Harry Langford
hjel2@cam.ac.uk

1. Inductive Definitions

Exercise 1.1 Let L be the subset of $\{a,b\}^*$ inductively defined by the axiom $\frac{}{\varepsilon}$ and the rule $\frac{u}{aub}$ (for any $u \in \{a,b\}$).

We need to use rule induction.

(a) Use rule induction to prove that every string in L is of the form $a^n b^n$ for some $n \in \mathbb{N}$.

Prove $P(n) : \forall u \in L.|u| \leq 2n \implies \exists k \in \mathbb{N}.k \leq n.u = a^k b^k$

At $n = 0$, there is only one possible string of length 0: ε. ε is of the form $a^0 b^0$. So $P(0)$ is true.

Assume now that $P(m)$ for some $m \in \mathbb{N}$.
This means that every string in L which is shorter than $2k$ must be of the form $a^k b^k$ for some $k \leq m$.

Since $P(m)$ is true, this means that the first string which is not of the form $a^k b^k$ for some $k \in \mathbb{N}$ must be of length $2m+1$ or $2m+2$. Using the rules, the only way to make a string longer than any known string is by deriving aub from u. This makes a string two longer than the previous string. So to derive a string of length $2(m+1)$ or $2m+1$ we must first take a string u of length $2m-1$ or $2m$ if it exists and then consider aub. By assumption if there is a string of length $2m$ then it must be of the form $a^m b^m$. So $aub = aa^m b^m b = a^{m+1} b^{m+1}$. This does not violate $P(m+1)$. Consider now strings of length $2m-1$. Since by assumption every string is of length $a^k b^k$ for some $k \in \mathbb{N}$, there can be no odd numbered string. And so there are no strings of length $2m-1$. So $P(m) \implies P(m+1)$.

Since $P(0)$ is true and $P(m) \implies P(m+1)$, by induction $P(n)$ must be true for all $n \in \mathbb{N}$. This means that every string in L is of the form $a^n b^n$ for some $n \in \mathbb{N}$.

(b) Use mathematical induction to prove that $\forall n \in \mathbb{N}.a^n b^n \in L$. Conclude that $L = \{a^n b^n | n \in \mathbb{N}\}$.

We are required to prove that for all $n \in \mathbb{N}.a^n b^n \in L$.

At $n = 0$: $a^0 b^0 = \varepsilon$ – which is an axiom of L and so is in L.

Now assume that $a^k b^k \in L$ for some $n \in \mathbb{N}$.
By the rule $\frac{u}{aub}$, $a^k b^k \in L \implies aa^k b^k b \in L \iff a^{k+1} b^{k+1} \in L$.
So $a^k b^k \in L \implies a^{k+1} b^{k+1} \in L$.

Since $a^0 b^0 \in L$ and $a^k b^k \in L \implies a^{k+1} b^{k+1} \in L$ we can conclude by induction that for all $n \in \mathbb{N}.a^n b^n \in L$.

Since we have proven in (a) that $L \subseteq \{a^n b^n | n \in \mathbb{N}\}$ and we have just proven that $\{a^n b^n | n \in \mathbb{N}\} \subseteq L$, we can conclude that $L = \{a^n b^n | n \in \mathbb{N}\}$ as required.

(c) Suppose we add the string a to L, that is consider $L' = L \cup \{a\}$. Is L' closed under the axiom and rule? If not, characterise the strings that would be the smallest set containing L' that is closed under the axiom and rule.

L' is not closed under the axiom and rule. The smallest set containing L' that would be closed under the axiom and rule is $\{a^n a^k b^n | n \in \mathbb{N}.k \in [2]\}$.

Exercise 1.2 Suppose $R \subseteq X \times X$ is a binary relation on a set X. Let $R^\dagger \subseteq X \times X$ be inductively defined by the following axioms and rules:

$$\frac{}{(x,x) \in R^\dagger}(x \in X), \qquad (1) \qquad \frac{(x,y) \in R^\dagger}{(x,z) \in R^\dagger}(x \in X \text{ and } (y,z) \in R). \qquad (2)$$

(a) Show that $R^\dagger$ is reflexive and that $R \subseteq R^\dagger$.

Assume $x \in X$.

By rule 1:
$(x, x) \in R^\dagger$. This is the definition of reflexive and so $R^\dagger$ is reflexive. ✓

To show that $R \subseteq R^\dagger$ I will show that $(x, y) \in R \implies (x, y) \in R^\dagger$.

Assume $(x, y) \in R$. Since $R^\dagger$ is reflexive, $(x, x) \in R$.

By rule 2:
$\frac{(x,x)\in R^\dagger}{(x,y)\in R^\dagger}(x \in X \text{ and } (x, y) \in R)$ and so $(x, y) \in R^\dagger$. Since x and y were arbitrary, we have shown that for any x, y: $(x, y) \in R \implies (x, y) \in R^\dagger$ and so $R \subseteq R^\dagger$.

(b) Use rule induction to show that $R^\dagger$ is a subset of

$$S \triangleq \{(y, z) \in X \times X | \forall x \in X.(x, y) \in R^\dagger \Rightarrow (x, z) \in R^\dagger\} \tag{3}$$

Deduce that $R^\dagger$ is transitive

Need to use rule induction

Will discuss.

By reflexivity, assume $\forall x \in X.(x, x) \in R^\dagger$.

Setting $x = y$ implies:

$$\begin{aligned} \{y, z)|(y, y) \in R^\dagger \implies (y, z) \in R^\dagger\} \subseteq S \\ \{(y, z)|(y, z) \in R^\dagger\} \subseteq S \\ R^\dagger \subseteq S \end{aligned} \tag{4}$$

So $R^\dagger \subseteq S$ as required.

Since $R^\dagger \subseteq S$, this means that $\forall x \in X.(x, y) \in R^\dagger \wedge (y, z) \in R^\dagger \implies (x, z) \in R^\dagger$. This is the definition of transitivity and so $R^\dagger$ is transitive.

(c) Suppose $S \subseteq X \times X$ is a reflexive and transitive binary relation and that $R \subseteq S$. Use rule induction to show that $R^\dagger \subseteq S$.

Deduce $R^\dagger$ is transitive.

Assume that $R \subseteq S$ and S is reflexive-transitive relation.
Hence:

$$\begin{aligned} \forall x \in X.(x, x) \in S \wedge (x, y) \in S \wedge (y, z) \in S \implies (x, z) \in S \implies \\ \forall x \in X.(x, x) \in S \wedge (x, y) \in R \wedge (y, z) \in R \implies (x, z) \in S \implies \\ R^\dagger \subseteq S \end{aligned} \tag{5}$$

(d) Deduce from (a) - (c) that $R^\dagger$ is equal to R^*, the reflexive-transitive closure of R.

Note the rules of $R^\dagger$.

$$\frac{x \in X}{(x, x) \in R^\dagger} \tag{6}$$

$$\frac{(x, y) \in R^\dagger \wedge (y, z) \in R}{(x, z) \in R^\dagger} \tag{7}$$

From (a), $13 \implies ((x, z) \in R \implies (x, z) \in R^\dagger)$.

So we can add this to the rules without changing the definition of R.

The new rules of R are as follows:

$$\frac{x \in X}{(x, x) \in R^\dagger} \tag{8}$$

$$\frac{(x, y) \in R}{(x, y) \in R^\dagger} \tag{9}$$

i'm not following why are we changing the rules?

$$\frac{(x,y) \in R^\dagger \wedge (y,z) \in R}{(x,z) \in R^\dagger} \tag{10}$$

Note that since $(x,z) \in R \implies (x,z) \in R^dagger$ we can change the definition further without affecting the set $R^\dagger$.

$$\frac{x \in X}{(x,x) \in R^\dagger} \tag{11}$$

$$\frac{(x,y) \in R}{(x,y) \in R^\dagger} \tag{12}$$

$$\frac{(x,y) \in R^\dagger \wedge (y,z) \in R^\dagger}{(x,z) \in R^\dagger} \tag{13}$$

This is the definition of the reflexive-transitive closure of R. So $R^\dagger$ is the reflexive-transitive closure of R.

Exercise 1.3 Let L be the subset of $\{a,b\}^*$ inductively defined by the axiom $\frac{}{ab}$ and the rules $\frac{au}{au^2}$ and $\frac{ab^3u}{au}(\forall u \in \{a,b\}^*)$

(a) Is ab^5 in L? Give a derivation, or show there isn't one.

$ab^5 \in L$. The derivation is below.

$$\frac{}{ab} \implies \frac{ab}{ab^2} \implies \frac{ab^2}{ab^4} \implies \frac{ab^4}{ab^8} \implies \frac{a(b^3)b^5\cancel{)}}{ab^5} \implies ab^5 \in L \tag{14}$$

(b) Use rule induction to show that every $u \in L$ is of the form ab^n with $n = 2^k - 3m \geq 0$ for some $k, m \in \mathbb{N}$.

Let $P(n)$ mean after n conclusions, every string we have seen is of the form $ab^k - 3m$ for some $k, m \in \mathbb{N}$.

At 0 we have made no conclusions and so the only string is the axiom ab.
ab is of the form $ab^k - 3m$ where $k = 1, m = 0$.
So $P(0)$ is true.

Now assume that $P(q)$ for some $q \in \mathbb{N}$. This means that every string we can reach with q or fewer conclusions is of the form $ab^k - 3m$ for some $k, m \in \mathbb{N}$. Let us take an arbitrary string in L after q conclusions. This string is of the form $ab^{2^k - 3m}$ by assumption. L is derived by two rules. So there are two rules we need to consider.

$$\frac{ab^{2^{k'} - 3m'}}{ab^{2 \cdot 2^{k'} - 2 \cdot 3m'}} \implies ab^{2^{k'+1} - 3(2m'} \tag{15}$$

This derives a string of the form $ab^{2^k - 3m}$ where $k, m \in \mathbb{N}$.

If $2^k - 3m \geq 3$:

$$ab^{2^{k'} - 3m'} \implies ab^3 b^{2^{k'} - 3(m'+1)} \implies \frac{ab^3 b^{2^{k'} - 3(m'+1)}}{ab^{2^{k'} - 3(m'+1)}} \implies ab^{2^{k'} - 3(m'+1)} \tag{16}$$

This also derives a string of the form $ab^{2^k - 3m}$ where $k, m \in \mathbb{N}$.

So $P(n) \implies P(n+1)$. Since $P(0)$, by induction, $P(n) \forall n \in \mathbb{N}$. This means that every string in L is of the form $ab^{2^k - 3m}$ for some $k, m \in \mathbb{N}$.

Rule induction! Will discuss.

(c) is ab^3 in L. Give a derivation or show there isnt one.

From the proof above we know that every $u \in L$ is of the form ab^n where $n = 2^k - 3m \geq 0$.

By (b) for a string u to be in L it must be of the form ab^n for some n such that $n = 2^k - 3m$. We have that $ab^3 = ab^n$ where $n = 3$. So a necessary condition for ab^3 to be in L is that 3 can be represented as $2^k - 3m$ for some $m, k \in \mathbb{N}$.

Assume 3 can be expressed in the form $2^k - 3m$.

$$\begin{aligned} 3 = 2^k - 3m &\implies \\ 3(m+1) = 2^k &\implies \\ 3 | 2^k & \end{aligned} \tag{17}$$

However, this is absurd. So 3 cannot be expressed in the form $2^k - 3m$ and so $ab^3 \notin L$.

(d) Can you characterize exactly which strings are in L?

All strings of the form ab^n where $n \not\equiv_3 0$ are in L. Proof ? :) (or use form in (b))

2. Regular Expressions

Exercise 2.1 Find regular expressions over $\{0, 1\}$ that determine the following languages.

(a) $\{u \mid u$ contains an even number of 1's$\}$

$$L = (0^*10^*10^*)^* \tag{18}$$

Doesn't accept 000

(b) $\{u \mid u$ contains an odd number of 0's$\}$

$$L = 1^*0(1^*01^*01^*)^* \tag{19}$$

Doesn't accept 101

Exercise 2.3 Show that $b^*a(b^*a)^*$ and $(a|b)^*a$ are equivalent regular expressions, that is, a string matches one iff it matches the other.

I will prove this by showing $u \in b^*a(b^*a)^* \implies u \in (a|b)^*a$ and then showing that $u \in (a|b)^*a \implies u \in b^*a(b^*a)^*$. OK !

Assume $u \in b^*a(b^*a)^*$:
This means that u matches $b^*a(b^*a)^n$ for some $n \in \mathbb{N}$. By associativity, $u = (b^*a)^n b^*a$. Now consider splitting u into $u'a$. So $u' = (b^*a)^n$. Trivially a matches a. Every letter in $u' \in \{a, b\}$. So u' matches $(a|b)^*$. So $u \in (a|b)^*a$ as required. So $b^*a(b^*a)^* \subseteq (a|b)^*a$.

Assume $u \in (a|b)^*a$:

Trivially there is a finite number of a's in u (say n) each of which separated by some finite number of b's. Since the last letter in u is a, this number of a's must be nonzero. Hence we can represent u as $(b^{x_i}a)^{n-1}(b^{x_n}a)$ for $0 \leq i < n$.

Since we know that $x_i \geq 0$ and $n - 1 \geq 0$, this is of the form $(b^*a)^*b^*a$. This is equivalent to $b^*a(b^*a)^*$ and so $u \in b^*a(b^*a)^*$ as required. This means that $(a|b)^*a \subseteq b^*a(b^*a)^*$.

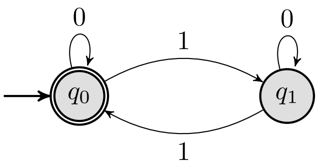
Since both languages are subsets of eqch other, they must be equal. Hence $b^*a(b^*a)^* = (a|b)^*a$

Harry Langford
hjel2@cam.ac.uk

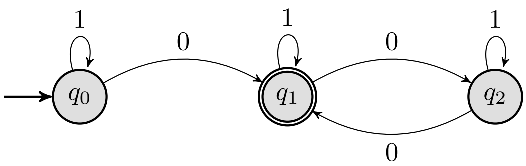
3. Finite Automata

Exercise 3.1 For each of the two languages mentioned in Exercise 2.1 find a DFA that accepts exactly that set of strings.

The following DFA accepts L the subset of $\{0, 1\}$ which which has an even number of 1's.



The following DFA accepts L the subset of $\{0, 1\}$ which which has an odd number of 0's.



Can we simplify this?

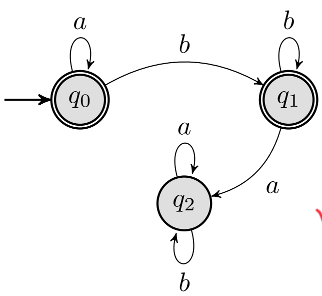
Exercise 3.2 Given an NFA$^\varepsilon$ $M = (Q, \Sigma, \Delta, s, F, T)$, we write $q \stackrel{u}{\Longrightarrow} q'$ to mean that there is a path in M from state q to state q' whose non-ε labels form the string $u \in \Sigma^*$. Show that $\{(q, u, q') | q \stackrel{u}{\Longrightarrow} q'\}$ is equal to the subset of $Q \times \Sigma \times Q$ inductively defined by the axioms and rules

$$\begin{array}{l} \overline{(q, \varepsilon, q} \\ \dfrac{(q, u, q')}{(q, u, q'')} \text{ if } q' \xrightarrow{\varepsilon} q'' \in M \\ \dfrac{(q, u, q')}{(q, ua, q'')} \text{ if } q' \xrightarrow{a} q'' \in M \end{array} \tag{20}$$

Will discuss.

Exercise 3.3 The example of the subset construction given in the lecture notes constructsa DFA with eight states whose language of accepted strings happens to be $L(a^*b^*)$. Give a DFA with the same language of accepted strings, but fewer states. Give an NFA with even fewer states that does the same job.

DFA to recognise a^*b^*:



NFA to recognise a^*b^*: