Local Structure: Subgraph Counts I

Graph Theory (Fall 2011)
Rutgers University
Swastik Kopparty

Until now we studed conditions which guaranteed the existence of one copy of a fixed graph in another. Now we turn to questions about the number of copies of a fixed graph in another.

A lot of the subject simplifies if the notion of counting is chosen well. The most convenient one is in the language of *graph homomorphisms*.

1 Graph Homomorphisms

Definition 1. *Let F, G be graphs. A* **homomorphism** *from F to G is a function $\varphi : V_F \to V_G$ such that for every edge $\{u, v\}$ of F, $\{\varphi(u), \varphi(v)\}$ is an edge of G (i.e., it is an edge preserving map).*

The set of homomorphisms from F to G is denoted $\mathsf{Hom}(F, G)$. The number of homomorphisms from F to G is denoted $\hom(F, G)$.

Let us look at some examples:

- $\hom(K_1, G)$ is the number of vertices of a graph G.
- $\hom(K_2, G)$ is twice the number of edges in a graph G.
- Let G be any bipartite graph. Then there is a homomorphism from G to K_2.
- $\hom(K_3, K_3) = 6$.
- $\hom(K_r, G)$ counts the number of ordered tuples of vertices $(u_1, \dots, u_r)$ of G such that $u_1, \dots, u_r$ form a K_r.
- $\hom(G, K_r)$ equals the number of r-colorings of G.
- $\hom(P_k, G)$ counts the number of walks of length k in G.
- $\hom(C_k, G)$ counts the number of closed walks of length k in G.

If F is small and G is big, then $\hom(F, G)$ counts (in a controversial way) the number of copies of F in G.

Other than counting homomorphisms, there are other (perhaps more natural) ways one could consider counting subgraphs of a graph which are also important. We discuss some of them now.

2 Homomorphism counts, Injective homomorphism counts, and Induced subgraph counts

A homomorphism $\varphi \in \mathsf{Hom}(F, G)$ is called injective if $\varphi(u) \neq \varphi(v)$ for $u \neq v$. We let $\mathsf{Inj}(F, G)$ denote the set of all injective homomorphisms from F to G, and we let $\mathsf{inj}(F, G)$ denote its cardinality.

An induced copy of F in G is a homomorphism $\varphi \in \mathsf{Hom}(F, G)$ that also preserves non-edges (i.e., $\{\varphi(u), \varphi(v)\} \in E_G$ iff $\{u, v\} \in E_F$). We let $\mathsf{Ind}(F, G)$ denote the set of all induced copies of F in G, and we let $\mathsf{ind}(F, G)$ denote its cardinality.

An automorphism is a bijective map that preserves both edges and non-edges (or equivalently, φ is a bijection of vertex sets, and both φ and φ^{-1} are homomorphisms. The number of automorphisms of F is denoted $\mathsf{aut}(F)$. For finite graphs, $\mathsf{aut}(F) = \mathsf{inj}(F, F) = \mathsf{ind}(F, F)$.

For a given graph G, the functions $\hom(\cdot, G)$, $\mathsf{inj}(\cdot, G)$ and $\mathsf{ind}(\cdot, G)$ all determine each other. Specifically, we have the following relations:

- $$\hom(F, G) = \sum_{\Pi \in \mathsf{Partitions}(V_F)} \mathsf{inj}(F/\Pi, G),$$

  where for a partition Π of V_F, F/Π denotes the graph which is obtained from F by identifying all vertices in a part of Π, and removing duplicate edges and self loops.

  This expresses $\hom(F, G)$ as a linear combination of $\mathsf{inj}(F, G)$ and $\mathsf{inj}(F_i, G)$ for some smaller graphs F_i. This allows us to invert the relation, and deduce that:

  $$\mathsf{inj}(F, G) = \hom(F, G) - \sum_{H \mid |V_H| < |V_F|} a_{F,H} \hom(H, G),$$

  for some coefficients $a_{F,H}$ (these $a_{F,H}$ are related to the Mobius function for the partial order of partitions, basically some careful inclusion-exclusion).

- $$\mathsf{inj}(F, G) = \sum_{H=(V,E) \mid V = V_F, E \subseteq E_F} \mathsf{ind}(H, G).$$

  Again, this allows us to express $\mathsf{ind}(F, G)$ as a finite linear combination of $\mathsf{ind}(F_i, G)$ for various graphs F_i.

But the main reason homomorphism counts are well behaved and have many interesting properties is their algebraic structure. We discuss this now.

3 Sums and Products

For graphs G_1, G_2, the sum $G_1 + G_2$ denotes the disjoint union of G_1 and G_2. The product $G_1 \times G_2$ denotes the graph (V, E) where $V = V_{G_1} \times V_{G_2}$, and $E = \{\{(u_1, u_2), (v_1, v_2)\} \mid \{u_1, v_1\} \in E_{G_1}, \{u_2, v_2\} \in E_{G_2}\}$.

These graph operations interact with hom very nicely:

- $\hom(F_1 + F_2, G) = \hom(F_1, G) \cdot \hom(F_2, G)$.
- $\hom(F, G_1 + G_2) = \hom(F, G_1) + \hom(F, G_2)$ **provided** F is connected.
- $\hom(F, G_1 \times G_2) = \hom(F, G_1) \cdot \hom(F, G_2)$.

This last property gives us the very powerful **powering trick**: Suppose we know that graphs F_1 and F_2 are such that $\hom(F_1, G) \leq 1000 \cdot \hom(F_2, G)$ for every graph G. Then automatically from this we can deduce the improved inequality $\hom(F_1, G) \leq \hom(F_2, G)$. Indeed, for every n we have:

$$\hom(F_1, G) \leq \hom(F_1, G^n)^{1/n} \leq (1000 \cdot \hom(F_2, G^n))^{1/n} \leq 1000^{1/n} \cdot \hom(F_2, G),$$

which implies that $\hom(F_1, G) \leq \hom(F_2, G)$.

We will see several homomorphism inequalities in the course, and in particular, we will see situations where the above powering trick can be used to improve some loose inequalities.