Grafos cordales y árboles clique.

Silvia B. Tondato.
Departamento de Matemática. Facultad de Ciencias Exactas.
tondato@mate.unlp.edu.ar

Director: Marisa Gutierrez. Codirector: Jayme Szwarcfiter.

Resumen

Los grafos cordales fueron caracterizados como aquellos grafos que admiten una representación por medio de árboles cliques. En este trabajo se estudian las relaciones entre los árboles cliques y subgrafos de un grafo cordal. Además, se probará que alrededor de todo vértice de grado 3 de un árbol clique con menor número de hojas, existe un tripla asteroidal en las ramas que inciden en el mencionado vértice.

Palabras claves: grafos cordales, árboles cliques, tripla asteroidal.

Introducción

Los grafos cordales fueron definidos como aquellos grafos que no poseen ciclos inducidos de 4 o más vértices.
Se sabe que para cada grafo cordal existe un árbol (árbol clique), cuyos vértices son los cliques del grafo y para cada vértice en el grafo, el conjunto de cliques que lo contienen forman un subárbol del árbol clique.
Los árboles cliques no sólo aparecen en la literatura de teoría de grafos, sino también el contexto de esquemas acíclicos de base de datos [3] y [5], matrices esparcidas [1], biología [2], entre otros. Como ejemplo los árboles cliques pueden utilizarse para estudiar la interacción de proteínas [10]. Esta situación puede ser representada por un grafo cuyos vértices son las proteínas y dos vértices son adyacentes si las correspondientes proteínas interactúan. Complejos proteicos pueden verse como cliques de este grafo. Luego, cuando el grafo es cordal, un árbol clique y su familia de subárboles, representando los vértices, dan una buena herramienta para ver la actividad de las proteínas en diferentes complejos.
En este trabajo, se estudia la relación entre los árboles cliques de un grafo cordal y sus subgrafos. Se probará que todos los árboles cliques de un grafo pueden obtenerse de árboles cliques de subgrafos.
En particular, se estudia el mínimo número de hojas de los árboles cliques de grafos cordales.
Por otro lado, triplas asteroidales fueron introducidas en [6], como tripla de vértices tales que entre dos de esos vértices existe un camino no atravesando la vecindad del tercero. En este artículo, se probará que alrededor de cada vértice de grado 3, en un árbol clique con menor número de hojas, hay una tripla en cada una de las ramas que inciden en él.

Preliminares

En este artículo, todos los grafos son finitos, no dirigidos, simples, conexos y cordales.
Un *grafo* G es un par (V, E), siendo V un conjunto de vértices y E un conjunto de aristas. Dos vértices *u* y *v* son ***adyacentes*** o vecinos, si existe un arista que los tiene como extremos.

Si *V’* es subconjunto de V, **G[*V’*]** es el ***subgrafo inducido*** de G cuyo conjunto de vértices es *V’*.
Si *G1* y *G2* son grafos, la ***unión*** es el grafo denotado por ***G1UG2*** cuyo conjunto de vértices es la unión de V(*G1*) con V(*G2*) y cuyo conjunto de aristas es la unión de E(*G1*) con E(*G2*).
Un subconjunto *C* de V es un ***clique*** de G, si G[*C*] es un subgrafo completo maximal de G.
Se denota por: ***C*(G)** al conjunto de cliques de G, y para cada *v* de V, ***Cv***={C€*C*(G): *v*€C}.
Un vértice *s* se dice ***simplicial***, si el conjunto formado por *s* y sus vecinos es un clique.
El ***grafo de intersección*** de una familia de conjuntos *F=(Fi)i€I,* es un grafo cuyo conjunto de vértices es *I*, dos vértices *i* y *j* son adyacentes siempre que la intersección de *Fi* y *Fj* no sea vacía.
El ***grafo clique*** de un grafo G, denotado por ***K(G)***, es el grafo de intersección de la familia de cliques de G.
El ***grafo clique valuado*** de un grafo G, denotado ***Kv(G)***, es el grafo clique donde cada arista *CiCj* esta valuada por el cardinal de la intersección de los cliques *Ci* y *Cj.*
Un grafo G se dice ***cordal*** si cada ciclo en G de longitud al menos 4 tiene una cuerda, es decir, una arista uniendo dos vértices no consecutivos del ciclo. De esta definición resulta claro que cualquier subgrafo inducido de un grafo cordal también es cordal.
Buneman [3], Gavril [4] y Walker [9] prueban independientemente que cada grafo es cordal si y sólo si es grafo de intersección de una familia de subárboles de un árbol. Es fácil verificar, que la familia de subárboles de un árbol tiene la propiedad ***Helly***, es decir toda subfamilia intersecante tiene un vértice común. Luego minimizando el número de vértices del árbol, preservando las intersecciones en la familia de subárboles, se obtiene un árbol cuyo conjunto de vértices es *C(G)* y cuyo familia de subárboles es *(Cv)v€V(G)*. Por este motivo, este árbol es denominado usualmente ***árbol clique***. Fue probado que esos árboles son exactamente los árboles generadores maximales de *Kv(G)*.
Una ***representación canónica*** de un grafo cordal G es un par ***(T, F)*** donde T es un árbol generador maximal de Kv(G), usualmente llamado árbol clique de G y F = (Cv)v€V(G). Es fácil probar que (Cv)v€V(G) es una ***familia separadora***, esto es la intersección de los miembros que contienen a un elemento *x* es {*x*}. Más aún, esta propiedad caracteriza a las representaciones canónicas.
Por este motivo, estudiar grafos cordales es equivalente a estudiar sus representaciones canónicas o sus árboles cliques.

Subgrafos de grafos cordales y sus árboles cliques.

Como ya fue mencionado previamente, los subgrafos de un grafo cordal son grafos cordales. Pero los árboles cliques de subgrafos podrían no ser subárboles de árboles cliques.
Existe una forma simple de obtener subgrafos inducidos de un grafo cordal, tal que cada árbol clique del subgrafo es un subárbol de algún clique del grafo.
Primero, veamos algunos resultados triviales respecto de árboles maximales.

<u>Propiedades de árboles generadores maximales</u>.
En lo que sigue, se recordaran propiedades usuales en árboles, que relacionaran árboles cliques de un grafo con árboles cliques de subrafos inducidos.
Sea T un árbol clique y T1 un subárbol de T, tal que E(T)-E(T1) inducen un subárbol T1° en T. En este caso, ***T1°*** es el ***complemento*** de T1 en T. Claramente T = T1° U T1.

Si G es un grafo cuyas aristas son valuadas por un número real positivo, el siguiente teorema implica que el árbol generador maximal puede verse localmente.
Se notará **w(T)** a la suma de las valuaciones de las aristas de T.
Teorema 1: Si T es un árbol generador maximal de G y T1 es un subárbol de T que tiene complemento, T1°. Entonces w(T) = w(T1) + w(T1°) y T1 es un árbol generador maximal de G[V(T1)]. Si T1' es otro árbol generador maximal de G[V(T1)] entonces T1' U T1° es también un árbol generador maximal de G.

Este teorema, nos permite obtener subgrafos de grafos cordales.

<u>Subgrafos y subárboles cliques</u>.
Sea G un grafo cordal y T un árbol clique de G. Si T1 es un subárbol de T, define V1 = {v€V(G): T1 tiene intersección no vacía con Cv} y GT1 = G[V1].

Teorema 2: Sea G un grafo cordal. Si T es un árbol clique de G y T1 es un subárbol de T con complemento entonces G[V1] es un grafo cordal y T1 es un árbol clique de G[V1].
Demostración: Se sabe que para cada v€V(G), Cv es una familia separadora de subárboles de T entonces la intersección de Cv con V(T1) es, para cada v, una familia separadora de subárboles de T1 cuya intersección es G[V1]. Por el Teorema 1, T1 es un árbol generador maximal de Kv(G[V1]) y entonces es un árbol clique de G[V1].

El siguiente resultado, prueba que todo árbol clique puede obtenerse mediante corte y pegado de árboles cliques.

Teorema 3: Sean G un grafo cordal, T un árbol clique de G y T1 un subárbol de T con complemento.
Si T1' es un árbol clique de G[V1] entonces T1' U T1° es un árbol clique de G.

Observar que los Teoremas 1, 2 y 3 pueden generalizarse, aún en el caso que T1 sea un subárbol sin complemento, razonando inductivamente.

Corolario: Sean G un grafo cordal, T un árbol clique de G. Si T1 es un subárbol de T entonces cada árbol clique de G[V1] es un subárbol de un árbol clique de G.

Número de hojas de grafos cordales.

Naturalmente, parámetros de los árboles, como número de hojas (***ln***), definen parámetros de grafos cordales.
Si G es un grafo cordal, el ***número de hojas de G*** es ***l(G)*** = min{ln(T), T un árbol clique de G}.
Un árbol clique es llamado ***l-óptimo*** de G si ln(T)=l(G).
Claramente grafo cordales conexos con número de hojas igual a 2 son exactamente grafos de intervalos conexos. Es fácil ver que el número de hojas de un grafo, puede incrementarse considerando subgrafos inducidos no conexos de un grafo de intervalos conexo.
En esta sección, se probará que si T es l-óptimo de G, alguno de sus subárboles de los correspondientes subgrafos, son l-óptimos.

Lema: Sean T1 un subárbol de T con complemento y *v* un vértice de T1 y de T1°.

i) $ln(T) = ln(T1) + ln(T1°)$, si v no es hoja de T1 o de T1°.
ii) $ln(T) = ln(T1) + ln(T1°) - 1$, si v es hoja de sólo uno de ellos.
iii) $ln(T) = ln(T1) + ln(T1°) - 2$, si v es hoja de ambos.

Teorema 4: Sean G un grafo cordal y T un árbol clique l-óptimo de G. Si T1 es un subárbol de T con complemento, y v es un vértice de T1 y su complemento, que no es hoja de T1 entonces T1 es un árbol clique l-óptimo de G[V1].
Demostración: T1 es un árbol clique de G entonces $l(G[V1]) \leq ln(T1)$.
Supongamos que $l(G[V1]) < ln(T1)$. Luego existe un árbol clique de G[V1] tal que $ln(T1') = l(G[V1]) < ln(T1)$.
Por otro lado, T' = T1' U T1° es un árbol clique de G.
Se analizarán varios casos.
Caso 1: v es hoja de T1° y T1'.Luego $ln(T') = ln(T1') + ln(T1°) – 2 < ln(T1) + ln(T1°) –2 < ln(T1) - ln(T1°) – 1 = ln(T)$.
Caso 2: v es hoja de T1' pero no de T1°. Luego $ln(T') = ln(T1') + ln(T1°) – 1 < ln(T1) - ln(T1°) – 1 = ln(T)$.
Caso 3: v es hoja de T1° pero no de T1'. Luego $ln(T') = ln(T1') + ln(T1°) – 1 < ln(T1) - ln(T1°) – 1 = ln(T)$.
Caso 4: v no es hoja de T1° ni de T1'. Luego $ln(T') = ln(T1') + ln(T1°) < ln(T1) - ln(T1°) = ln(T)$.
Por lo tanto, $ln(T') < ln(T)$, contradiciendo que T es l-óptimo. Luego $l(G[V1]) = ln(T1)$.

3-asteroidales en grafos cordales.

Otro parámetro interesante de grafos cordales, y de grafos en general, es el número asteroidal.
Un conjunto de k vértices de un grafo G es un ***k-asteroidal*** de G si para cada a€A, los vértices de A-{a} están en la misma componente conexa de G-N[a]. Esto es equivalente a decir que cada tripla de vértices de A son una tripla asteroidal de G
Este concepto aparece en [7]; en este artículo ***a(G)*** denota el máximo tamaño de un conjunto asteroidal de G.
Lin, McKee y West prueban en [7] que $a(G) \leq l(G)$. Por motivos de espacio, no se presenta la demostración que utiliza el Teorema 4 e inducción.

Teorema 5: Si G es un grafo cordal entonces $a(G) \leq l(G)$.

Se sabe que los grafos de intervalos son grafos cordales sin triplas asteroidales. Se generalizará este resultado probando que para cada vértice de grado al menos 3 en un árbol clique l-óptimo existe un tripla que incide en dicho vértice.

Teorema 6: Sean G un grafo cordal, T un árbol clique de G l-óptimo, t un vértice de T de grado al menos 3 y *R1, R2, R3* tres ramas de T incidiendo en t. Entonces, existe una tripla asteroidal {*x1, x2, x3*} de G con *xi€Ri* para todo i€{1,2,3}.
Demostración: Por inducción en l(G). Si $l(G) = 3$ entonces G no es de intervalos y tiene una tripla asteroidal. Como $l(G) = 3$, existe un único vértice de grado 3 en T. Es fácil comprobar que ningún par de vértices de la tripla puede estar en la misma rama de T que inciden en t. Luego el teorema es válido.
Se supone por hipótesis inductiva que el teorema vale si $l(G) < k$. Ahora, sea $l(G) = k$. Claramente existe un árbol clique T de G, tal que $ln(T) = l(G)$. Sean t un vértice de grado al menos 3 y *R1*, *R2*, *R3* tres ramas de T, que inciden en el. Considerar T1 = *R1* U

R2 U *R3*. Si T1 ≠ T, por el Teorema 2, se tiene que G[V1] es un grafo cordal y T1 es un árbol clique. Como *t* no es hoja de T1 y T ≠ T1, se tiene que l(G[V1]) < k. Por hipótesis inductiva, como *R1*, *R2, R3* son tres ramas en T1, existe una tripla en G[V1] con cada vértice en *Ri* para *i* €{1,2,3}. Luego, existe un tripla en G como era requerido.
En caso que T1 = T, sean *t* y *t'* vértices de grado 3. Sin perder generalidad supongamos que *t'*€R3 y sea *R* una rama de T que incide en *t'* y no tiene a *t*. Ahora, sea T1° = *R*. Nuevamente, G[V1] resulta un grafo cordal y T1 su árbol clique. Como t' no es hoja de T1, por el Teorema 4, l(G[V1]) = ln(T1) < ln(T) = k. Por hipótesis inductiva, como *t* es un vértice de grado 3 en T1 y *R1*, *R2*, *R'3* = *R3* ∩ T1, tres ramas de T1 que inciden en *t*, existe una tripla de G[V1}]. Luego, existe un tripla en G.

Corolario: Sean G un grafo cordal, T un árbol clique de G l-óptimo, *t* un vértice de T de grado al menos 3 y *R1, R2, R3* tres ramas de T incidiendo en *t*. Entonces, existe una tripla asteroidal {*x1, x2, x3*} de vértices simpliciales de G con *xi*€*Ri* para todo i€{1,2,3}.
Demostración: Por el Teorema 6, existe una tripla asteroidal, *x1, x2, x3* de G. Si *x1* no es un vértice simplicial de G, existen al menos 2 clqiues adyacentes en T que lo contienen. Sea *C1,* el primer clique en esas condiciones, más próximo a *t* que lo contiene. Sea *s1* el primer vértice simplicial, cuya representación en T es más próxima a *C1* en dirección a la hoja a la hoja que se ubicada sobre *R1*.
Claramente, *s1*, *x2*, *x3* es un tripla asteroidal de G. Análogamente, si *xi* con i€{2,3} no son vértices simpliciales de G entonces existen respectivamente *s2*, *s3* vértices simpliciales tales que *s1*, *s2*, *s3* es un tripla asteroidal de simpliciales de G.

Bibliografía.

1. J. Blair and B. Peyton. An introduction to chordal graphs and clique trees. Graph Theory and Sparce Matrix Computation, 56:1-29, 1993.
2. A. Berry, A. Sigayret, C. Sinoquet. Maximal sub-triangulation as improving phylogenetic Data. Proccedings of JIM'03. Soft Computing-Recent Advances in Knowledge Discovery, G. Govaert, R. Haenle and Nadif(eds), 1900:01, 2005.
3. P. Buneman, A characterization of rigid circuit graphs, Discrete Mathematics 9 (1974), 205-212.
4. F. Gavril, The intersection graphs of subtrees in trees are exactly the chordal graphs, Journal of Combinatorial Theory (Series B) 16 (1974), 47-56.
5. B. Korter, L. Lovász and R. Schrader. Greedois. Number 4 in Algorithms and Combinatorics. Springer. Verlag, 1991.
6. C. Lekkerkerker and J. Boland, Representation of a finite graph by a set of intervals on the real line, Fundamenta Mathematicae 51 (1962), 45-64.
7. I. Lin, T. McKee and D. B. West, The leafage of a chordal graphs, Maht/C/9807022v/3 Jul 1998.
8. T. McKee and F. McMorris, Topics in Intersection Graph Theory, Monographs on Discrete Mathematics and Applications, vol. 2 SIAM, Philadelphia, 1999.
9. J. Walter, Representations of chordal graphs as subtrees of a tree, Journal of Graph Theory 2 (1978), 265-267.
10. E. Zotenko, K. Guimaraes, R. Jothi and T. Przytycka. Decomposition of overlapping protein complexes: A graph theorical method for analysing static and dynamic protein associations. Algorithms for Molecular Biology 1:7, 2006.