V2000: Notes for Weeks 3-4

Reading assignment

Dumas-McCarthy: (DM) Ch 1, sections 2.5, 3.1– 3.4
Daepp-Gorkin (DG): Ch 27, 3, 4

1. Equivalence relations

Of particular importance among relations are *equivalence relations*, those that satisfy properties (a), (b), and (d) – reflexive, symmetric, and transitive. The prototype of an equivalence relation is equality. We check that equality in a set X is an equivalence relation:

(reflexive) If $x \in X$ then $x = x$ (and x is the only element of X that is equal to x).
(symmetric) If $x = y$ then $y = x$.
(transitive) If $x = y$ and $y = z$ then $x = z$.

An equivalence relation is a way to make a new set out of a given set. For example, let X be the set of people and let R be the relation "is a relative of." We admit that R is reflexive (this is a convention). Then by the (conventional) meaning of "relative" if x is a relative of y and y is a relative of z then x is a relative of z. This wouldn't work for the relation "cousin (or for that matter "blood relative").

Definition 1.1. *Let X be a set and R an equivalence relation on X. Let $x \in X$ and let $[x]_R = \{y \in X \mid yRx\}$. This is a subset of X called the "equivalence class" of x with respect to R.*

Thus for the relation "is a relative of" the equivalence class that contains you is your *(extended) family*. This is a very extended family, since it includes not only all your cousins and their cousins and so forth, but also in-laws and *their* in-laws and indeed any concatenation of conventional family relations. An interesting question (not for this course) is whether everyone belongs to the same family in this extended sense.

Question 1.2. *Etymologically, the relevance of "relative" to equivalence relations is clear. What about the relation "is in a relationship with"?*

1.1. **Partitions.** A *partition* of a set X is a collection ("*family*" in (DM)) of subsets Y_i indexed by some other set $i \in I$ such that

(a) If $i \neq j$ then $Y_i \cap Y_j = \emptyset$.

(b) Every $x \in X$ belongs to some Y_i (which is unique by (a); equivalently $X = \cup_{i \in I} Y_i$.

Note that set I that was introduced inconspicuously – the *index set* of the partition; it is the set without which it would not be possible to talk about partitions in the first place.

Theorem 1.3. *(a) Let $X = \cup_{i\in I} Y_i$ be a partition of X. Define a relation R on X by xRy if x and y belong to the same Y_i. Then R is an equivalence relation.*

(b) Let R be an equivalence relation on X. Then the collection of equivalence classes $[x]_R$ forms a partition of X.

In other words, equivalence relations and partitions are the same thing, viewed in a different way: as relations or as collections of subsets. The proof is given in (DM). I add a

Set-theoretic warning. One needs to know that each $[x]_R$ is a subset of X, and that the collection of equivalence classes forms a set. Bertrand Russell observed in 1903 that not every collection of sets that one can define is itself a set. Namely, let S be the collection of all sets that are not elements of themselves (the collection of all x such that $x \notin x$). Is S an element of S? If it is, then it is not; but if it is not, then it is. That is *Russell's Paradox* and set theory has to be axiomatized in a way that will not allow the definition of such an S.

In the case of equivalence classes, the existence of a set is guaranteed by the Zermelo-Fraenkel axioms (see Appendix B to (DM)), spectifically the **Power Set Axiom** that asserts that the collection of subsets of X forms a set (the *power set* $P(X)$ of X) and the **Axiom Schema of Separation**, which asserts that any subset of a set Y (like $P(X)$) defined by a reasonable formula is itself a set. Dumas-McCarthy don't worry about this detail, and neither will we.

Point (a) of the theorem is easy. The main point of the proof of (b) is to show that if x is not equivalent to y then $[x]_R \cap [y]_R = \emptyset$. Suppose not; i.e., suppose $z \in [x]_R \cap [y]_R$. In other words zRx and zRy. By symmetry, xRz and zRy. By transitivity, xRy, contradiction. A similar argument shows that if xRy then $[x]_R \subset [y]_R$; then by symmetry, $[y]_R \subset [x]_R$, and thus $[x]_R = [y]_R$. (Yes, one has to *prove* that if U and V are subsets of X with $U \subset V$ and $V \subset U$, then $U = V$. If this does not look obvious to you, you should write down a proof. If it does look obvious, you should also write down a proof.)

One also uses the more convenient notation $x \sim y$ for an equivalence relation; thus R is the subset $\{(x, y) \mid x \sim y\} \subset X \times X$. The set of equivalence classes for $\sim$ is often denoted $X/\sim$.

Level sets. Let $f : X \to Y$ be a function. Define the relation xRx' if $f(x) = f(x')$. This is shown in (DM) to be an equivalence relation. The corresponding partition, denoted X/f in (DM), is indexed by y in the image of f: $X/f = \{f^{-1}(y) \mid y \in image(f)\}$. We denote the equivalence class of x by $[x]_f$.

Lemma 1.4. *Let $Z \subset Y$ denote the image of f. There is a bijective map $\hat{f} : X/f \to Z$, with the property that*

$$\hat{f}([x]_f) = f(x)$$

and this property uniquely determines $\hat{f}$.

The proof is in (DM) and will be given in class.

2. Modular arithmetic

Any positive integer N defines an equivalence relation *congruence modulo* N on $\mathbb{Z}$. This is of interest because the arithmetic operations on $\mathbb{Z}$ carry over to the equivalence classes for this operation, and this is the prototype for some of the basic constructions in algebra.

Let $c, d \in \mathbb{Z}$. We say c divides d, and write $c \mid d$, if $c \neq 0$ and $\frac{d}{c} \in \mathbb{Z}$.

Definition 2.1. *Let $N \in \mathbb{N}$, $N > 0$. Let $a, b \in \mathbb{Z}$. Write $a \equiv b \pmod{N}$, and say a is congruent to b modulo N, if $N \mid (b - a)$.*

Theorem 2.2. *For any N, congruence modulo N is an equivalence relation on $\mathbb{Z}$.*

Proof. Exercise 2.5 (or given in class if time permits). □

Thus $\equiv \pmod{N}$ defines a partition of $\mathbb{Z}$ into *congruence classes* modulo N. How many? Well, if $0 \leq a < b \leq N - 1$, then $0 < b - a \leq N - 1$ and so N can't divide $b - a$. So $[0], [1], \ldots, [N - 1]$ are all in distinct congruence classes. But $[0] = [N]$ and indeed any integer that is not in $[0, N - 1]$ is congruent to an integer in $[0, N - 1]$. So there are N congruence classes modulo N. The set of congruence classes is written $\mathbb{Z}_N$ (by topologists) or $\mathbb{Z}/N\mathbb{Z}$ (by number theorists).

Telling time. Telling time relative to a 24-hour system amounts to addition modulo 24. For the moment, assume we use a 24-hour system as in European trains: so 5 PM is 17:00. You take a train at 17:00 and you arrive 14 hours later. What time is it? The answer:

$$17 + 14 = 31 \equiv 7 \pmod{24}$$

so you arrive at 7:00, or 7 AM. Here we have used the following fundamental fact about modular arithmetic:

Proposition 2.3. *Let $a, b, c, d \in \mathbb{Z}$. Suppose $a \equiv c \pmod{N}$ and $b \equiv d \pmod{N}$. Then*

$$a + b \equiv c + d \pmod{N}$$

and

$$ac \equiv bd \pmod{N}.$$

Thus if we define the following operations on $\mathbb{Z}/N\mathbb{Z}$:

$$[a] + [c] = [a + c]; [a][c] = [ac]$$

then the operations are well-defined functions from $\mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z} \to \mathbb{Z}/N\mathbb{Z}$.

The proof is in the book (and given in class). We need to explain what is meant by *well-defined*. Say X is a set and R is an equivalence relation on X, with $X/\sim$ the set of equivalence classes. Let $f : X \to Y$ be a function to some other set Y. We say X is *well-defined* modulo R if, whenever xRx', $f(x) = f(x')$. In the present case, $X = \mathbb{Z} \times \mathbb{Z}$, R is congruence modulo N on both factors, and $f : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}/N\mathbb{Z}$ is addition or multiplication.

Exercise 2.4. *Check that addition and multiplication in $\mathbb{Z}/N\mathbb{Z}$ satisfy the commutative, associative, and distributive laws.*

2.1. **Final digits.** The compatibility of multiplication with congruences can be used to compute the final digits of complicated expressions. If a is an integer, its final digit is the number $i \in [0, 9]$ such that $a \equiv i \pmod{10}$.

Powers of 5. For any $n > 0$, $5^n \equiv 5 \pmod{10}$. How do we know this? It's true for $n = 1$. Suppose we know it for n. Thus $5^n \equiv 5 \pmod{10}$. Now multiply both sides by 5: $5^{n+1} \equiv 5^2 = 25 \equiv 5 \pmod{10}$. This is an argument by *induction*, which we will study more systematically later.

Powers of 3. . We know that $3^2 = 9 \equiv -1 \pmod{10}$. Thus $3^4 \equiv (-1)^2 = 1 \pmod{10}$. So to find the last digit of 3^{103}, we observe that $103 \equiv 3 \pmod 4$, thus $103 = 3 + c\dot{4}$ for some integer c (which happens to be 25 but that doesn't matter).

Now

$$3^{103} = 3^{4c+3} = 3^{4c} \cdot 3^3 = (3^4)^c \cdot 3^3 \equiv 1^c \cdot 27 \equiv 7 \pmod{10}.$$

So the last digit is 7. Note that $7 = (-1) \cdot 3$ and $3^3 = 3^2 \cdot 3 \equiv (-1) \cdot 3 \pmod{10}$.

We can continue this indefinitely. What is the last digit of 3^{6^4}? Well, 6^4 is even, and even powers of 3 have final digit either 1 or -1, the former if the power is divisible by 4. So is 6^4 divisible by 4? Yes, because $6^4 = 2^4 \cdot 3^4$ and 2^4 is divisible by 4. Thus the last digit of 3^{6^4} is $+1$.

3. Propositional logic

3.1. **Forming propositions.** We return to the consideration of *statements* from the first week. (Review the discussion of connectives and truth tables in the first week's notes.) Specifically, we are interested in examining how operations on statements affect truth. To emphasize that propositional logic is a form of calculation with truth, we give truth and falsity numerical values (Boolean logic): 0 means "false" (or F), 1 means "true." These are the only two possibilities; however one can work mathematically with more sophisticated logical systems, where a statement can take intermediate values ("maybe") or the truth of a statement depends on other circumstances ("it's raining").

Recall the four propositional connectives $\neg, \wedge, \vee, \Rightarrow$. The first week's notes worked out the truth tables for $\vee$ and $\Rightarrow$. Here are the definitions in Boolean logic. In what

follows, P and Q are statements, and $T(P)$ denotes the truth value of P.

$$T(\neg P) = 1 - T(P).$$

$$T(P \wedge Q) = T(P)T(Q).$$

$$T(P \vee Q) = T(P) + T(Q) - T(P)T(Q).$$

$$T(P \Rightarrow Q) = 1 - T(P) + T(P)T(Q).$$

The verification that $T(P \Rightarrow Q) = T(Q \vee \neg P)$ is simpler using these formulas.

$$T(Q \vee \neg P) = T(Q) + T(\neg P) - T(Q)T(\neg P)$$
$$= T(Q) + 1 - T(P) - T(Q)(1 - T(P)) = 1 - T(P) + T(Q)T(P).$$

In the implication $P \Rightarrow Q$, one calls P the *hypothesis* and Q the *conclusion*; or more formally, P is the *antecedent* and Q is the *consequence*. In logic as well as in mathematics, we need to be able to work with statements of this form without assuming their truth. Thus we can write

$$[P \Rightarrow Q] \Rightarrow [\neg R \vee S].$$

This makes sense if P is "It will snow today," Q means "Trains are delayed," R means "I am coming to work" and S means "I will listen to the radio.": A literal translation would be *If if it will snow today then trains are delayed, then either I am not coming to work or I will listen to the radio (or both)."*

Normal English doesn't know how to read a sentence that starts with "If if," so one has to adapt the structure. A translation into normal speech might be

If snow today will cause a delay in the trains, I will either not come to work or I will listen to the radio (or both)."

Note the placement of the brackets, which, as with simple arithmetic, indicates the order of the operations. If we move the brackets we get

$$P \Rightarrow [Q \Rightarrow \neg R] \vee S$$

which means *If it snows today then either if the trains are delayed I will not come to work, or I will listen to the radio (or both)."*

Don't worry about the specific choices for adaptation into natural language. It is a useful exercise to rewrite sentences in natural language using propositional connectives; normally there should be no ambiguity when translating in that direction. First, some vocabulary:

- An *atomic statement* is a statement with no visible propositional connectives.
- A *well-formed statement* is a statement obtained

  (a) from a collection of atomic statements using propositional connectives according to the natural rules:

$$\neg P$$

$$P \vee Q$$

$$P \wedge Q$$
$$P \Rightarrow Q;$$

(b) more generally, by the same rules applied recursively to well-formed statements.

- A *compound statement* is a well-formed statement composed of atomic statements and propositional connectives.

In other words, one has a vocabulary consisting of atomic statements and the symbols $\neg, \wedge, \vee, \Rightarrow$, and one forms statements consistently with the rules above, but not every string of symbols is well formed. For example, $\neg \wedge P$ is not a well-formed statement.

Examples.

1. *If X is elected then I will move to Australia.*
   becomes
   $$P \Rightarrow Q$$
   where P is "X is elected" and Q is "I will move to Australia."
2. *Zika causes microcephaly in mice.* (The title of an article in *Science*, May 11, 2016.)
   could become
   $$[(x \in fM) \wedge (x \in Z) \Rightarrow (x \in m)$$
   where fM is the set of fetal mice, Z is the set of animals infected by Zika, and m is the set of animals with microcephaly. This can in turn be written more abstractly as
   $$[P \wedge Q] \Rightarrow R$$
   where P is "x is a fetal mouse", Q is "x is infected by Zika" and R is "x has microcephaly". Of course the article only claims this is true with a certain degree of probability and is not inevitable. Logic can also accomodate probabilistic statements, but we won't see this in the course.

We will see more examples after introducing quantifiers.

3.2. **Propositional equivalence.**

Definition 3.1. *Two well-formed statements are* **propositionally equivalent** *if their truth values are equal for any choice of truth values of their constituent atomic statements.*

Compare: two functions f and g from $\mathbb{R}^n$ to $\mathbb{R}$ are equal if $f(x_1, x_2, \ldots, x_n) = g(x_1, x_2, \ldots, x_n)$ for all $x_1, x_2, \ldots, x_n \in \mathbb{R}^n$. Here the atomic statements are treated as variables and you are allowed to substitute 1 (for true) or 0 (for false) to get a truth value according to the rules given above.

Examples in the text: **De Morgan's Laws**:

$$\neg[P \vee Q] = [\neg P \wedge \neg Q];\ \neg[P \wedge Q] = [\neg P \vee \neg Q]$$

3.3. **Converse, contrapositive, inverse.** Syntax allows us to take an implication and turn it around in several ways. Consider the typical implication *If P then Q*, or equivalently $P \Rightarrow Q$. Here are the three natural ways to transform the sentence.

a. (Original statement) $P \Rightarrow Q$ (*If P then Q*)
b. (Converse) $Q \Rightarrow P$ (*If Q then P*)
c. (Inverse) $\neg P \Rightarrow \neg Q$ (*If not P then not Q*)
d. (Contrapositive) $\neg Q \Rightarrow \neg P$ (*If not Q then not P*)

So if P is *The month is October* and Q is *The season is fall* then the original sentence is the true statement

If the month is October then the season is fall
the converse is
If the season is fall then the month is October
which is not true; the inverse is
If the month is not October then the season is not fall
also not true; and the contrapositive is
If the season is not fall then the month is not October,
which is again true. We have seen a special case of

Proposition 3.2. *(a) $P \Rightarrow Q$ and $\neg Q \Rightarrow \neg P$ are propositionally equivalent.*
(b) $Q \Rightarrow P$ and $\neg P \Rightarrow \neg Q$ are propositionally equivalent.

This can be proved by truth tables or by computation. Obviously (b) is just (a) with P and Q switched, but the meaning relative to the original statement is different. Point (a) says that the contrapositive is true if and only if the original statement is true. Often it is more natural to prove the contrapositive. Here we have to admit the *Law of the Excluded Middle* and the *Law of Non-contradiction*. Suppose we want to prove

$$P \Rightarrow Q \tag{1}$$

and we know

$$\neg Q \Rightarrow \neg P \tag{2}$$

but we have never heard of truth tables. How to conclude? Well, assume P. Suppose Q is not true; then necessarily $\neg Q$ is true – because of the *Law of the Excluded Middle*. But then (2) tells us that $\neg P$ is true. So we have both P and $\neg P$. But by the *Law of Non-contradiction* P and $\neg P$ can't both be true. We have assumed P, so $\neg P$ is not true. Thus by (2) $\neg Q$ is not true, and by the *Law of the Excluded Middle*, Q is true.

If you find this reasoning suspect, you may prefer intuitionism, which does not assume the *Law of the Excluded Middle* and the *Law of Non-contradiction*.

Example of proof by contrapositive.

Theorem 3.3 (Hippasus of Metapontum?)**.** *The square root of* 2 *is irrational.*

The proof appears in Book X of Euclid's *Elements*, Proposition 117. We want to write this in the form $P \Rightarrow Q$. One way is to take P to be the atomic statement $x^2 = 2$ and Q to be $x \notin \mathbb{Q}$. Then the contrapositive is $\neg Q \Rightarrow \neg P$. Now $\neg Q$ is the sentence "It is not true that $x \notin \mathbb{Q}$" or more simply $x \in \mathbb{Q}$. We must then prove that $x^2 \neq 2$. So suppose $x = \frac{a}{b}$ with $a, b \in \mathbb{Z}$. One then concludes the proof as in Example 3.23 of (DM). This proof requires a discussion of formulas and quantifiers, which are the next topics.

3.4. **Formulas and sets.** The book (DM) uses the word "universe" without explanation. In practice, it means a set that contains all the other sets under discussion. We have seen (Russell's paradox) that there is no set that contains all sets, but when we speak of a universe we will mean a set for which we want to describe subsets. For example, we are used to describing a curve in $\mathbb{R}^2$ by a *formula*: $x^2 - y^2 = 1$ describes a hyperbola in $\mathbb{R}^2$.

Logical formulas do not need to be arithmetic. Let $U_1, \ldots, U_n$ be universes (i.e., sets). A *formula* $P(x_1, \ldots, x_n)$ is an expression that says something about the elements $x_i \in U_i$. For example, if $U_1 = U_2 = \mathbb{R}$, $P(x, y)$ is the formula $x^2 - y^2 = 1$. (The book (DM) uses a circle rather than a hyperbola). Now this is not true for all $(x, y) \in \mathbb{R}^2$. The formula $P(x, y)$ is then true for certain pairs (x, y) and false for others.

One could just as easily take $U = \mathbb{R}^2$ and write $P(v)$ where $v = (x, y) \in \mathbb{R}^2$. This is more convenient in most situations.

Definition 3.4. *The characteristic set χ_P of the formula $P(x)$ where U is the universe of the variable x, is the subset of $x \in U$ for which $P(x)$ is true; alternatively,*

$$\chi_P = \{x \in U \mid T(P(x)) = 1\}.$$

Thus $U = \chi_P \cup \chi_{\neg P}$ and $\chi_P \cap \chi_{\neg P} = \emptyset$; this is thus a partition of U.

Suppose P and Q are two formulas in x where x is a variable in U. Then

$$\chi_{P \wedge Q} = \chi_P \cap \chi_Q, \ \chi_{P \vee Q} = \chi_P \cup \chi_Q.$$

Exercise 3.5. *Prove that $T(P \Rightarrow Q) = 1$ if and only if $\chi_P \subset \chi_Q$.*

In the last exercise, we are saying that $T(P \Rightarrow Q) = 1$ when $P(x) \Rightarrow Q(x)$ for all x. This leads us to the use of

3.5. **Quantifiers.** To make useful sentences in set theory we want to work with assertions that are true for all elements of a set, or for some elements of a set (without necessarily specifying them). For example, if U is $\mathbb{C}$, the statement

$$a + b = b + a$$

is true for all $a, b \in \mathbb{C}$. On the other hand, if f is a polynomial of degree > 0 with coefficients in $\mathbb{C}$, then the equation

$$f(x) = 0$$

has a solution with $x \in \mathbb{C}$; this is the *Fundamental Theorem of Algebra*, but the proof doesn't give you a rule to find x.

Definition 3.6. *Let X be a set (in a universe U). Let $P(x)$ be a formula with universe U. The* universal quantifier $\forall$ *is used to assert truth of P for all $x \in X$:*

$$(\forall x \in X)P(x)$$

which means

$$(\forall x)([x \in X] \Rightarrow [P(x)]).$$

For example, if $U = \mathbb{C}$, $X = \mathbb{C} \setminus \{0\}$, and $P(x)$ is the formula $\frac{x}{x} = 1$, then the statement

$$(\forall x \in X)P(x)$$

is true. In the version

$$(\forall x)([x \in X] \Rightarrow [P(x)]),$$

the when x is first mentioned we are just saying $x \in \mathbb{C}$; but then $P(x)$ is not true unless $x \neq 0$.

Exercise 3.7. *Show that*

$$(\forall x \in X)P(x) \Leftrightarrow X \subset \chi_P.$$

Definition 3.8. *Let X be a set (in a universe U). Let $P(x)$ be a formula with universe U. The* existential quantifier $\exists$ *is used to assert truth of P for some (not necessariily specified) $x \in X$:*

$$(\exists x \in X)P(x)$$

which means

$$(\exists x)([x \in X] \Rightarrow [P(x)]).$$

For example, if U is the set of things in the universe (the actual universe) and X is the surface of the earth, the sentence "What goes up must come down" can be written as a formula with an existential quantifier: $\exists x \in X$ where the thing that goes up comes down. Before we turn to more familiar sentences in calculus, a few properties of quantifiers:

3.5.1. *Multiple quantifiers.* Suppose $U_i, i = 1, \ldots, n$ are universes, $X_i \subset U_i$ subsets, x_i a variable in U_i, $P(x_1, \ldots, x_n)$ a formula. We can write

$$(\forall x_1 \in X_1)(\exists x_2 \in X_2)P(x_1, x_2)$$

For example, if $X_1 = \mathbb{R}$ and $U_2 = X_2 = \mathbb{C}$, and $P(x_1, x_2)$ is $x_2^2 = x_1$, this says that every real number has a square root that is also a complex number. It's not true if $X_2 = \mathbb{R}$.

We can also write statements like

$$(\exists x_2 \in X_2)P(x_1, x_2)$$

which is a formula in the variable x_1. Thus in the square root example, with now $U_1 = X_2 = \mathbb{R}$, the characteristic set of this formula is the set of x_1 with real square root, in other words, the set $[0, \infty)$.

In this example, x_2 is a *bound variable*, x_1 is an *open* (or *unbound*) variable.

Warning! The sentences

$$(\forall x_1 \in X_1)(\exists x_2 \in X_2)P(x_1, x_2)$$

and

$$(\exists x_2 \in X_2)(\forall x_1 \in X_1)P(x_1, x_2)$$

are not equivalent! There are examples in the book and in the homework.

3.5.2. *Negation of quantifiers.* The negation of the sentence

$$(\forall x \in X)P(x),$$

i.e. $\not(\forall x \in X)P(x)]$.

$$(\exists x \in X)\neg P(x).$$

Similarly,

$$\neg[(\exists x \in X)P(x)]$$

is equivalent to

$$(\forall x \in X)\neg P(x)].$$

So loosely, $\neg$ exchanges $\forall$ and $\exists$.