Árbol

Un **árbol** es un modelo abstracto de una estructura jerárquica, impuesta por existencia de un camino único entre cualquier nodo a la raíz.

En esta estructura, la raíz se encuentra en lo más alto del árbol y los otros nodos están organizados en niveles, de acuerdo con su distancia hasta la raiz.

Definition

Formalmente un **árbol libre** es un grafo no-orientado, conexo y sin ciclos denotado por $T = (V, E)$, donde V es un conjunto de nodos y E es un conjunto de arcos o aristas.

Árborescencia

Si se permite que un nodo $u \in V$ sea distinguido como la **raíz** de T, entonces se puede decir que T es una **arborescencia**, definida por la tripleta $T = (V, E, r)$, donde :

- E describe la relación binaria de **parentesco** entre los nodos de T, es decir, cada arista $(u, v) \in E$ es una relación *(padre, hijo)*,
- La raiz no tiene un padre, es decir, $\nexists (u, r) \in E$ tal que $u, r \in V$,
- Cada nodo $u \in V \setminus \{r\}$ tiene un padre único, y
- Sólo existe un camino único $[r, \ldots, v]$ para todo $v \in V$.

Árbol vs Árborescencia

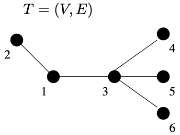


FIG. 2.2 – Arbre libre

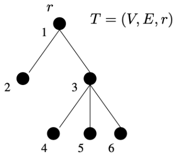


FIG. 2.3 – Arborescence

La raíz r de un árbol T es denotada por $root(T)$.

Propiedades de lo árboles

Dado un árbol T de k nodos (diremos que es un k-árbol o un árbol sobre $[k]$) su tamaño es determinado por $|T| = |V|$, igualmente escrito como $size(T) = |V|$.

Si un camino $[r, \ldots, u]$, de la raíz a un nodo $u \in V$, está formado por k arcos, entonces la *profundidad* de u es igual a k y es denotada por $depth(u) = k$.

La profundidad de un árbol T está determinada por la máxima profundidad de sus hojas. Formalmente, $depth(T) = max\{depth(u) | u \in V\}$.

Relaciones de parentesco

Dado un árbol $T = (V, E, r)$, se pueden definir las siguientes relaciones en función de un camino único $[r, \ldots, u]$, para todo $u \in V$:

- El *padre* de un nodo $u \in V \setminus \{r\}$, denotado por $parent(u)$, es el nodo adyacente a u en el camino $[r, \ldots, u]$ tal que $parent(u) = \{v \in V | (v, u) \in E\}$. El padre de la raíz es indefinido, por lo tanto $parent(r) = \emptyset$.
- Si para una arista $(v, u) \in E$, decimos que v es el padre de u, entonces u es el *hijo* de v. El conjunto de hijos de un nodo $v \in V$ es definido por $children(v) = \{u \in V | (v, u) \in E\}$. Si $children(u) = \emptyset$ entonces el nodo u es llamado una *hoja* , de lo contrario es referido como un *nodo interno*.

Relaciones ancestrales

- Las relaciones ancestrales de un árbol T son codificadas por la clausura transitiva E^+ de sus aristas :

$$\begin{aligned} E^0 &= \{(u,u) | u \in V\}, \\ E^{n+1} &= \{(u,w) | \exists v \in V \text{ tal que } (u,v) \in E \wedge (v,w) \in E^n\}, \text{ para } n \geq 0, \\ E^{\leq i} &= \cup_{n>0}^{i} E^n, \\ E^+ &= \cup_{n>0} E^n \end{aligned}$$

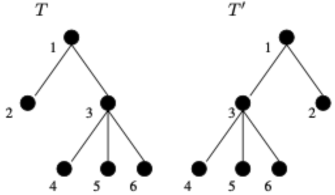
  Los *ancestros* de un nodo $u \in V \setminus \{r\}$, son todos los nodos del camino $[r, \ldots, u]$ excepto u. El conjunto de ancestros de u está definido por la función $ancestors(u) = \{v \in V | (v,u) \in E^+\}$.

- Si v es un ancestro de u, entonces decimos que u es un *descendiente* de v. El conjunto de descendientes de v está definido por la función $descendants(v) = \{u \in V | (v,u) \in E^+\}$.

Relaciones de orden

Decimos que un árbol es *ordenado* si se establece una relación de orden parcial $\preceq$ entre los hijos de un nodo $u \in V$ tal que ellos puedan ser ordenados del *primero* al *último* de los hijos.

Dicho de otra forma, existe un orden relativo $\preceq$ entre los subárboles $T_1, \ldots, T_n$ que componen a un árbol T. Considerando esta relación de orden, un *árbol ordenado* es definido por $T = (V, E, r, \preceq)$.



Relaciones de hermandad

Dado un árbol ordenado $T = (V, E, r, \preceq)$, los nodos $w \in V$ y $u, v, y \in V \setminus \{r\}$, y las relaciones de parentesco $(w, u), (w, v), (w, y) \in E$, es posible definir las funciones siguientes para establecer una distinción entre los hijos de w :

- Los nodos $u, v \in V \setminus \{r\}$ sont *hermanos* si $parent(u) = parent(v)$. El conjunto de hermanos de un nodo, es definido por $siblings(u) = \{v \in V | parent(u) = parent(v)\}$.
- El hijo de *más a la izquierda* de w está definido por $first(w) = \{u | \nexists v, (w, v) \in E \wedge v \preceq u\}$.
- El hijo de *más a la derecha* de w esta definido por $last(w) = \{v | \nexists u, (w, u) \in E \wedge v \preceq u\}$.

Relaciones de hermandad ...

- El *hermano siguiente* de u, está definido $next(u) = \{v|(w,v) \in E \wedge u \preceq v \wedge \nexists y$ tal que $(w,y) \in E \wedge u \preceq y \preceq v\}$.

- El nodo u es el *hermano anterior* de v, denotado por $previous(v)$, si $next(u) = \{v\}$.

- La *rama más a la izquierda* de T, denotada $lmb(T)$, corresponde al camino $[v_1, v_2, \ldots, v_k]$, tal que $v_1 = r$, y $first(v_{i-1}) = \{v_i\}, \forall i, 2 \leq i \leq k$ y que $deg(v_k) = 1$.

- El nodo terminal v_k dentro de $lmb(T)$ es llamado la *hoja más a la izquierda* y es denotada por $lml(T)$.

Relaciones de hermandad . . .

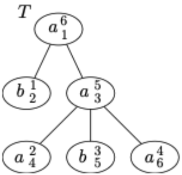
- La *rama más a la derecha* de T, denotada por $rmb(T)$, corresponde al camino $[v_1, v_2..., v_k]$, tal que $v_1 = r$ y que $last(v_{i-1}) = \{v_i\}, \forall i, 2 \leq i \leq k$ et $deg(v_k) = 1$.

- El nodo terminal v_k dentro de $rmb(T)$ es llamado la *hoja más a la derecha* y es denotada por $rml(T)$.

Étiquetas

Dado un *alfabeto* $\Sigma = \{a, b, c, ...\}$, se dice que un árbol es *etiquetado* si a cada nodo $u \in V$ le corresponde una etiqueta $l \in \Sigma$.

Es definido por $T = (V, E, r, \lambda)$ donde $\lambda : V \to \Sigma$ tal que $\lambda(u) = l$.



Arbre étiqueté T sur $\Sigma = \{a, b\}$

Representación recursiva

Un árbol puede ser definido de forma recursiva, en función de los sub-árboles que lo componen. Decimos que una árbol es un conjunto finito de nodos si:

1. Existe un nodo especial, llamado raíz.
2. Los nodos restantes son divididos en conjuntos considerados a sí mismos como árboles.

Lo anterior permite realizar una representación de un árbol utilizando paréntesis anidados. Consideremos un árbol T con sus sub-árboles inmediatos $< T_1, T_2, \ldots, T_n > n \geq 0$ cuya etiqueta en la raiz es $\lambda(root(T)) = l$, entonces el árbol T es representado por $T = (l(T_1)\ldots(T_n))$.

El árbol de la figura anterior es representado por $T = (a(b)(a(a)(b)(a)))$

Recorrido de un árbol

- Un proceso para visitar los nodos en un cierto orden es llamado *recorrido*.

- Un recorrido que visita cada nodo de un árbol exactamente una vez, es llamado *enumeración*.

- Los nodos de los árboles pueden ser enumerados, recorriendo de forma vertical (en profundidad) u horizontal (en amplitud).

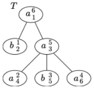
- Un primer método permite recorrer el árbol de sub-árbol en sub-árbol, mientras que un segundo método lo hace de nivel en nivel.

- Un árbol $T = (V, E, r)$ se recorre estableciendo una biyección de orden sobre n nodos, tal que : $V \to \{1, ..., k\}$, de manera que el orden para el primer nodo u sea $ordre[u] = 1$, para el segundo nodo v, $ordre[v] = 2$, y así continuando hasta el último nodo w con un orden $ordre[w] = k$.

Recorrido de un árbol en pre-orden

pre(T)
Data: Dado un $T = (V, E, r)$ sobre k nodos
Result: Una biyección de orden $pre(T) : V \rightarrow \{1, \ldots, k\}$
tal que para todo $u \in V$

$pre(r) = 1;$
if *u no es una hoja* **then**
  $pre(first(u)) = pre(u) + 1;$
else if *u no es el último hijo* **then**
  $pre(next(u)) = pre(u) + size(T[u]);$



Arbre étiqueté T sur $\Sigma = \{$