Sogic and Let Theory

January 27, 2020

1 Propositional Logic

Let P be a set of ***primitive propositions***, i.e. P is a set of symbols with $(,), \bot, \implies \notin P$. Unless stated otherwise (i.e. that P is uncountable), we may assume that $P = \{p_1, p_2, \ldots\}$.

The set of ***propositions***, denoted by $L(P)$ or simply just L, is defined inductively as follows:

1. $P \subset L$
2. $\bot \in L$, called FALSE
3. if $p, q \in L$, then $(p \implies q) \in L$

Each proposition is a string of symbols from $P \cup \{(,), \bot, \implies\}$, for instance we have the propositions $p_1, (p_1 \implies p_1), ((p_1 \implies p_2) \implies (p_2 \implies (\bot \implies p_3)))$. For readability, we often draw symbols $(,)$ in different ways, for instance as $[, (, \big($.

Sometimes we omit the outside pair of parentheses when writing down propositions, for instance $p_1 \implies p_2$ is shorthand for $(p_1 \implies p_2)$.

Also we use some abbreviations, e.g.:

NOT: $\neg p$ to mean $(p \implies \bot)$

OR: $p \vee q$ to mean $(\neg p \implies q)$

AND: $p \wedge q$ to mean $\neg(\neg p \vee \neg q)$

What do we mean by L "defined inductively"? Define $L_0 = P \cup \{\bot\}$. Then, given L_n, we can define $L_{n+1} = L_{n-1} \cup \{(p \implies q) : p, q \in L_{n-1}\}$. Then we set $L = \bigcup_{n=0}^{\infty} L_n$. Note: if $p \in L \setminus (P \cup \{\bot\})$, then it is easy to show that there are ***unique*** $q, r \in L$ with $p = (q \implies r)$.

1.1 Semantic Entailment

A ***valuation*** is a function $v : L \to \{0, 1\}$ satisfying:

1. $v(\bot) = 0$
2. For all $p, q \in L, v(p \implies q) = \begin{cases} 0 & v(p) = 1, v(q) = 0 \\ 1 & \text{otherwise} \end{cases}$.

If $p \in L$ and $v(p) = 1$ for every valuation, we say that p is a ***tautology***, and write $\vDash p$.

Examples:

1. $\vDash (p \implies p)$

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

So this is a tautology.

2. $\vDash (p \implies (q \implies p))$

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

So this is a tautology.

3. Is $\vDash (p \implies (q \implies r)) \implies ((p \implies q) \implies (p \implies r))$?

Suppose not. Then for some p, q, r and valuation v we have:

$$v(p \implies (q \implies r)) = 1$$
$$v((p \implies q) \implies (p \implies r)) = 0.$$

So $v(p \implies q) = 1, v(p \implies r) = 0$. Hence $v(p) = 1, v(r) = 0, v(q) = 1$. But then $v(q \implies r) = 0$, and so $v(p \implies (q \implies r)) = 0 \lightning$.

4. $\vDash ((p \implies \bot) \implies \bot) \implies p$, i.e. $\neg\neg p \implies p$, i.e. $(\neg p \vee p)$. This is the Law of the Excluded Middle, and is also a tautology.

Note that a valuation is entirely determined by its values on the primitive propositions.

Proposition 1.1.

1. *Let $v, w : L \to \{0, 1\}$ be valuations with $v|_P = w|_P$. Then $v = w$.*
2. *Let $f : P \to \{0, 1\}$. Then there is a valuation $v : L \to \{0, 1\}$ with $v|_P = f$.*

Proof.

1. We prove this by induction on n, so that $v|_{L_n} = w|_{L_n}$. For the base case of $n = 0$, $v|_P = w|_P$, and $v(\bot) = 0 = w(\bot)$. Then for the induction step, $v|_{L_{n-1}} = w|_{L_{n-1}}$. Let $p \in L_n \setminus L_{n-1}$. Then $p = (q \implies r)$ for some $q, r \in L_{n-1}$. We know that $v(q) = w(q), v(r) = w(r)$, and so $v(p) = w(p)$.
2. We define v successively on $L_0, L_1, L_2, \ldots$

   L_0: Let $v|_P = f$ and let $v(\bot) = 0$

   L_n: If $p \in L_n \setminus L_{n-1}$, then $p = (q \implies r)$, and so set $v(p)$ to be 0 if $v(q) = 1, v(r) = 0$, and 1 otherwise. Since propositions are built up in a unique way, this is indeed a valuation.

□

Let $S \subset L$. We say that v is a ***model*** of S if v is a valuation with $v(x) = 1$ for all $x \in S$. If $S = \{p\}$, we say that v is a model of p. If every model of $S \subset L$ is a model of $p \in L$, we say that S ***semantically entails*** p, and write $S \vDash p$. Note that $\emptyset \vDash p$ is exactly the same as $\vDash p$.

For example, $\{p, p \implies q\} \vDash q$.

1.2 Syntactic Entailment (Provability)

Our proof system will have axioms as follows for all $p, q, r \in L$:

A1 $p \implies (q \implies p)$

A2 $(p \implies (q \implies r)) \implies ((p \implies q) \implies (p \implies r))$

A3 $((p \implies \bot) \implies \bot) \implies p$

Our proof system also has a ***deduction rule*** known as ***modus ponens*** (MP): for all $p, q \in L$, from p and $(p \implies q)$ we can deduce q.

Note that each axiom is a tautology. For MP, see the last example of **§1.1**

Let $S \subset L$ and $p \in L$. A ***proof*** of p from S is a sequence $t_1, t_2, \ldots, t_n \in L$ of finite length with $t_n = p$ such that, for each i, either t_i is an axiom, or $t_i \in S$ (a ***hypothesis***), or there exist $j, k < i$ with $t_k = (t_j \implies t_i)$.

If there exists a proof of p from S, we say that S ***syntactically entails*** p, or S ***proves*** p, and we write $S \vdash p$. If $S = \emptyset$, we say p is a ***theorem*** and write $\vdash p$.

Example: $\vdash (p \implies p)$
Use A2, with $r = p$, to get $(p \implies (q \implies p)) \implies ((p \implies q) \implies (p \implies p))$. Now the first bracket is a theorem by A1, and if we take $q = (p \implies p)$ in the second, we can use modus ponens twice with A1 to deduce the final bracket, that $(p \implies p)$. We will write this formally:

Lemma 1.2. *For all $p \in L, \vdash (p \implies p)$*

Proof.

1. $(p \implies ((p \implies p) \implies p)) \implies ((p \implies (p \implies p)) \implies (p \implies p))$ (A2)
2. $p \implies ((p \implies p) \implies p)$ (A1)
3. $(p \implies (p \implies p)) \implies (p \implies p)$ (MP on 1, 2)
4. $p \implies (p \implies p)$ (A1)
5. $p \implies p$ (MP on 3, 4)

□

Proposition 1.3 (The Deduction Theorem)**.** *Let $S \subset L$ and $p, q \in L$. Then $S \vdash (p \implies q)$ if and only if $S \cup \{p\} \vdash q$.*

Proof. Suppose $t_1, \ldots, t_n$ is a proof of $p \implies q$ from S. Then $t_1, \ldots, t_n, p, q$ is a proof of q from $S \cup \{p\}$. Suppose that $t_1, \ldots, t_n$ instead is a proof of q from $S \cup \{p\}$. We show by induction on i that $S \vdash (p \implies t_i)$ for each i, and then we will be done since $t_n = q$.

1. If $t_i \in S$:

- $t_i \implies (p \implies t_i)$ (A1)
- t_i (hypothesis)
- $(p \implies t_i)$ (MP)

2. If $t_i = p$, use Lemma **1.2**
3. If $t_j = (t_j \implies t_i)$ for some $j, k < i$, then write down proofs of $(p \implies t_j), (p \implies t_k)$ from S. Then append:
   - $(p \implies (t_j \implies t_i)) \implies ((p \implies t_j) \implies (p \implies t_i))$ (A2)
   - $(p \implies t_j) \implies (p \implies t_i)$ (MP)
   - $p \implies t_i$ (MP)

□

1.3 The Completeness Theorem and Applications

The key result of this section will be that $\vDash$ and $\vdash$ coincide. There will be two directions to prove:

1. ***Soundness:*** If $S \vdash p$ then $S \vDash p$.
2. ***Adequacy:*** If $S \vDash p$ then $S \vdash p$.

Proposition 1.4 (Soundness Theorem)**.** *Let $S \subset L$ and $p \in L$ with $S \vdash p$. Then $S \vDash p$.*

Proof. Let $t_1, \ldots, t_n$ be a proof of p from S. Let v be a model of S. We show by induction on i that $v(t_i) = 1$ for $1 \leq i \leq n$.

If $t_i \in S$ then $v(t_i) = 1$. If t_i is an axiom then $\vDash t_i$ so $v(t_i) = 1$. Otherwise, $t_k = (t_j \implies t_i)$ for some $j, k < i$. By the induction hypothesis, $v(t_j) = v(t_j \implies t_i) = 1$, so $v(t_i) = 1$. □

For adequacy, first consider the special case $p = \bot$, i.e. "If $S \vDash \bot$ them $S \vdash \bot$". We will prove the contrapositive: "If $S \nvdash \bot$ then $S \nvDash \bot$". If $S \nvdash \bot$ we say that S is ***consistent***. '$S \vDash \bot$" means "if v is a model of S then $v(\bot) = 1$". But $v(\bot) = 0$ for every valuation v, so this says "S has no model." Hence "$S \nvDash \bot$" says "S has a model".

Theorem 1.5 (Model Existence Lemma)**.** *Let $S \subset L$ be consistent. Then S has a model.*

Proof in the case P is countable. L is countable, as each $p \in L$ is a finite string of symbols from $P \cup \{(,), \bot, \implies\}$.

We write $L = \{x_1, x_2, \ldots\}$. We shall recursively construct sets $S_n \subset L$ with $S = S_0 \subset S_1 \subset \ldots$ and S_n consistent.

The base case is trivial, as $S_0 = S$ is consistent by hypothesis. Then for $n > 0$, we have S_{n-1} consistent. If $S_{n-1} \cup \{\neg x_n\}$ is consistent, let $S_n = S_n \cup \{\neg x_n\}$. Otherwise, $S_{n-1} \cup \{\neg x_n\} \vdash \bot$, and by the deduction theorem, $S_{n-1} \vdash (\neg x_n \implies \bot)$, i.e. that $S_{n-1} \vdash \neg\neg x_n$. But $S_{n-1} \vdash (\neg\neg x_n \implies x_n)$ by (A3), and so $S_{n-1} \vdash x_n$ by (MP). But S_{n-1} is consistent, so let $S_n = S_{n-1} \cup \{x_n\}$.

Then let $\bar{S} = \bigcup_{n=1}^{\infty} S_n$. Firstly, S_n is consistent - suppose $t_1, \ldots, t_n$ is a proof of $\bot$ from $\bar{S}$. Then there is some collection $i_1, \ldots, i_m \in \mathbb{N}$ such that the hypotheses used in the proof come

from $S_{i_1}, \ldots, S_{i_m}$. Let $I = \max\{i_1, \ldots, i_m\}$. Then every hypothesis comes from S_I, and so $t_1, \ldots, t_n$ is a proof of $\bot$ from S_I. $\lightning$

Also, for every $p \in L$ we have $p \in \bar{S}$ or $\neg p \in \bar{S}$. Moreover, $\bar{S}$ is ***deductively closed*** (d.c): if $\bar{S} \vdash p$ then $p \in \bar{S}$. Indeed, suppose that $\bar{S} \vdash p$ but $p \notin \bar{S}$. Then $\neg p \in \bar{S}$. Now $\bar{S} \vdash p$ and $\bar{S} \vdash \neg p$, i.e. $\bar{S} \vdash (p \implies \bot)$. So by (MP), $\bar{S} \vdash \bot$ $\lightning$.

Now let $v : L \to \{0, 1\}$ be the indicator function of $\bar{S}$. We must check that v is a valuation. As $\bar{S}$ is consistent, it is certainly true that $\bot \notin \bar{S}$, and so $v(\bot) = 0$.

Let $p, q \in L$. We want to think about $(p \implies q)$:

Case 1. Suppose $v(q) = 1$. Then $q \in \bar{S}$, so $\bar{S} \vdash (p \implies q)$, but $\bar{S}$ is deductively closed, and so $(p \implies q) \in \bar{S}$, and $v(p \implies q) = 1$.

Case 2. Suppose $v(p) = 0$. Again, we must show that $v(p \implies q) = 1$, i.e. that $\bar{S} \vdash (p \implies q)$. By the Deduction Theorem, this is equivalent to $S \cup \{p\} \vdash q$, and $p \notin S$, so $\neg p \in S$ and it will be enough to show that $\{p, \neg p\} \vdash q$. We have:

1. $(p \implies \bot)$ (hyp)
2. p (hyp)
3. $\bot$ (MP on 1,2)
4. $((q \implies \bot) \implies \bot) \implies q$ (A3)
5. $\bot \implies ((q \implies \bot) \implies \bot)$ (A1)
6. $(q \implies \bot) \implies \bot$ (MP on 3,5)
7. q (MP on 4,6)

Case 3. $v(p) = 1, v(q) = 0$. We want to show that $v(p \implies q) = 0$. Suppose instead that $v(p \implies q) = 1$, so that $(p \implies q) \in \bar{S}, p \in \bar{S}$. But then by (MP) $\bar{S} \vdash q$, so $q \in \bar{S}$, so $v(q) = 1$. $\lightning$

We have now shown that v is a valuation. Moreover, $S \subset \bar{S}$ so $v(p) = 1$ for all $p \in S$. Hence v is a model of S. □

Corollary 1.6 (The Adequacy Theorem)**.** *Let $S \subset L$ and $p \in L$ with $S \vDash p$. Then $S \vdash p$.*

Proof. Suppose v is a model of $S \cup \{\neg p\}$. Then $v(p) = 1$ and $v(\neg p) = 1$, so $v(\bot) = 1$ $\lightning$. So $S \cup \{\neg p\}$ has no model, and so by the model existence lemma it is inconsistent. That is, $S \cup \{\neg p\} \vdash \bot$. Then by the deduction theorem, $S \vdash (\neg p \implies \bot)$, i.e. $S \vdash \neg\neg p$), and hence $S \vdash p$. □

Theorem 1.7 (The Completeness Theorem)**.** *Let $S \subset L$ and $p \in L$. Then $S \vDash p$ if and only if $S \vdash p$.*

Proof. Soundness and adequacy. □

Two important applications:

Corollary 1.8 (Compactness Theorem)**.** *Let $S \subset L$ such that every finite subset of S has a model. Then S has a model.*

Proof. Not at all obvious a priori, but obvious if we replace "has a model" by "is consistent". If S is not consistent, then $S \vdash \bot$, so, as proofs are finite, $T \vdash \bot$ for some finite $T \subset S$, so T is inconsistent. □

Corollary 1.9 (The Decidability Theorem)**.** *Let $S \subset L$ be finite and $p \in L$. Then there is an algorithm to determine in finite time whether or not $S \vdash p$.*

Proof. Obvious if we replace $\vdash$ by $\models$, and then do a truth table. □

1.4 What happens when P is uncountable?

This will be just a sketch - it will be a while before we can do this properly in chapter 3. We have only proved the completeness theorem under the assumption that P is countable. However, we only used this when showing that, if S is consistent, then there is a consistent $\bar{S} \supset S$ with $p \in \bar{S}$ or $\neg p \in \bar{S}$ for all $p \in L$.

We needed P to be countable so that $L = \{x_1, x_2, \ldots\}$ is countable and we can consider the x_is in turn, deciding if $x_i \in \bar{S}$ or $\neg x_i \in \bar{S}$.

Can we do without this assumption? Now allow P, and hence L, to be uncountable. Let $S \subset L$ be consistent and look for $\bar{S} \supset S$ consistent with $p \in \bar{S}$ or $\neg p \in \bar{S}$ for all $p \in L$. We could try $\bar{S} = S$, and if it works then we are done.

Otherwise, there is some $x_0 \in L$ with $x_0 \notin \bar{S}$ and $\neg x_0 \notin \bar{S}$. Exactly as in the countable case, either $\bar{S} \cup \{x_0\}$ or $\bar{S} \cup \{\neg x_0\}$ is consistent. So add x_0 or $\neg x_0$ to $\bar{S}$, keeping it consistent. If $\bar{S}$ works now, then we're done, otherwise there is some $x_1 \in L$ with $x_1 \notin \bar{S}$ and $\neg x_1 \notin \bar{S}$, and so on and so forth. If this never terminates, then we keep on going forever.

If after we've done this infinitely many times, if we are done the stop. Otherwise, we have x_ω with $x_\omega \notin \bar{S}$ and $\neg x_\omega \notin \bar{S}$, so either add x_ω or $\neg x_\omega$ to $\bar{S}$. If we're not done there is another $x_{\omega+1}$ and so on.

Either eventually we finish, or we have to go on forever, until we get to $x_{\omega \cdot 2}$, and then keep on going…

If this never terminates, we get loads and loads of x_is. What are they being indexed by? It looks to be some sort of extension of $\mathbb{N}$.:

$$0, 1, 2, 3, 4, \ldots, \omega, \omega+1, \omega+2, \ldots, \omega 2, \omega 2+1, \ldots, \omega 3, \ldots, \omega 4, \ldots,$$

$$\omega^2, \omega^2+1, \ldots, \omega^2+\omega, \ldots, \omega^2 2, \ldots, \omega^3, \ldots, \omega^4, \ldots, \omega^\omega, \ldots, \omega^{\omega^\omega}, \ldots, \omega^{\omega^{\omega^{\cdot^{\cdot^{\cdot}}}}} = \epsilon_0, \ldots$$

This is only countably many ordinals, but in fact if we keep on going, we can get to uncountably many numbers. We call these indices ***ordinals***, and we will define them in chapter 2 in such a way that:

- 0 is an ordinal
- For any ordinal α there is a least ordinal larger than it
- Given any set of ordinals, there is a least ordinal bigger than all of them

We will use Hartog's Lemma to show that in fact, eventually we do run out of stuff in L if we use this indexing, which says that, for any set X, there are more ordinals than there are elements of X.

Note that here we are using the axiom of choice, because this is part of the maths tripos which uses the axiom of choice. Later on we, will think about what might happen if we don't have the axiom of choice, but we won't worry about that for now.

2 Ordinals

2.1 Functions and Relations

A ***function*** $f : X \to Y$ from a set X to a set Y is a subset $f \subset X \times Y$ such that for all $x \in X$ there is a unique $y \in Y$ with $(x, y) \in f$. We write $f(x) = y$ to mean $(x, y) \in f$. If $Z \subset X$, the ***restriction*** of f to Z is the function $f|_Z : Z \to Y$ given by $f|_Z = f \cap (Z \times Y)$.

If $f^{-1}(\{y\}) = \{x \in X : f(x) = y\}$ has at most one element for every $y \in Y$, we say f is an ***injection*** (i.e. no two x's get mapped to the same y). If $f^{-1}(\{y\})$ has at least one element for every $y \in Y$, we say f is a ***surjection*** (i.e. every y gets mapped to by some x). If f is both an injection and a surjection, it is called a ***bijection***.

A ***relation*** R on a set X is a subset $R \subset X \times X$. We write xRy to mean $(x, y) \in R$. If R, S are relations on sets X, Y respectively, an ***isomorphism*** from (X, R) to (Y, S) is a bijection $f : X \to Y$ such that, for all $x, y \in X, xRy \iff f(x)Sf(y)$. If such an isomorphism exists, we say that (X, R) and (Y, S) are ***isomorphic***.

Note: this is not the right way to think about functions and relations, but it is convenient sometimes. Keep thinking of a function as "something that associates a unique element of Y with each element of X". Note that this does give us a way to define functions and relations in such a way that they live in the universe of sets.

2.2 Well-Ordering

A ***total order*** on a set X is a relation $<$ on X satisfying:

- For all $x, y \in X$, precisely one of $x = y, x < y, y < x$ holds. (trichotomy)
- Fro all $x, y, z \in X$, if $x < y$ and $y < z$ then $x < z$. (transitivity)

If $<$ is a total order on X, we can define a relation $\leq$ on X by $x \leq y$ if $x < y$ or $x = y$. This satisfies:

- For all $x, y \in X, x \leq y$ or $y \leq x$.
- For all $x, y \in X$, if $x \leq y$ and $y \leq x$, then $x = y$.
- For all $x, y, z \in X$, if $x \leq y, y \leq z$, then $x \leq z$.

Conversely, given a relation $\leq$ satisfying these conditions, we can define a total order $<$ on X by $x < y$ if $x \leq y$ and $x \neq y$. Some sources will define $\leq$ to be a total order, but here we will use the $<$ version.

A total order of X is a ***well-ordering*** of X if every non-empty subset $Z \subset X$ has a least element (i.e. an element $y \in Z$ such that for all $x \in Z, y \leq x$.

An ***ordinal*** is a well-ordered set with isomorphic sets considered to be the same. Given a well-ordered set X, the ***order-type*** of X is the corresponding ordinal.

A total order $<$ is a well-ordering if and only if there is no infinite descending sequence $x_1 > x_2 > \ldots$.

Examples:

1. $X = \{a, b, c, d\}$ with $a < b, c, d; b < c, d; c < d$ well orders X.
2. The usual order on $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ defines a total order. $\mathbb{N}$ is well ordered by this, $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ are not, as $-1 > -2 > -3 > \ldots$.
3. $x < y$ if $|x| - \frac{1+\operatorname{sgn} x}{4} < |y| - \frac{1+\operatorname{sgn} y}{4}$ is however a well ordering on $\mathbb{Z}$, with $0 < 1 < -1 < 2 < -2 < \ldots$.
4. Take $\{\frac{1}{n} : n \in \mathbb{N} \setminus \{0\}\}$ with the usual ordering. This is not a well ordering, as $1 > \frac{1}{2} > \frac{1}{3} > \ldots$.
5. Take $\{-\frac{1}{n} : n \in \mathbb{N} \setminus \{0\}\}$ with the usual ordering. This is a well ordering, as there are only finitely many elements less than any given element. This has the same order type as $\mathbb{N}$ with the usual order.
6. What about $X = \left\{-\frac{1}{n} : n \in \mathbb{N} \setminus \{0\}\right\} \cup \left\{1 - \frac{1}{n} : n \in \mathbb{N} \setminus \{0\}\right\}$ with the usual order. Even though there are infinitely many elements less than say $\frac{2}{3}$, we can see that any descending sequence can have only finitely many positive terms, before a negative term, and then only finitely many negative terms. Hence this is in fact a well-ordering, but it is not the same order-type as $\mathbb{N}$. This should remind you of the discussion at the end of section 1, where we went through infinitely many elements, and then did it again.