New Geometries

When Mathematical Logic wasn't even born yet, the concept of mathematical reality was already awakening. Two thousand years earlier, Euclid laid down the axioms for Geometry and actually this initiated this new awakening of realities.
His axiom for parallelity was pretty ugly and already he thought maybe it could be derived from the other much simpler axioms. Later many tried continually and the search was indirectness.
They simply added the negation of the parallelity axiom to the others and started to search for contradiction. Had they found one then they could easily reverse the arguments into a direct proof of the parallelity axiom. But instead of contradiction, strange but beautiful possibilities opened. Janos Bolyai wrote: "From nothing I created another new world." Gauss went on the same road and wrote a very rude reply to Janos who turned to him with his results. But as rude Gauss was, he was right. These mere possibilities meant nothing tangible yet. The coin didn't drop neither for Gauss nor Bolyai because they envisioned their new worlds as alternate physical reality behind geometry.
Few decades later the German mathematicians realized that using not real lines as the "lines", rather intervals or circles, the parallelity axiom can be false. The most embarrassingly simple example was actually known for hundreds of years as spherical geometry. Perfectly visual on the surface of our globe. Here the lines are the main circles like the equator or the time lines. The smaller circles remain circles.
The crucial tricky "point" is to regard the opposite points of the globe as single points.
Then it becomes true that two lines cross in a point, two points determine a line, and so on.
But voila, there are no parallels at all! All main circles cross in "a point".
Unfortunately, here other simpler axioms fail too. Namely, that lines must be infinite long.
But there are models where distorting the lengths, this fact can remain too formally.
And yet the coin still didn't drop after these complicated models.
Finally, it was Beltrami who realized that they had already found what they were looking for.
Indeed, the origin of the whole contradiction searches was to see if the parallelity axiom could be derived from the others. Now if there are models where all the other axioms are true but the parallelity is not, then "game's over"! Indeed, if the parallelity axiom would follow from the others then this would mean that in all realities where the other axioms are true the parallelity is true too. But some of these inner models showed the opposite.
So for reality, mathematics doesn't have to go to physics, it has its own realities inside other realities. Probably this new vision of reality was fermenting the concept of sets.
Strangely, this stayed unconscious because Cantor didn't see this role of sets.
This is typical for the deep geniuses. They are like prophets and receive messages not fully comprehensible for themselves. In science this means discovering things that go way beyond their own visions. The three such deep geniuses were Cantor, Einstein and Turing.
As opposed, the three super geniuses Newton, Gauss and Gödel were able to see what the already achieved facts necessitate as a new vision.
As history proved, there are always depths that the widest visions still miss.

Variables and Quantors

We turn to the most crucial application of Taboo Avoidance explained at the end of book Sets.
First we need a bit of observation about the used "or" meaning in V and its negativity as falsity of their "and". In truth, we have a deeper layer of "and" already in every v because the values in a window have to be all correct. So actually in the finite subsets of V we have a double layer of finite many "or"-s and "and"-s. This combination as fundamental came out only pretty late in Logic. So now we go back to Aristotle who made an amazing system that remained hollow. We mean Formal Logic which by the time of enlightenment was regarded as truly just a "formal" emptiness even by religious scholars.
Arguments like "If every horse is an animal and some horses are blah blah blah . . " were indeed irrelevant for deeper thinking. So it's quite astonishing that when mathematical logic started, the same "If-then" sentences were formed.

But that was only the first stage and later by the new name of Proof Theory, the language shifted towards using "and" and "or". So what stayed and what changed?
Beside this mentioned "If-then" obsession, the even more important déjà vu of the new logic with Formal Logic were the "every" and "some", appearing also in our mockery of Aristotle.
Without getting into metaphysics, we must realize that though "some" is obviously much weaker than "every", "some" actually hides a deeper essence of "every". Simply because when we claim that "some . . . " then we actually claim that this some exists. So the best is to regard "some" as "there is". Existence or "being" and negatives as "nothing" were parts of not only old metaphysical arguments but in almost all twentieth century titles of existentialism.
It's a tragedy that most of these thinkers never really understood the crucial new realization of Mathematical Logic and so they still talked with an ancient blindness.
But what was this blindness and also the crucial failing of Aristotle in spite of his incredible foresight to regard the "every" and "there is" operations as vital?
The answer in one word is, "variables"!
You may say, come on, variables were part of mathematics since ancient times!
But you are wrong! A consistent variable usage in equations and mathematical operations was not present even at the beginning of the nineteenth century. Even by Gauss!
The confusion stems from the fact that we rewrite history and quote the old mathematicians with new formalism! Today we serve mathematics on the silver platter of this relatively new abstraction and kids already in elementary school use variables. Unfortunately, there were some idiots quite recently who tried to go against this and used blank squares, triangles and circles as variables. Thinking that this "helps the kids". But these idiots died out. (I hope!)
The fact that $x , y , z , \ . \ . \ .$ can mean any number is an a priory ability of human understanding but it simply never got to this pure level earlier. It is a mystery why it took so long, because there is no biological evolution involved. An ancient Egyptian boy brought back to the present would learn math just as well in elementary school as the others. Probably would hate it as most kids do at present too, but this is a different story. It's the socially communicated language that carries the evolution! So were only these $x , y , z , \ . \ . \ .$ missing from the "If every horse is an animal, . . " kind of arguments too? YES! Natural language is a maze of "cover up"-s to hide what lies underneath what we say. But what we mean and how we think, involves these!
The use of variables! We don't hide the use of "everybody" or "somebody" but we apply these quantity prefixes or so called "quantors" directly to the nouns, adjectives and verbs that hide the properties or relations of objects. So drawing attention to that "being a bicycle" or "being red" or "owning something" are merely properties or relations of objects is not enough!
We could say that B is abbreviating the bicycle feature, R is being red and O the ownership.
But applying the quantors to these directly as Aristotle did becomes hollow.
What we need is $B(x)$, $R(x)$ and $O(x,y)$. We might think that this is just stupid over complication because these variables are irrelevant. But they are not! Their choices are yes but how they interrelate with each other is not.
Plus there is a crucial application of these variable depending relations when we use them not with variables rather with fix concrete objects or persons. My bicycle is a concrete object but I don't name it. I only regard myself as concrete named object as "I" hidden in "my".
So my bicycle should be merely an x , for which $B(x)$ and $O(I , x) = B(x) \wedge O(I , x)$ is true.
But this naming of myself as "I" is actually a still not perfectly solved problem because everybody's sentence means a different "I"! Using Peter or P as name is much clearer!
Obviously, there are many Peters too but when I talk about Peter, I usually mean a concrete person. To say that somebody is called Peter is again a controversial claim because we can not use the Peter in place of him yet. So the name assignments and subjective "I" sentences are a complicated and excluded field in our Mathematical Logic. Luckily, most claims about the world are not such and can be perfectly described.
So Mathematical Logic is not just mathematical at all! It should be called Grammatics and taught in elementary school second grade before even getting to math! This ain't gonna happen for still few hundred years! But let's continue in the present! The claim that Peter owns a red bicycle is this: $\exists y \ (\ O \ (\ P \ , \ y \ ) \wedge B(y) \wedge R(y) \ )$. The bracketed three claims combined with "and" is clear but what is this ancient hieroglyph in the front?

Well, it is the “there is” quantor applied for the y variable.
And that was the whole point I was on about what Aristotle missed.
Now the quantors come alive! In fact, they kill the variable they use! So y is not a variable anymore! It was needed merely to be quantized. As we say, the full expression is closed.
The other closing was by using P = Peter. So if we had started with that open too, then we get the $(O (x , y) \wedge B(y) \wedge R(y))$ two variable expression.
This could be closed by other ways and thus make other statements about the world.
$\exists x \exists y (O (x , y) \wedge B(y) \wedge R(y))$ for example claims that there is somebody who owns a red bicycle. Now we’re cooking, especially if we introduce the other quantor $\forall$.
$\forall x \exists y (O (x , y) \wedge B(y) \wedge R(y))$ means that everybody owns a red bicycle.
Clearly not true in our world and even the $\forall x \exists y (O (x , y) \wedge B(y))$ claim that everybody owns a bicycle is false.
The $\forall x \exists y (O (x , y) \wedge R(y))$ claim that everybody owns something red feels to be true but actually if we allow the $O (x , y)$ ownership to mean wider objects than people in its x variable then it is also false. Now we can come to the newest shift I talked about, the shift from the “implication” obsession to the and-or supremacy!
The “or” is abbreviated as $\vee$ or $|$ or ∇ because we use “or” in three different sense.
The first most frequent means that we only regard it to be false if neither members are true.
In short: $f \vee f = f$ but $t \vee f$, $f \vee t$, $t \vee t = t$. The $|$ exclusion “or” on the other hand is only false if both members are true: $t | t = f$ while $t | f$, $f | t$, $f | f = t$. And finally the ∇ “either-or” is only true if exactly one member is true, so: $t \nabla t = f$, $t \nabla f$, $f \nabla t = t$, $f \nabla f = f$.
For the controversial $\rightarrow$ implication, strangely there is no controversy at all!
It is only to be false if from true assumption we get false consequence, so:
$t \rightarrow f = f$ but $t \rightarrow t$, $f \rightarrow t$, $f \rightarrow f = t$. Thus these statements are true:
If the Pope is a woman then the world is round.
If the Pope is a woman then flies can talk.
If the Pope is a man then then the world is round
And this last shows that the truth doesn’t imply any cause and effect relation.
Observe that $\_ ? \_ = \_$ can have sixteen possibility as truth definition for ? because for the four possible input combinations we can have this many different t , f assignments.
We only encountered five: $\wedge$, $\vee$, $|$, ∇ , $\rightarrow$ and a sixth is the frequently used $\leftrightarrow$ which suggest it’s meaning, “if and only if” or “then only then” which are the same.
Clearly, $A \leftrightarrow B$ should be $A \rightarrow B \wedge B \rightarrow A$ and this raises the question of really how many such operations are needed to express the others. But we still didn’t mention the most important truth operation which uses single input, the negation, abbreviated as $\neg$.
Amazingly, any single truth or false valued operation would be enough to express all others with using negations and this would suggest the implication obsession and indeed there were formalizations using only $\rightarrow$ and $\neg$. But then again we could use only $\vee$ and $\neg$ too.
The replacement of $\rightarrow$ with $\vee$ is interesting: $A \rightarrow B = \neg A \vee B$ says that the assumption is false so we can not use it or the consequence has to be true too.
$A \rightarrow B = \neg (A \wedge \neg B)$ is much better in spite of being longer because it really says what $\rightarrow$ means, being false only if a true assumption implies something false.
But the really important rules for us are these:

$\neg (A \wedge B) = \neg A \vee \neg B$ and $\neg (A \vee B) = \neg A \wedge \neg B$

$\neg \forall x \ldots = \exists x \neg \ldots$ and $\neg \exists x \ldots = \forall x \neg \ldots$

$A \wedge (B_1 \vee B_2 \vee \ldots \vee B_n) = (A \wedge B_1) \vee (A \wedge B_2) \vee \ldots \vee (A \wedge B_n)$ and

$A \vee (B_1 \wedge B_2 \wedge \ldots \wedge B_n) = (A \vee B_1) \wedge (A \vee B_2) \wedge \ldots \wedge (A \vee B_n)$

The first four, allow to drive $\neg$ inward, reaching finally the basic relations.
The basic relations or their negations are also called literals.

The last two allow to drive either one of $\wedge$, $\vee$ inward again and so we can end up with the already mentioned double layer of them either one being in the bottom.
Just like at the predictions, we'll prefer an "or" of "and"-s.
There is also a way by changing variables to drive the quantors outward.
These two, bringing all quantors in front and moving the $\wedge$-s in, are not essential for the system that I will show, only the moving of all $\neg$ in, to form literals.
The above rules of quantor negations forces me to tell this story of my first Analysis class at uni in Budapest. We got a new lecturer Laszlo Czach who studied in the Soviet Union.
He wrote on the blackboard the following sentence:
"Every woman has a moment in her life when she'd like to do that's not alright."
It sounds much more rhythmical in Hungarian and as Czach explained, it was from an old song.
He also said that this will be crucial to understand analysis which made everybody giggle.
Then he asked who could tell what the negation of the sentence is.
Many didn't know but soon we all agreed that it is:
"There is woman who has no moments when she'd like to do that is not alright."
Let's formalize! W (x) = x is woman, A (x) = x is alright to do,
L (x , y , z) = x would like to do at y time z . Thus the original statement is:

$$\forall x \ \{ W(x) \ \rightarrow \ \exists y \, \exists z \, [L(x,y,z) \ \wedge \neg A(z)] \}$$

Let's see the negative by our rules:

$$\neg \ \forall x \ \{ W(x) \ \rightarrow \ . \ . \ . \ . \ \} \ = \ \exists x \ \neg \ \{ W(x) \ \rightarrow \ . \ . \ . \ . \ \} \ =$$

$$\exists x \ \neg \ \neg \ \{ W(x) \ \wedge \ \neg \ \exists y \exists z \, [L(x,y,z) \ \wedge \ \neg \ A(z)] \ \} \ =$$

$$\exists x \ \{ W(x) \ \wedge \ \forall y \forall z \ \neg \ [L(x,y,z) \ \wedge \ \neg \ A(z)] \ \} \ =$$

$$\exists x \ \{ W(x) \ \wedge \ \forall y \forall z \ [L(x,y,z) \ \rightarrow \ A(z)] \ \}$$

There is a woman and at any time for anything if she'd like to do it, it's alright.
We of course prefer not to use $\rightarrow$ so:

$$\exists x \ \{ W(x) \ \wedge \ \forall y \forall z \ [\neg L(x,y,z) \ \vee \ A(z)] \ \}$$

There is a woman that at any time for anything, she wouldn't like to do, it is alright.

Situation matrix

The system of Logic I will reveal, uses this language but as I said we can go even further.
We can bring all quantors to the front and all the appearing $\wedge$-s down to the literals.
These $\wedge$-s of literals could be called scenarios because their $\vee$ combining is the whole claim which we call a situation. Actually, we can avoid to use $\wedge$ and $\vee$ altogether if we separate the literals in the scenarios by commas and write the scenarios under each other.
So the full claim becomes a situation matrix preceded by the quantors:

$$\forall \ x \ \exists \ y \ . \ . \ \begin{bmatrix} b_4(x,y) \ , \ \neg b_3 \\ \neg b_2(z) \ , \ b_3(x) \\ b_4(y) \end{bmatrix}$$

The b-s are some basic relations or properties of our language.
Now I show a classic statement of Euclid to be expressed in this manner.
It claims that there are infinite many primes.
The primes are numbers above 1 that are not composites and to be composite means to be a product with members above 1.
So we'll clearly need the $>$ bigger relation and the $x \cdot y = z$ multiplication relation.
The claimed infinity can be expressed with $>$ by saying that there is prime above any number.
In fact, we can avoid the condition of primes to be bigger than 1 too because being above any value clearly implies to be above 1 too.
So we merely have to claim that there is a non composite y above any x :

$$\forall x\, \exists y \left\{ y > x \wedge \neg\, (\exists z_1 \exists z_2 [z_1 > 1 \wedge z_2 > 1 \wedge z_1 \cdot z_2 = y]) \right\} =$$

$$\forall x\, \exists y \left\{ y > x \wedge \forall z_1 \forall z_2 \neg\, [z_1 > 1 \wedge z_2 > 1 \wedge z_1 \cdot z_2 = y] \right\} =$$

$$\forall x\, \exists y\, \forall z_1 \forall z_2 \left\{ y > x \wedge [\neg\, (z_1 > 1) \vee \neg (z_2 > 1) \vee \neg (z_1 \cdot z_2 = y)] \right\} =$$

$$\forall x\, \exists y\, \forall z_1 \forall z_2 \left\{ [y > x \wedge \neg (z_1 > 1)] \vee [y > x \wedge \neg (z_2 > 1)] \vee [y > x \wedge \neg (z_1 \cdot z_2 = y)] \right\}$$

This is ready to be written in matrix form but can be written nicer if we realize that not being greater than 1 actually means being 1 and we use the $\neq$ symbol for $\neg (z_1 \cdot z_2 = y)$:

$$\forall x\, \exists y\, \forall z_1 \forall z_2 \left\{ [y > x \wedge z_1 = 1] \vee [y > x \wedge z_2 = 1] \vee [y > x \wedge z_1 \cdot z_2 \neq y)] \right\} =$$

$$\forall x\, \exists y\, \forall z_1 \forall z_2 \begin{bmatrix} y > x & , & z_1 = 1 \\ y > x & , & z_2 = 1 \\ y > x & , & z_1 \cdot z_2 \neq y \end{bmatrix}$$

Amazingly, every mathematical statement is such quantized matrix, using of course other basic relations than our $y > x$ and $x \cdot y = z$ were and using other names than our 1 was.
In fact, names are the key to develop a system of logical rules that guarantees truths without knowing the meanings. Truths should be valid in all realities but to call them logical necessities suggests that we need a derivation of these without checking realities.
And amazingly, the two quantor's trivial meanings are enough to find such derivation system if we combine this with name replacements, that is concretizations.

<u>Consistency Logic Of Concretizations</u>

Imagine we have a set of initial statements and using only two possible concretizations, we add new statements to our set one by one, allowing to use the added ones too.
The two allowed rules are:

$\forall x \,..\, (x) \quad \Rightarrow \quad ..\, (n)$ for any n already used or new name.
$\exists x \,..\, (x) \quad \Rightarrow \quad ..\, (n)$ for any new n name that we have not used yet.

The dots stand for further possible quantors. Left of the $\Rightarrow$ symbol is the used statement being already a member in our widening set and the statement to the right is the new added one.
In the bracket, abbreviating the matrix we mean all replacements of x appearances by n.
These are the safest claims that definitely must follow from already established truths.

Indeed, the meaning of $\forall$ is that all objects obey it.
For an existing object we can introduce a new name.
Quite unbelievably, these two safest rules are actually enough to derive all truths!
But first something different is our claim:
Suppose we find finite many $C_1, C_2, \ldots, C_n$ members in our widened set that:
1. These are all totally concretized, that is contain no quantors in front at all.
2. No matter how we choose a line from each of these concretized matrixes, there will be two lines that contain both $b_k(\ldots)$ and $\neg b_k(\ldots)$.
Then our initial statements are impossible to be true together.

In 2. the dots in the parenthesis abbreviate the same exact names.
And we claim the existence of such contradicting pairs with every possible line selections.
Only if all scenario selections can not escape contradiction do we have a real contradiction. Finally, observe that we said "there will be two lines that contain both . . ".
This is a bit confusing as if we had suggested that each will contain one. And indeed, this is the typical but we allowed that already one line contains both. That of course means that that line is already contradictory in that matrix. But then how come we didn't avoid it by choosing other?
Well, maybe we couldn't because the others would bring about other contradictions.
So the possibilities are complex but one thing is sure. This observation makes our initial statement set impossible. Now we must go from impossibility to necessity.
The idea is this: We use as initial set, statements that all obey the meanings of $\forall$ and $\exists$ among our names. Then we want to use concretizations so that this obeying remains.
So when we decide to make a $\forall$ statement's concretization, we must make it for all already used names and when we create a new name for $\exists$ we must make concretizations for all already existing $\forall$ statements using this new name.
Doing one, at once necessitates the other because new quantors become the first.
Of course the number of quantors decrease and sometimes they all disappear again and again.
This is good because we'll have a surprise about these fully concretized statements.
If our initial set is a single statement then the start is almost perfect because we only have to take care of the appearing finite many names that initiate the widening sequence.
We'll widen step by step but the perfect widenings require to do finite groups and returns.
So it becomes a back and forth game. At any point we'll have only finite many added ones that definitely do not obey all the $\forall$ meanings. Yet the full sequence will do because any partial concretization only uses finite many names. In essence, we have to apply this to each statement in a bigger initial set but make sure that the new name is new in the total set and the application of all earlier names is regarded in the full set too.
If the initial set is a single sequence, we can go in a similar back and forth way through all of them. If they are more, we need to be even trickier.
Now comes the surprise, namely an application of the Compactness Theorem.
The fully concretized matrixes are the s elements and their lines are the $k(s)$ possible choices.
The taboos are any finite sets of contradicting lines, that is ones that contain contradicting literals. The taboo avoidance on finite sets is our assumption of no logical contradiction.
And then the Compactness Theorem implies a choice function, that is a full selection of lines without contradiction from all fully concretized matrixes.
The literals of these tell how to define a reality because there are no contradicting ones.
What's more, this reality will satisfy our quantor meanings step by step through the partial quantifications way back to the original set of statements. So we obtained a model, proving that our method of showing the impossibility of an initial set is perfect. If we can not get a finitely contradicting $C_1, C_2, \ldots, C_n$ set then the initial set is possible.
What's more, if we can turn the just mentioned idea that the taboo avoidance on finite sets should define our Logic, then we can also prove that this Logic is Complete.
Meaning that it proves anything that could be by any Logic that obeys Models.
We must go back a little to fully understand this:
Models obey Logic as Beltrami realized it even before Logic was defined.

So a statement that can be derived by a correct Logic should be true in all models.
A negative, reversed version means that what is not true in a model that shouldn't be derivable.
Actually this was Beltrami's recognition for the parallelity axiom. Logic can not derive it from the other axioms because there are models where those axioms are true but the parallelity axiom is not. But a true reversal of the models obeying Logic means that if a statement is true in all models then it must be derivable too. Or again in negative form: If a statement is not derivable then it must be false in some model. So this is what we meant by Logic obeying models.
Now we can have an other twist and regard the statement itself as a negated one.
This way we can simplify the situation because we don't have to regard some assumed axioms and the statement being consequence of them. Indeed, if we accept indirectness as an always usable method then the negative of a statement leading to contradiction should always imply a derivability of the statement directly too. Or again in negative: A non derivability must mean that adding the negated statement to the other assumptions we can not get a contradiction.
And then we can also say that a Logic obeys models if for any non contradicting or in other words consistent set of statements there is a model where all the statements are true.
So in a sense, we can define our Logic by first defining our consistency. Namely, as not being able to derive finite many concretized matrixes where all line choices are contradicting.
And indeed, not to be able to derive such, implies a model where our initial statements are all true. We simply must do our extensions by the two concretizations smartly as we'll show for a single statement. Thus a consistency guarantees a possible model. So our Logic is complete.

So let our example of a single initial statement be: $\forall x\ \exists y\ \exists z\ \forall w\ (\ 1 , x , y , z , w\ )$.
Its matrix already contains the 1 name. As new names I will continue to use 2 , 3 , . .
To make the notation even simpler I will avoid the x , y , z , w variables and the matrix too.
Instead, I just write name and quantor quadruples that show which variables are replaced by the names under the original quantors.

∀	∃	∃	∀	abbreviating the $\forall x\ \exists y\ \exists z\ \forall w\ (\ 1 , x , y , z , w\ )$ statement.
1	∃	∃	∀	Using all names 1 in frontal ∀ -s.
1	2	∃	∀	New names 2 , 3 for frontal ∃ -s of the previous group.
1	2	3	∀	
2	∃	∃	∀	
3	∃	∃	∀	

1	2	3	1	Using new names 2 , 3 in frontal ∀ -s before the last group. Using all names 1 , 2 , 3 in frontal ∀ -s of the last group.
1	2	3	2	
1	2	3	3	
2	4	∃	∀	
3	5	∃	∀	New names: 4 , 5 , 6 , 7 for frontal ∃ -s of the previous group.
2	4	6	∀	
3	5	7	∀	

4	∃	∃	∀
5	∃	∃	∀
6	∃	∃	∀
7	∃	∃	∀
1	2	3	4
1	2	3	5
1	2	3	6
1	2	3	7
------	------	------	------
2	4	6	1
2	4	6	2
2	4	6	3
2	4	6	4
2	4	6	5
2	4	6	6
2	4	6	7
3	5	7	1
3	5	7	2
3	5	7	3
3	5	7	4
3	5	7	5
3	5	7	6
3	5	7	7

Using new names 4 , 5 , 6 , 7 in frontal ∀ -s before the last g.
Using all names 1, 2, 3, 4, 5, 6, 7 in frontal ∀ -s of the last g.

4	8	∃	∀
5	9	∃	∀
6	10	∃	∀
7	11	∃	∀

.
.

New names: 8 , 9 , 10 , 11 , 12 , 13 , 14 , 15 for frontal ∃ -s of the previous group.

Logic Of Formulas

Variables were the magic wand that rejuvenated Formal Logic. Then Logic went overboard and realized that variables are also usable to replace the artificial names we used so successfully.
So the Logic Of Concretizations got buried and most text books reflect this denial.
The plus side is that indirectness, the contradiction search also stepped back and instead from the axioms we go step by step to derive new Formulas.
So this seems to correspond to the old classical, unconscious Logic. But this is a mirage.
The reason is that I had to say Formulas above! Indeed, due to the avoidance of artificial logical names, we must derive not only statements but open formulas too, that imitate the partial concretizations. So to define a Logic, using formulas is not only the most illogical thing, it is actually a didactical "crime" in view of the crystal clear system of concretizations we showed.
But the truth is even worse because the previously introduced matrixes were only crystallized even later. So for a long time the open logic was formulated using implications.
I will not even go into this "ancient relic" of Hilbertian quantor axioms because the and-or supremacy is unquestionable by today. Strangely, the new went back to the old buried treasure.
Namely, a simple system of Tait can be very naturally derived by our earlier results.
The fundamental fact is that we'll regard as Logical axioms all finite $\vee$ expressions that contain contradicting literals. This corresponds to our earlier finite many matrix lines containing contradicting literals but now it can contain anything else too.
Earlier it was a contradiction because we had $\wedge$-s but now with $\vee$-s this means a trivial necessity. Indeed, an "or" that contains opposite members must definitely be true.
The big extra baggage is all the other formulas included. But this extra baggage is exactly the secret treasure that will make our final statement that we derive. To give a nicer feel, we could say that only two opposite literals with a single $\vee$ are to be regarded as axioms and then allow as first rule of our Logic to widen these with an arbitrary expression. In fact, such widening should be allowed from any derived expression, as a more generally useful rule.
Our new $\exists$ introduction rule is in perfect harmony with this $\vee$ widening! Totally unrestricted!
As complete opposites, the $\forall$ and $\wedge$ introductions will be very restricted.
$\forall$ can only be introduced if the variable doesn't appear anywhere else.
Before I come to $\wedge$, I tell the simplest rule, appropriately called as the rule of simplification.
We start from $\vee$-s in our logical axioms, so we merely need simplifying the built ones if we reach an $\vee$ combination with more same members. Obviously, we can omit the repeats and only keep one of them. The $\wedge$ introduction is a twist on this simplification rule. Namely:
If we derive two almost identical $\vee$ combinations except one member, A_i being B in one, so looking like $A_1 \vee A_2 \vee .. B .. \vee A_n$ and A_i being C in the other, so looking like $A_1 \vee A_2 \vee .. C .. \vee A_n$, then we can derive $A_1 \vee A_2 \vee .. (B \wedge C) .. \vee A_n$.
This corresponds in concretizations to continue separately for two statements connected by $\wedge$.
We didn't see this in our matrix method because all $\wedge$-s were at the bottom.
So this system allows much more natural expressions than our quantized matrixes.
Namely $\wedge$, $\vee$, $\forall$, $\exists$ used freely and only assuming the $\neg$ negations at the bottom.
The $\wedge$ introductions mean that we don't just have to start from an ad hoc $\vee$ combination as logical axiom, but actually a whole set of such that then will merge by these $\wedge$ introductions.
Now comes the special beauty of this particular Tait Formula Logic that I will show.
Our earlier ingenious concretization sequence can be incorporated by starting from any statement suspected to be derivable and obtain all the needed ad hoc logical axioms. Namely:
We'll regard the unimportant order of the $\vee$-s in our suspected statement as very important!
And so:
We apply variable replacements corresponding to the reversed quantor introductions always at the first appearing quantor in our $\vee$ expressions.
At $\forall$ we simply use a next new variable.
At $\exists$ we replace the "next" yet not applied variable and add the used $\exists$ sub expression with an $\vee$ to the end. Amazingly, these added $\exists$ expressions become first, infinite many times.
Encountering an $\wedge$ we must apply the whole process for both members.

As an example I will regard the $\forall x \exists y \forall z [B(x,y) \vee \neg B(x,z)]$ logical necessity.
This contains no $\wedge$ so we have a single process.
I will use 1 , 2 , 3 , . . . as new variables and before I do the replacements I will prepare this by putting the new variable after the old original but still keep the quantification.
This way the process will become perfectly clear:

$\forall x \exists y \forall z [B(x,y) \vee \neg B(x,z)]$

$\forall x1 \exists y \forall z [B(x,y) \vee \neg B(x,z)]$

$\exists y \forall z [B(1,y) \vee \neg B(1,z)]$

$\exists y1 \forall z [B(1,y) \vee \neg B(1,z)]$

$\forall z [B(1,1) \vee \neg B(1,z)] \vee \exists y \forall z [B(1,y) \vee \neg B(1,z)]$

$\forall z2 [B(1,1) \vee \neg B(1,z)] \vee \exists y \forall z [B(1,y) \vee \neg B(1,z)]$

$B(1,1) \vee \neg B(1,2) \vee \exists y \forall z [B(1,y) \vee \neg B(1,z)]$

$B(1,1) \vee \neg B(1,2) \vee \exists y2 \forall z [B(1,y) \vee \neg B(1,z)]$

$B(1,1) \vee \neg B(1,2) \vee \forall z [B(1,2) \vee \neg B(1,z)] \vee \exists y \forall z [B(1,y) \vee \neg B(1,z)]$

$B(1,1) \vee \neg B(1,2) \vee \forall z3 [B(1,2) \vee \neg B(1,z)] \vee \exists y \forall z [B(1,y) \vee \neg B(1,z)]$

$B(1,1) \vee \neg B(1,2) \vee B(1,2) \vee \neg B(1,3) \vee \exists y \forall z [B(1,y) \vee \neg B(1,z)]$

$B(1,1) \vee \neg B(1,2) \vee B(1,2) \vee \neg B(1,3) \vee \exists y3 \forall z [B(1,y) \vee \neg B(1,z)]$

$B(1,1) \vee \neg B(1,2) \vee B(1,2) \vee \neg B(1,3) \vee \forall z [B(1,3) \vee \neg B(1,z)] \vee$
$\exists y \forall z [B(1,y) \vee \neg B(1,z)]$

$B(1,1) \vee \neg B(1,2) \vee B(1,2) \vee \neg B(1,3) \vee \forall z4 [B(1,3) \vee \neg B(1,z)] \vee$
$\exists y \forall z [B(1,y) \vee \neg B(1,z)]$

$B(1,1) \vee \neg B(1,2) \vee B(1,2) \vee \neg B(1,3) \vee B(1,3) \vee \neg B(1,4) \vee$
$\exists y \forall z [B(1,y) \vee \neg B(1,z)]$

$B(1,1) \vee \neg B(1,2) \vee B(1,2) \vee \neg B(1,3) \vee B(1,3) \vee \neg B(1,4) \vee$
$\exists y4 \forall z [B(1,y) \vee \neg B(1,z)]$

$B(1,1) \vee \neg B(1,2) \vee B(1,2) \vee \neg B(1,3) \vee B(1,3) \vee \neg B(1,4) \vee$
$\forall z [B(1,4) \vee \neg B(1,z)] \vee \exists y \forall z [B(1,y) \vee \neg B(1,z)]$

$B(1,1) \vee \neg B(1,2) \vee B(1,2) \vee \neg B(1,3) \vee B(1,3) \vee \neg B(1,4) \vee$
$\forall z5 [B(1,4) \vee \neg B(1,z)] \vee \exists y \forall z [B(1,y) \vee \neg B(1,z)]$

$B(1,1) \vee \neg B(1,2) \vee B(1,2) \vee \neg B(1,3) \vee B(1,3) \vee \neg B(1,4) \vee B(1,4) \vee$
$\neg B(1,5) \vee \exists y \forall z [B(1,y) \vee \neg B(1,z)]$

This goes on infinitely!

But at the eleventh line we could have stopped because $\neg B(1,2) \vee B(1,2)$ appeared, making that line a logical axiom. From this line we can derive our statement backwards, using quantor introductions, corresponding to the variable replacements and simplifications corresponding to the used end repetition trick:

$$B(1,1) \vee \neg B(1,2) \vee B(1,2) \vee \neg B(1,3) \vee \exists y \forall z [B(1,y) \vee \neg B(1,z)]$$

$$B(1,1) \vee \neg B(1,2) \vee \forall z [B(1,2) \vee \neg B(1,z)] \vee \exists y \forall z [B(1,y) \vee \neg B(1,z)]$$

$$B(1,1) \vee \neg B(1,2) \vee \exists y \forall z [B(1,y) \vee \neg B(1,z)] \vee \exists y \forall z [B(1,y) \vee \neg B(1,z)]$$

$$B(1,1) \vee \neg B(1,2) \vee \exists y \forall z [B(1,y) \vee \neg B(1,z)]$$

$$\forall z [B(1,1) \vee \neg B(1,z)] \vee \exists y \forall z [B(1,y) \vee \neg B(1,z)]$$

$$\exists y \forall z [B(1,y) \vee \neg B(1,z)] \vee \exists y \forall z [B(1,y) \vee \neg B(1,z)]$$

$$\exists y \forall z [B(1,y) \vee \neg B(1,z)]$$

$$\forall x \exists y \forall z [B(x,y) \vee \neg B(x,z)]$$

The beauty of this system is the seemingly crazy thing what we did above, that is to continue the process in spite of founding a truth! Indeed, what if we had not found such truth?
The reversals in negative show that then we found something even better! A sequential reality where the opposite of our statement is true. So then we can be sure that it wasn't a wild goose chase, the statement is definitely not a necessity!
To use axioms to derive something, the axioms have to be broken down to open expressions and those used. The break down rules are easy and the usage, that is mixing with the logical axioms needs only one rule. Not surprisingly, this is a twist on the $\wedge$ introduction rule.
Namely, suppose that the mentioned B and C only differing members of two derived otherwise same $\vee$ expressions are such that $C = \neg B$. Then $B \wedge C = f$ false and so can simply be omitted from the combined $A_1 \vee A_2 \vee .. (B \wedge C) .. \vee A_n$.
So we can say instead of the $\wedge$ introduction rule, that deriving both
$A_1 \vee A_2 \vee .. B .. \vee A_n$ and $A_1 \vee A_2 \vee .. \neg B .. \vee A_n$ allows one to derive
$A_1 \vee A_2 \vee .. \vee A_{i-1} \vee A_{i+1} \vee .. \vee A_n$. So we can "cut out" the opposite members.
Observe that this "cut" rule could not be simplified by some f rule for falsity.
Indeed, $B \wedge C = f$ is not a derived statement, in fact if it were, we reached a contradiction.
The formal opposition of the parts is only that allows the cut.
What's more, in our system this formal opposition is not even formal because we have no $\neg$-s at expression levels. So our cut rule must involve a verification first that the two expressions are opposite. This of course is not that hard.
This cut rule should be called Pair Cut Rule, referring to the cut out contradicting pair.
A simpler meaning of it comes out by regarding all the remaining A-s as a C.
Then from deriving $B \vee C$ and $\neg B \vee C$ we derive C and using implications instead, we derived $\neg B \rightarrow C$ and $B \rightarrow C$ so indeed, the C consequence must be true regardless of B being true or false. The Pair Cut name is important because other "cuts" are possible too. The weaker Single Cut derives a C expression from the derived B and $\neg B \vee C$. Its meaning comes out by realizing that this second is $B \rightarrow C$. So this is the good old chain rule or Modus Ponens. A cut that cuts out more, yet is equivalent with the Pair Cut is the Pair and Or Cut:
Here we derive again C but now from three derived expressions: $\neg A \vee C$, $\neg B \vee C$, $A \vee B$.
Again, the better meaning comes out by realizing the first two as $A \rightarrow C$, $B \rightarrow C$.
And this suggests at once a big jump to cut out many expressions if $A_1 \rightarrow C$, . . . , $A_n \rightarrow C$ are all derived together with $A_1 \vee .. \vee A_n$. Indeed, these should allow to derive C.

This Multi Cut is the most general form for avoiding finite many cases as conditions.
But Pair Cut is already the most frequent logical form to apply some “ad hoc” B as trick.
A nice geometrical example is the proof that the chromatic number of the plane is at least 4.
This is defined as the minimum number of colors that assigned to the points, we can avoid some distance between same colored points. It’s fairly easy to create a coloring with seven colors that avoids same colored ones apart from a certain distance. It uses a honey combing of the plane.
The exact number is probably 4 but this has not been proven yet.
What we’ll show now is that 3 is definitely not enough.
So with using only 3 colors we must have same colored pairs apart with any d distance.
This is our C conclusion. But what should be B? The ad hoc-ness is striking here indeed.
Indeed, B should be that there are no two different colored points $d\sqrt{3}$ away!
If this is true then picking any point and drawing a $d\sqrt{3}$ radius circle around, all points of this circle are same colored as the picked center. This obviously implies same colored pints d away, namely as any d string of this circle.
Now comes the other even more interesting half of the proof that $\neg B \to C$ too.
That’s where the ad hoc-ness of $\sqrt{3}$ will come in.
$\neg B$ means that there are two points $d\sqrt{3}$ away, with different colors.
An equal d sided triangle has $\frac{d\sqrt{3}}{2}$ height, and that’s the reason behind our $\neg B$.
So we pick a pair that $\neg B$ claims and erect perpendicular $\frac{d}{2}$ distances from the middle of our $d\sqrt{3}$ connector. We get two points that are d distanced but are also d away from our picked two points. So we have two equal sided triangles formed by the four points.
Thus of course either one of these two new points have same color as one of the picked ones or they must be both different from those two colors but then be the third color and thus be same.
Here the truths of B or $\neg B$ were depending on the coloring of the plane.
But Pair Cut is useful for even when B , $\neg B$ are statements in a theory but hard to decide.
An amazing algebraic example is to prove C that there are irrational numbers so that their exponentiation gives a rational. B should say that $\sqrt{2}^{\sqrt{2}}$ is rational. Then $B \to C$ is proved if we derive that $\sqrt{2}$ is irrational. But then $\neg B \to C$ is derivable too because if $\neg B$ is assumed, that is $\sqrt{2}^{\sqrt{2}}$ being irrational then $(\sqrt{2}^{\sqrt{2}})^{\sqrt{2}} = \sqrt{2}^{\sqrt{2}\sqrt{2}} = \sqrt{2}^{2} = 2$ gives C with using $\sqrt{2}^{\sqrt{2}}$ as base and $\sqrt{2}$ as exponent.
So we obtained C without the truth of $\neg B$ that $\sqrt{2}^{\sqrt{2}}$ is irrational, which is hard to prove.
The used easy truth that $\sqrt{2}$ is irrational follows from the Unique Prime Factorization Theorem:

$2n^2 \neq m^2 \to \sqrt{2}\, n \neq m \to \sqrt{2} \neq \frac{m}{n}$. Surprisingly, the same way we can get that:

$2^m \neq 3^n \to 2^{\frac{m}{n}} \neq 3 \to [\, 2^E = 3 \to E$ is irrational $\to 2E$ is irrational]. Then:

$\sqrt{2}^{2E} = 2^E = 3$ so the $\sqrt{2}$ and $2E$ irrationals prove our original claim because 3 is rational.

Back to our subject, already the Single Cut or Modus Ponens allows to use multiple conditions.
Indeed, from $B_1 \wedge B_2 \wedge \ldots \wedge B_n$ and $(B_1 \wedge B_2 \wedge \ldots \wedge B_n) \to C$ derived we get C.
This second of course is the same as $\neg(B_1 \wedge B_2 \wedge \ldots \wedge B_n) \vee C = \neg B_1 \vee \ldots \vee \neg B_n \vee C$.
So repeated single cuts can be used. Especially interesting is if these $B_1, B_2, \ldots B_n$ conditions are merely some members from a set of $A_1, A_2, \ldots$ axioms.

Then if $\neg B_1 \vee \, .. \, \vee \neg B_n \vee C$ is a logical necessity we can use our axioms with Single Cut.
And indeed, single cut is enough because by Compactness, finite many axioms must imply any C that is true in all realities of $A_1, A_2, \ldots$
Still the Pair Cut is a better choice than the Single Cut when we'll see the true motivation of Gentzen to create his new system to replace the already accepted one created by Hilbert.
Gentzen's system was much more complicated than what we introduced because he didn't restrict the expressions to have all $\neg$-s at the bottom.
Ours was discovered by Tait as a simplification of Gentzen's.
The heuristic importance of Tait's system is that in spite of requiring the negations at the bottom, allowing both quantors, $\wedge$ and $\vee$ in any order, the natural statements of mathematics or even every day sentences are very naturally translatable. So having the seen algorithm that gives a definite search for logical necessities is a crowning of this system.
But observe that the crucial point here is the "definite". Namely, that a failure of the search directly gives a model for the non necessity, that is for a falsity of the suspect statement.
Indeed, merely a search for necessity can be obtained from any system of Logic!

<u>Generation</u>

The even wider fact is that any derivation system is also a generation system!
The simple reason is that there are only finite many rules. A hidden problem is that we need axioms too and usually these are infinite many. But these are generable too as a sequence.
So going up to a beginning in this list and using the finite many rules for only a restricted many times, we create only a finite number of derivations. Now if we increase the beginnings, that is the used axioms and the number of rule applications simultaneously in any manner, we will get all possible derivations. This is an ugly but effective sequence of all derivable theorems that mathematicians haven't even discovered yet. This feels paradoxical because it suggests that a machine could replace mathematicians. But this is not the case for two very different reasons.
The practical reason is that such generation would spit out all derivable statements.
Most of these would not be regarded as new theorems, merely as trivial consequences.
Some intelligent person would have to choose the meaningful real theorems.
The more important theoretical reason comes out if we only want to find the proofs for the already existing claims that have not been proved yet. These will be found by the machine but without any guarantee when. Simply because this generation by derivations can give short theorems much later than quite long derivable monstrosity statements.
And this is the essence with an other angle mingled! Namely, whether we could also generate those statements that can not be derived. First of all, observe that with two such generations we could decide if something is theorem or not in finite time even if that time is long due to our machines being slow. We simply have to watch both lists and spot our statement that we are interested in. But if it would turn out that such second list doesn't even exist, then no matter how fast machines we build, the wait at only one list for a statement could be arbitrary long.
Also observe that if the generation of statements were increasing not just by the lengths of derivations but by the lengths of the derived statements too, then at once we had a perfect second derivation of the non-theorems.
Indeed, as we surpass a length in our list, we can find all the missing ones with same length.
So the above mentioned very long wait for even short statements is a theoretical point too.
Even before the generations as potentially wild goose chase were so clearly seen as we just described it, those who believed in some effective methods to replace derivations were envisioning such dual systems that would decide both the theorems non-theorems.
Hidden behind this lied the assumption that maybe with good axioms all statements were in one side, so the axioms were complete. But obviously not all axiom systems can be complete, some are just framework type theories and as a best example is Logic the widest framework system.
So Hilbert very smartly, had only a personal belief in perfect axiom systems but apart from that raised the question if for Logic there could be a decision method of necessities.

This was an official problem on the famous list published at the Paris math congress.
The nightmare that some sets are impossible to list was not conceived yet but the fact that short theorems can have very long derivations was clear by the role of logical axioms.
Indeed, the logical rules alone would keep some length increase. In our system for example, all the rules increase length except the cuts. But the axioms which are just a second list, when mix with the logical rules, make the possible shortening of the derived theorems unpredictable.

Gentzen

Returning from this little detour to the Tait system is very educational.
We didn't merely derive the necessities but had the reverse method to search for the necessary logical axioms too. By what we explained, the lengths of the found axioms can not be predictable by the length of the initial statement that turns out to be a necessity.
Indeed, otherwise we could tell in finite steps if a statement is not necessity by using only axioms under certain lengths and use all the increasing rules and no cuts.
So the avoidability of cuts was very useful here. This relates to what above we hinted as the "true motivation of Gentzen" to create his system. We might think that this motivation was the same we called as crowning result of Tait's system, the direct verification of necessity.
Except that the crown is bigger because Gentzen didn't have to use negations at the bottom.
But this wasn't Gentzen's motivation at all, "merely" the elimination of cut rules.
We had to use quotation mark because more is behind and it is in the used word "elimination".
The crowning result proves that cut rules are not needed, but this is not an elimination as such!
Let's start from an other angle: The fact that for logical necessities we don't need cut rules, suggests a dilemma. What if we derive some logical necessities and then use those to derive a new one by using cut rules. This obviously will be shorter than deriving the new necessity from the bottom without cuts. So is there an algorithm that gives the long and tedious road from the simpler one? This is a true elimination of the cuts and this was Gentzen's goal.
To find such cut elimination algorithm, the best is to use the Pair Cut rule which is more powerful than Single Cut but much simpler than Pair and Or Cut or Multi Cut would be.
That's why the Pair Cut became simply called as the Cut rule.
But this idea, to be able to find a cut elimination process hid an even bigger goal, beyond necessities, that is when we have non logical axioms, namely the axioms of Number Theory.
Derivation methods that are complete do not imply that the axioms have a reality.
Indeed, that would be absurd because we can start with intentionally contradictory axioms.
Remember, we also said that the completeness of our logic shows that statements true in all realities are derivable. But we didn't mention if there are such realities or not.
Our sequential reality construction too was relative depending on the beginnings avoiding contradictions. Without the cut rule it is very evident from the rules themselves that a contradiction can not be derived and indeed we only derive necessities true in all realities.
So if a similar cut elimination could be found for Number Theory then this would show that Number Theory can not lead to contradiction.
You may say that this is an even crazier goal than to eliminate cut itself! Indeed, the natural numbers as reality surely satisfy the axioms quite visibly. Also, a contradiction would mean that all statements are derivable. This feels absurd if we think about those hard to prove theorems of Number Theory? Unfortunately both arguments have holes.
The set of naturals obeying the axioms is only visible from the outside.
A strict Number Theory should only talk about numbers, that is finite sets.
For the complicated proofs we can reply that maybe contradictions are even harder to prove.
The fact that we have no contradiction in Number Theory is a quite complicated claim.
Namely, that there can not be derivation of both A and $\neg A$ for any A statement.
Of course, if this were the case then we could derive any S statement by widening A to $A \vee S$ also $\neg A$ to $\neg A \vee S$ and then cut to S. So formally we must show that not every S is derivable which seems so easy. But to see this, would mean to analyze all derivations.

We might even use number theoretical concepts along the line and in fact use Number Theory to prove its own non contradictoriness or “consistency”.
Gödel proved that this is impossible! First of all, a system has to be complex enough to be able to talk about derivations translated into its own language. Number Theory is complex enough and to show this was Gödel’s first and longer achievement. But then a quite simple argument showed that if this non contradictoriness is proved then it actually means a contradiction!
So if you believe in that Number Theory is consistent, then you also must accept that this can not be proven. Consistency is a question of “faith”.
But Gentzen wanted to challenge this bleak picture and see how far one can see the consistency of the axioms of the naturals. And he succeeded! Without “breaking the law”, that is contradicting Gödel’s result. What’s more, his system became the start of Proof Theory.
He was a strange person too! A Nazi who corresponded with Jews and died in a Russian labor camp. Gentzen did something else, what a less famous person Jaskowski did at the same.
They introduced new, “more natural” deduction systems already for quantorless expressions.
The way I explained, the quantorless expressions are just stupid combination of the possible sixteen operations. So every such expression can be evaluated by the possible t , f combinations and seen to be always t or not. End of story!
I maintain this position but I also realize that such evaluation is actually a tremendous task.
To simplify this process is a nice goal and as computability became a subject, this side gave a real importance of these “natural” systems. But some are championing them as philosophically important subject. I just regard it as new version of the bad old Formal Logic.
Thinking, whether mathematical or other, is much too complicated to be formalized!
In my opinion it is not even material! I was passionately telling that Grammatics, the introduction of the true hidden nature of statements should be elementary school subject!
But if you propose a deduction system that reflects thinking, I just say “you are crazy”.
So how can I regard the forms of all everyday claims to be simple and yet regard the thinking about them as transcendental? Because we don’t think through these claims.
To make exact claims is important! To make exact proofs is also important but it doesn’t drive thinking either! We think in visions and they should be transferred too as well as possible!
So we must come to the historical background of the most important aspect of derivations that we only touched upon in the previous section.

Turing

Gentzen was not the only one who tried to defy Gödel! A young English mathematician Turing, who later became the mind behind breaking the German code system in the War, realized that Gödel’s vision was narrow. Gödel’s above mentioned result, the non provability of a system’s own non contradictoriness was only a culmination of his first similar result. This showed that the ability of Number Theory to talk about its own derivations causes the system to have undecidable S statements for which neither S nor $\neg$S can be derived. The truth is this:
The ability to talk about itself indeed causes the system to be complex as derivation method.
But then this complexity plus the unemphasized fact that the infinite many induction axioms are given by a simple scheme and thus effectively, is the real cause of undecidability.
In fact, it is not the undecidability that is basic at all! Rather that the derivable statements, the theorems as a set and the underivable statements, the non-theorems are different kind of sets!
That’s why the effective listability or generability of the axioms is crucial.
Continuing with derivations from them in some fix order that goes through all possible derivations, these two, the simple generation of the axioms and the complex generation of all possible theorems, form a single complex generation of all theorems. And amazingly, then the complement set, the non-theorems is an even more complex set that can not be generated by any method at all! This is what forces the existence of undecidable statements!
Indeed, if there were no such undecidable statements, then we would have an immediate method of generating all non-theorems by simply generating the theorems and then just formally negate them. There is a simple reason why this grand vision is not emphasized.

This reality of the non-theorems of Number Theory being a "non generable set by any means" is not a mathematically provable claim. Simply because "by any means" is not a definite claim. Effectivity is a mystery, still outside mathematics! Gödel's proof on the contrary was totally exact. It didn't mention Effectivity, generations and such semi physical concepts. It looked at the derivabilities of Number Theory, coded by numbers. Number Theory can then talk about its own derivations and this implies undecidable statements.
Daring to talk about the reality that some sets are simply non generable, means that Gödel not only kept in dark the fact that the set of non-theorems is a non generable set but that those related number sets that he used as codes are such too.
Turing went to the bigger, semi physical vision and yet he also stayed within mathematics!
By replacing Number Theory with something that at least suggested the wider concept of Effectivity, namely with Computers. In fact, he used the "computable numbers" expression already in his title. We might think that he was merely sloppy to say computable sets of numbers, meaning the generable code numbers that have non computable complements.
But this was not the case! He meant real numbers in the title, which are infinite binaries and so correspond to set separations of the naturals. By computable then he meant both half being generable. So this first half of his title was unfortunate. Indeed, exactly the generations as naturally not implying generable complements was his crucial discovery.
But then the second half of his title was making the first part an even worse choice.
The assumption that all effective methods are doable by his machines was really behind this second part. Namely, referring to the Hilbert problem in German. But calling it in German or in English as decision method for necessities, makes no difference in the essence, that this problem was not exact because the term "method" is not defined.
Still, Turing's assumption could only be used because he stepped from the particularness of deriving theorems to generate objects in general. Gödel did not make the same claim of refuting Hilbert's problem though he knew it to be false! Why? Because unlike Turing, he still hoped for Effectivity to be graspable inside mathematics! And here we mean old mathematics about numbers. Now of course, Turing's outside world of machines became part of mathematics.
The assumption by Turing that all Effectivity is machine doable can be called as external universality. But his proof that there are generable sets with non generable complements rested upon the inner universality of some machines. This was the real essence and so his article title should have been: "Universal computing machines".

Rice very briefly but very strongly

The most important result of mathematics after the war was a conceptual widening of Turing's results and is called Rice's Theorem. I regard it as an equal of the Well Ordering Theorem.
Similarly to that, it is not hard to prove, yet it changed our whole vision. Mine anyway!
It finally lifted the veil from our eyes about the real meaning of programs as data.
Programs as data was the key for Turing to show how "universal" machines exist that can simulate any others. This then implies machines that generate sets that have complements not generable by any machine. Unfortunately, always some diagonal tricks have to be used.
By Rice's Theorem, if we separate the programs according to how they run, then if one side is effective, the other can not be. It's that simple! Yet nobody realized it before.
The Formalist parrots regard this theorem only as a trivial consequence of results known earlier.
Or regard abstract versions of it. They don't want to deal with visions and understanding.
So then why did Rice discover this only in 1953? To answer this, is a taboo today.

Turing Falsifications and games

In the sixties in our math special high school a very smart decision introduced computer classes.
So the girls who couldn't get into uni went to be programmers. But that's about it! The name of Turing or what he discovered was never mentioned. And that was typical not just in Hungary.

Then Turing became the “in thing”. A very influential text book writer in America had a “vision” of rewriting history. He claimed that Turing was at once accepted by Gödel.
A complete lie but his goal to use Gödel’s acceptance was even worse. Namely, to suggest that “computability” is the perfect word for the new essence that Turing discovered.
Beside appealing to Turing’s obvious attachment to the word computer, it was very much tapping into the main stream because “home computers” became everyday household items.
Strangely, this is paradoxical already and I wonder if a mother ever asked her teenage daughter: “Why are you always at that damn computer, when you have nothing to compute?”
By the way, before the home computer age the “Turing Machine” name already became an “in thing” and was falsely assumed as the discovery of computers. Some have absolutely no idea about Effectivity, but like saying “Turing Machine”, that indeed sounds “cool”.
Had poor Turing’s father been called Smith, would these parrots love to say “Smith Machine”?
The newest and stupidest Turing falsification is throwing his name into human consciousness.
And again he gave reason for this lie too because he wrote about this.
Of course his machine has nothing to do with his arguments about that.
Computer, machine, derivation, generation, recursive were the words I used so far.
Now the mentioned text book writer wanted and succeeded to rename this recursive adjective to computable. The really illogical “enumerable” adjective in “recursively enumerable” didn’t bother him so just translated it to “computably enumerable”.
But really, the deeper problem is twofold! Firstly, regarding computers as the only form of effectivity and so eliminate the word Effective altogether. This is by the way impossible by even those who stood in line because their articles are filled with this word.
So why don’t we just accept it as a perfect word for the bigger thing itself.
Effective systems are much wider than computers. And games for example are the most ancient effective systems! Some allowed steps are given and we apply them to win.
So Turing’s discovery has to be true here too. And it is! In fact, it is possible to see the naturalness of non generable complements, without knowing about universal machines or understanding the breakthrough of Rice’s Theorem, by simply regarding a game like chess.
Most chess players only want to win and so they never think about these following things:
We put some pieces on a board randomly and ask a good player “what is here?”. Observe that for some situations the obvious reaction could be “this is impossible”. For example, the pawns go only forward and start from the second line so they couldn’t be on the first line.
But a more usual reaction is that “this is crazy”. But then we can ask: “Could this be the result of a game, regardless how crazy the players were?”. To establish this, we must find a game history, the actual steps from start. So to find this for a given setup can be quite hard.
But now I’ll ask something more important that is not hard to answer at all.
Don’t worry about individual situations! Instead, tell me which situations as a class is more complicated, the achievable ones or the not? And then the coin can drop! The achievable ones must be much simpler! Indeed, these each have a game history, a finite sequence of situations.
But to establish that a situation is not obtainable by a game, we must check all possible arbitrary long games and only then, after this infinite check can we tell that it didn’t come about.
That’s at the bottom too for the fact that the non-theorems as a set is much more complex than the set of theorems.
But now the second element of stupidity in these Turing falsifications is the saddest!
Namely, missing to see what Turing really achieved as first.
Text generating systems were known before him and were developed later too as Grammars.
But these do not reveal the crucial universalities. Only a system bringing the duality of halting or running forever to the bottom can do that. Turing tables as text alterators do exactly this.
This then becomes a fountain of truths. The text collections as “recognition” by halt and thus ignoring the results altogether is one such amazing truth. And then the use of results in spite of this, for not separability and text complexity is even more surprising.
But the simplest truth is what I already mentioned in the first section:
The super geniuses Newton, Gauss, Gödel could not see what the deep geniuses Cantor, Einstein, Turing received. Sets, Time and Effectivity became fountains of new truths.
The new wider visions are still missing for all these three revolutions.

Gödel, Robinson, Presburger, Turing

About my view how Turing went beyond Gödel by recognizing Effectivity, some might argue and say that Gödel had the perfect vision just simply tried to be exact.
Well, just a bit later history proved without any doubt that my claim is true!
Namely, with a better vision, he could have done much easier, much more.
Here is Peano Arithmetic that Gödel used:
Zero will be regarded as natural number, in fact 0 will be our only name.
But we can name other numbers too because the next number function denoted as $'$ and addition and multiplication will be basic functions too.
So the naturals can be expressed as $0'$, $0''$, $0'''$, . .
Six axioms and a seventh system of infinite many axioms are as follows:

1. $x' \neq 0$
2. $x' = y' \;\rightarrow\; x = y$
3. $x + 0 = x$
4. $x + y' = (x + y)'$
5. $x \cdot 0 = 0$
6. $x \cdot y' = x \cdot y + x$

7(∞). $[\, P(0) \text{ and } \forall x\, (\, P(x) \rightarrow P(x')\,)\,] \;\rightarrow\; \forall z\, P(z)$
$P(0)$ is called the 0 condition and $\forall x\, (\, P(x) \rightarrow P(x')\,)$ the step condition.

The appearing unquantified open variables of course all mean universality.
The meanings are pretty obvious.
A logical question is why a reversal of axiom 2. is not stated.
Simply because it is part of Logic.

Omitting the 7(∞). induction axioms we get a weak "framework" theory like say algebra.

But even as such, this system is almost useless because for example the following trivial four facts are not provable in it:

8. $x' \neq x$
7. $x \neq 0 \;\rightarrow\; \exists y\, (y' = x)$
9. $x + y = y + x$
10. $x \cdot y = y \cdot x$

The strange numbering will be clear soon.
As a little detour, we show how 7(∞) can prove these.
An interesting detour because these proofs are far from trivial applications of 7(∞).

8.'s 0 condition follows from 1.
Its step condition $x' \neq x \;\rightarrow\; x'' \neq x'$ follows from $x'' = x' \;\rightarrow\; x' = x$ a case of 2.

7.'s 0 condition is true because $0 \neq 0$ is false.
Its step condition $(\, x \neq 0 \rightarrow \exists y\, (y' = x)\,) \;\rightarrow\; (\, x' \neq 0 \rightarrow \exists y\, (y' = x')\,)$ has in its second implication a truth as condition by 1. and also a truth as consequence by using $y = x$.
So this whole consequence is true regardless of the condition. So the step condition is true too.

9. is going to be a nightmare. (So 10. we'll not even show.)
First we prove that 9. is true with $y = 0$, that is $x + 0 = 0 + x$.
By 3. we only have to prove $x = 0 + x$ which we do with induction.
The 0 condition says $0 = 0 + 0$ which is 3. again with $x = 0$.
The step condition says $x = 0 + x \;\rightarrow\; x' = 0 + x'$. Which follows by: $x' = (0 + x)' = 0 + x'$.
The first equality is true by the assumption, the second by 4.

Now we prove that a 4*. $x' + y = (x + y)'$ first member incrementation is valid for addition.
We use induction on y. The 0 condition is trivial by 3. again and the step condition is:
$x' + y = (x + y)' \rightarrow x' + y' = (x + y')'$. Which follows by:
$x' + y' = (x' + y)' = (x + y)'' = (x + y')'$.
The first equality is true by 4. the second by the assumption and the third by 4. again.
Now we finally prove $x + y = y + x$ with induction on y again.
The 0 condition is what we proved first, so we only need to show the step condition:
$x + y = y + x \rightarrow x + y' = y' + x$. Which follows by: $x + y' = (x + y)' = (y + x)' = y' + x$.
The first equality is 4. the second is true by the assumption and the third by 4*.

Claim: 8. , 7. , 9. , 10. don't imply each other and can not be obtained without $7(\infty)$.
Yet now we omit $7(\infty)$. and instead add to the first six axiom 7.
This new 1. – 7. axiom system is Robinson Arithmetic. The seven statements can be combined into a single Q statement, thus denoting Robinson Arithmetic.

Now we show why $x' \neq x$, $x + y = y + x$, $x \cdot y = y \cdot x$ are not derivable from Q.

Enough to show models where Q is true but $x' \neq x$, $x + y = y + x$, $x \cdot y = y \cdot x$ are false.

But even before that, we can have little glimpse though not a real justification why Robinson chose 7. as new axiom to replace induction. Indeed, the full justification will only be the actual magic of his system to which we'll come soon. But even now we can already see that without 7. things are boring. Namely, if we regard the S set of all those s numbers that are not 0 and yet they have no previous then they are just simple new staring values taking the role of 0.
Just regarding one s new start, our numbers are: $0 , 0' = 1 , 0'' = 2 , \ . \ . \ s , s' , s'' , \ . \ .$
Of course, we must tell also how the two operations behave but this is quite predictable too.
Accepting 1. – 7. and seeing that $x' \neq x$, $x + y = y + x$, $x \cdot y = y \cdot x$ are still not derivable are more interesting. We might even think that now a previous simple s restart is impossible.
But the first case, that is finding a model where for some x we have $x' = x$ allows an even simpler situation as: $0 , 0' = 1 , 0'' = 2 , \ . \ . \ . \ s = s'$. So we truly need a single new number.
Of course, beside $s = s'$ we must tell again how the operations behave if they involve s.
Using 4. $x + s = x + s' = (x + s)'$ which implies $x + s = s$ if only s is successor of itself.
Similarly, but using 6. $x \cdot s = x \cdot s' = (x \cdot s)'$ so $x \cdot s = s$ too.
We might jump to say that then simply s is a self preserving infinity. But by 5. $s \cdot 0 = 0$.
So the operation values involving s are always s except $s \cdot 0 = 0$.
Easy to see that then indeed all seven axioms will be obeyed.
Interestingly, since we derived that $x \cdot s = s$ must be for all x and thus $0 \cdot s = s$ too, but as we just saw $s \cdot 0 = 0$, thus in this model $x \cdot y = y \cdot x$ will not be true either.

To make a model where $x + y = y + x$ fails, we need two new numbers u , v that are successors of themselves. Then naturally we must have $u + n = u$, $v + n = v$ too but surprisingly, we must also have $x + u = v$, $x + v = u$.
And then of course, $u + n \neq n + u$, $v + n \neq n + v$, so indeed $x + y = y + x$ will be violated.
Actually in all cases where either u or v occurs once.
To define multiplication we obviously set $u \cdot 0 = 0$ and $v \cdot 0 = 0$.
For other u , v involved cases if u is first member the result is v and same way in reverse.
For the remaining n named first member cases the result is same as the second member.
An amazing postscript to all this is that if we accept $x' \neq x$ as axiom 8. and thus exclude our simple models just described then the possibility of weird models with new elements still remains. It is evident that now a new element s can not work alone because it needs previous ones by 7. but it can not be itself by 8. It's not that hard to see either that finite many new members are not enough now and actually two sets of forward and backward successors from ω must exist. So a full model must look like:

$$0 , 1 , 2 , 3 , \ . \ . \ .$$
$$. \ . \ . \ . , -3 , -2 , -1 , s , +1 , +2 , +3 , \ . \ . \ .$$

The first line is the normal natural numbers abbreviated from their 0’’’’ forms and the second line are the extra non standard numbers.
And indeed, these objects together satisfy 1. , 2. , 7. , 8. because: 0 has no previous, the previous is unique, a non 0 has previous and nothing is previous of itself.
The really shocking fact is that including axioms 3.– 6. and even $7(\infty)$, such strange models will still exist. Not with a single such double infinite sequence but with infinite many added.
The big question we ask naturally is how the hell the additions and multiplications will look like. Not surprisingly, it can be shown that they can not be given effectively.
Only their existence follows by Zemelo’s Axiom Of Choice.

A benefit of the weakness of Robinson Arithmetic is that it is visibly non contradictory.
Indeed, by our Logic in a contradictory system everything is derivable and here we saw non derivable statements by being false in some models while being true among the real naturals.
But this visibility came from Set Theory. For both sides! Even the truths among the real naturals. So we can not see from the axioms that they couldn’t derive two opposing statements!

Amazingly, in spite of the simplicity of this system, the non-theorems are non generable.
To understand how this can be, we must remember that there are logical axioms too.
As I explained in my little detour at the Tait system, these mixing through the logical rules can make the derivable theorems very complex.
But now comes the real magic of Robinson’s system:

Adding new consistent axioms, can not make all statements to become decidable.

This feels unbelievable because the non-theorems will obviously shrink by adding new axioms.
The simple point is that some non-theorems will not shrink at all.
Namely, some are non-theorems because they are the negations of theorems and so deriving these we would get instant contradictions and thus able to derive everything. These negatives of theorems should be called the “intentional non-theorems” or “anti-theorems”.
We don’t want these to be theorems. More concretely, if we only regard consistent extensions then these earlier anti-theorems remain.
Now we can pinpoint the crucial special feature of Robinson’s system:
For those effective E sets for which the $\neg E$ complement set is also effective, that is are dually effective, we can represent these pairs dually, that is with also opposite P , $\neg P$ property pairs.
These cases of representations are actually theorem and anti-theorem pairs and so these remain in any consistent extensions.
And so for seeing the non generability of the theorems in such extensions we would only have to see that the representations of the complementing generable pairs implies this.
But this is not true! We’ll only get a bit weaker claim that at least one of the theorems or non-theorems is not generable. This is very logical too! Indeed, a simpler argument shows that the representation of all generable sets implies non generable non-theorems.
So since our condition is weaker we should get a weaker consequence too.
But it will be sufficient anyway if an extension has generable theorems.
Since I mentioned this simpler argument, I will show that too.
In fact, I will go through all seven arguments that lead to incompleteness.
The first is the simplest and doesn’t use effectivity at all only a very strong diagonality:

1. Undecidable Argument:
Suppose that for every $P(x)$ property we can assign a $< P(x) >$ code number so that that a $G(x,y)$ “Gödel relation” can represent the property case derivations:

$$\vdash P(n) \qquad \text{iff} \qquad \vdash G(n, < P(x) >)$$

Then if $\vdash$ is consistent then there is undecidable statement pair by $\vdash$.

The proof is surreal!

We claim that $G(<\neg G(x,x)>, <\neg G(x,x)>)$ and its negative are an undecidable pair.

Indeed, using $\neg G(x,x)$ as $P(x)$ and n as $<P(x)> = <\neg G(x,x)>$:

$\vdash \neg G(<\neg G(x,x)>, <\neg G(x,x)>)$ iff $\vdash G(<\neg G(x,x)>, <\neg G(x,x)>)$.

If both sides are true then our system is inconsistent which we assumed not to be.
So both sides must be false, meaning exactly that the two statements are an undecidable pair.

If you are scratching your head I am not surprised. Actually, this was didactically bad on two levels. It avoided effectivity which is the real essence but it was also a totally ad hoc trick.
And this second more formal error can be tamed a bit as follows:
We can say that two $P(x)$ and $P'(x)$ properties share an n number derivably if:

$$\vdash P(n) \quad \text{iff} \quad \vdash P'(n)$$

Obviously, if our axiom system is consistent then a P and its negative $\neg P$ can only share an n derivably if neither is derivable at n and so these cases are an undecidable pair of statements. Thus to show that there is undecidable pair, it is enough to exhibit a P_0 property that shares derivably some n number with every P property. Indeed, then $\neg P_0$ must be among these P-s and so the derivably shared two cases are an undecidable pair of statements.
Then we can reveal that:
Such P_0 is $G(x,x)$. and The derivably shared value with a P is $<P>$.
Indeed: $\vdash P(<P>)$ iff $\vdash G(<P>, <P>)$ = $\vdash P_0(<P>)$.

As a complete opposite, the next two arguments use the concept of effectivity as an outside field with the result that there is E_0 generable set that the $\neg E_0$ complement is not generable.

2. True Representation Argument
Here in addition, we use that all E generable number sets can be represented in the language of our Number Theory with suitable properties being true: $n \in E$ iff $P(n)$ is true.
Which also means that: $n \in \neg E$ iff $P(n)$ is false.
Then if E_0 is a generable set with non generable complement and P_0 represents it in the above sense, then the false cases of this P_0 is a non generable statement set. So then the set of all false statements is non generable either because otherwise we could recognize the P_0 cases among these and they would have to be generable too. And then the truths are not generable either since they are perfectly negated pairs of the falsities. Then we can realize that in an axiom system where the axioms are generable, the theorems are generable too because the logical steps can be used systematically. So since the theorems are generable but the truths are not, these two sets can not be identical. If we assume that our axioms do not imply any false statements, that is the theorems are a proper subset of the truths then all non derivable truths are undecidable statements.
Indeed, they are not derivable and their negatives either because they are false.
If we do not assume that our system only derives truths, the situation becomes quite complex.
First of all, this existence of theorems that are false does not mean a contradiction yet.
Indeed, a contradiction is only if some S , $\neg S$ pair are both theorems.
Logic is such that this automatically means that all statements are derivable theorems.
So the assumption of the non existence of such contradictory pair, also called as consistency, is really the minimal assumption about any axiom system.
The previous result of undecidable pair meant that neither S nor $\neg S$ were theorems.
The opposite, that is decidability thus merely means that S or $\neg S$ is a theorem and thus allows both of them to be, which of course is a contradiction. But with assuming consistency the or is actually an either or. This is also called as the completeness of our axiom system.
Then if the axioms and thus the theorems too are generable then the non theorems are too.
Indeed, we can just generate the theorems and then simply negate them.

So we obtained by consistency that a decidability of all statements would imply that while the truths and falsities are both non generable, the theorems and non-theorems are both generable.

3. Derivable Representation Argument

We can avoid the previous complications by making representations of the effective sets not by truth but derivability. We defined such representation of a T set already as:
$P(x)$ represents a T set of naturals in an $\mathbf{A}$ axiom system if: $n \in T$ iff $\mathbf{A} \vdash P(n)$.
Assuming that this exists for every E generable set, our new argument is this:
Let E_0 be again a generable number set that has a non generable $\neg E_0$ complement and let P_0 be a property that represents it. Then $\{ n ; \mathbf{A} \neg \vdash P_0(n) \}$, the set of those n numbers for which $P_0(n)$ are non-theorems is not generable. Thus the set of the $P_0(n)$ non-theorems is not generable either. And thus all the non-theorems are neither because we could recognize these.

We showed in the previous argument how derivable axioms and completeness implies derivable non-theorems. So now we get that: If in a consistent generable axiom system the generable sets have derivable representations then there has to be undecidable statement pair.

From now on we avoid the external assumption of an E_0 with non effective complement.
Simply because we can establish that inside by using diagonality.
So we return to a continuation of the Undecidability Argument.
Diagonality and so ad hoc-ness will remain in these and I will not use similar tamings as in 1.

4. Not Representable Argument:
If $\vdash$ is consistent then $\neg D = \neg \{ <P> ; \vdash P(<P>) \}$ is not representable.

The name of D should be "diagonal derivability code set". Now the proof:
Enough to show a contradiction from $\neg D$ being represented by a P_0.
Such representation would mean: $n \in \neg D$ iff $\vdash P_0(n)$.
Using n as $<P_0>$: $<P_0> \in \neg D$ iff $\vdash P_0(<P_0>)$
But observe also that:
$\neg D = \{ n ;\ n \neq <P>$ for any P or $(n = <P>$ for a P and $\neg \vdash P(<P>)) \}$. And so:
$n \in \neg D$ iff $n \neq <P>$ for any P or $(n = <P>$ for a P and $\neg \vdash P(<P>))$.
Using again n as $<P_0>$ the first or possibility is false and the second's first and part is true with $P = P_0$, so $<P_0> \in \neg D$ iff $\neg \vdash P_0(<P_0>)$.
$<P_0> \in \neg D$ is either true or false but both imply $\vdash P_0(<P_0>)$ and $\neg \vdash P_0(<P_0>)$ both to be true and thus having a contradiction in $\vdash$.

5. Not Generable Argument:
If every E generable number set can be represented by a $P(x)$ property: $n \in E \leftrightarrow \vdash P(n)$ and $\vdash$ is consistent then the non-theorems is not generable.

By our two conditions and the Not Representable Argument, $\neg D$ is not effective.
Thus the corresponding $\neg \vdash P(<P>)$ diagonal non-theorems and thus all the non-theorems are not generable either because the diagonal ones were recognizable among them.

The heuristic continuation is the same as in 3. If in a consistent generable axiom system the generable sets have derivable representations then there has to be undecidable statement pair.

From now on we use “selectable” as a short for “dually effective”.
That is being effective and having a complement that is also effective.

6. Not Selectable Argument:
If every E , $\neg E$ selectable number sets can be represented by P_1 , P_2 properties:

$$n \in E \quad \leftrightarrow \quad \vdash P_1(n) \qquad \text{and} \qquad n \in \neg E \quad \leftrightarrow \quad \vdash P_2(n)$$

then the theorems and non-theorems are not selectable.

The condition implies consistency.
So by the Not Representable Argument $\neg D$ can not be represented.
Thus D and $\neg D$ are not selectable and so the diagonal theorems and diagonal non-theorems and thus all theorems and non-theorems are not selectable either.

This argument is especially suited for consistent extensions of Q.
Yet strangely, it didn’t use the dual representability that is true in Q and remains to be true in all consistent extensions. We only used that all selectable sets are represented.
But not necessarily dually!
And indeed, we got something weaker than what is true in Q that the theorems are generable and the non-theorems are not. Only that one of them is not generable.
So we might think that then this weakness is a fault of not using the full truth of dual representability in these extensions. So using the full truth we would get a similar situation what is true in Q. But this is not the case because there can be consistent extensions of Q where it is not true that the theorems are generable. In fact, this Argument can show this.
If for example we add all true statements as axioms to Q, we only get by our argument that at least one of the theorems or non-theorems is not effective.
Of course, now they are perfect negatives, so if one were generable, the other had to be too.
So we get eventually that they are both not effective.
In “normal”, not only consistent but also effective extensions like adding the $7(\infty)$ scheme back again, we do get that the non-theorems must be the non effective.

But now we show that this argument is not necessary to see that the consistent extensions of Robinson inherit the non effectivity of the non-theorems.
In fact, we get a much better vision for what really inherits. It is the not separability.
So beside remaining in consistent extension, the theorems and anti-theorems can not be separated by any effective complementing supersets.
What’s more, we’ll show special theorems and anti-theorems that are already not separable.

7. Not Separable Argument:
If every E , $\neg E$ selectable sets can be dually represented:

$$n \in E \quad \leftrightarrow \quad \vdash P(n) \qquad \text{and} \qquad n \in \neg E \quad \leftrightarrow \quad \vdash \neg P(n)$$

then the theorems and anti-theorems are not separable.

Not surprisingly, the mentioned special theorems will be the diagonal ones:
$D = \{ P(<P>) \,;\ \vdash P(<P>) \}$. But now we don’t form the complement that was:
$\neg D = \{ P(<P>) \,;\ \neg \vdash P(<P>) \}$ rather the “total opposite” or “anti diagonal” set:
$D\neg = \{ P(<P>) \,;\ \vdash \neg P(<P>) \}$.
We can prove that D and $D\neg$ are not separable.

Suppose they were. Then we had some E , $\neg E$ that:

$$\vdash P(<P>) \text{ implies } <P> \in E \quad , \quad \vdash \neg P(<P>) \text{ implies } <P> \in \neg E.$$

Regarding the P_0 , $\neg P_0$ that dually represent E , $\neg E$ we would have:

$$\vdash P(<P>) \text{ implies } \vdash P_0(<P>) \text{ and } \vdash \neg P(<P>) \text{ implies } \vdash \neg P_0(<P>).$$

At $P = \neg P_0$ we get:

$\vdash \neg P_0(<\neg P_0>)$ implies $\vdash P_0(<\neg P_0>)$ and $\vdash \neg\neg P_0(<\neg P_0>)$ implies $\vdash \neg P_0(<\neg P_0>)$.

Assuming $\vdash P_0$ implies $\vdash \neg\neg P_0$ we would have $\vdash P_0(<\neg P_0>)$ iff $\vdash \neg P_0(<\neg P_0>)$.

So P_0 could not be a dually representing property.

Back to the bigger picture:
Undecidability of the full Number Theory came out from a sub system, the Robinson Arithmetic, where undecidability is trivial from the outside.
All this shows clearly that the point is not undecidability, rather non generable complements, that is non selectable sets, or even more, sets that are not separable.
This makes also perfect sense of a historical prelude to Gödel.
Just a year before Gödel realized the undecidable statements in arithmetic, Presburger proved that without multiplication, this "Baby Genius" Arithmetic is complete.
It is Baby like because it uses only addition but it is Genius because it uses all inductions.
First of all, by the True representation Argument it follows that the language of addition is not enough for true representations of all generable sets. Indeed, otherwise any generable axioms that only derive truths could not derive all truths. But Presburger Arithmetic does all that.
The proof is not trivial at all. So even this Baby Genius Arithmetic is not trivially complete.
Above I said that the infinity of logical axioms explains how the derived theorems can be complex in a seemingly simple system too. So complex that the complement is not generable at all by any means. But "explains" doesn't mean that it always causes it!
Presburger's result shows that the potential complexity can be avoided and we can have generable complement. Thus the opposite case, that is actual complexity so that the complement is not generable, is not trivially visible. Only the representabilities show this.

This leads to a final and most important consequence of Robinson's result and also a need for historical correctness. Explaining also why I added Turing's name as last in the section title.
Let's start with history. His ground breaking 1936 article has a very misleading title.
The first part in it the "Computable numbers" is misleading already because refers not to his crucial new recognition that machines can collect the effective sets of objects.
These effective sets are what earlier were called as recursively enumerable sets.
Turing grasped these directly, without using special selectable, recursive sets.
So computing machines, as he also called his machines, would suggest that he'll also use the word computable for the machine collectable sets.
But he kept the computable adjective for the selectable collections.
And so he meant not computable sets of natural numbers, rather computable real numbers.
As those infinite binaries where both the 0-s and 1-s have effectively recognizable places.
Regarding real numbers, reveals a bigger vision that later with Randomness Theory turned out to be valid. Also, this might seem as a generalization of Cantor's earlier big step when he showed that the set of all Algebraic numbers is merely a sequence and so can not be all reals.
All computable real numbers are merely a sequence too. But this hides confusing further facts:
Indeed, the real numbers that have effective 0 or 1 places are sequencable too.
But while these one sided sequencings can be effective, the both sided ones can not be.

The second part of his title, "application to Hilbert's Entscheidungs Problem" is confusing too.
Because he tries to show by the external universality of his system as effectivity, that the logical necessities are not selectable. But Presburger's result shows that Turing's "proof" is faulty.
Indeed, the Baby Genius Arithmetic could be Logic! The true argument why Logic is not selectable, comes out from Robinson Arithmetic! Namely, that it has only finite many axioms and, so it is simply the Q statement. Indeed if Logic were selectable then Robinson Arithmetic would have to be too because the consequences of Q are at once recognizable.
Of course, remember that while Robinson Arithmetic is inheriting the non effectivity of the non theorems in consistent effective extensions, Logic is not! As Presburger Arithmetic proves it.

Also observe that this non selectability of Logic only came out for the language of Robinson Arithmetic, that is using two operations or two three variable relations instead.
The obvious question is what the minimal language conditions are for such non selectability.
The answer is simple, we must have an at least two variable relation or function or at least two one variable functions. To prove that these make an empty Logic non selectable is not easy.
We'll only prove three selectability results.
The first will be that not using any non logical symbols and non logical axioms, that is the empty Logic is selectable. To see why the no non logical symbols doesn't automatically means no non logical axioms, we must remember that equality is part of Logic but only with the three axioms: $x = x$, $x = y \rightarrow y = x$, $x = y \wedge y = z \rightarrow x = z$.
So we can have axioms about equality beyond these when selectability is not true anymore.
We'll show how this can happen by a counter example after our last theorem in the next section. That last theorem will instantly imply that the empty Logic is selectable.
Then in the section after the next the last theorem will imply that allowing only basic properties with finite many axioms keeps the selectability.
The idea for these two results is the same that we just used to show that Logic is not selectable for the Robinson language. So leaving out axioms, that is narrowing an axiom system.
This is an illogical idea because selectability does not inherit backwards even leaving out a single axiom. So if $\mathbf{A}$+B is a non selectable axiom system then $\mathbf{A}$ is not necessarily such.
Our third result will be actually an example for this, by showing a selectable finite system using a two variable $<$ relation. Indeed, then these axioms are actually a single B statement and dropping it, the remaining empty Logic plus $<$ is not selectable by the just mentioned result that a single two variable relation is already enough to make Logic non selectable.

Of course, we do know a condition when omitting a single B axiom keeps selectability.
Namely, if not only $\mathbf{A}$+B is selectable but $\mathbf{A}+\neg$B too. Indeed for all B:
$\mathbf{A}+B \vdash C$ and $\mathbf{A}+\neg B \vdash C$ implies $\mathbf{A} \vdash B \rightarrow C$ and $\mathbf{A} \vdash \neg B \rightarrow C$ implies $\mathbf{A} \vdash C$.
Also in reverse: $\mathbf{A} \vdash C$ implies $\mathbf{A}+B \vdash C$ and $\mathbf{A}+\neg B \vdash C$. In negative form:
$\mathbf{A}+B \neg\vdash C$ or $\mathbf{A}+\neg B \neg\vdash C$ implies $\mathbf{A} \neg\vdash C$.
So we can generate C as theorem in $\mathbf{A}$ if C is generated as theorem in both $\mathbf{A}$+B , $\mathbf{A}+\neg$B.
And we can generate C as non theorem in $\mathbf{A}$ if C is generated as non theorem in at least one of $\mathbf{A}$+B , $\mathbf{A}+\neg$B.
Observe that dual effectivity implies consistency and so the theorems and non theorems are disjoint and thus exactly one of the two generation cases must happen.
This is the ground idea behind the two narrowings too.
But for the wider idea of narrowing itself we also have a wider explanation.
Namely, that only very simple or very complicated conditions have simple consequences in Math. And usually to go back from the very complicated is easier.

Quantor Narrowing

Quantor Narrowing Theorem:
If $\mathbf{A} \cup \Delta$ is selectable and regarding all C statements, the truths of the:
$\exists B \in \Delta$ that $\mathbf{A}+\neg B \neg\vdash C$ or $\forall B \in \Delta$ is such that $\mathbf{A}+\neg B \vdash C$ truths are selectable, then $\mathbf{A}$ is selectable too.

Enough to show that:
$\mathbf{A} \cup \Delta \neg\vdash C$ or $\exists B \in \Delta$ that $\mathbf{A}+\neg B \neg\vdash C$ implies $\mathbf{A} \neg\vdash C$. And
$\mathbf{A} \cup \Delta \vdash C$ and $\forall B \in \Delta$ is such that $\mathbf{A}+\neg B \vdash C$ implies $\mathbf{A} \vdash C$.

Indeed, then since the two conditions are exact opposites:

If C is generated as non-theorem of $\mathbf{A} \cup \Delta$ or the $\exists B \in \Delta$ that $\mathbf{A} + \neg B \neg \mid\!- C$ truth is generated then we should generate C as non-theorem of $\mathbf{A}$.
If C is generated as theorem of $\mathbf{A} \cup \Delta$ and also the $\forall B \in \Delta$ is such that $\mathbf{A} + \neg B \mid\!- C$ truth is generated then we should generate C as theorem of $\mathbf{A}$.

The first implication is trivial and for the second observe that:
$\mathbf{A} \cup \Delta \mid\!- C$ implies $A \mid\!- (B_1 \wedge \ . \ . \ \wedge B_n) \rightarrow C$ and also that:
$\forall B \in \Delta$ is such that $\mathbf{A} + \neg B \mid\!- C$ implies $(\mathbf{A} + \neg B_1 \mid\!- C$ and . . and $\mathbf{A} + \neg B_n \mid\!- C)$ implies $\mathbf{A} \mid\!- (\neg B_1 \vee \ . \ . \ \vee \neg B_n) \rightarrow C$ implies $\mathbf{A} \mid\!- \neg (B_1 \wedge \ . \ . \ \wedge B_n) \rightarrow C$.
The two achieved consequences then together indeed imply $\mathbf{A} \mid\!- C$.

Corollary:

If $\mathbf{A}$ is generable then for it to be selectable we need only that the non-theorems are generable. And for this it is enough if regarding all C statements, the truths of $\exists B \in \Delta$ that $\mathbf{A} + \neg B \neg \mid\!- C$ are generable.

Now we need three theorems about models.

Cardinality Theorem:
If $\mathbf{A}$ has an infinite model then it has models with arbitrary infinite cardinality.

This result goes back to the very basics!
Consistency implying a model is the completeness of Logic. This was only presented in 1930 by Gödel. The Lövenheim Skolem Theorem was a result in 1920.
It used the existence of a model and showed that if the used names are finite or a sequence then there is a sequential sub-model inside. Now we might say "so what" by remembering the inner models of the non Euclidian geometries. But here there was no need for weird interpretations of the basic relations. Everything remains as was, we simply restrict our individuals to a single sequence where every axiom will remain to be valid. Now regarding the axioms of Set Theory that were already stated by 1920, this then means that if there is a model for Set Theory then there is sequential sub-model too. But the whole point of Set Theory is the cardinalities, starting with Cantor's original heuristic proof that the real numbers are not sequencable. How could that be in a sequential model. Well, one Gordian solution is that Set Theory is special, it is about everything and so its axioms have no model at all and thus there is no such sub-model either.
Actually, as all Gordian solutions it reveals something big that can be destroyed by going into the details where the devil runs free. Here we have the Completeness of Logic saying that any consistent axiom system must have a model. So Set Theory if not useless should have a model. And that, a sequential sub-model in which it is a true statement that there are non sequential sets. Then we could have a new Gordian truth that maybe this just showed that the axioms for Set Theory are not really sufficient, they can not reveal that Sets are everything and so can not be locked inside a single model.
Unfortunately, we can destroy this big truth again by saying that even if the present axioms for Sets are not a final perfect Set Theory, they still must be true for sets. They already claim non sequential sets to exist and wherever they are true there is sequential sub-model.
So finally we can only say that maybe saying that there is non sequential set is itself a mirage!
A linguistic extravagancy.
By the way, Skolem in 1922 when the Completeness of Logic was not yet proved, reformulated his result and actually created the sequential model not inside an original model, rather from the natural numbers. Exactly as we did to show the Completeness of Logic itself.
This was of course only a partial proof and indeed, he didn't claim this result either.

A strange detail is that while in his original sub-model claim he needed to choose the members by the Axiom Of Choice, in the second version this axiom was not needed but then in the 1930 full proof of Completeness, Gödel again needed the Axiom Of Choice. This clearly shows that this particular axiom is not a controversial member in the axioms as most crazy parrots claim.
It is a particularly important one that is almost a Logical axiom among sets but has nothing to do with the dilemma we dwelling on now, the adequacy of Logic to claim truths.
An other interesting detail is that the original sub-model proof though needed the Axiom Of Choice was actually a simpler version of how we showed that the concretizations can be done so smartly that the widened set obeys the quantors. Using an already existing model and regarding all the true situations in it, we can keep only the named ones extended by sequences so that the quantor meanings remain. The simplification is due to that we don't have to use alternating introductions for the quantors rather handle all universal quantors simultaneously.
We still need successive introductions because new depending names come about for the existential quantors. In fact, this goes back to the heart of quantors themselves.
Our consecutive quantor usage can be challenged and instead start with only universalities and dependences of variables. This at once shows that only some existences truly claim individuals to be named. These are the frontal ones in our consecutive quantor notation. While those that come after a universality, actually claim function values for universalities.
Not surprisingly, with this more general quantor system actually the proof becomes much easier. We simply use a definition that goes deeper. Actually, the well accepted consecutive quantor system is a special class of the wider quantor definition. It became the only surviving class because among the general ones the negation is not always possible while the negative of a consecutive quantor sequence is again such.

All the aboves have nothing to do with why we'll use the Cardinality Theorem!
But an opposite use of the Completeness of logic gradually became clear too.
Here we add names of bigger infinites with axioms claiming that they are not equal.
Then of course the existence of a model must be proved.
So actually this lead to Gödel's final proof of the Completeness Theorem.
This opposite "upward" meaning is just as surprising as was the "downward", implying sequential models for the universe of sets with arbitrary big cardinalities in it.
Indeed, by this upward meaning the natural numbers must have models with arbitrary big cardinalities. This was known years before the precise proof presented by Gödel in 1930.
But this date has an even deeper importance. Indeed, Gödel in that math congress already told some understanding ears that in spite of Logic being Complete, the axiom system of naturals can not decide every statement. This later became called as an Incompleteness of the axioms.
The earlier already suspected bigger infinite models for the naturals may seem to suggest this.
Indeed, such weird models with weird new numbers beside the real natural numbers should mean new truths too that are not true among the real ones only. So then these statement must be undecidable from the axioms because the derivable ones must be true in all models.
But just a year earlier, Presburger proved that the Baby Genius Arithmetic using only addition but all the induction axioms is complete. This Baby Genius Arithmetic also has weird models but Presburger's result meant that in these weird models there are no weird truths at all.
So many, including Hilbert were crossing their fingers that maybe the same is true for the full Number Theory. But multiplication destroys this situation. One would expect that Gödel then found some weird models where multiplication allows weird new truths. But that didn't happen at all. Instead he showed that multiplication allows some weird self asserting statements that simply can not be decided without contradiction. So then consistency implies undeciadability.
Then Turing realized that behind these self asserting statements actually the simple fact lies that while the theorems are generable, the left over complement set, the non-theorems are not.
They can not be effectively listed by any effective method.
In an even more plausible way this means that there can not be a counter Logic and counter axioms for the naturals that would derive exactly the non-theorems.
So the moral of the story could be that models turned out to be useless while a new world of effectivities opened. This section actually shows that models are not totally useless.

Los Vaught Theorem:
If **A** has only infinite models and a c cardinality ones have same truths, then **A** is complete.

The completeness of **A** means that for any C exactly one of C or $\neg C$ is theorem.
The assumption that the c cardinality models have same truths, silently assumes that there are such models and thus **A** is consistent. That is, both C and $\neg C$ can not be derivable.
So a non completeness would mean that a C would have to be undecidable and thus by indirectness, $\mathbf{A}+C$ and $\mathbf{A}+\neg C$ were both consistent.
Then these both had models by the completeness of Logic.
These models were both infinite by the first assumption.
So by the Cardinality Theorem $\mathbf{A}+C$ and $\mathbf{A}+\neg C$ both had c cardinality models too.
But this contradicts the second assumption.

By the previous two theorems, our third main one is:
If **A** is finite and an infinite c cardinality models have same truths, then **A** is selectable.

The second condition reminds us the Los Vaught Theorem, so not surprisingly, we want to extend **A** into an $\mathbf{A}\cup\Delta$ that would have only infinite models. Indeed, then the c models of $\mathbf{A}\cup\Delta$ are also c models of **A** and so the c models of $\mathbf{A}\cup\Delta$ must have same truths too.
Thus then indeed, we can use Los Vaught and so $\mathbf{A}\cup\Delta$ is complete and thus selectable too.
This reminds us the Quantor Narrowing Theorem, so if **A** and Δ satisfy the conditions of that then the selectability of **A** at once follows.
The Corollary is enough of course, because **A** is trivially generable. So:

For every natural number we define a statement about equality:
$|M| \geq n \ := \ \exists x_1, . . , \exists x_n (x_1 \neq x_2 \ , \ . \ . \ . \ , x_{n-1} \neq x_n).$
The naming reflects that having this statement as axiom makes sure that a model has at least n many members.
$|M| = n \ := |M| \geq n \wedge \neg(|M| \geq n+1)$ as axiom makes sure that a model has exactly n members.
Then $\Delta = \{ |M|\neq 1 , |M|\neq 2 , \ . \ . \ . \ \}$ makes indeed sure that the models of $\mathbf{A}\cup\Delta$ can only be infinite. So we must only check the conditions of the Corollary.
And indeed, the $\mathbf{A} + \neg(|M|\neq n) \neg\vdash C$ that is $\mathbf{A} + |M| = n \neg\vdash C$ that is

$\mathbf{A}\neg\vdash |M| = n \rightarrow C$ truths are generable by checking the finite models.

This theorem becomes false if **A** is not finite only generable. Namely:
Let Δ' be a generable subset of Δ with $\Delta - \Delta'$ being non generable and let **A** have Δ' as non logical axioms.

The $|M| = n \rightarrow C$ formed non-theorems are now not generable for sure.
This is the example I promised at the end of the previous section after listing the three equality axioms. Here indeed there are no non logical symbols but there are extra non logical axioms that destroy the selectability of the three equality axioms.

And now as our special application of the last theorem we can come to the just mentioned empty Logic, that is having only the equality axioms. We claim that his system is selectable!
Indeed, it has finite many axioms and any infinite cardinality models are isomorph and so have same truths.

Selection Narrowing

Now we'll narrow now not from a fix Δ rather a whole set of them.
Our goal is selectability and strangely we'll use a certain selection for this.
They will be Γ_s selections made from a Γ set of opposing statement pairs.

First an easy result:
If $\mathbf{A} \cup \Gamma_s \vdash C$ for every s selection made from Γ then $\mathbf{A} \vdash C$ too.

Let Γ_M denote the selection of Γ as the statements true in the M model.
Let M be a model of $\mathbf{A}$. C must be true in M. Indeed, $\mathbf{A} \cup \Gamma_M \vdash C$ and all models obey Logic, meaning that if some assumptions are true then the logical consequences must be too.
The $\mathbf{A} \cup \Gamma_M$ statements are true in M so the C consequence must be true too.
But Logic obeys models too, meaning that if a C statement is true in all models where some $\mathbf{A}$ assumptions are true then C is derivable by Logic from $\mathbf{A}$, that is $\mathbf{A} \vdash C$.

We call a Γ_s as $\mathbf{A}$ consistent or $\mathbf{A}$ complete according to $\mathbf{A} \cup \Gamma_s$ being consistent or complete.
We call a Γ an $\mathbf{A}$ decider if every Γ_s selection that is $\mathbf{A}$ consistent, is $\mathbf{A}$ complete too.

Our main theorem is:
For every $\mathbf{A}$ if Γ is an $\mathbf{A}$ decider then for every C statement made in the language of $\mathbf{A}$ there is a B $\wedge$-$\vee$ combination made from Γ statements so that:
$\mathbf{A} + B \vdash C$ and $\mathbf{A} + \neg B \vdash \neg C$.

Enough to show that if $\vee\Gamma$ denotes the set of $\vee$ combinations made from Γ and $(\vee\Gamma)_{A+C}$ denotes the $\mathbf{A}$+C consequences in $\vee\Gamma$ then $\mathbf{A} \cup (\vee\Gamma)_{A+C} \vdash C$.
Indeed, then for some $B = B_1 \wedge .. \wedge B_n$ made from $(\vee\Gamma)_{A+C}$, we have $\mathbf{A} + B \vdash C$.
But also $\mathbf{A} + C \vdash B_i$ thus $\mathbf{A} + C \vdash B$ and so $\mathbf{A} \vdash C \leftrightarrow B$ thus $\mathbf{A} + \neg B \vdash \neg C$.
To prove $\mathbf{A} \cup (\vee\Gamma)_{A+C} \vdash C$ we'll use the previous theorem.
So we'll show that for all Γ_s selections: $\mathbf{A} \cup (\vee\Gamma)_{A+C} \cup \Gamma_s \vdash C$.
This follows by examining the three natural cases and amazingly, they each will imply that two of the three members are already enough to derive C. Indeed:
If Γ_s is not $\mathbf{A}$ consistent then $A \cup \Gamma_s \vdash \text{false} \vdash C$.
If Γ_s is $\mathbf{A}$ consistent then by the assumption of being an $\mathbf{A}$ decider it is complete too.
So we have two sub cases: $\mathbf{A} \cup \Gamma_s \vdash C$ or $\mathbf{A} \cup \Gamma_s \vdash \neg C$. The first gives two members at once.
In the second case, for some $B_1 \wedge .. \wedge B_n$ made from Γ_s, $\mathbf{A} + B_1 \wedge .. \wedge B_n \vdash \neg C$.
So $\mathbf{A} + C \vdash \neg(B_1 \wedge .. \wedge B_n) = \neg B_1 \vee .. \vee \neg B_n$ which is thus in $(\vee\Gamma)_{A+C}$.
So $(\vee\Gamma)_{A+C} \cup \Gamma_s \vdash (\neg B_1 \vee .. \vee \neg B_n) \wedge (B_1 \wedge .. \wedge B_n) = \text{false} \vdash C$.

No we can show how this theorem can be used as a selector.
If $\mathbf{A}$ and Γ are both generable, Γ is an $\mathbf{A}$ decider and the theorems and non-theorems of $\mathbf{A}$ are selectable among $\vee\Gamma$ then $\mathbf{A}$ is selectable.
So all theorems and non-theorems of $\mathbf{A}$ are selectable.

For a given C statement this is how we can tell if C or $\neg C$ or neither is theorem of $\mathbf{A}$.
By the generability of Γ we can generate all $\wedge$-$\vee$ combinations from Γ.

Observe that if $(B_1 \vee .. \vee B_n) \wedge .. \wedge (\quad \vee \quad) = D_1 \wedge .. \wedge D_m = F_i$ then
$\neg F_k = \neg D_1 \vee .. \vee \neg D_m = (\neg B_1 \wedge .. \wedge \neg B_n) \vee .. \vee (\quad \wedge \quad) = F_j$.
We can even generate so that $\neg F_i = F_{i+1}$.
By **A** being generable we can also generate all the pairs of theorems.
So we can see if $F_i \rightarrow C$ and $\neg F_i \rightarrow \neg C$ both appear.
By our previous theorem this must happen for some F_i.

Having found $F_i = D_1 \wedge .. \wedge D_m$ and $\neg F_i = E_1 \wedge .. \wedge E_k$, by the selectability of the theorems among $\vee \Gamma$, we can check if all D-s or E-s are theorems.
If all D-s are then F_i and C is theorem too.
If all E-s are then $\neg F_i$ and $\neg C$ is theorem too.
Assuming consistency of **A**, both of these will not happen.
If neither all D-s nor all F-s are theorems then neither C nor $\neg C$ is theorem.

And now as an application, we can show that a finite **A** using only $p_1, ..., p_m$ properties is selectable. Obviously enough to show that Logic using only these properties is selectable.
Let's form all possible $\wedge$ combinations of p-s and $\neg$p-s :

$q_1 = p_1 \wedge .. \wedge p_m$

.

.

$q_M = \neg p_1 \wedge .. \wedge \neg p_m$

Also let $\exists n\, q_i$ denote $\exists x_1 .. \exists x_n [x_1 \neq x_2 \wedge .. \wedge x_{n-1} \neq x_n \wedge q_i(x_1) \wedge .. \wedge q_i(x_n)]$.

Then $\exists 1\, q_1, \exists 2\, q_1, ...$
$\exists 2\, q_2, \exists 2\, q_2, ...$

.

.

$\exists 1\, q_M, \exists 2\, q_M, ...$

and their negated statements as a Γ is an **A** decider. Indeed:

Consistent splits are only all non negated versions or all negated from a q_i.
And such splits are complete because either all lines start negation and then the models are all finite or there is a line without negative and then only infinite models can be and the sequencable ones are isomorph.
So in the first case trivially and in the second by Los Vaught we get the completeness.
So the conditions of our selection method stand too:
Our Γ is generable, **A** is generable since it was assumed to be finite, we just showed that Γ is an **A** decider and finally a $B_1 \vee .. \vee B_n$ made from Γ is theorem if and only if there is a B_j, B_k pair among the members that $B_j = \exists n\, q_i$ and $B_k = \neg \exists N\, q_i$ so that $n < N$.
Indeed, if such pair exists then $\neg B_j \rightarrow B_k$ means $\neg B_k \rightarrow B_j$ so $B_j \vee B_k =$ true.
And if such pair doesn't exist then $\neg(B_1 \vee .. \vee B_n) = \neg B_1 \wedge .. \wedge \neg B_n$ can be realized in a finite model by simply taking as many from each q_i combination's $\neg B_1, .., \neg B_n$ members the $\neg\neg \exists N\, q_i$ formed one as the biggest N claims to exist.

Beyond the basic operations

To examine these most clearly, we should avoid the successor function $x' = x + 1$ used as basic symbol and rather stick to relations. So then $x \lhd y$ denotes consecutiveness.
The price is that we can not get the values of $1, 2, \ldots$ as $0', 0'', \ldots$ so we have instead $1, 2, \ldots$ as infinite many basic names with the infinite many name axioms: $1 \lhd 2 \lhd \ldots$
An inclusion of 0 as natural number was first accepted by Peano to simplify his rules and there is an anecdote about this: Once at the opera when they gathered their coats after the show, his wife said "did you get all pieces?". Peano said "yes, zero, one, two, three". The wife looked at the four coats and said "and you are the mathematician". I agree with his wife and will not include 0 as natural. The three basic claims about $\lhd$ are very "simple" again:
For every x there is a single y that $x \lhd y$.
For every y except 1 there is a single x that $x \lhd y$.
There is no x that $x \lhd 1$.
Why I wrote the "simple" in quotation mark is the already mentioned non standard models that apply here again. But now coming to the more important operational axioms:

$$x \lhd z \rightarrow x + 1 = z \quad , \quad x + (y+1) = (x + y) +1$$

$$x \cdot 1 = x \quad , \quad x \cdot (y + 1) = x \cdot y + x$$

$$x^1 = x \quad , \quad x^{(y+1)} = x^y \cdot x$$

Or regarding the operations as relations and avoiding the functional tricks in the second axioms:

$$x \lhd z \rightarrow x + 1 = z \quad , \quad x + y = z \wedge y \lhd y' \wedge z \lhd z' \rightarrow x + y' = z'$$

$$x \cdot 1 = x \quad , \quad x \cdot y = z \wedge y \lhd y' \wedge z + y = z' \rightarrow x \cdot y' = z'$$

$$x^1 = x \quad , \quad x^y = z \wedge y \lhd y' \wedge z \cdot y = z' \rightarrow x^{y'} = z'$$

So the operations are now three variable relations each "defined" by two axioms.
The first is an initial statement, the second an implication that makes us able to get the cases for all names by using the name axioms too. To derive $4 + 3 = 7$ for example, we go like this:
$4 + 1 = 5$ by the initial axiom using the $4 \lhd 5$ name axiom and the Logic rule Modus Ponens.
Then by the implicative axiom $4 + 1 = 5 \wedge 1 \lhd 2 \wedge 5 \lhd 6 \rightarrow 4 + 2 = 6$ so we get $4 + 2 = 6$.
Then $4 + 2 = 6 \wedge 2 \lhd 3 \wedge 6 \lhd 7 \rightarrow 4 + 3 = 7$ gives $4 + 3 = 7$.
For multiplication cases of course we need to derive the needed addition cases and for exponentiation both addition and multiplication cases.
We get all true operational cases and so the "definition" was perfect but we still had to use quotation mark because a definition as such means explicit definition, using earlier defined or accepted basic relations only. Here the crucial implicative statements had the relation in the condition already. So we actually had a circularity. That's why the usual name is Peano rules.
Which of course doesn't really reveal that these are perfectly okay axioms.
Quite amazingly, if we want to continue our operations infinitely by similar rules, the power of general recursion allows to get all those as a number dependent $x\ [n]\ y = z$ relation.
So $n=1$ should give addition, $n=2$ multiplication and so on. The rules are then:

$$x \lhd z \rightarrow x\,[1]\,1 = z$$

$$n \lhd n' \rightarrow x\,[n']\,1 = x$$

$$n \lhd n' \wedge x\,[n']\,y = z \wedge z\,[n]\,x = z' \wedge y \lhd y' \rightarrow x\,[n']\,y' = z'$$

The first rule simply says that for $y = 1$ value, addition is same as $x \lhd z$.
The second says that for the other [n] operations the initial value is always x.
The third rule says that to step in y, we need not just the value of the present n' operation but the previous n too, applied for the old result z with x.
And indeed, to increase the multiplication value from y to the next y' we just have to add x.
To increase the exponentiation value from y to y' we must multiply by x and to increase the [4] value from y to y' we must raise the old result to the x power.

So x [4] 1 is x by the second rule and then x [4] 2 is x^x.

Then x [4] 3 is $(x^x)^x = x^{(x^2)}$ and we shouldn't write simply x^{x^x} for this!

Indeed, the $x^{(x^x)}$ value is different and could be understood for x^{x^x} just as well.

This is a consequence of exponentiation not being arbitrary in its order unlike addition and multiplication. So, for multiplication we can simply say that it is repeated addition and for exponentiation that it is repeated multiplication but for x [4] y we shouldn't just say that it is repeated exponentiation.
A simpler fact is that the exchangeability of order already fails for exponentiation.
Indeed, $2^3 = 8$ but $3^2 = 9$ which feels trivial but same rules created addition and multiplication and for those the order is irrelevant. This is a mystery!
Observe something else that's interesting! For the old normal operations the operational value z was always bigger than the x , y input values. Now with x [n] y this fails.
For example, for all n values we get merely x as initial value at $y = 1$.
This fact that we can get small z values for big n hides the fact how fast z grows.
This caused big havoc at the early recursion period of Effectivity.
Then the normal operations were first generalized in a different direction as f functions that are given initial values and then defined for step by step increasing inputs.
These primitive recursive functions can not grow as fast as our x [n] y and this was regarded as a complexity of x [n] y = z. The truth is quite the opposite.
Our rules only give the impression that the collectable tuples are complex because we have the freedom to choose derived cases and stupid choices can become an infinity of cases that are only a very small subset of all derivable cases. But we can be much smarter too.
We can regard an m fixed value and try all numbers up to m as inputs that is case conditions in all our finite many defining conditions. Obviously as start we'll only be able to use these numbers in the $\lhd$ basic relations there. Then we get some target tuples and we again only regard the ones using values up to m. These can now be used in our second try of all conditions. We again get new tuples with values up to m and we use these again. Repeating this, we get a stage where no more new targets can be obtained. Simply because we couldn't get infinite many target tuples by using only numbers up to m. So we derived all target tuples with using values up to m. Observe that this m will also be the maximal value in all of our derivable targets above. Indeed, in every target in our rules above, the maximal variable value was always at least as big as other values in the conditions. Namely, the maximal was always z' except at x [n'] y' = z'. And here if n' is the same or bigger than z' then the only variable not up to z in the conditions is n.
m appearing as maximal in the targets then implies that a target with values up to m can not come about by increasing m to an M. Simply because then M will be the maximal.
This means the same that the tuples not in our targets up to m will remain outside tuples as we increase m. So we can derive the outside tuples too as follows:
We derive the inside ones in the above systematic manner using m = 1 , 2 , 3 , . . and at every such m value once we have all the derivable ones, we check all tuples formable up to m and list the ones not derived.
We can get rules that avoid this maximality of the targets by simply dropping out variables.
But such rules will still not produce a relation that trivially has no generable complement.
To see why, we must remember Rice's Theorem. We must collect programs!
The use of variables in logical derivations do the same in hidden form.

Gödel really

Amazingly, for Gödel's discovery we must go opposite to the generalization we "wasted" so much time just now. In fact, we must omit even exponentiation. So we need only seven axioms beside the infinite many name axioms. Three for $\lhd$ and four for addition and multiplication.
This is Robinson's system that as I said could have been also enough for Gödel's discovery.
Suppose we have a $(r_1, r_2, \ldots, r_n)$ so called tuple of natural numbers!
Can we create a code for it? We instantly think of a single number but a pair of c, d would be just as good. Indeed, for arbitrary long tuples just a double is a perfect result.
The more important requirement is that we should be able to recover our tuple from c, d using multiplication with logical symbols. This luckily allows division or rather dividability and even remainders. And indeed, with these it is possible to define an explicit $F(c, d, i) = r$ expression so that for any $(r_1, r_2, \ldots, r_n)$ tuple there are c, d codes that:
For all i up to n we have that $F(c, d, i) = r$ is derivable for a single r.
This is great because this way we don't need to decode n, rather just go up to n many members in our sequence of the defined r-s.
Now let's see what such expression can do because it is unbelievable.
We can explicitly define exponentiation! Indeed:
$x^y = z$ means that there is a $(r_1, r_2, \ldots, r_y)$ tuple that $r_1 = x$, $r_2 = x^2, \ldots, r_y = x^y = z$.
Here we still have the dots and y as subscript but we can avoid all that by saying instead:

$$\exists c \exists d \left\{ F(c,d,1) = x \ \wedge \ \forall k \,[\, k < y \rightarrow F(c, d, k) \cdot x = F(c, d, k+1)] \ \wedge \ F(c, d, y) = z \right\}.$$

We can see that any other repeatedly calculated tuples can be formalized if the calculation is expressible by addition and multiplication. And indeed, by such tuples we can describe anything effective. So not just exponentiation but all generable collections just became explicit.
Then we can show that the cases of these explicit definitions are even derivable in our system.
So the generable sets are not just explicitly definable in our language but representable in our axiom system.
The theorem we need for the concrete $F(c, d, i) = r$ is called the Chinese Remainder Theorem:

Let $d_1, d_2, \ldots, d_n$ be real, that is non zero naturals and all relative primes to each other!
Then for every $r_1, r_2, \ldots, r_n$ values, each under the corresponding d, that is $r_i < d_i$,
we can find a c so that these r-s are all the remainders of the corresponding d-s in c.
And these r values may include zeroes as indeed, remainders may be such.

The notation reveals that our tuple will be the remainders in c but how do we get the mentioned d_i values from a single d. First of course, let's find our d.
Let $r_1 + r_2 + \ldots + r_n + n = m$ and then $d = 2 \cdot 3 \cdot \ldots \cdot m = m\,!$
Then $d_i = i\,d + 1$ will satisfy their conditions in the Chinese Remainder Theorem.
The first condition, the r values being under them is trivially true because :
$r_i < m < m\,! = d < i\,d + 1 = d_i$
The second condition, the relative primness means that:
For every $i \leq n$ the $i\,d + 1$ values are all relative primes to each other. First of all:
They can not have a non 1 divider of d as divider, because such leaves 1 remainder.
Now if two of them $j\,d + 1$ and $k\,d + 1$ had a common p prime factor then p would divide $(k - j)\,d$ too. But p divides separately and can't divide d since it is a divider of d too.
But neither can divide $k - j < n < m$ because $d = m\,!$, so $k - j$ is a divider of d too.
Thus we can use the theorem and claim that such c exists.
This c is a concrete number, though the theorem doesn't give it explicitly.
For our purpose it is enough that it exists.

Then the explicit expression for $F(c, d, i) = r$ is :

" r is the remainder of $i\,d+1$ in c " $=$ $r < i\,d+1 \wedge \exists\, q < c\ [\, c = q\,(i\,d+1) + r\,]$

This is always defined for all i values up to a point, namely when $i\,d+1$ exceeds c.

Some people define remainders even beyond there, by regarding the q quotient as 0.
Then the expression is giving a full infinite sequence of r-s.

For a proof sketch of the Chinese Remainder Theorem observe these:
Let $P = d_1\, d_2\, ..\, d_n$. The $0, 1, 2, \dots, P-1$ values under P mean P many choices to try as c and as easy to see, they all give different remainder tuples. But the possible under valued tuple combinations are also this many combinatorically as product of choice numbers.
Thus, at least one tried value under P must be a c, giving our tuple.
There is an uglier method to find an $F(c, d, i) = r$ expression without the Chinese Remainder Theorem and it is using a c that is explicit from the tuple.