1. Trees; context free grammars.

1.1. Trees.

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

\[ T = (N, o, p) \]

such that

(i) \( N \) is a set;

(ii) \( o \in N \);

(iii) \( p : N \sim \{o\} \rightarrow N \);

(iv) if \( \nu \in N \sim \{o\} \) there is \( d \in \mathbb{N}^+ \) such that \((\nu, o) \in p^d\).

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

The members of \( N \) are called **nodes**. \( o \) is called the **root node**. If \( n \in N \sim \{o\} \) we call \( p(n) \) the **parent of** \( n \). If \( n \in N \) we let \( c(n) = p(n) \) = \( \{ \mu \in N \sim \{o\} : p(\mu) = n \} \) and call the members of this set the **children of** \( n \). 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 \( p^{-n}[\{o\}] \subset N \sim \{o\} \).

**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

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

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 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 \( N \).
Suppose \( A \) and we let \( T = (I, D) \).

**Proof.** \( o \in \text{dmn} \, d \) since \((o, 0) \in d\). If \( \nu \in \mathcal{N} \sim \{o\} \) there is by (iv) \( n \in \mathbb{N}^+ \) such that \((\nu, o) \in p^n \) so \( \nu \in p^{-n}[\{o\}] \subset \text{dmn} \, d \).

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

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

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 \( o \in \text{dmn} \, p^{n_2-n_1} \subset \text{rng} \, p \) which is incompatible with Lemma 1.1(iii). \( \square \)

**Corollary 1.1.** \( \{p^{-d}[\{o\}] : d \in \mathbb{N}\} \) is a partition of \( \mathcal{N}. \)

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

**Proposition 1.1.** \( A(o) = \emptyset. \)

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

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

**Proof.** \( \mu \in \text{dmn} \, d \) by Theorem ??.

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 \( \nu \in p^{n-d(\mu)}[\{o\}]. \) Since \( o \notin \text{dmn} \, p \) we find that \( n - d(\mu) < 0 \) and \( d(\nu) = d(\mu) - n \in \|0, d(\mu)\|. \) \( \square \)

### 1.2. Subtrees.

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

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

### 1.3. Isomorphisms.** Suppose \( \mathcal{T}_i = (\mathcal{N}_i, o_i, p_i), \ i = 1, 2 \) are rooted trees and \( \iota: \mathcal{N}_1 \to \mathcal{N}_2. \) We say \( \iota \) is an **isomorphism from \( \mathcal{T}_1 \) to \( \mathcal{T}_2 \)** if \( \iota \) is univalent, \( \text{rng} \, \iota = \mathcal{N}_2 \)
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 \( \mathcal{T}_2 \) to \( \mathcal{T}_1. \)
1.4. Tree orderings. Suppose $T = (N, o, p)$ is a tree. We say $\prec$ is a tree ordering of $T$ if

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

**Proposition 1.3.** Suppose $W$ is a function with domain $N$ whose value at $n \in N$ 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 $n \in N$ and $\mu, \nu \in d^{-1}\{n + 1\}$. Then

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

is a tree ordering of $T$.

**Theorem 1.2.** Suppose for each $\xi \in N$ 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 $\xi \in N$ and each $\mu, \nu \in c(\xi)$,

$$\mu \prec_\xi \nu \iff \mu \prec \nu.$$  

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

$$\prec = \{(\mu, \nu) \in N \times N : d(\mu) < d(\nu)\} \cup \left(\bigcup_{n \in N} W_n\right).$$

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

Suppose $O_i = (N_i, o_i, p_i, \prec_i)$, $i = 1, 2$ are ordered trees and $\iota : N_1 \to N_2$. We say $\iota$ is an isomorphism from $O_1$ to $O_2$ if $\iota$ is an isomorphism from $(N_1, o_1, p_1)$ to $(N_2, o_2, p_2)$ and

$$\mu, \nu \in N_1 \text{ and } \mu \prec_1 \nu \Rightarrow \iota(\mu) \prec_2 \iota(\nu).$$

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

$$\prec$$

$$\{(\mu, \nu) \in O \times O : \mu \prec \nu\}$$

and note that $(O, \sigma, q, \prec)$ is an ordered tree which we call the **ordered tree associated to $U$**. In particular, if $\nu \in N$, $O = N_\nu$, $\sigma = \nu$ 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 $P : I \sim \{0\} \to I$ such that

\[(3) \quad P^{-1}[\{0\}] \neq \emptyset;\]
\[(4) \quad i \in I \sim \{0\} \Rightarrow P(i) < i;\]
\[(5) \quad i, j \in I \sim \{0\} \text{ and } i < j \Rightarrow P(i) \leq P(j).\]

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

$$W : I[0, N) \to N$$

be such that

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

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 $\mu < \nu$. 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 $P(i) < P(j)$. □

**Theorem 1.3.** Suppose, for $i = 1, 2$, $T_i = (N_i, p_i, o_i)$ 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 $\iota : N_1 \to N_2$ 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

$$P_2 = W_2^{-1} \circ p_2 \circ W_2$$
$$= W_2^{-1} \circ o \circ p_1 \circ o^{-1} \circ W_2$$
$$= W_2^{-1} \circ o \circ W_1 \circ p_1 \circ W_1^{-1} \circ o^{-1} \circ W_2$$
$$= (W_2^{-1} \circ o \circ W_1) \circ p_1 \circ (W_1^{-1} \circ o^{-1} \circ W_2);$$

Since $W_2^{-1} \circ o \circ W_1$ and $W_1^{-1} \circ o^{-1} \circ W_2$ 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

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

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 \( \mathcal{T} \);

(ii) for each \( \nu \in \text{I}(\mathcal{T}) \), \((\mathcal{O}(\nu), \nu, \mathcal{P}(\nu)) \) is a tree;

(iii) the family

\[ \{ \text{i}(\mathcal{T}) \} \cup \{ \mathcal{O}(\nu) : \nu \in \text{I}(\mathcal{T}) \} \]

is disjoint.

Let

\[ \mathcal{U} = \text{i}(\mathcal{T}) \cup \left( \bigcup \{ \mathcal{O}(\nu) : \nu \in \text{I}(\mathcal{T)} \} \right) \]

and let

\[ q = p \cup \left( \bigcup \{ \mathcal{P}(\nu) : \nu \in \text{I}(\mathcal{T)} \} \right) . \]

We leave it as an exercise for the reader to verify that

\[ (\mathcal{U}, \text{o}, q) \]
is a tree; that \( \mathcal{N} \subseteq \mathcal{U} \); and that, for each \( \nu \in \text{I}(\mathcal{T}) \), \((\mathcal{O}(\nu), \nu, \mathcal{P}(\nu)) \) is the subtree associated to the node \( \nu \) of \( \mathcal{U} \).

Now let us suppose that

\[ (\mathcal{N}, \text{o}, p, <) \]
is an ordered tree and that, for each \( \nu \in \text{I}(\mathcal{T}) \), \(<_\nu\) 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) \( (\mathcal{U}, \text{o}, q, <) \) is an ordered tree;

(ii) if \( \mu, \xi \in \mathcal{N} \) and \( \mu < \xi \) then \( \mu < \xi \);

(iii) if \( \nu \in \text{I}(\mathcal{T}) \), \( \mu, \xi \in \mathcal{O}(\nu) \) and \( \mu <_\nu \xi \) then \( \mu < \xi \).

2. **Context free grammars.**

**Definition 2.1.** By a **context free grammar** ordered triple

\[ \mathcal{G} = (T, N, s, \mathcal{P}) \]
such that

(i) \( T \) is a set;

(ii) \( N \) is a set, \( T \cap N = \emptyset \) and \( s \in N \).

(iii) \( \mathcal{P} \subset N \times (T \cup N)^* \);
The members of $T$ are called **tokens** or **terminal symbols**. The members of $N$ are called **nonterminals** or **nonterminal symbols**. $s$ is called the **start symbol**.

The members of $P$ are called **productions**. Instead of writing $(r, s) \in P$ one often writes
\[ r := \epsilon \quad \text{if } |s| = 0 \]
and
\[ r := s_0 \ s_1 \cdots s_{|s|-1} \quad \text{if } |s| > 0. \]
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
\[ (r, s_i) \in P, \quad i \in I(n). \]
Obviously, if $(T, N, s, P)$ is a context free grammar then so is $(T, N, t, P)$ if $t \in N$.

**Definition 2.2.** A parse tree $Q$ for the context free grammar $G = (T, N, s, P)$ 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 $\nu \in i(T)$ and $\nu$ has $m$ children
        \[ \mu_0 < \mu_2 < \cdots < \mu_{m-1} \]
        then
        \[ (f(\nu), (f(\mu_0), \ldots, f(\mu_{m-1}))) \in P. \]
Notice that the notion of parse tree is independent of $s$.
We define
\[ <Q> \in (T)^* \]
as follows. Let
\[ \mathcal{L} = \{ \nu \in i(T) : f(\nu) \neq \epsilon \} \]
and let
\[ n = |\mathcal{L}|. \]
If $n = 0$ we let
\[ <Q> = \epsilon \]
and if $n > 0$ and
\[ \nu_0 < \nu_1 < \cdots < \nu_{n-1} \]
are the members of $\mathcal{L}$ we let
\[ <Q> = (f(\nu_0))(f(\nu_1)) \cdots (f(\nu_{n-1})). \]
For each $t \in N$ we let
\[ \mathbf{L}(G, t) \]
be the set of $<Q>$ as above where $f(o) = t$. We let
\[ \mathbf{L}(G) = \mathbf{L}(G, s) \]
and we call this language on the alphabet $T$ the **language generated by $G$**.
Definition 2.3. We say the context free grammar $G$ is **good** if $Q_i = (N_i, o_i, p_i, <_i, f_i)$, $i = 1, 2$, are parse trees such that $<_1 = <_2$ 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 $G$ is **bad** if it is not good.

2.1. A bad grammar. Let $B$ be the context free grammar defined as follows. Let $T = \mathbb{N} \cup \{-, +, \ast\}$

and let

$$N = \{ \text{expr} \}.$$ 

Let $\text{expr}$ be the start symbol. Let the productions be given by

- $\text{expr} := j$ for $j \in \mathbb{N}$
- $\text{expr} := -$ \text{expr}
- $\text{expr} := \text{expr} + \text{expr}$
- $\text{expr} := \text{expr} \ast \text{expr}$

Note that there are an infinite number of productions.

This grammar derives the string $27 + 5 \ast 298$ 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.