1. Trees; context free grammars.

1.1. Trees.

Definition 1.1. By a rooted tree we mean an ordered triple

\[ T = (\mathcal{N}, o, p) \]

such that

(i) \( \mathcal{N} \) is a set;
(ii) \( o \in \mathcal{N} \);
(iii) \( p : \mathcal{N} \sim \{o\} \to \mathcal{N} \);
(iv) if \( \nu \in \mathcal{N} \sim \{o\} \) there is \( d \in \mathbb{N}^+ \) such that \( (\nu, o) \in p^d \).

One calls \( o \) the root of \( T \).

The members of \( \mathcal{N} \) are called nodes. \( o \) is called the root node. If \( \nu \in \mathcal{N} \sim \{o\} \) we call \( p(\nu) \) the parent of \( \nu \). If \( \nu \in \mathcal{N} \) we let

\[ c(\nu) = p^{-1}[\{\nu\}] = \{ \mu \in \mathcal{N} \sim \{o\} : p(\mu) = \nu \} \]

and call the members of this set the children of \( \nu \). A node which has children is called an interior node. A node which has no children is called a leaf node. We let

\[ i \quad \text{and} \quad l \]

be the set of interior nodes of \( T \) and the set of leaf nodes of \( T \), respectively.

Lemma 1.1. Suppose \( n \in \mathbb{N} \). Then \( p^{-n}[\{o\}] \subset \mathcal{N} \sim \{o\} \).

Proof. Suppose, contrary to the Lemma, \( o \in p^{-n}[\{o\}] \). Since \( p^{-n} = (p^n)^{-1} \) we have \( (o, o) \in p^n \) so

\[ o \in \text{dmm} p^n = \begin{cases} \{ (\nu, \nu) : \nu \in \mathcal{N} \sim \{o\} \} & \text{if}\ n = 0, \\ \subset \text{dmm} p & \text{if}\ n > 0 \end{cases} \]

which is incompatible with 1.1(iii). \( \square \)

Lemma 1.2. Suppose \( m, n \in \mathbb{N} \) and \( p^{-m}[\{o\}] \cap p^{-n}[\{o\}] \neq \emptyset \). Then \( m = n \).

Proof. Suppose, contrary to the Lemma, \( \mu \in p^{-m}[\{o\}] \cap p^{-n}[\{o\}] \) and \( m > n \). Since \( p^{-n} = (p^n)^{-1} \) we have \( (o, \mu) \in p^{-n} \); also, \( (\mu, o) \in p^{m} \) so that

\[ (o, o) \in p^{-n} \circ p^{m} = p^{m-n} \]

which is incompatible with 1.1(iii). \( \square \)

Definition 1.2. Let

\[ d = \{ (o, 0) \} \cup \left( \bigcup_{n \in \mathbb{N}^+} p^{-n}[\{o\}] \times \{n\} \right) \]

here \( d \) stands for “depth”.

Obviously,

\[ p^{-n}[\{o\}] = d^{-1}[\{n\}] \quad \text{for}\ n \in \mathbb{N}^+. \]

Theorem 1.1. \( d \) is a function with domain \( \mathcal{N} \).
Proof. $o \in \text{dmn} \ d$ since $(o, 0) \in d$. If $\nu \in N \sim \{o\}$ there is by (iv) $n \in N^+$ such that $(\nu, o) \in p^n$ so $\nu \in p^{-n}[\{o\}] \subseteq \text{dmn} \ d$.

Suppose, contrary to the Theorem, there are $\nu \in N$ and $n_i$, $i = 1, 2$, such that $n_2 > n_1$ and $(\nu, n_i) \in d, i = 1, 2$.

In case $n_1 = 0$ we would have $\nu = o$ and $\nu \in p^{-n_2}[\{o\}]$ which is excluded by Lemma 1.2.

In case $n_1 > 0$ we would have $\nu \in p^{-n_1}[\{o\}], i = 1, 2$. Then $(\nu, o) \in p^{-n_1}$ and $(o, \nu) \in (p^{-n_2})^{-1} = p^{n_2}$ so $(o, o) \in p^{-n_1} \circ p^{n_2} = p^{n_2-n_1}$. Thus $o \in \text{dmn} p^{n_2-n_1} \subseteq \text{rng} p$ which is incompatible with Lemma 1.1(iii).

Corollary 1.1. \(\{p^{-d}(\nu) : d \in \mathbb{N}\}\) is a partition of $N$.

For each $\mu \in N$ we let
\[
\mathbf{A}(\mu) = \bigcup_{n \in \mathbb{N}^+} \{\nu : (\mu, \nu) \in p^n\}
\]
and we let
\[
\mathbf{D}(\mu) = \bigcup_{n \in \mathbb{N}^+} \{\nu : (\nu, \mu) \in p^n\};
\]
The members of $\mathbf{A}(\mu)$ are called ancestors of $\mu$ and the members of $\mathbf{D}(\mu)$ are called descendants of $\mu$.

Proposition 1.1. $\mathbf{A}(o) = \emptyset$.

Proof. Suppose $n \in \mathbb{N}^+$, $(o, \nu) \in p^n$. Then $o \in \text{dmn} p^n \subset \text{dmn} p = N \sim \{o\}$ which is impossible.

Proposition 1.2. Suppose $\mu \in N \sim \{o\}$. Then $\mu \in \text{dmn} \ d$, $d(\mu) > 0$, $\mu \in p^{-d(\mu)}[\{o\}]$,
\[
\mathbf{A}(\mu) = \{p^n(\mu) : n \in I(1, d(\mu))\}
\]
and $d|A(\mu)$ is a univalent function with range $[0, d(\mu))$.

Proof. $\mu \in \text{dmn} \ d$ by Theorem 1.1. Since $\mu \neq o$ we have $d(\mu) > 0$ and $\mu \in p^{-d(\mu)}[\{o\}]$; in particular, $(\mu, o) \in p^{d(\mu)}$ and $(o, \mu) \in p^{d(\mu)}$. Suppose $n \in \mathbb{N}^+$ and $(\mu, \nu) \in p^n$. Then, as $(\nu, mu) \in p^{-n}$ we have $(\nu, o) \in p^{-n} \circ p^{d(\mu)} = p^{d(\mu)-n}$ so $\nu \in p^{-d(\mu)}[\{o\}]$. Since $o \notin \text{dmn} p$ we find that $n-d(\mu) < 0$ and $d(\nu) = d(\mu) - n \in [0, d(\mu))$.

1.2. Subtrees.

Definition 1.3. We say the rooted tree $U = (\mathcal{O}, \sigma, q)$ is a rooted subtree of $T = (N, o, p)$ if $\sigma \in \mathcal{O} \subseteq N$ and $q = p|(\mathcal{O} \sim \{\sigma\})$.

Given $\nu \in N$ one easily verifies that
\[
\{\nu\} \cup \mathbf{D}(\nu), \nu, p|\mathbf{D}(\nu)
\]
is a rooted subtree of $T$ which we call the rooted subtree associated to the node $\nu$.

1.3. Isomorphisms. Suppose $T_i = (N_i, o_i, p_i), i = 1, 2$ are rooted trees and $\iota : N_1 \to N_2$. We say $\iota$ is an isomorphism from $T_1$ to $T_2$ if $\iota$ is univalent, $\text{rng} \ i = N_2$ and $p_2 \circ \iota = \iota \circ p_1$;

it follows that $\iota(o_1) = o_2$ and that $\iota^{-1}$ is an isomorphism from $T_2$ to $T_1$. 
1.4. Tree orderings. Suppose \( T = (\mathcal{N}, o, p) \) is a tree. We say \( \prec \) is a tree ordering of \( T \) if

(i) \( \prec \) is well ordering of \( \mathcal{N} \);
(ii) if \( \mu, \nu \in \mathcal{N} \) and \( d(\mu) < d(\nu) \) then \( \mu \prec \nu \).
(iii) if \( \mu, \nu \in \mathcal{N} \sim \{o\} \), \( d(\mu) = d(\nu) \) and \( p(\mu) \prec p(\nu) \) then \( \mu \prec \nu \).

**Proposition 1.3.** Suppose \( W \) is a function with domain \( \mathbb{N} \) whose value at \( n \in \mathbb{N} \) is a well ordering of \( d^{-1}\{n\} \) such that

\[
(p(\mu), p(\nu)) \in W(n) \Rightarrow (\mu, \nu) \in W(n + 1)
\]

whenever \( n \in \mathbb{N} \) and \( \mu, \nu \in d^{-1}\{n + 1\} \). Then

\[
\bigcup_{n=0}^{\infty} W(n) \cup \{ (\mu, \nu) \in \mathcal{N} \times \mathcal{N} : d(\mu) < d(\nu) \}
\]

is a tree ordering of \( T \).

**Theorem 1.2.** Suppose for each \( \xi \in \mathcal{N} \) we are given a well ordering \( \prec_\xi \) of \( c(\xi) \). Then there is one and only one tree ordering \( \prec \) of \( T \) such that, for each \( \xi \in \mathcal{N} \) and each \( \mu, \nu \in c(\xi) \),

\[
(2) \quad \mu \prec_\xi \nu \Leftrightarrow \mu \prec \nu.
\]

**Proof.** With regard to existence, we begin by constructing by induction on \( n \in \mathbb{N} \) a well ordering \( W_n \) of \( d^{-1}\{n\} \) as follows. We let \( W_0 = \emptyset \). If \( n \in \mathbb{N} \) and \( W_n \) is constructed we let

\[
U_{n+1} = \{ (\mu, \nu) \in d^{-1}\{n + 1\} : \mu \neq \nu \text{ and } (p(\mu), p(\nu)) \in W_n \};
\]

\[
V_{n+1} = \{ (\mu, \nu) \in d^{-1}\{n + 1\} : \mu \neq \nu, p(\mu) = p(\nu) \text{ and } (\mu, \nu) \in \prec_{p(\mu)} \};
\]

and we let \( W_{n+1} = U_{n+1} \cup V_{n+1} \). We let

\[
\prec = \{ (\mu, \nu) \in \mathcal{N} \times \mathcal{N} : d(\mu) < d(\nu) \} \cup \bigcup_{n \in \mathbb{N}} W_n.
\]

It should be clear that \( \prec \) is the unique tree ordering of \( T \) such that (2) holds. \( \square \)

Suppose \( \mathcal{O}_i = (\mathcal{N}_i, o_i, p_i, \prec_i) \), \( i = 1, 2 \) are ordered trees and \( \iota : \mathcal{N}_1 \rightarrow \mathcal{N}_2 \). We say \( \iota \) is an isomorphism from \( \mathcal{O}_1 \) to \( \mathcal{O}_2 \) if \( \iota \) is an isomorphism from \( (\mathcal{N}_1, o_1, p_1) \) to \( (\mathcal{N}_2, o_2, p_2) \) and

\[
\mu, \nu \in \mathcal{N}_1 \text{ and } \mu \prec_1 \nu \Rightarrow \iota(\mu) \prec_2 \iota(\nu).
\]

Suppose \( \mathcal{O} = (\mathcal{N}, o, p, \prec) \) is an ordered tree and \( \mathcal{U} = (\mathcal{O}, \sigma, q) \) is a subtree of \( T = (\mathcal{N}, o, p) \). Let

\[
\prec = \{ (\mu, \nu) \in \mathcal{O} \times \mathcal{O} : \mu \prec \nu \}
\]

and note that \( (\mathcal{O}, \sigma, q, \prec) \) is an ordered tree which we call the ordered tree associated to \( \mathcal{U} \). In particular, if \( \nu \in \mathcal{N} \), \( \mathcal{O} = \mathcal{N}_\nu \), \( \sigma = \nu \) and \( q = q_\nu \) we call this ordered tree the ordered tree associated to \( \nu \).
1.5. **Tree codes.** Suppose $I$ is an initial segment of $\mathbb{N}$. We let

$$T(I)$$

be the set of functions $P : I \sim \{0\} \rightarrow I$ such that

- $P^{-1}([0]) \neq \emptyset$;
- $i \in I \sim \{0\} \rightarrow P(i) < i$;
- $i, j \in I \sim \{0\}$ and $i < j \Rightarrow P(i) \leq P(j)$.

**Proposition 1.4.** Suppose $I$ is an initial segment of $\mathbb{N}$ and $P \in T(I)$. Then $T = ((I, 0, P), P, 0)$ is a tree and the standard well ordering of $I$ is a tree ordering of $T$.

**Proof.** Suppose $T = (\mathcal{N}, o, p)$ is a tree and $\prec$ is tree ordering of $T$. Let $\mathcal{N} = \text{card} \mathcal{N}$ and let

$$W : I[0, N) \rightarrow \mathcal{N}$$

be such that

$$i, j \in I[0, N) \text{ and } i < j \Rightarrow W(i) \prec W(j).$$

Note that $W(0) = o$ and $W$ is univalent; in particular, $W^{-1}$ is a function. Let

$$P = W^{-1} \circ p \circ W.$$

**Proposition 1.5.** $P \in T(N)$.

**Remark 1.1.** We call $P$ the tree code for $(T, \prec)$.

**Proof.** Suppose $i \in I(1, N)$. Let $\mu = W(i)$. Then $W(P(i)) = p(W(i)) = p(\mu) \prec \mu = W(i)$ so $P(i) < i$.

Suppose $i, j \in I(1, N), i < j$ and $P(i) \neq P(j)$. Let $\mu = W(i)$ and let $\nu = W(j)$. Then $\mu \prec \nu$. We have $W(P(i)) = p(W(i)) = p(\mu)$ and $W(P(j)) = p(W(j)) = p(\nu)$. Since $P(i) \neq P(j)$ we have $\mu \neq \nu$ so $p(\mu) < p(\nu)$ so $P(i) \prec P(j)$.

**Theorem 1.3.** Suppose, for $i = 1, 2$, $T_i = (\mathcal{N}_i, p_i, \alpha_i)$ is a tree and $\prec_i$ is a tree ordering of $T_i$. Then $T_i, i = 1, 2,$ are $(\prec_1, \prec_2)$ isomorphic if and only if they have the same tree codes.

**Proof.** Let $\iota : \mathcal{N}_1 \rightarrow \mathcal{N}_2$ be a $(\prec_1, \prec_2)$ isomorphism and let $T_i, i = 1, 2,$ be the $\prec_i$ tree codes of $T_i$ respectively. Then

$$P_2 = W_2^{-1} \circ p_2 \circ W_2$$

$$= W_2^{-1} \circ \iota \circ p_1 \circ \iota^{-1} \circ W_2$$

$$= W_2^{-1} \circ \iota \circ W_1 \circ P_1 \circ W_1^{-1} \circ \iota^{-1} \circ W_2$$

$$= (W_2^{-1} \circ \iota \circ W_1) \circ P_1 \circ (W_1^{-1} \circ \iota^{-1} \circ W_2)$$

Since $W_2^{-1} \circ \iota \circ W_1$ and $W_1^{-1} \circ \iota^{-1} \circ W_2$ are order preserving maps of $I[1, N)$ to itself they are equal to the identity map of $I[1, N)$. □
1.6. Growing trees. Suppose 
\[ T = (N, o, p) \]

is a tree and 
\[ \mathcal{O} \quad \text{and} \quad \mathcal{P} \]
satisfy the following conditions:

(i) \( \mathcal{O} \) and \( \mathcal{P} \) are functions with domain the leaf nodes of \( T \);
(ii) for each \( \nu \in I(T) \), \( (\mathcal{O}(\nu), \nu, \mathcal{P}(\nu)) \) is a tree;
(iii) the family 
\[ \{ i(T) \} \cup \{ \mathcal{O}(\nu) : \nu \in I(T) \} \]
is disjointed.

Let 
\[ U = i(T) \cup \left( \bigcup \{ \mathcal{O}(\nu) : \nu \in I(T) \} \right) \]
and let 
\[ q = p \cup \left( \bigcup \{ \mathcal{P}(\nu) : \nu \in I(T) \} \right) . \]

We leave it as an exercise for the reader to verify that 
\[ (U, o, q) \]
is a tree; that \( N \subset U \); and that, for each \( \nu \in I(T) \), \( (\mathcal{O}(\nu), \nu, \mathcal{P}(\nu)) \) is the subtree associated to the node \( \nu \) of \( U \).

Now let us suppose that 
\[ (N, o, p, <) \]
is an ordered tree and that, for each \( \nu \in I(T) \), \( <_{\nu} \) is such that 
\[ (\mathcal{O}(\nu), \nu, \mathcal{P}(\nu), <_{\nu}) \]
is an ordered tree. We leave it to the reader to verify that there is one and only one 
\[ < \]
such that

(i) \( (U, o, q, <) \) is an ordered tree;
(ii) if \( \mu, \xi \in N \) and \( \mu < \xi \) then \( \mu < \xi \);
(iii) if \( \nu \in I(T) \), \( \mu, \xi \in \mathcal{O}(\nu) \) and \( \mu <_{\nu} \xi \) then \( \mu < \xi \).

2. Context free grammars.

For a set \( A \) we let \( A^0 = \{ \text{null} \} \) and we let 
\[ A^* = \bigcup_{n=0}^{\infty} A^n \]
. For \( a \in A^* \) we let 
\[ |a| \in \mathbb{N} \]
be such that \( |a| = n \) if \( a \in A^n \).

If \( a, b \in A^* \) we define the **concatenation**
\[ ab \in A^* \]
by requiring that
\[ ab = b \quad \text{if} \quad a = \text{null}; \quad ab = a \quad \text{if} \quad b = \text{null}; \]
and that
\[(a b)_i = \begin{cases} a_i & \text{if } i \in [1, M], \\ b_{i-M} & \text{if } i \in (M, M+N]. \end{cases}\]

Obviously,
\[|a b| = |a||b|.

**Definition 2.1.** By a context free grammar ordered triple 
\[G = (T, N, \mathcal{P})\]
such that
(i) \(T\) is a set, \(N\) is a set and \(T \cap N = \emptyset\);
(ii) \(\mathcal{P} \subseteq N \times (T \cup N)^*\); 
(iii) \(\mathcal{P} \subseteq (T \cup N)^*\).

The members of \(T\) are called tokens or terminal symbols. The members of \(N\) are called nonterminals or nonterminal symbols. The members of \(\mathcal{P}\) are called productions. For \(A, B \in (T \cup N)^*\) we write 
\[A \Rightarrow_B B\]
if there are a production \((n, s) \in \mathcal{P}\) and \(A_L, A_R, B_L, B_R \in (T \cup N)^*\) such that
\[A = A_L n A_R \quad \text{and} \quad B = B_L s B_R.

Suppose \(T_A = (N_A, o_A, p_A)\) is a tree such that there is a univalent function \(L_A\) with domain \(\text{rng} A\) and range the leaf nodes of \(T_A\). We Let \(\nu = L_A(n)\); let \(N_B = N_A \cup (\text{rng} s \times \{\nu\})\); let \(o_B = o_A\); let \(p_B = p_A \cup ((\text{rng} s \times \{\nu\}) \times \{\nu\})\); and let
\[L_B = (L_A|N_A \sim \{\nu\}) \cup (\text{rng} s \times \{\nu\}).\n
We say the nonterminal \(n\) generates \(t \in T^*\) if there is a finite sequence \(A_1, \ldots, A_N\) such that 
\[n = A_1 \Rightarrow_{G} \cdots \Rightarrow_G A_N = t.\n
For each

Instead of writing \((r, s) \in \mathcal{P}\) one often writes 
\[r := \epsilon \quad \text{if } |s| = 0\]
and
\[r := s_0 s_1 \ldots s_{|s|-1} \quad \text{if } |s| > 0.\n
If \(r \in N, n \in \mathbb{N}^+\) and \(s_0, \ldots, s_{n-1} \in (T \cup N)^*\) one often writes 
\[r := s_0 | s_1 | \cdots | s_{n-1}\]
instead of 
\[(r, s_i) \in \mathcal{P}, \quad i \in I(n).\n
Obviously, if \((T, N, s, \mathcal{P})\) is a context free grammar then so is \((T, N, t, \mathcal{P})\) if \(t \in N\).

**Definition 2.2.** A parse tree \(Q\) for the context free grammar \(G = (T, N, s, \mathcal{P})\) is an ordered quintuple 
\[(N, o, p, <, f)\]
such that
(i) \((N, o, p, <)\) is an ordered tree;
(ii) \(f : N \to T \cup N\);
(a) if \(\nu \in I(T)\) then \(f(\nu) \in N\);
(b) if \(\nu \in I(T)\) then \(f(\nu) \in T\);
(iii) if \( \nu \in \mathbf{i}(T) \) and \( \nu \) has \( m \) children
\[
\mu_0 < \mu_2 < \ldots < \mu_{m-1}
\]
then
\[
(f(\nu), (f(\mu_0), \ldots, f(\mu_{m-1}))) \in \mathcal{P}.
\]
Notice that the notion of parse tree is independent of \( s \).
We define
\[
< \mathcal{Q} > \in (T)^*
\]
as follows. Let
\[
\mathcal{L} = \{ \nu \in \mathbf{i}(T) : f(\nu) \neq \epsilon \}
\]
and let
\[
n = |\mathcal{L}|.
\]
If \( n = 0 \) we let
\[
< \mathcal{Q} > = \epsilon
\]
and if \( n > 0 \) and
\[
\nu_0 < \nu_1 < \ldots < \nu_{n-1}
\]
are the members of \( \mathcal{L} \) we let
\[
< \mathcal{Q} > = (f(\nu_0))(f(\nu_1)) \cdots (f(\nu_{n-1})).
\]
For each \( t \in N \) we let
\[
L(\mathcal{G}, t)
\]
be the set of \( < \mathcal{Q} > \) as above where \( f(o) = t \). We let
\[
L(\mathcal{G}) = L(\mathcal{G}, s)
\]
and we call this language on the alphabet \( T \) the language generated by \( \mathcal{G} \).

**Definition 2.3.** We say the context free grammar \( \mathcal{G} \) is **good** if \( \mathcal{Q}_i = (N_i, o_i, p_i, <_i, f_i), i = 1, 2 \), are parse trees such that \( < \mathcal{Q}_1 > = < \mathcal{Q}_2 > \) then the ordered trees \((N_i, o_i, p_i, <_i)\) are isomorphic; recall that this is the case if and only if they have the same tree codes. We say \( \mathcal{G} \) is **bad** if it is not good.

2.1. A **bad grammar**. Let \( \mathcal{B} \) be the context free grammar defined as follows. Let
\[
T = N \cup \{ -, +, * \}
\]
and let
\[
N = \{ \text{expr} \}.
\]
Let \( \text{expr} \) be the start symbol. Let the productions be given by
\[
\begin{align*}
\text{expr} & := j \quad \text{for } j \in N \\
\text{expr} & := - \text{ expr} \\
\text{expr} & := \text{expr} + \text{expr} \\
\text{expr} & := \text{expr} * \text{expr}
\end{align*}
\]
Note that there are an infinite number of productions.
This grammar derives the string \( 27 + 5 * 298 \) with two nonisomorphic parse trees so it is bad.

There are a number of ways to deal with this problem. One is to introduce parentheses which we now do. As we shall see, there are other ways to deal with this problem.