions, with the latter modeling the content of *sentences* on that of *singular terms*.

i. *Begriffsschrift*.
ii. *Grundlagen.*
iii. *Funktion und Gegenstand.*
iv. *Uber Sinn und Bedeutung*.
v. *Grundgesetze*
vi. *The Thought*, and *Negation*

21. One will see, particularly from last time and this time, how I think we should read the mighty dead philosophers: by thinking *with* them, about the important topics they are addressing. This is worth doing in a *philosophical* spirit, as opposed to an *antiquarian* one, only if one is convinced they have lessons to teach (that they understood something other folks didn't or don't), and that we have not yet fully digested the consequences of those lessons.
I am principally interested in what we can learn from Frege, what lessons we can draw from his discussion that can be applied in thinking about issues today.