Logic and Set Theory

January 29, 2018

Contents

0 Miscellaneous

Some introductory speech

1 Propositional logic

Let P denote a set of *primitive proposition*, unless otherwise stated, $P = \{p_1, p_2, ...\}$.

Definition. The *language* or *set of propositions* $L = L(P)$ is defined inductively by:
(1) $p \in L\ \forall p \in P$;
(2) $\perp \in L$, where $\perp$ is read as 'false';
(3) If $p, q \in L$, then $(p \implies q) \in L$. For example, $(p_1 \implies L)$, $((p_1 \implies p_2) \implies (p_1 \implies p_3))$.

Note that at this point, each proposition is only a finite string of symbols from the alphabet $(,), \implies, \perp, p_1, p_2, ...$ and do not really mean anything (until we define so).
By *inductively define*, we mean more precisely that we set $L_1 = P \cup \{\perp\}$, and $L_{n+1} = L_n \cup \{(p \implies q) : p, q \in L_n\}$, and then put $L = L_1 \cup L_2 \cup ....$

Each proposition is built up *uniquely* from 1) and 2) using 3). For example, $((p_1 \implies p_2) \implies (p_1 \implies p_3))$ came from $(p_1 \implies p_2)$ and $(p_1 \implies p_3)$. We often omit outer brackets or use different brackets for clarity.

Now we can define some useful things:
• $\neg p$ (not p), as an abbreviation for $p \implies \perp$;
• $p \vee q$ (p or q), as an abbreviation for $(\neg p) \implies q$;
• $p \wedge q$ (p and q), as an abbreviation for $\neg(p \implies (\neg q))$.

These definitions 'make sense' in the way that we expect them to.

Definition. A *valuation* is a function $v : L \to \{0, 1\}$ s.t.
(1) $v(\perp) = 0$; (2)

$$v(p \implies q) = \begin{cases} 0 & v(p) = 1, v(q) = 0 \\ 1 & else \end{cases} \quad \forall p, q \in L$$

Remark. On $\{0, 1\}$, we could define a constant $\perp$ by $\perp = 0$, and an operation $\implies$ by $a \implies b = 0$ if $a = 1, b = 0$ and 1 otherwise. Then a valuation is a function $L \to \{0, 1\}$ that preserves the structure ($\perp$ and $\implies$), i.e. a homomorphism.

Proposition. (1) If v, v' are valuations with $v(p) = v'(p)\ \forall p \in P$, then $v = v'$ (on L).
(2) For any $w : P \to \{0, 1\}$, there exists a valuation v with $v(p) = w(p)\ \forall p \in P$.
In short, a valuation is defined by its value on p, and any values will do.

Proof. (1) We have $v(p) = v'(p)\ \forall p \in L_1$. However, if $v(p) = v'(p)$ and $v(q) = v'(q)$ then $v(p \implies q) = v'(p \implies q)$, so $v = v'$ on L_2. Continue inductively we have $v = v'$ on $L_n \forall n$.
(2) Set $v(p) = w(p)\ \forall p \in P$ and $v(\perp) = 0$: this defines v on L_1. Having defined v on L_n, use the rules for valuation to inductively define v on L_{n+1} so we can extend v to L. □

Definition. We say p is a *tautology*, written $\vDash p$, if $v(p) = 1\ \forall$ valuations v.
Some examples:
(1) $p \implies (q \implies p)$: a true statement is implies by anything. We can verify this by:

$v(p)$	$v(q)$	$v(q \implies p)$	$v(p \implies (q \implies p))$
1	1	1	1
1	0	1	1
0	1	0	1
0	0	1	1

So we see that this is indeed a tautology;
(2) $(\neg\neg p) \implies p$, i.e. $((p \implies \perp) \implies \perp) \implies p$, called the "law of excluded middle";

(3) $[p \implies (q \implies r)] \implies [(p \implies q) \implies (p \implies r)]$.
Indeed, if not then we have some v with $v(p \implies (q \implies r)) = 1$, $v(\implies (p \implies q) \implies (p \implies r)) = 0$. So $v(p \implies q) = 1$, $v(p \implies r) = 0$. This happens when $v(p) = 1$, $v(r) = 0$, so also $v(q) = 1$. But then $v(q \implies r) = 0$, so $v(p \implies (q \implies r)) = 0$.

Definition. For $S \subset L$, $t \in L$, say S *entails* or *semantically implies* t, written $S \vDash t$ if $v(s) = 1 \forall s \in S \implies v(t) = 1$, for each valuation v.
("Whenever all of S is true, t is true as well.")

For example, $\{p \implies q, q \implies r\} \vDash (p \implies r)$. To prove this, suppose not: so we have v with $v(p \implies q) = v(q \implies r) = 1$ but $v(p \implies r) = 0$. So $v(p) = 1$, $v(r) = 0$, so $v(q) = 0$, but then $v(p \implies q) = 0$.

If $v(t) = 1$ we say t is true in v or that v is a model of t.

For $S \subset L$, v is a model of S if $v(s) = 1\ \forall s \in S$. So $S \vDash t$ says that every model of S is a model of t. For example, in fact $\vDash t$ is the same as $\phi \vDash t$.

2 Syntactic implication

For a notion of 'proof', we will need axioms and deduction rules. As axioms, we'll take:
1. $p \implies (q \implies p)\ \forall p, q \in L$;
2. $[p \implies (q \implies r)] \implies [(p \implies q) \implies (p \implies r)]\ \forall p, q, r \in L$;
3. $(\neg\neg p) \implies p\ \forall p \in L$.

Note: these are all tautologies. Sometimes we say they are 3 axiom-schemes, as all of these are infinite sets of axioms.

As deduction rules, we'll take just *modus ponens*: from p, and $p \implies q$, we can deduce q.

For $S \subset L$, $t \in L$, a *proof* of t from S cosists of a finite sequence $t_1, ..., t_n$ of propositions, with $t_n = t$, s.t. $\forall i$ the proposition t_i is an axiom, or a member of S, or there exists $j, k < i$ with $t_j = (t_k \implies t_i)$.

We say S is the *hypotheses* or *premises* and t is the *conclusion*.

If there exists a proof of t from S, we say S *proves* or *syntactically implies* t, written $S \vdash t$.

If $\phi \vdash t$, we say t is a *theorem*, written $\vdash t$.

Example. $\{p \implies q, q \implies r\} \vdash p \implies r$.
we deduce by the following:
(1) $[p \implies (q \implies r)] \implies [(p \implies q) \implies (p \implies r)]$; (axiom 2)
(2) $q \implies r$; (hypothesis)
(3) $(q \implies r) \implies (p \implies (q \implies r))$; (axiom 1)
(4) $p \implies (q \implies r)$; (mp on 2,3)
(5) $(p \implies q) \implies (p \implies r)$ (mp on 1,4);
(6) $p \implies q$; (hypothesis)
(7) $p \implies r$. (mp on 5,6)

Example. Let's now try to prove $\vdash p \implies p$. Axiom 1 and 3 probably don't help so look at axiom 2; if we make $(p \implies q)$ and $p \implies (q \implies r)$ something that's a theorem, and make $p \implies r$ to be $p \implies p$ then we are done. So we need to take $p = p, q = (p \implies p), r = p$. Now:
(1) $[p \implies ((p \implies p) \implies p)] \implies [(p \implies (p \implies p)) \implies (p \implies p)]$; (axiom 2)
(2) $p \implies ((p \implies p) \implies p)$; (axiom 1)
(3) $(p \implies (p \implies p)) \implies (p \implies p)$; (mp on 1,2)
(4) $p \implies (p \implies p)$; (axiom 1)
(5) $p \implies p$. (mp on 3,4)

Proofs are made easier by:

Proposition. (2, deduction theorem)
Let $S \subset L$, $p, q \in L$. Then $S \vdash (p \implies q)$ if and only if $(S \cup \{p\}) \vdash q$.

Proof. Forward: given a proof of $p \implies q$ from S, add the lines p (hypothesis), q (mp) to optaion a proof of q from $S \cup \{p\}$.
Backward: if we have proof $t_1, ..., t_n = q$ of q from $S \cup \{p\}$. We'll show that $S \vdash (p \implies t_i) \forall i$, so $p \implies t_n = q$.
If t_i is an axiom, then we have $\vdash t_i \implies (p \implies t_i)$, so $\vdash p \implies t_i$;
If $t_i \in S$, write down t_i, $t_i \implies (p \implies t_i)$, $p \implies t_i$ we get a proof of $p \implies t_i$ from S;
If $t_i = p$: we know $\vdash (p \implies p)$, so done;
If t_i obtained by mp: in that case we have some earlier lines t_j and $t_j \implies t_i$. By induction, we may assume $S \vdash (p \implies t_j)$ and $S \vdash (p \implies (t_j \implies t_i))$. Now we can write down $[p \implies (t_j \implies t_i)] \implies [(p \implies t_j) \implies (t_i)]$ by axiom 2, $p \implies (t_j \implies t_i)$, $p \implies t_j) \implies (p \implies t_i)$ (mp), $p \implies t_j$, $p \implies t_i$ (mp) to obtain $S \vdash (p \implies t_i)$.

These are all of the cases. So $S \vdash (p \implies q)$. □

This is why we chose axiom 2 as we did – to make this proof work.

Example. To show $\{p \implies q, q \implies r\} \vdash (p \implies r)$, it's enough to show that $\{p \implies q, q \implies r, p\} \vdash r$, which is trivial by mp.

Now, how are $\vdash$ and $\vDash$ related? We are going to prove the *completeness theorem*: $S \vdash t \iff S \vDash t$.

This ensures that our proofs are sound, in the sense that everything it can prove is not absurd ($S \vdash t$ then $S \vDash t$), and are adequate, i.e. our axioms are powerful enough to define every semantic consequence of S, which is not obvious ($S \vDash t$ then $S \vdash t$).

Proposition. (3)
Let $S \subset L$, $t \in L$. Then $S \vdash t \implies S \vDash t$.

Proof. Given a valuation v with $v(s) = 1 \ \forall s \in S$, we want $v(t) = 1$.
We have $v(p) = 1 \ \forall p$ axiom as our axioms are all tautologies (proven earier); $v(p) = 1 \ \forall p \in S$ by definition of v; also if $v(p) = 1$ and $v(p \implies q) = 1$, then also $v(q) = 1$ (by definition of $\implies$). So $v(p) = 1$ for each line p of our proof of t from S. □

We say $S \subset L$ consistent if $S \nvdash \perp$. One special case of adequacy is: $S \vDash \perp \implies S \vdash \perp$, i.e. if S has no model then S inconsistent, i.e. if S is consistent then S has a model. This implies adequacy: given $S \vDash t$, we have $S \cup \{\neg t\} \vDash \perp$, so by our special case we have $S \cup \{\neg t\} \vdash \perp$, i.e. $S \vdash ((\neg t) \implies t)$ by deduction theorem, so $S \vdash \neg\neg t$. But $S \vdash ((\neg\neg t) \implies t)$ by axiom 3, so $S \vdash t$ (mp).

Theorem. (4)
Let $S \subset L$ be consistent, then S has a model.
The idea is that we would like to define valuation v by $v(p) = 1 \iff p \in S$, or more sensibly, $v(p) = 1 \iff S \vdash p$.
But maybe $S \nvdash p_3, S \nvdash \neg p_3$, but a valuation maps half of L to 1, so we want to 'grow' S to contain one of p or $\neg p$ for each $p \in L$, while keeping consistency.

Proof. Claim: for any consistent $S \subset L$, $p \in L$, $S \cup \{p\}$ or $S \cup \{\neg p\}$ consistent. *Proof of claim.* If not, then $S \cup \{p\} \vdash \perp$ and $S \cup \{\neg p\} \vdash \perp$, then $S \vdash (p \implies \perp)$ (deduction theorem), i.e. $S \vdash \not p$, so $S \vdash \perp$ contradiction.

Now L is countable as each L_n is countable, so we can list L as $t_1, t_2, ...$ Put $S_0 = S$; set $S_1 = s_0 \cup \{t_1\}$ or $s_0 \cup (\neg t_1\}$ so that S_1 is consistent. Then set $S_2 = S_1 \cup \{t_2\}$ or $S_1 \cup \{\neg t_2\}$ so that S_2 is consistent, and continue likewise. Set $\bar{S} = S_0 \cup S_1 \cup S_2 \cup ...$ Then $\bar{S} \supset S$, and $\bar{S}$ is consistent (as each S_n is, and each proof is finite). $\forall p \in L$, we have either $p \in S$ or $(\neg p) \in S$. Also, $\bar{S}$ is *deductively closed*, meaning that is $\bar{S} \vdash p$ then $p \in \bar{S}$: if $p \notin \bar{S}$ then $(\neg p) \in \bar{S}$, so $\bar{S} \vdash p$, $\bar{S} \vdash (\not p)$ so $\bar{S} \vdash \perp$ contradiction.
Define $v : L \to \{0, 1\}$ by $p \to 1$ if $p \in \bar{S}$, 0 otherwise. Then v is a valuation: $v(\perp) = 0$ as $\perp \notin \bar{S}$; for $v(p \implies q)$:
If $v(p) = 1$, $v(q) = 0$: We have $p \in \bar{S}$, $q \notin \bar{S}$, and want $v(p \implies q) = 0$, i.e. $(p \implies q \notin \bar{S}$. But if $9p \implies q) \in \bar{S}$ then $\bar{S} \vdash q$ contradiction;
If $v(q) = 1$: have $q \in \bar{S}$, and want $v(p \implies q) = 1$, i.e. $(p \implies q) \int \bar{S}$. But $\vdash q \implies (p \implies q)$ so $\bar{S} \vdash (p \implies q)$;
If $v(p) = 0$: have $p \notin \bar{S}$, i.e. $(\neg p) \in \bar{S}$ and want $(p \implies q) \in \bar{S}$. So we need $(p \implies \perp) \vdash (p \implies q)$, i.e. $p \implies \perp, p \vdash q$ (deduction theorem). Thus it's enough to show that $\perp \vdash q$. But $(\neg\neg q) \implies q$, and $\vdash (\perp \implies (\neg\neg q))$ (axiom 3 and 1 – to see the second one, write $\neg$ explicitly using $\implies$ and $\perp$), so $\vdash (\perp \implies q)$, i.e. $\perp \vdash q$. □

Remark. Sometimes this is called 'completeness theorem'. The proof used P being countable to get L countable; in fact, result still holds if P is uncountable (see chapter 3).

By remark before theorem 4, we have

Corollary. (5, adequacy)
Let $S \subset L$, $t \in L$. Then if $S \vDash t$ then $S \vdash t$.

And hence,

Theorem. (6, completeness theorem)
Let $S \subset L$, $t \in L$. Then $S \vdash t \iff S \vDash t$.

Some consequences:

Corollary. (7, compactness theorem)
Let $S \subset L$, $t \in L$ with $S \vDash t$. Then $\exists$ finite $S' \subset S$ with $S' \vDash t$.
This is trivial if we replace $\vDash$ by $\vdash$ (as proofs are finite).

Special case for $t = \perp$: If S has no model then some finite $S' \subset S$ has no model. Equivalently,

Corollary. (7', compactness theorem, equivalent form)
Let $S \subset L$. If every finite subset of S has a model then S has a model.
This *isi* equivalent to corollary 7 because $S \vDash t \iff S \cup \{\neg t\}$ has no model and $S' \vDash t \iff S' \cup (\neg t)$ has no model.

Corollary. (8, decidability theorem)
There is an algorithm to determine (in finite time) whether or not, for a given finite $S \subset L$ and $t \in L$, we have $S \vdash t$.
This is highly non-obviuos; however it's trivial to decide if $S \vDash t$ just by drawing a truth table, and $\vDash \iff \vdash$.

3 Well-Orderings and Ordinals

Definition. A *total order* or *linear order* on a set X is a relation $<$ on X, such that
(1) Irreflexive: Not $x < x \ \forall x \in X$;
(2) Transitive: $x < y, y < z \implies x < z \ \forall x, y, z \in X$;
(3) Trichotomous: $x < y$ or $x = y$ or $y < x \ \forall x, y \in X$.
Note: two of (iii) cannot hold: if $x < y$, $y < x$ then $x < x$ by transitivity.
Write $x \leq y$ if $x < y$ or $x = y$, and $y > x$ if $x < y$.

We can also define total order in terms of $\leq$:
(1) Reflexive: $x \leq x \ \forall x \in X$;
(2) Transitive: $x \leq y, y \leq z \implies x \in z \ \forall x, y, z \in X$;
(3) Antisymmetric: $x \leq y, y \leq x \implies x = y \ \forall x, y \in X$;
(4) 'Tri'chotomous (although it's only two): $x \leq y$ or $y \leq x \ \forall x, y \in X$.

Example. $\mathbb{N}, \mathbb{Q}, \mathbb{R}$ with the usual orders are all total orders.
$\mathbb{N}^+$ the relation 'divides' is not a total order: for example we don't have any of $2|3, 3|2$ or $2 = 3$.
$\mathcal{P}(S)$ for some S (with $|S| \geq 2$ to be rigorous), with $x \leq y$ if $x \subseteq y$ is not a total order for the same reason.

A total order is a *well-ordering* if every (non-empty) subset has a least element, i.e. $\forall S \subset X, S \neq \phi \implies \exists x \in S, x \leq y \forall y \in S$.

Example. 1.$\mathbb{N}$ with the usual $<$ is a well ordering.
2.$\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ with the usual $<$ are not well orderings.
3.$\mathbb{Q}^+ \cup \{0\}$ with the usual $<$ is not a well ordering (e.g. $(0, \infty) \subset \mathbb{Q}^+ \cup \{0\}$).
4.The set $\{1 - \frac{1}{n} : n = 2, 3, ...\}$ as a subset of $\mathbb{R}$ with the usual ordering is a well ordering. 5.The set $\{1 - \frac{1}{n} : n = 2, 3, ...\} \cup \{1\}$ as a subset of $\mathbb{R}$ with the usual ordering is a well ordering. 6.The set $\{1 - \frac{1}{n} : n = 2, 3, ...\} \cup \{2 - \frac{1}{n} : n = 2, 3, ...\}$ (same assumption) is a well ordering.

Remark. X is well-ordered iff there is no $x_1 > x_2 > x_3 > ...$ in X.
Clearly if there is such a sequence then $S = \{x_1, x_2, ...\}$ has no least element.
Conversely, if $S \subset X$ has no least element, then for each element $x \in S$ there exists a $x' \in S$ with $x' < x$, so we can just pick $x, x', ...$ inductively.

Definition. We say total orders X, Y are *isomorphic* if there exists a bijection $f : X \to Y$ that is order-preserving, i.e. $x < y \iff f(x) < f(y)$.
For example, 1 and 4 above are isomorphic; 5 and 6 are isomorphic; 4 and 5 are not isomorphic (one has a greatest element, and the other doesn't).

Here comes the first reason why well orderings are useful:

Proposition. (1, Proof by induction)
Let X be well-ordered, and let $S \subset X$ be s.t. if $y \in S \ \forall y < x$ then $x \in S$ (each $x \in X$). Then $S = X$.
Equivalently, if $p(x)$ is a property s.t. $\forall x$: if $p(y) \forall y < x$ then $p(x)$, then $p(x) \forall x$.

Proof. If $S \neq X$ then let x be the least element of $X \setminus S$. Then $x \notin S$. But $y \in S\ \forall y < x$, contradiction. $\square$

A typical use:

Proposition. Let X, Y be isomorphic well-orderings. Then there is a *unique* isomorphism from X to Y.

Proof. Let f, g be isomorphisms. We'll show $f(x) = g(x)\ \forall x$ by induction. Thus we may assume $f(y) = g(y)\ \forall y < x$, and want $f(x) = g(x)$. Let a be the least element of $Y \setminus \{f(y) : y < x\}$. Then we must have $f(x) = a$: if $f(x) > a$, then some $x' > x$ has $f(x') = a$ by surjectivity, contradiction. The same shows $g(x)$ =least element of $Y \setminus \{g(y) : y < x\}$, but this is the same as a. So $f(x) = g(x)$. $\square$

Remark. This is false for total orders in general. One example is, consider from $\mathbb{Z} to \mathbb{Z}$, we could either take identity, or $x \to x - 5$; or from $\mathbb{R}$ to $\mathbb{R}$ we could take identity or $x \to x - 5$ or $x \to x^3$...

Definition. In a total order X, an *initial segment* I is a subset of X such that $x \in I, y < x \implies y \in I$.

Example. For any $x \in X$, set $I(x) = \{y \in X : y < x\}$. Then this is an initial segment.
Obviously, not every initial segment is of this form: for example, in $\mathbb{R}$ we can take $\{x : x \leq 3\}$; or in $\mathbb{Q}$, take $\{x : x^2 < 2\} \cup \{x < 0\}$ (this cannot be written as above form as $\sqrt{2} \notin \mathbb{Q}$.
Note: in a well-ordering, every proper initial segment *is* of the above form: let x be the least elemnt of $X \setminus I$. Then $y < x \implies y \in I$. Conversely, if $y \in I$, then we must have $y < x$: otherwise $x \in I$, contradiction.