Intermediate Logic

Lecture notes

Sandy Berkovski

Bilkent University
Fall 2010

Chapter 1

Paraphrase

1.1 Paraphrase: connectives

Let us rehearse some topics from the introductory course. We paraphrase declarative sentences that are either true or false into sentences of formal logic. Many ordinary sentences are not declarative, and many declarative sentences are not either true or false.

Example 1.1. 'Would you like some tea?' is not a declarative sentence. 'I am hot' is a declarative sentence, but does not have a truth value. 'Ankara is pretty' is a declarative sentence, but again does not have a truth value. why? ...

In order for a (declarative) sentence to have a truth value it must be supplemented with different contextual parameters. Those may include time, place, or the author of its utterance. Thus 'Ankara is pretty' must be transformed into 'Ankara is pretty in 2010', or perhaps even into 'Ankara is pretty-according-to-UN-standard-of prettiness in 2010'. why? ...

These complications are significant philosophically, but are less so logically. We shall in fact ignore them and treat all grammatically declarative sentences as if they were supplemented with contextual parameters and thus truth-evaluable.

The paraphrase of ordinary language sentences into the sentences of formal logic is to a very large extent arbitrary. The 'correctness' depends on the purposes at hand.

1.2 Paraphrase: quantifiers

Sometimes paraphrase is straightforward.

Example 1.2. Let us paraphrase the sentence 'Everyone is happy' into the language of quantifiers. Let P be the predicate '① is happy'. Then we have:

$$\forall x P x.$$

Let us paraphrase the sentence 'There is someone unhappy'. We have:

$$\exists x \neg P x.$$

But already here we have a choice. Let Q be the predicate '① is unhappy'. Then:

$$\exists x Q x.$$

Which option is better? The better option is the one which reflects negation in the original sentence.

In most cases paraphrase is not quite trivial.

Example 1.3. 'Every philosopher is virtuous.' We immediately have a choice. We can take the domain of objects over which our variables range to be the set of philosophers. Then we have simply: Domain of variables

$$\forall x V x.$$

There is nothing logically wrong with this decision. But this is not a healthy policy. Normally we want to represent the linguistic structure of the original sentence in our paraphrase. And the original sentence has the predicate '① is a philosopher'. Now we cannot paraphrase this sentence as: Why?

$$\forall x(Px \wedge Vx).$$

We must use implication:

$$\forall x(Px \supset Vx).$$

Example 1.4. 'Every Greek philosopher is virtuous.' If we take the domain to include Greek people, then we get the paraphrase as in the previous example. Naturally we want a different paraphrase. So our domain will include all people:

$$\forall x((Px \wedge Gx) \supset Vx).$$

Let us try something more complicated.

Example 1.5. 'Whenever the Fed cuts the interest rate by more than quarter a percent, the market shares rise and the dollar weakens in relation to all major currencies.' We make the following stipulations:

P : ① cuts the interest rate by ② at ③
Q : ① rises at ②
R : ① weakens in relation to ② at ③
S : ① is bigger than ②
T : ① is a moment of time
U : ① is a major currency
V : ① is a market share
a : 0.25%
b : US Dollar
c : Federal Reserve.

Then we have:

What is our domain?

$$\forall x(Tx \supset \forall y(Sya \supset (Pcyx \supset (\forall z(Vz \supset Qzx) \wedge \forall u(Uu \supset Rbux))))).$$

Now to existential quantification.

Example 1.6. 'Some politicians are virtuous.' The paraphrase is achieved thus:

$$\exists x(Px \wedge Vx).$$

Why aren't we using implication here? The sentence:

$$\exists x(Px \supset Vx)$$

in effect says that someone is either not a politician or is virtuous. This is true even when all politicians are evil.

Example 1.7. 'Nobody in Bilkent invests in Swedish krona.' We have:

$$\neg\exists x(Bx \wedge Ixa).$$

Note that this is equivalent to:

$$\forall x(Bx \supset \neg Ixa).$$

We move to consider cases involving both existential and universal quantification.

Example 1.8. 'Those who have tall girlfriends are tall.' The domain is the set of men. We have:

$$\forall x(\exists y(Gyx \wedge Ty) \supset Tx).$$

Interestingly, that the sentence 'Those who have *only* tall girlfriends are tall' is rendered differently:

$$\forall x(\exists y(Gyx \supset Ty) \supset Tx).$$

Example 1.9. 'Policemen know what criminals are up to.' The domain is the set of people and plans. We make the following stipulations:

Opaque context!

P : ① is a policeman
C : ① is a criminal
U : ① is up to ②
K : ① knows that ② is up to ③.

It is ambiguous what the sentence asserts. First guess:

$$\forall x(Px \supset \forall y \forall z((Cy \wedge Uyz) \supset Kxyz)).$$

But this sentence says that every policeman knows what *every* criminal is up to. Second guess:

$$\forall x(Px \supset \exists y \forall z((Cy \wedge Uyz) \supset Kxyz)).$$

This sentence says that every policeman knows what some criminal is up to. Yet one can take the original sentence to imply that all the criminals are under the control of policemen. Third guess:

$$\forall x(Cx \supset \exists y(Py \wedge \forall z(Uyz \supset Kyxz))).$$

Now this sentence fares better. But why to assume that one policeman must know *everything* the criminal is up to? Fourth guess:

$$\forall x(Cx \supset \forall y(Uyx \supset \exists y(Py \wedge Kzxy))).$$

It seems that this sentence conveys most adequately the meaning of the original English sentence.

Example 1.10. 'All the passengers on that flight lost some of their luggage.' Here it is important to symbolise the relations properly, even though not all of them are reflected in the surface structure of the sentence. We stipulate:

P : ① is a passenger on ②
L : ① is a piece of luggage
B : ① belongs to ②
R : ① lost ②
a : That flight.

Then we get:

$$\forall x(Pxa \supset \exists y(Ly \wedge Byx \wedge Rxy)).$$

1.3 Interpretation

When we interpret a formula of predicate calculus, we generally do the following:

1. Fix the domain.
2. Correlate each sentence parameter with a sentence.
3. Correlate each individual parameter with an individual in the domain.
4. Correlate each predicate parameter with a predicate.

One may notice that we have simply reversed the above procedure of translating sentences of a natural language (English) into the language of predicate calculus.

An uninterpreted formula cannot be assigned a truth value. It is only after we have fixed an interpretation that we can enquire whether a formula is true or false *in* an interpretation.

Example 1.11. Consider the formula $\forall x Px$. Consider the interpretation:

M : the set of natural numbers
P : ① is a number.

Given this interpretation, the formula says that whatever natural number you select, that natural number is a number. Clearly, the formula is true on this interpretation. Consider another interpretation:

M : the set of Bilkent students
P : ① is tall.

Given this interpretation, the formula says that whatever Bilkent student you select, that student is tall. Clearly, the formula is false on this interpretation.

Sometimes we wish to know whether a formula is valid—that is, whether it is true under *every* interpretation. We can use a more advanced method, such as semantic tableaux. Or we may simply try to find an interpretation under which the formula is false. If there is any one such interpretation, then the formula is invalid.

Example 1.12. Consider the formula $\forall x \exists y Pxy$. Consider the interpretation:

M : the set of people registered for our course
P : ① is taller than ②.

Given this interpretation, the formula says that whatever person in our course you select, there is a person taller than him. Clearly, in our course there will a person who is taller than everyone else in the course. So the formula is false on this interpretation. And therefore, it is invalid.

Similarly, sometimes we wish to know whether a formula is satisfiable—that is, whether it is true under *at least one* interpretation.

Example 1.13. Consider the formula $Pa \supset \forall x Px$. Consider the interpretation:

M : the set of French kings
P : ① is a king
a : Louis XIV.

Clearly the formula is true in this interpretation. Hence it is satisfiable.

Example 1.14. Consider the formula $\forall x(Px \supset Q) \supset \neg\exists x Px$. Consider the interpretation:

M : the set of French kings
P : ① is a king
Q : Moscow is the capital of France.

The formula is true in this interpretation. Hence it is satisfiable. Why?

Chapter 2

Preliminaries

2.1 Philosophical and mathematical motivation

Perhaps one need not any special motivation to study a certain subject. We study logic, one may argue, because it is out there. It seems, however, useful to understand a little more what it is that we shall study, and how the field of logic interacts with other intellectual activities.

We cannot hope to provide a definition of logic. The reason is simple. Logic has been around for more than two thousand years. But even in the past one hundred years activities so multifarious were termed 'logic', that they cannot be covered by a one-line definition. A similar and more familiar problem exists with the definition of mathematics. The issue is interesting philosophically, because a substantive definition may affect the way you think about mathematics, your views on mathematical ontology or mathematical proofs. But clearly, no one-line definition can capture all activities we recognise as 'mathematical'. An interesting attempt was made by the algebraist Serge Lang. He defined mathematics as 'everything published in mathematical journals in the past one hundred and forty years.' Clearly this is not a serious definition (it is circular), but it gives us a hint of how to elucidate the subject of mathematics.

Another hint was given in 1844 by the Russian logician Platon Poretzky who defined formal logic—the kind of logic we are going to study in this course—as logic in its content, but mathematics in its method. Let us, therefore, try to clarify the intended contrast by looking at our logic's historical development, at what was 'published in logical journals in the past two thousand years.' Any such excursus would by necessity contain difficult logical concepts to be refined later in the course; it should be read with care. The earliest European logical treatise belongs to Aristotle. He perceived logic as the science of consequence, determining what follows from what. Why was that important? Suppose the sentence S *entails* the sentence S'. Presumably that means that the truth of S entails the truth of S'. To know such a consequence is useful, as it allows us to know the truth of one sentence just by knowing the truth of another. A question immediately arises as to how we justify our inference. Aristotle collected numerous data on inferences in natural language and assembled them into different forms, the so-called 'syllogistic figures'. They were classified according to their subject-predicate structure. Once the correct forms were established, one could use them for building correct particular arguments. That was the chief utility of logic both for Aristotle and for the Scholastic philosophers of the Middle Ages. The limitations of the Aristotelian treatment have soon become apparent. Syllogistic figures were few in number and involved some very simple sentences. More complex inferences, involving more complex sentences, had to be translated into simpler sentences before any logical analysis of them was possible.

Partly in reaction to these shortcomings a different approach to logic was taken by Leibniz (and some other earlier thinkers, notably by Lull). Leibniz hoped to create a *formalised* universal language. Such a language, using artificial symbols, would have replaced natural language in conducting philosophical and scientific arguments. In the universal language a mathematics of reason will be possible. Every term and every sentence of that language is to be assigned a certain 'character', a number, which will eventually determine the validity of inferences where they occur. Therefore, Leibniz muses, any paradox is really a result of a miscalculation, to be remedied by arithmetical laws of our new grammar. Any philosophical discussion can be solved by computation; instead of engaging in fruitless polemics the opponents should say to each other, 'Let's calculate!'

Example 2.1. Here is an illustration of Leibniz' idea. In modern terms, we investigate the sentences' behaviour in accordance with their quantificational structure:

Every wise man is righteous.

$$+70 - 33 \quad +10 - 3.$$

This is a true universal statement. Every term—that is, subject and predicate—is assigned a pair of numbers, each with a different sign. Its truth is reflected in the fact that each number in the characteristic pair of the subject can be divided by a number with the same sign in the characteristic pair of the predicate. When this is not the case, the statement is false.

Formalisation of language received further boost with the work of de Morgan, Boole, and Peirce in the mid-nineteenth century. They were the first to classify inferences according to their sentential structure. But it is only with Frege that logic begins to play a novel role as the language of mathematics. In fact, Frege attempted to reduce all mathematics to logic.

Another crucial step in the development of logic was the emergence of formal axiomatic systems, where major contribution was made by Hilbert. When mathematical theories were represented in the form of axiomatic systems, it became possible to explore the properties of those systems. The properties in question, *e.g.* consistency, were logical.

Since then foundational issues, justification of the mathematical discipline as a whole, that animated Frege and Hilbert, have lost their appeal for mathematicians. They remain popular with some philosophers. But the methods of model theory, an offspring of the axiomatic approach, have proved exceptionally fruitful in investigating the properties of various mathematical theories, notably set theory, abstract algebra, geometry, and analysis.

What is the significance of formal logic for philosophers? First of all, many issues in the philosophy of language, such as the theory of meaning and theory of reference have direct connection to the systematic formalisation of natural language. They become inseparable from a deep analysis of logical constants, quantifiers, and predicates that can be conducted rigourously with the machinery of first-order logic. Secondly, many metaphysical and epistemological topics, such as conceptual knowledge or analyticity, are best explored by formal tools. (*Cf.* Dummett's remark...) Then there is the philosophical import of Gödel's theorems, the nature of Turing machines, and their impact on the theory of truth, the conception of rationality, the mind/body problem, and so forth.

That said, there is no need to constantly question the utility of logic for specific philosophical and mathematical concerns, the scope of its application in solving various puzzles. Our interest in it must be pure and selfless, much like our interest in philosophy or mathematics.

2.2 General concepts

Let us introduce some concepts which will be of use in the future. We presuppose familiarity with the notation of sentence logic. The signs $\wedge$, $\vee$, $\neg$, $\rightarrow$ will stand for conjunction, disjunction, negation, and implication respectively. One should note that our logical symbols are used as meta-linguistic abbreviations, that is, they are not part of the language of the set theory we develop. (In subsequent chapters we shall employ a different sign for implication.)

2.2.1 Sets

The concept of a set can be thought of as a generalisation of the concept of collection. Any collection of objects will constitute a set. The fact that a certain object x is *included* in the set Z will be symbolised as $x \in Z$.[1] The notion of inclusion is a primitive notion. Sets are identified by the objects they include. Thus, the *axiom of extensionality* states that if any two sets include the same objects, they are identical. That is, to specify a set it suffices to specify all of its elements.

The notion of inclusion allows us to define some further set-theoretic operations as follows:

- A set X is a subset of a set Y—symbolically, $X \subseteq Y$—if every element included in X is also included in Y. The empty set $\varnothing$ includes no elements at all. By those definitions we conclude that for any set X, $X \subseteq X$ and $\varnothing \subseteq X$.
- A set X is a *proper subset* of Y—symbolically, $X \subset Y$—just in case $X \subseteq Y$ and $X \neq Y$.

[1] The sign $\in$ derives from the Greek verb *esti*, 'to be.'

- The *union* of X and Y is the set $Z = X \cup Y$ which includes all the elements of X and Y.
- The *intersection* $Z = X \cap Y$ is the set including the elements belonging both to X and Y.
- The *difference* $Z = X - Y$ is the set including the elements of X which are not in Y.

When $Y \subseteq X$, it is also useful to introduce the *complement* of Y in X which is just the difference $X - Y$. We shall designate the complement of Y as Y'.

Set-theoretic operations can be usefully explained with the aid of Venn's diagrams. The idea is to take as the basic set (the universe set) the collection of points on the plane. We represent the sets X and Y as circles (that is, containing points within their circumferences). Then the operations just discussed can be depicted as in Figure 2.1 where they are represented by hatched regions (the definition of $X \oplus Y$ is left as an exercise).

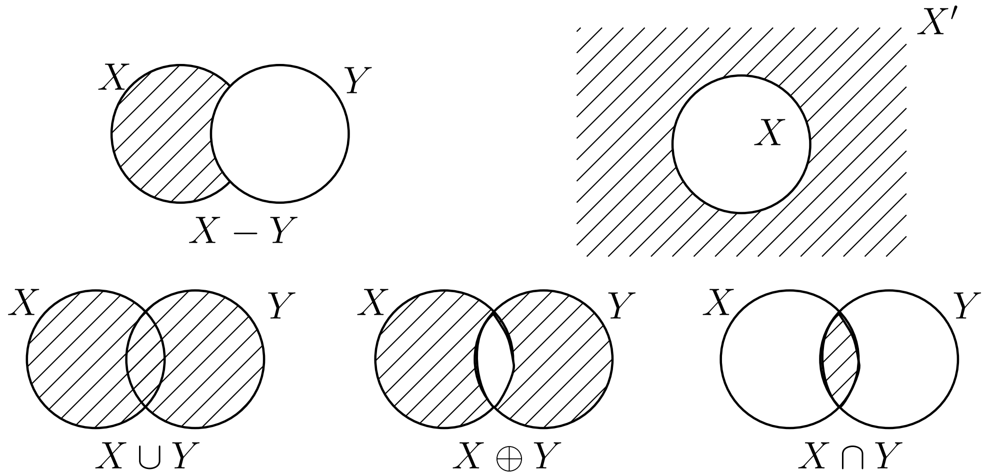

Figure 2.1: Set-theoretic operations

Cartesian product

We now move on to introduce relations and mappings. The *cartesian product* $Z = X \times Y$ of the sets X and Y is the collection of all pairs $\langle a, b\rangle$ such that $a \in X$ and $b \in Y$. Note that $X \times Y$ is not equal to $Y \times X$. If $A_1 = \cdots = A_n$, then the cartesian product $A_1 \times \cdots \times A_n$ will be designated as A^n.

Example 2.2. Let $X = \{7, 8\}$ and $Y = \{5, 6\}$. Then $Z = X \times Y = \{\langle 7, 5\rangle, \langle 7, 6\rangle, \langle 8, 5\rangle, \langle 8, 6\rangle\}$.

One way of defining binary relations is to draw their graphs on the Cartesian plane.

Example 2.3. Consider $\alpha = \{\langle x, y\rangle \mid x, y \in \mathbb{R} \wedge x > 2 \wedge y < 3\}$.

Example 2.4. Consider $\alpha = \{\langle x, y\rangle \mid x, y \in \mathbb{R} \wedge x = y\}$.

Each subset α of $X \times Y$ is called a relation on X and Y. If the pair $\langle a, b\rangle \in \alpha$, then we shall say that a stands in relation α to b. We can introduce the following operations on relations:

$$a(\alpha \cup \beta)b \iff a\alpha b \text{ or } a\beta b;$$
$$a(\alpha \cap \beta)b \iff a\alpha b \text{ and } a\beta b;$$
$$a\alpha' b \iff \text{ not } a\alpha' b.$$

We may therefore speak about disjunction, conjunction, and negation of relations.

Example 2.5. The relation $=$ of equality in the set $\mathbb{N}$ of natural numbers can be thought of as a collection of all pairs $\langle 0, 0\rangle, \langle 1, 1\rangle, \ldots$. The complement of this relation is the inequality relation. The relation $<$ is the set of all pairs $\langle a, b\rangle$ such that $a < b$. The relation $\leq$ is the same as the relation $< \cup =$. The relation $< \cap =$ is empty—that is, it is a necessarily false relation. By contrast, the relation $\leq \cup >$ is necessarily true.

Two more operations on relations should be mentioned. The *inverse* of α is the relation α^{-1} such that $\langle b, a\rangle \in \alpha^{-1}$ just in case $\langle a, b\rangle \in \alpha$. Suppose now that α is a relation of X to Y and β is a relation of Y to Z. The *composition* $\alpha\beta$ of X to Z is a relation such that $\langle a, b\rangle \in \alpha\beta$ just in case there is an element x such that $a\alpha x$ and $x\beta b$. The relations α and β are *permutable* if $\alpha\beta = \beta\alpha$. The *identity relation* in a set X consists of all pairs $\langle a, a\rangle$, where $a \in X$, and is denoted by ι_X.

2.2.2 Mappings

Among various kinds of relations several have particular importance.

Definition 2.6. A relation α defined on X and Y is called a *mapping* of X into Y if for each $a \in X$ there is exactly one $b \in Y$ such that $a\alpha b$. The element $b = a\alpha$ is the *image* of a, and a is the pre-image of b.

Definition 2.7. The set of all x for which there is y such that $x\alpha y$ is called the *domain* of α. The set of all y for which there is x such that $x\alpha y$ is called the *range* of α.

The fact that α is a mapping of X into Y is designated as $\alpha\colon X \to Y$. The fact that $b = a\alpha$ may sometimes be convenient to designate as $\alpha\colon a \mapsto b$. One may verify that a familiar notion of function, as understood, *e.g.*, in analysis, coincides with the notion of mapping. An n-ary *operation* is the mapping of A^n into A.

Definition 2.8. A mapping α of X into Y is called a *mapping of X onto Y* if for each $b \in Y$ there is at least one $x \in X$ such that $x\alpha b$. A mapping α of X onto Y is a *bijection* (or *one-to-one*) if the inverse relation α^{-1} is a mapping of Y onto X. Bijection

Remark. The definition of mapping-into guarantees that each element in the range of a mapping-onto has *exactly one* pre-image.

Example 2.9. The function $f(x) = 4x$ defined on reals is a bijection. The function $f(x) = x^2$ defined on reals is not a bijection. why? ...

Consider now a bijection α of X onto Y. Its inverse α^{-1} is also a bijection. Since for any $x \in X$ and $y \in Y$ we have $(x\alpha)\alpha^{-1} = x$ and $(y\alpha)\alpha^{-1} = y$, it follows that $\alpha\alpha^{-1} = \iota_X$ and $\alpha^{-1}\alpha = \iota_Y$. When α is a bijection of X onto X, then $\alpha\alpha^{-1} = \alpha^{-1}\alpha$. This in fact gives us a necessary and sufficient condition for a relation to be a bijection.

Cardinality

The notion of bijection allows us to introduce cardinalities. With each set X we associate a *cardinal* (or a *cardinal number*, or *cardinality*, or *power*) denoted by $|X|$ such that $|X| = |Y|$ just in case there is a bijection between X and Y. Thus, the empty set $\varnothing$ is assigned the cardinal 0, while the set $\{x_1, \ldots, x_n\}$ is assigned the cardinal n. The cardinal of the set of all natural numbers is denoted by $\aleph_0$. Where $|X| = |Y|$ the sets X and Y are called *equinumerous*. Further, if there is a one-to-one mapping of X into Y, then $|X| \leq |Y|$. A set X is *finite* if it is either empty, or else equinumerous with the set $\{1, 2, \ldots, n\}$ for some natural number n. Sets equinumerous with the set of all natural numbers are called *denumerable*. Sets which are either finite or denumerable are called *countable*. Equinumerousity of X and Y is denoted by $X \approx Y$.

Example 2.10. The sets $X = \{0, 1, 2, 3, \ldots\} = \mathbb{N}$ and $Y = \{0, 2, 3, 4, \ldots\} = \mathbb{N} - \{1\}$ are equinumerous, since it is possible to find a bijective mapping of X into Y.

2.2.3 Equivalence relations

We shall now introduce a further important class of relations.

Definition 2.11. A binary relation α in a set X is an *equivalence relation* on X if it satisfies the following conditions: Equivalence relation

- Reflexivity: $x\alpha x$,
- Symmetry: if $x\alpha y$, then $y\alpha x$,
- Transitivity: if $x\alpha y$ and $y\alpha z$, then $x\alpha z$,

for all $x, y, z \in X$.

Our discussion above allows to express the properties of equivalence relations also in the following form: why? ...

- Reflexivity: $\iota \subseteq \alpha$;
- Symmetry: $\alpha^{-1} \subseteq \alpha$;
- Transitivity: $\alpha^2 \subseteq \alpha$.

A family M of non-empty subsets of a set X is called a *partition* of X if each element of X belongs exactly to one element of M. The elements of M are called *classes* of the partition.

Example 2.12. Let R be the set of pairs $\langle x, y\rangle$ such that x and y are lines on the Euclidean plane and x is parallel to y. Then R is an equivalence relation.

Example 2.13. The set $X = \{x, y\}$ has two partitions: $M_1 = \{\{x, y\}\}$ and $M_2 = \{\{x\}, \{y\}\}$. The set $X = \{x, y, z\}$ has five partitions: $M_1 = \{\{x, y, z\}\}$, $M_2 = \{\{x\}, \{y\}, \{z\}\}$, $M_3 = \{\{x\}, \{y, z\}\}$, $M_4 = \{\{x, z\}, \{y\}\}$, $M_5 = \{\{x, y\}, \{z\}\}$. See Figures 2.2 and 2.3.

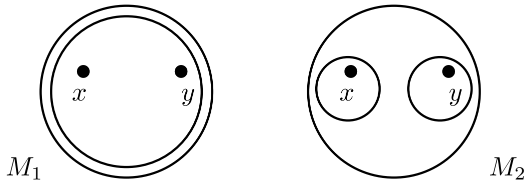

Figure 2.2: Partition for $\{x, y\}$

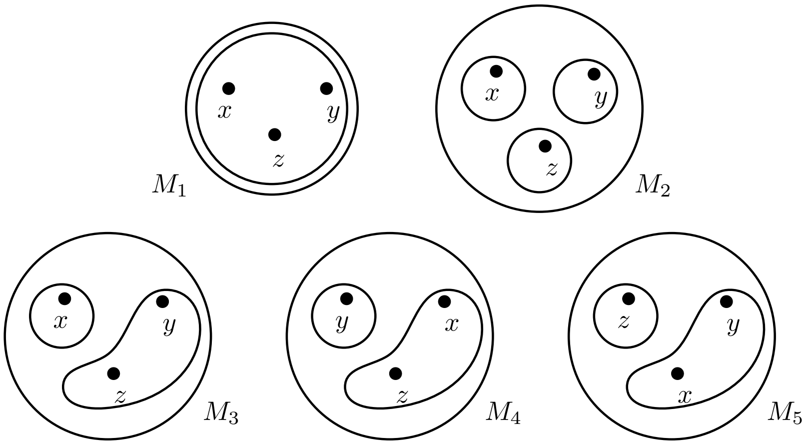

Figure 2.3: Partition for $\{x, y, z\}$

With each partition M of the set X we can associate a relation ϕ such that for any $x, y \in X$, $x\phi y$ just in case x and y belong to the same class of partition. Clearly the relation ϕ is an `equivalence relation`. We say that ϕ is *induced* by M. why? by definition

We can show that each equivalence relation ϕ induces a partition. For each $x \in X$, let $[x]$ be the set of all elements u such that $u\phi x$. The sets $[x]$ are called *equivalence classes* (see Figure 2.4). We further note that

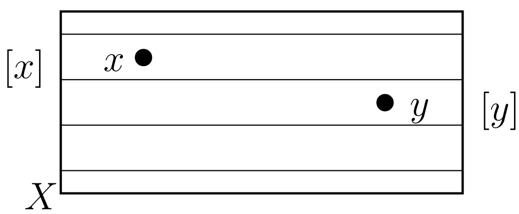

Figure 2.4: Equivalence classes

since ϕ is reflexive, $x \in [x]$. We must show that if $y \in [x]$, then $[y] = [x]$:

$y \in [x]$	Ass.	(1)
$y\phi x$	(1), def. of equivalence classes	(2)
$u \in [x]$	Ass.	(3)
$u\phi x$	(3), def. of equivalence classes	(4)
$x\phi y$	(2), ϕ is symmetrical	(5)
$u\phi y$	(4), (5), ϕ is transitive	(6)
$u \in [y]$	(6)	(7)
$[x] \subseteq [y]$	(3), (7)	(8)
$u \in [y]$	Ass.	(9)
$u\phi y$	(9), def. of equivalence classes	(10)
$u\phi x$	(9), (2), ϕ is transitive	(11)
$[y] \subseteq [x]$	(9), (11)	(12)
$[x] = [y]$	(8), (12).	(13)

Therefore, the family of all equivalence classes induced by ϕ is a partition of X. This family is called the *factor set of X by ϕ* and is designated as X/ϕ. We have shown that $x\phi y$ just in case there is a set $M \in X/\phi$ such that $x \in M$ and $y \in M$.

2.2.4 Trees

An *unordered tree* $\mathscr{T}$ is a triple $\langle S, \ell, R\rangle$ consisting of the following:

1. A set S of *points*;
2. A function ℓ assigning to each $x \in S$ a natural number $\ell(x)$ called the *level* of x;
3. A binary relation ①R② defined in S which is interpreted as '① is a *predecessor* of ②' or '② is a *successor* of ①'. This relation satisfies the following conditions:

   C_1: There is a unique point a_1 of level 1 called the *origin* of the tree;

   C_2: Every point save for the origin has a unique predecessor;

   C_3: For any $x, y \in S$, if xRy, then $\ell(y) = \ell(x) + 1$.

A point x is an *end point* if it has no successors; it is a *simple point* if it has just one successor; and it is a *junction point* if it has more than one successor. A *path* is any ordered sequence of points containing the origin such that each of its terms is the predecessor of the next (that is, except the last one, if there is such). A *branch* is a path whose last term is an end of the tree, or else it is an infinite path.

The conditions on R mean that that for any $x \in S$, there exists a unique path P_x whose last term is x. If y lies on P_x, then y *dominates* x. In that case, and if $x \neq y$, then y lies *above* x. The points x and y are *comparable* if one of them dominates the other. The point y lies *between* x and z if y is above one element of the pair $\langle x, z\rangle$ and below the other.

An *ordered tree* $\mathscr{T}$ is a quadruple $\langle S, \ell, R, \theta\rangle$, where θ is a function assigning to each point x an ordered sequence $\theta(x)$ containing all the successors of x. We shall then be able to speak about the first, second, and so on, successors of x.

Sometimes we shall have to add new points to the given ordered tree $\mathscr{T} = \langle S, \ell, R, \theta\rangle$. This will be done by adding them as successors to end points. Namely, to add distinct elements $y_1, y_2, \ldots, y_n$ to an end point $x \in S$ we add each y_i to S, add each $\langle x, y_i\rangle$ to R, let $\ell(y_1) = \ell(y_2) = \cdots = \ell(y_n) = \ell(x) + 1$, and let $\theta(x) = \langle y_1, \ldots, y_n\rangle$.

Most frequently we shall use *dyadic* ordered trees, in which each junction point has exactly two successors. The first successor will be called the *left successor*, and the second will be called the *right successor*.

A tree is called *finitely generated* if each point has a finite number of successors. A tree is *finite* if it has only finitely many points. A tree is *infinite* if it has infinitely many points. Note that a finitely generated tree may be infinite.

Example 2.14. The abstract notion of a tree we have given has a simple graphical representation. Consider a dyadic tree $\mathscr{T} = \langle S, \ell, R, \theta \rangle$, where we let $S = \{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$, $\ell(\mathbf{A}) = 1$, $\ell(\mathbf{B}) = 2$, $\ell(\mathbf{C}) = \ell(\mathbf{D}) = 3$, and $\theta(\mathbf{B}) = \langle \mathbf{D}, \mathbf{C} \rangle$. We draw the tree by placing the origin at the top, the successor y of each point x below x, and connect x and y with a line. The successors are ordered from left to right. The resulting tree is shown on Figure 2.5.

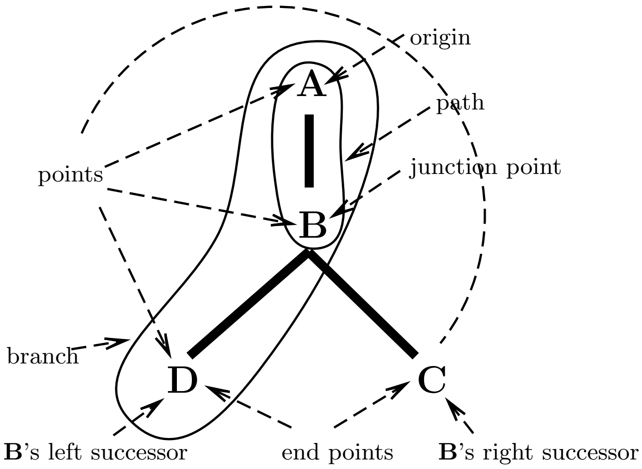

Figure 2.5: A dyadic tree

2.3 Logic and metamathematics

2.3.1 Axiomatic systems

The axiomatic method in mathematics goes back to Euclid's treatise on geometry. It can be characterised as follows. At the beginning we introduce primitive terms with the aid of definitions. Examples will include 'point', 'straight line', 'plane'. Definitions establish the meaning of the terms. We then formulate propositions involving those terms whose truth is seen as intuitive and in no need of proof. The source of their truth is supposed to lie in the very meaning of the terms. Euclid differentiates between *postulates* and *axioms*, but we may agree on calling all such propositions 'axioms'. Using primitive terms we may define new terms, and from the axioms logically derive new propositions, the *theorems* of our axiomatic system.

Such a system, in which the meaning of primitive terms is fixed from the start, may be termed 'informal'. Historically, the impetus for developing an alternative conception of axiomatic system, a 'formal' one, was given by the continuing uncertainty over Euclid's Fifth Postulate. The principal role of that postulate was in proving the theorem that through a point P not lying on a given line l it is possible to draw exactly one line parallel to the original line l. Both the original postulate and the theorem are not quite intuitive. How would one know, for instance, that the two allegedly parallel lines do not meet in some distant point in space?

A lot of effort was invested in proving the Fifth Postulate from the rest of the axioms. All these attempts were destined to fail, since in the early 19th century there was built a system in which through the point P, described as above, it is possible to draw infinitely many lines parallel to l.

Whereas in the past, prior to the 19th century, the axioms were required to be self-evident, there is no such demand in the modern conception of axiomatic systems. Arbitrary sentence may serve as axioms. This view is directly linked to the emergence of non-Euclidean geometries.

2.3.2 Language and metalanguage

Suppose we have an axiomatic system. We wish to investigate its properties. Then, on the one hand, we have a language in which our system is expressed, and, on the other hand, a language in which we talk about it. Those two languages do not need to coincide.

This is a perfectly familiar situation. A Turkish language textbook for English speakers will have Turkish as object-language and English as metalanguage. We similarly will study formal languages with the aid of English as the metalanguage.

The distinction generalises into the use-mention distinction. Consider the statement:

Ankara is pretty.

Following Quine, we say that the name ‘Ankara’ is *used* here to *mention* the city named by it. If, however, we write:

‘Ankara’ is pretty,

then we express our admiration for the name, not for the city. That is to say, the name is mentioned, whereas the quotation is used to refer to the name. Quotation-marks, then, are a device for jumping the level in the use-mention hierarchy.

For our later discussion it will be important to keep in mind the distinction between variables and meta-variables. Variables typically take as their values different objects. For examples:

Variables and meta-variables

Let there be two people, x and y, and x loves y.

Meta-variables act just like variables, except that they are designed to occur in the meta-language. Since meta-language is used to make claims about an object-language, it is to be expected that the values of meta-variables will typically be linguistic items—which themselves belong to the object-language. For example:

Let A be a formula of the Euclidean geometry.

Here the letter ‘A’ is used as a name of some formula of the Euclidean geometry. The formula itself belongs to the object-language, *i.e.* the language of the Euclidean geometry. The letter ‘A’ belongs to the meta-language and should be regarded as a meta-variable. (Compare: ‘Let ‘Tibbs’ be the name of our cat.’)

2.3.3 Set-theoretic constructions

Every discipline pays at least attention to its own language. But the focus on language in mathematics is unique. This can be partly explained by essential epistemological problems. In natural science the object of investigation is given. There is no need to doubt its reality unless one is engaged in a global sceptical enquiry. In mathematics the object of investigation is elusive. It is not clear in what sense numbers and sets are real. It is similarly unclear, for the same reason precisely, in what sense mathematical claims can be true.

This difficulty has important consequences. In building the edifice of formal logic we must use intuitive, informal notions. Metalanguage is their medium. Also, the focus on truth prompts deep enquiries into its behaviour in mathematical theories. They result in claims about incompleteness, independence, undecidability of various axiomatic systems.

Various schools in philmath....

Developments in set theory had considerable influence on the methods of modern logic. We give several illustrations.

Cantor’s diagonal method

Proposition 2.15 (Cantor)**.** *There is no 1-1 mapping between the set* $\mathbb{N}$ *and the set* $A = \{x \in \mathbb{R} \mid 0 < x < 1\}$.

Proof. We represent the members of A in the form $0, \alpha_1\alpha_2 \cdots \alpha_n \cdots$, where α_i is a decimal digit.

Suppose that A is denumerable. Then there is a bijection between $\mathbb{N}$ and A, and we can enumerate the members of A. We have:

$$
\begin{aligned}
&a_1 = 0, \alpha_{11}\alpha_{12} \cdots \alpha_{1j} \cdots \\
&a_2 = 0, \alpha_{21}\alpha_{22} \cdots \alpha_{2j} \cdots \\
&\cdots \\
&a_i = 0, \alpha_{i1}\alpha_{i2} \cdots \alpha_{ij} \cdots \\
&\cdots
\end{aligned}
$$

For each $j \in \mathbb{N}$ we can select a decimal digit β_j such that $\beta_j \neq \alpha_{jj}$, and β_j is not either a the digit ‘9’ or ‘0’. Consider the number:

$$b = 0, \beta_1\beta_2 \cdots \beta_j \cdots .$$

Now $b \in \mathbb{R}$ and $0 < b < 1$. Therefore, there is $i \in \mathbb{N}$ for which $b = a_i$. Then $\beta_j = \alpha_{ij}$ for every $j \in \mathbb{N}$. In particular, $\beta_i = \alpha_{ii}$. But this is impossible, since we selected b precisely to satisfy the condition $\beta_i \neq \alpha_{ii}$. We obtained a contradiction. Therefore, A is not denumerable. □

Russell's Paradox

Definition 2.16. A set X is *good* if it is not an element of itself. X is *bad* otherwise.

Consider the set Σ that contains all the good sets. Is Σ good or bad?

1. Suppose that Σ is good. Then Σ must contain itself by the definition of Σ. But then Σ is bad by the definition of badness. Contradiction.
2. Suppose that Σ is bad. Then Σ must not contain itself by the definition of Σ. But then Σ is good by the definition of goodness. Contradiction.

A natural conclusion of this reasoning is that Σ does not exist.

Grelling's Paradox

Definition 2.17. An adjective E of the English language is called 'autological' if E describes itself. E is *heterological* otherwise.

So, for instance, 'short' is autological, but 'long' is heterological. Now is 'heterological' heterological?

Berry's Paradox Each of the English names of natural numbers contains a number of words. For example, the first number requiring at least two words is 21. Let us consider the name (or more exactly, the definite description) of a certain number: 'The first number whose name requires at least eleven words.' This is an adequate name (description) of a number. Yet it contains just ten words.

2.3.4 The universal compiler

We shall now see how the methods just described may be applied to a logical problem. It is related with Gödel's incompleteness results.

Suppose that a computing centre receives texts written in a programming language, such as PASCAL. A special program translates these texts into programmes written in machine code. Some texts contain errors (e.g. syntactic errors), and the translator is made to react to them by halting.

A question arises if it is possible to have a translator which would discover *all* errors in any input text.

To fix the terminology, we consider a programming language L. Its alphabet contains a finite number of symbols. Any finite sequence of those symbols will be a *text*. The text is a 'correct program' if it is possible to translate it into machine code and if the translated code allows the computer to work for infinitely long time printing as its output the table of the form:

$$\begin{array}{ccccc} 1 & 2 & 3 & 4 & \cdots \\ f(1) & f(2) & f(3) & f(4) & \cdots \end{array}$$

where $f(n) \in \mathbb{N}$.

Definition 2.18. The program COMPILER is a program whose input is any text and whose output is 1 if the text is a correct programme, and 0 otherwise.

Proposition 2.19. *The* COMPILER *does not exist.*

Proof. Suppose that the COMPILER exists. We shall now write a certain Big Program which will work as follows.

1. Arrange all symbols of L in alphabetic order.
2. A special subroutine of BP generates all the texts of L arranged by length, while the texts of the same length are arranged in lexicographical order.
3. A text is received on the input of the COMPILER. If the output is 0, this text is discarded. If the output is 1, then the text is a correct program, and it is assigned the index i. A subsequent correct program gets the index $i + 1$. In this way we generate all and only all correct programs in lexicographic order.
4. Let the nth correct program compute the function $f_n(k)$. The final subroutine of BP computes for every $n \in \mathbb{N}$ the number $F(n) = f_n(n) + 1$. It generates the pair $(n, F(n))$ in the output.

Now the program BP is itself a correct program. So the function $F(n)$ must be contained in the array $\langle f_1, f_2, \ldots, f_m, \ldots\rangle$ of *all* the functions computed by correct programs.

But this is impossible. F cannot be in the array, since $F(n) \neq f_n(n)$ for every $n \in \mathbb{N}$. $\square$