Árboles

Clase 25

IIC 1253

Prof. Pedro Bahamondes

Outline

¿Por qué estudiar árboles?

- **Sistemas de archivos**: Carpetas y archivos jerárquicamente
- **Estructuras de datos**: Árboles binarios de búsqueda, tries, ...
- **Abstract Syntax Trees (AST)**: Representación de expresiones y programas en compiladores y análisis estático.
- **Árboles de decisión**: Para inteligencia artificial y aprendizaje automático.
- **Árboles de juego**: De posibles jugadas en juegos como ajedrez o Go.
- **Árboles filogenéticos**: Relaciones evolutivas entre especies.
- **Desglose estructurado del trabajo (WBS)**: Herramienta de planificación de proyectos.
- **Protocolos de enrutamiento**: Algunas redes usan árboles para determinar rutas óptimas.

Comprender árboles permite abordar problemas complejos con estructuras claras y eficientes

Objetivos de la clase

- Conocer definiciones de árbol y demostrar que son equivalentes
- Conocer elementos esenciales de un árbol
- Demostrar propiedades que relacionan elementos de un árbol

Outline

Árboles

Definición

Un grafo $T = (V, E)$ es un **árbol** si para cada par de vértices $x, y \in V$ existe un único camino entre ellos.

Árboles

Definición

Un grafo $T = (V, E)$ es un **árbol** si para cada par de vértices $x, y \in V$ existe un único camino entre ellos.

Por lo tanto, siempre es *conexo*. Si relajamos esta condición:

Definición

Un grafo $T = (V, E)$ es un **bosque** si para cada par de vértices $x, y \in V$, si existe un camino entre ellos, este es único.

Un bosque se ve como un conjunto de árboles.

Árboles

En general hablaremos de árboles **con raíz**.

- Distinguimos uno de los vértices $r \in V$, al que llamaremos la **raíz** del árbol.
- Los vértices de grado menor o igual a 1 se llaman **hojas**.
- Los dibujamos con la raíz arriba y los demás vértices hacia abajo.

Árboles

Hay muchas definiciones equivalentes para los árboles:

Definición

Un grafo $T = (V, E)$ es un **árbol** si y sólo si es conexo y acíclico.

Definición

Un grafo $T = (V, E)$ es un **árbol** si y sólo si es conexo y todas sus aristas son de corte.

Ejercicio

Demuestre las definiciones anteriores.

Árboles

Definición

Un grafo $T = (V, E)$ es un **árbol** si y sólo si es conexo y acíclico.

<u>Demostración:</u>

($\Rightarrow$) Primero si T es un árbol es por definición conexo, nos falta demostrar entonces que un árbol no puede tener ciclos. Supongamos que T tuviese un ciclo, y sea C un ciclo en T que pasa por los vértices u y v. Supongamos que C parte (y termina) en u, entonces C es de la forma $(u, \ldots, v, \ldots, u)$, por lo que se puede dividir en dos porciones, una para ir de u a v, digamos p_1, y otra (distinta ya que un ciclo no repite aristas) para ir de v a u, digamos p_2. Resulta entonces que p_1 y p_2 son dos caminos distintos entre u y v en T, lo que contradice el hecho de que T es un árbol. Finalmente T no puede tener ciclos.

Árboles

($\Leftarrow$) Como T es conexo, para cada par de vértices existe un camino que los une. Falta demostrar que ese camino es único. Supongamos entonces que T no tiene ciclos pero que sin embargo existe un par de vértices con dos caminos distintos uniéndolos en T. Sean u y v estos vértices y sean p_1 y p_2 los dos caminos distintos en T que unen a u con v. Dado que estos caminos son distintos entonces ambos tienen al menos tres vértices.

Sea x el vértice anterior al primer vértice que diferencia a p_1 y p_2 (note que x está en p_1 y en p_2). Sea y el vértice siguiente a x que pertenece simultáneamente a p_1 y p_2. El camino entre x e y a través de p_1 junto con el camino entre x e y a través de p_2 forman un ciclo en T lo que contradice nuestra hipótesis de que T no tiene ciclos. Finalmente no pueden existir dos caminos distintos entre u y v, de donde concluimos que para todo par de vértices en T existe un único camino que los une y por lo tanto T es un árbol.

Árboles

Definición

Un grafo $T = (V, E)$ es un **árbol** si y sólo si es conexo y todas sus aristas son de corte.

Demostración:

En la sección anterior demostramos que una arista es de corte si y sólo si no pertenece a ningún ciclo en el grafo. Ahora, T es un árbol si y sólo si T es conexo y no tiene ningún ciclo, si y sólo si todas sus aristas cumplen con la propiedad de no pertenecer a un ciclo, si y sólo si, todas sus aristas son de corte.

Árboles

Vimos que un grafo es bipartito si y sólo si no tiene ciclos de largo impar.
De esto se deduce inmediatamente que:

Teorema

Todo árbol es un grafo bipartito.

Árboles

La siguiente propiedad nos permitirá hacer demostraciones por inducción sobre los árboles:

Teorema

Si T es un árbol y v es una hoja de él, entonces el grafo $T - v$ (el grafo que resulta de quitar el vértice y sus aristas incidentes) es un árbol.

Ejercicio

Demuestre el teorema.

Árboles

Teorema

Si T es un árbol y v es una hoja de él, entonces el grafo $T - v$ es un árbol.

Demostración:

Para demostrar que el grafo $T - v$ es un árbol debemos comprobar que para cualquier par de vértices en $T - v$, existe un único camino que los une. Sea u y w dos vértices en T distintos de v, y sea la secuencia $P = (u, u_1, u_2, ..., u_n, w)$ el único camino en T que une a u con w. Es claro que el vértice v no aparece en P ya que todos los vértices de P (excepto u y w) deben tener grado al menos 2, luego si eliminamos v de T no afecta al camino entre u y w, luego el camino $P = (u, u_1, u_2, ..., u_n, w)$ entre u y w también existe en $T - v$. Como la demostración la hicimos en general para un par de vértices cualquiera, en $T - v$ existe un único camino entre todo par de vértices y por lo tanto $T - v$ también es un árbol.

Árboles

Ahora podemos establecer una última definición muy simple de un árbol:

Definición

Un grafo $T = (V, E)$ con n vértices es un **árbol** si y sólo si es conexo y tiene exactamente $n - 1$ aristas.

Ejercicio

Demuestre que la definición anterior es equivalente a las demás.

Podemos determinar si un grafo conexo es o no un árbol.

Árboles

Definición

Un grafo $T = (V, E)$ con n vértices es un **árbol** si y sólo si es conexo y tiene exactamente $n - 1$ aristas.

Demostración:

($\Rightarrow$) Si T es un árbol con n vértices, entonces claramente es conexo, falta mostrar que tiene exactamente $n - 1$ aristas, lo haremos por inducción en n.

BI: Si $n = 1$ tenemos un árbol con sólo un vértice y sin aristas, por lo que se cumple la propiedad: $|E| = 0 = 1 - 1 = n - 1$.

HI: Supongamos que un árbol con n vértices tiene $n - 1$ aristas.

TI: Sea ahora T un árbol con $n + 1$ vértices, queremos demostrar que T tiene exactamente $(n + 1) - 1 = n$ aristas. Centrémonos en una hoja v cualquiera. Por el lema anterior $T - v$ también es un árbol y tiene exactamente n vértices por lo que se aplica la HI, luego $T - v$ tiene exactamente $n - 1$ aristas. Dado que v es una hoja, v tiene grado 1 en T y por lo tanto T tiene exactamente una arista más que $T - v$, o sea T tiene exactamente n aristas, por lo que se cumple la propiedad.

Árboles

($\Leftarrow$) En la sección anterior demostramos que un grafo con n vértices y k aristas tiene al menos $n-k$ componentes conexas. Si T es un grafo conexo con n vértices y exactamente $n-1$ aristas y tomamos una arista e cualquiera de T, entonces dado que $T-e$ tiene $n-2$ aristas, por el teorema mencionado, $T-e$ tiene al menos dos componentes conexas y por lo tanto e es una arista de corte. Dado que elegimos e como una arista cualquiera, T cumple con que todas sus aristas son de corte y por lo tanto T es un árbol.

Outline

Árboles

Las siguientes definiciones se usan mucho en aplicaciones de los árboles en computación.

Definición

Sea $T = (V, E)$ un árbol con raíz r y x un vértice cualquiera.

- La **profundidad** de x es el largo del camino que lo une con r (r tiene profundidad 0).
- La **altura** o **profundidad** del árbol es el máximo de las profundidades de sus vértices.
- Los **ancestros** de x son los vértices que aparecen en el camino entre él y r. Note que x es ancestro de sí mismo.
- El **padre** de x es su ancestro (propio) de mayor profundidad. Diremos que x es **hijo** de su padre.
- Dos vértices x e y con el mismo padre son **hermanos**.

Árboles binarios

Definición

Un árbol con raíz se dice **binario** si todo vértice tiene grado a lo más 3; o equivalentemente, si todo vértice tiene a lo más dos hijos.

Podemos distinguir entre hijos izquierdos y derechos

Teorema

La cantidad de vértices sin hijos de un árbol binario es la cantidad de vértices con exactamente dos hijos más 1.

Ejercicio

Demuestre el teorema.

Árboles binarios

Teorema

La cantidad de vértices sin hijos de un árbol binario es la cantidad de vértices con exactamente dos hijos más 1.

Demostración:

Por inducción en la cantidad de vértices del árbol binario.

BI: El caso base es un árbol compuesto por sólo un vértice, la raíz. Un árbol de estas características tiene sólo una hoja y ningún vértice con dos hijos, luego cumple la propiedad.

HI: Supongamos que un árbol binario con n vértices tiene una hoja más que vértices con dos hijos.

TI: Sea T un árbol binario con $n+1$ vértices. Sea v una hoja de T, sabemos que $T-v$ es también un árbol binario y tiene exactamente n vértices por lo que $T-v$ cumple con HI, o sea tiene una hoja más que vértices con dos hijos. Supongamos que $T-v$ tiene k vértices con dos hijos entonces por HI tiene $k+1$ hojas. Lo que podamos decir dependerá de si v tenía o no un hermano.

Árboles binarios

- Si v tiene un hermano en T, entonces el padre de V es un vérticee con dos hijos en T. Ahora, en el árbol $T - v$, el vértice que era padre de v tiene sólo un hijo. Lo anterior quiere decir que T tiene exactamente un vértice más con dos hijos que $T - v$, o sea que T tiene exactamente $k + 1$ vértices con dos hijos. Ahora también ocurre que T tiene exactamente una hoja más que $T - v$, o sea que T tiene $k + 2$ hojas. Hemos concluido que T tiene $k + 2$ hojas y $k + 1$ vértices con dos hijos y por lo tanto cumple con la propiedad.
- Si v no tiene hermano, entonces el vértice padre de v en T se convierte en una hoja en el árbol $T - v$, lo que quiere decir que T y $T - v$ tienen exactamente la misma cantidad de hojas, $k + 1$. El único vértice que ve afectado su cantidad de hijos en $T - v$ es el padre de v, este tiene exactamente un hijo en T y 0 hijos en $T - v$ por lo que la cantidad de vértices con dos hijos en T es también la misma que en $T - v$ e igual a k. Hemos concluido que T tiene $k + 1$ hojas y k vértices con dos hijos y por lo tanto cumple con la propiedad.

Árboles binarios

Teorema

La cantidad de vértices sin hijos de un árbol binario es la cantidad de vértices con exactamente dos hijos más 1.

Ejercicio

La ANFP está organizando la Copa Chile 2022. Si este año participan n equipos, ¿cuántos partidos se jugarán?

Respuesta: $n - 1$

Árboles binarios

Finalmente, podemos tomar una clase de árboles binarios que se usan mucho para establecer cotas para las aplicaciones de ellos.

Definición

Un **árbol binario completo** es un árbol binario tal que:

1. Todas las hojas están a la misma profundidad.
2. Todos los vértices que no son hojas tienen exactamente dos hijos.

Árboles binarios

Teorema

1. Un árbol binario completo de altura H tiene exactamente 2^H hojas.
2. Un árbol binario completo de altura H tiene exactamente $2^{H+1} - 1$ vértices.
3. Si H es la altura de un árbol binario completo con n vértices, entonces $H \leq \log_2(n)$.

Ejercicio

Demuestre el teorema anterior.

Árboles binarios

Teorema

Un árbol binario completo de altura H tiene exactamente 2^H hojas.

Demostración:
Sea $T = (V, E)$ un árbol binario completo, demostraremos la propiedad por inducción en la altura H.

BI: Si $H = 0$ entonces T corresponde un vértice sin aristas. Luego la cantidad de hojas es igual a $1 = 2^0 = 2^H$.

HI: Suponemos que todo árbol de altura H tiene 2^H hojas.

TI: Sea T un árbol de altura $H+1$ y raíz r. Si eliminamos r del árbol junto con sus aristas incidentes obtenemos un bosque de 2 árboles binarios completos de altura H. Luego, podemos aplicar la HI, con lo que cada árbol en $T - r$ tiene 2^H hojas. Es claro que la cantidad de hojas de T es igual a la suma de todas las hojas de los arboles inducidos al remover r. Con lo que T tendrá una cantidad de hojas igual a $2^H + 2^H = 2 \cdot 2^H = 2^{H+1}$.

Árboles binarios

Teorema

Un árbol binario completo de altura H tiene exactamente $2^{H+1} - 1$ vértices.

<u>Demostración:</u>
Sea T un árbol binario completo con altura H. Por el teorema anterior T debe tener 2^H hojas. Luego, por el otro teorema anterior sabemos que debe tener $2^H - 1$ vértices con exactamente 2 hijos. Dado que todo vértice en un árbol binario es hoja o tiene 2 hijos, concluimos que T debe tener $2^H + (2^H - 1) = 2 \cdot 2^H - 1 = 2^{H+1} - 1$ vértices.

Árboles binarios

Teorema

Si H es la altura de un árbol binario completo con n vértices, entonces $H \leq \log_2(n)$.

<u>Demostración:</u>

Sea T un árbol binario completo con n vértices y altura H. Sabemos que la cantidad de hojas (2^H) tiene que ser menor o igual a la cantidad total de vértices (n).

$$2^H \leq n \Rightarrow H \leq \log_2(n)$$

Outline

Objetivos de la clase

- Conocer definiciones de árbol y demostrar que son equivalentes
- Conocer elementos esenciales de un árbol
- Demostrar propiedades que relacionan elementos de un árbol