Clique

Claudia Puentes

December 2024

Un clique en un grafo $G = (V, E)$ es un subconjunto de vértices $C \subseteq V$ tal que todos los pares de vértices en C están conectados directamente por una arista.

El problema consiste en hallar un clique máximo en un grafo. Para demostrar que este problema es NP-Completo, debemos cumplir dos condiciones fundamentales: 1. El problema está en NP: Probar que una solución candidata puede verificarse en tiempo polinomial. 2. El problema es NP-Hard: Probar que cualquier problema en NP puede ser reducido al problema de encontrar un clique máximo en tiempo polinomial.

A continuación, abordamos cada paso en detalle.

Dado un grafo $G = (V, E)$ y un entero k, necesitamos verificar si existe un clique de tamaño k en G. Consideremos un conjunto candidato de k vértices en el grafo G. Verifiquemos si todos los $\binom{k}{2} = \frac{k(k-1)}{2}$ pares de vértices en el conjunto están conectados por una arista. Cada verificación toma tiempo $O(1)$, para cada par de vértices, solo necesitamos revisar si hay una arista entre ellos en el grafo. Esto se puede hacer en tiempo constante si el grafo está representado con una matriz de adyacencia o un diccionario de listas de adyacencia. Como hay $\binom{k}{2}$ pares, y cada verificación toma $O(1)$, el tiempo total es proporcional a:

$$T = \binom{k}{2} \cdot O(1) = \frac{k \cdot (k-1)}{2} \cdot O(1).$$

Descartamos constantes y términos menores para obtener $T \in O(k^2)$.

Dado que este tiempo es polinomial en el tamaño del grafo, concluimos que el problema de decidir si un conjunto dado de vértices forma un clique de tamaño k es verificable en tiempo polinomial. Por tanto, el problema está en NP.

Para demostrar que hallar un clique máximo es NP-Hard, reducimos un problema conocido como NP-Completo al problema de encontrar un clique máximo. Usaremos el problema 3-SAT como base para la reducción. El problema de 3-SAT consiste en determinar si una fórmula booleana en forma normal conjuntiva (CNF), donde cada cláusula tiene exactamente tres literales, es satisfacible. Una fórmula 3-SAT tiene la forma:

$$\phi = (x_1 \vee \neg x_2 \vee x_3) \wedge (\neg x_1 \vee x_4 \vee x_5) \wedge \dots$$

con m cláusulas y n variables. Aquí, cada cláusula es una disyunción ("or") de tres literales, y la fórmula es una conjunción ("and") de varias cláusulas.

Dada una fórmula ϕ con m cláusulas y n variables, construimos un grafo G de la siguiente manera:
Cada literal en ϕ (por ejemplo, x_i o $\neg x_i$) se convierte en un nodo del grafo. Si la fórmula tiene m cláusulas y cada cláusula contiene 3 literales, el grafo tendrá $3m$ nodos.
Cada nodo representa un literal en una cláusula específica y lleva una etiqueta que indica a qué cláusula pertenece. Por ejemplo, $x_1(C_1)$ representa el literal x_1 en la cláusula C_1.

Conectamos dos nodos con una arista si y solo si los nodos provienen de cláusulas diferentes, y los literales correspondientes no son contradictorios (por ejemplo, x_i y $\neg x_i$ no pueden estar conectados).

Aunque un literal como x_1 pueda aparecer en varias cláusulas, cada aparición es tratada como un nodo independiente porque representa su uso en el contexto específico de esa cláusula. Esto asegura que cada nodo capture la relación lógica única que un literal tiene en su cláusula. Además, este diseño permite conectar nodos de distintas cláusulas incluso si representan el mismo literal.

Un clique de tamaño m en este grafo corresponde a una asignación satisfactoria de la fórmula 3-SAT: cada nodo en el clique pertenece a una cláusula distinta, garantizando que cada cláusula tiene al menos un literal verdadero, y no hay literales contradictorios en el clique, lo que asegura consistencia en la asignación.

Que pasa si intentas incluir más de m nodos
Supongamos que intentas construir un clique de tamaño mayor a m, por ejemplo, $m + 1$. Esto implicaría seleccionar $m+1$ nodos. Como hay solo m cláusulas, al menos dos de los nodos seleccionados deben provenir de la misma cláusula. Los nodos dentro de la misma cláusula no están conectados (por diseño del grafo). Un clique requiere que todos los nodos estén conectados entre sí. Por lo tanto, no es posible construir un clique de tamaño mayor a m.

Un clique de tamaño m en este grafo corresponde a una asignación satisfactoria para ϕ. Esto se debe a que cada nodo en el clique pertenece a una cláusula diferente, garantizando que al seleccionar exactamente un literal, cada cláusula tiene al menos un literal verdadero. Los nodos en el clique están conectados, lo que implica que los literales seleccionados no son contradictorios. Por lo tanto, existe una asignación booleana consistente que satisface todas las cláusulas de ϕ.

La construcción del grafo y la verificación del clique pueden realizarse en tiempo polinomial. Como 3-SAT es NP-Hard, hemos demostrado que encontrar un clique máximo es al menos tan difícil como resolver 3-SAT. Por lo tanto, Max Clique es NP-Hard. Dado que el problema de hallar un clique máximo está en NP y es NP-Hard, podemos concluir que este problema es NP-Completo. Esta demostración se basa en la verificación eficiente de soluciones candidatas y en la reducción desde un problema NP-Completo como 3-SAT.