Graph homomorphisms: Open problems

László Lovász

June 2008

DRAFT

Contents

For definitions, see `http://www.cs.elte.hu/~lovasz/definitions.pdf`

For more bibliography, see `http://www.cs.elte.hu/~lovasz/homrefs.pdf`

1 Homomorphism functions

1.1 Connection rank

Problem 1 For which graph parameters have all the connection matrices finite rank?

Homomorphism functions $\hom(., G)$ are examples for every weighted graph G (here the nodeweights and edgeweights can be negative). Dual homomorphism densities $\hom(F, .)$ also have finite connection rank. Every evaluation of the Tutte polynomial is a further example.

One reason to be interested in this question is the fact that such a graph parameter can be evaluated in polynomial time for graphs with bounded treewidth [59]. Very recently Godlin and Makowski proved that all graph parameters which are evaluations of graph polynomials definable in Monadic Second Order Logic have finite connection rank.

Problem 2 What are the possible connection rank functions of graph parameters?

It was noticed in [41] that every finite connection rank function $r(f, k)$ is log-superadditive, i.e., $r(f, k + l) \geq r(f, k)r(f, l)$. The connection rank of homomorphism functions $\hom(., G)$ is also logconvex, i.e., it satisfies $r(f, k + 1)r(f, k - 1) \geq r(f, k)^2$.

Problem 3 Which graph parameters have a logconvex connection rank function?

The main theorem in [41] characterizes graph parameters of the form $\hom(., G)$ as reflection positive parameters for which the connection rank function is bounded by c^k. Lovász and Szegedy [70] proved that if a multiplicative reflection positive graph parameter f has finite connection rank for $k = 1, 2$, then $r(f, k)$ is finite for every k, and it can only grow like c^k or c^{k^2} (in the sense that either $(\log r(f, k))/k$ or $(\log r(f, k))/k^2$ tends to a finite nonzero limit); furthermore, if the parameter depends only on the simple graph underlying a multigraph, then its connection rank grows like c^k. For a generic substitution of the Tutte polynomial (which is not reflection positive), the connection rank grows like k^k.

Problem 4 Can we characterize multiplicative simple graph parameters with exponentially bounded connection rank?

An example of such a rank function is a generalized homomorphism function $\hom(., G)$, where G may have negative nodeweights and edgeweights (such functions are multiplicative but not necessarily reflection positive). Another example is the number $\mathsf{match}(G)$ of perfect matchings in the graph G: this is trivially multiplicative, and it is not hard to see that its connection rank function is 2^k [58]. A vague conjecture is that all simple graph parameters with exponentially bounded connection rank function are limits of such homomorphism functions.

Graph parameters of the form $f = \hom(F, .)$ satisfy $r(f, k) = O(k^{|V(F)|})$. These parameters are not multiplicative, in fact, $r(f, 0) = 2$ is F is connected.

Problem 5 Can we characterize simple graph parameters with polynomially bounded connection rank?

1.2 Characterizing homomorphism functions

The main theorem in [41] characterizes graph parameters that can be represented in the form hom(., H), where H is a weighted graph. Schrijver [81] characterizes such functions where H only has edgeweights (equivalently, all nodeweights are 1). In [62], such functions were characterized where H is an unweighted graph with loops allowed.

Problem 6 Is there a simple graph version of the theorem in [41] characterizing graph parameters hom(., H), where H is an (unweighted) simple graph (no loops or mjultiple edges)?

1.3 Graph algebras

Graph algebras are defined on formal linear combinations of a finte number of graphs (called quantum graphs).

Problem 7 Define algebras of infinite linear combinations of graphs with appropriate convergence properties, and find the structure of the resulting algebra. How is this related to graphons?

1.4 Dual homomorphism problems

The questions above have analogues in the case when we reverse the arrows in the category of graph homomorphism. In [62] simple graph parameters of the form hom(F, .) were characterized, where F is a simple graph, via right-reflection positivity and an appropriate right-rank condition. But many further problems remain:

Problem 8 Is it enough to assume the right-rank condition for, say, k=1,2?

Problem 9 Characterize weighted graph parameters of the form hom(F, .), where F is a simple graph or a weighted graph.

Problem 10 Characterize right-reflection positive simple graph parameters (without any rank condition).

Problem 11 Characterize graph parameters with finite right-connection rank.

In [61], the results of [41] are extended to a semigroup setting. The dual results suggest:

Problem 12 Is there a common generalization of the results in [41] and [62] in terms of categories?

2 Dense graphs

2.1 Graph distances

Let G and G' be two graphs on the same set of n nodes. For $S, T \subseteq V(G)$, let $e_G(S,T)$ denote the number of edges in G with one endnode in S and the other in T. We define

$$d_\square(G,H) = \frac{1}{n^2} \max_{S,T\subseteq V(G)} |e_G(S,T) - e_H(S,T)|.$$

Let G and H be graphs on the same number of n nodes. We define

$$\hat{\delta}_\square(G,H) = \min d_\square(\tilde{G},\tilde{H}), \qquad (1)$$

where $\tilde{G}$ and $\tilde{H}$ range over all labelings of G and G' by $1,\dots,n$, respectively.

Let G have n nodes and H have r nodes. Let $G[k]$ be obtained from G by replacing each node by k twins of it. We define

$$\delta_\square(G,H) = \lim_{k\to\infty} \hat{\delta}_\square(G[kr],H[kn]).$$

(For more details and motivation concerning these distances, see [22].

Conjecture 1 There is a constant $c \ge 1$ such that

$$\hat{\delta}_\square(G,H) \le c\delta_\square(G,H).$$

Perhaps $c = 3/2$ is good enough. Alon (unpublished) showed that $c = 1+\varepsilon$ is enough if $|V(G)|$ and $|V(H)|$ are large enough. In [24] it is shown that $\hat{\delta}_\square(G,H) \le 32\delta_\square(G,H)^{1/67}$.

We know that if two graphs satisfy $\delta_\square(G,G') < \varepsilon$, and G,G' are large enough, then the densities $t(C_k,G)$ and $t(C_k,G')$ are close for every k, which implies that the spectra of G and G' are close in the sense that for $k \le k(\varepsilon)$, the k-th largest [smallest] eigenvalue of G, divided by $V(G)$, is close to the k-th largest [smallest] eigenvalue of G', divided by $V(G')$.

Problem 13 Find the best bound on the distance of spectra in terms of the distance of the graphs.

The Weak Regularity Lemma is equivalent to saying that for every $k \ge 1$, for every graph G there is a graph H with at most k nodes such that $\delta_\square(G,H) < 10/\sqrt{\log\log k}$.

Problem 14 Is there a notion of distance such that the original Regularity Lemma (or an even stronger lemma like that due to Alon, Fisher, Krivelevich and Szegedy [5]) can be expressed as an approximation result?

2.2 Graphon dimension

In the definition of dimension, we could replace the norm L_1 by L_2 or other norms. Furthermore, it seems better to replace W by

$$W \circ W = \int_0^1 W(.,x)W(x,.)\,dx$$

(cf. [68]).

Problem 15 How much is the dimension modified by these changes?

2.3 Convergence

It was proved in [25] that if (G_n) is a convergent graph sequence, i.e., $t(F,G_n)$ is convergent for every simple graph F, then it is also convergent in the sense that $\ln \hom(G_n,H)/|V(G_n)|^2$ is convergent for every weighted graph H with positive edgeweights, and $\mathcal{E}(G_n,H)$ is convergent for every weighted graph H with total nodeweight 1. This means that knowing the numbers $t(F,G_n)$ approximately gives us the numbers $\hom(G_n,H)$ approximately.

Problem 16 Is there an explicit formula to relate left and right homomorphism numbers?

In the sparse case the answer is in the affirmative under some conditions, see [20].

2.4 Extremal graph theory

An *extremal problem* (for the purposes of this document) is to find the minimum, over all functions $W \in \mathcal{W}_0$, of $\sum_{i=1}^m a_i t(F_i,W)$ with some graphs $F_1,\dots,F_m$ and real numbers $a_1,\dots,a_m$. (We may add constraints on other subgraph densities, but this would not add real generality.) Classical extremal problems like Turán's Theorem can be cast this way.

Problem 17 (a) Does every algebraic inequality between the subgraph densities $t(F,W)$ that holds for all $W \in \mathcal{W}$ follow from a finite number of semidefiniteness inequalities?

(b) Does the Blakley–Roy inequality

$$t(P_k,W) \ge t(K_2,W)^{k-1}$$

follow from a finite number of semidefiniteness inequalities?

It is easy to see that the answer is positive for even k. I expect that the answer is negative, but it may be difficult to prove it.

Conjecture 2 (Sidorenko) For every bipartite graph F and every $W \in \mathcal{W}$,

$$t(F,W) \ge t(K_2,W)^{|E(F)|}.$$

Problem 18 Does semidefinite duality have any implication or combinatorial meaning in this setting?

The semidefinite method also extends to hypergraphs, which could be particularly interesting since even the simplest version of Turán's Theorem is unsolved in this case.

Problem 19 Find an application of the semidefinite method to hypergraphs.

2.5 Positivstellensatz for graphs

The following problem would give a complete solution of extremal graph problems in a sense:

Problem 20 Which quantum graphs x satisfy $t(x,W) \ge 0$ for every $W \in \mathcal{W}$ [or for every $W \in \mathcal{W}_0$]?

Let $y_1,\dots,y_k$ be a k-labeled quantum graphs for some k, then the quantum graph $x=\sum_i y_i^2$ (with the labels forgotten) satisfies $t(x,W)\ge 0$ for all W. Call such (unlabeled) quantum graphs *square-sums*. More generally, if y and z are squaresums and $y = zx + x$, then $t(x,W)\ge 0$ for all W.

Problem 21 Find a quantum graph x for which $t(x,W)\ge 0$ for all W, and which is not of this form.

The following version may be easier to settle.

Problem 22 Which simple graphs F satisfy $t(F,W)\ge 0$ for every $W\in\mathcal{W}$?

It is easy to see that if F_0 is a k-labeled simple graph with no edge between labeled nodes, then $F=F_0F_0$ (with the labeles deleted) satisfies $t(F,W)\ge 0$ for every $W\in\mathcal{W}$. I don't know any other such graph, but it is difficult to decide about a given graph if it has this property. For example,

Problem 23 Let F be obtained by doubling each node of a triangle. Does $t(F,W)\ge 0$ hold for every $W\in\mathcal{W}$?

For every even positive integer k, the functional $t(C_k,W)^{1/k}$ defines a norm on $\mathcal{W}$ (the Neumann-Schatten norm). This suggests the following question:

Problem 24 For which simple graphs F is $t(F,W)^{1/|E(F)|}$ a norm (or seminorm) on $\mathcal{W}$? For which simple graphs F is $t(F,|W|)^{1/|E(F)|}$ a norm on $\mathcal{W}$?

The second version is due to Hatami [50]. He showed that if a simple graph F has the second property in Problem 24, then it satisfies Sidorenko's conjecture 2. He also proved that all cubes have the second property.

2.6 Finitely forcible graphons

We call a function $W\in\mathcal{W}_0$ *finitely forcible* if there exist a finite list of graphs $F_1,\dots,F_m$ and real numbers $a_1,\dots,a_m$ such that the equations $t(F_1,U)=a_1,\dots,t(F_m,U)=a_m$ are satisfied by precisely those functions $U\in\mathcal{W}_0$ which arise from W by measure preserving transformation. It is easy to see that instead of a finite number of graphs, we could impose a single condition $t(x,W)=0$ for the quantum graph $x=\sum_{i=1}^m(F_iF_i-2a_iF_i+a_i^2K_1)$. This construction also shows that every finitely forcible graphon is the solution of an extremal problem. We conjecture the following converse:

Conjecture 3 Every extremal problem has a finitely forcible optimum.

Almost all classical extremal problems have a solution that is a stepfunction. It was shown by Lovász and Sós [65] that *every stepfunction is finitely forcible*, and it was conjectured that these are the only ones. Very recently B. Szegedy and Lovász found other finitely forcible graphons.

Problem 25 Characterize finitely forcible graphons.

A more general question is the following:

Problem 26 Which properties of a weighted graph H (or of a graphon W) can be expressed as $t(x, H) = 0$ (or $t(x, W) = 0$) for some quantum graph x?

A complete characterization even in Problem 25 appears to be difficult, but some more special conjectures may be accessible. One class of examples consists of functions $W : [0,1]^2 \to [0,1]$ defined by

$$W(x, y) = \begin{cases} 1, & \text{if } p(x, y) > 0, \\ 0, & \text{otherwise}, \end{cases} \tag{2}$$

where $p(x, y)$ is a symmetric real polynomial that is monotone increasing on $[0,1]^2$. We conjecture that monotonicity is not needed here:

Conjecture 4 The function defined by (2) is finitely for every symmetric real polynomial $p(x, y)$.

A different example is the lexicographic product of a countable sequence of C_4-s.

Conjecture 5 The lexicographic product of a countable sequence of copies of any graph is finitely forcible.

Conjecture 6 Every finitely forcible function has finite dimension.

We say that a graphon W is *weakly homogeneous* if there exists a nonzero 2-labeled quantum graph g such that $t_{xy}(g, W) = 0$ for all $x, y \in [0, 1]$. If g is restricted to linear combinations of paths (labeled at both ends), then the condition is equivalent to W having finite rank.

Th following necessary condition for finite forcibility was derived in [71]: *every finitely forcible graphon is weakly homogeneous.*

Problem 27 Characterize weakly homogeneous graphons.

Even the first nontrivial special case appears to be difficult. Let $K_3^{(2)}$ denote the triangle with two labeled nodes.

Problem 28 Characterize graphons W for which $t_{xy}(K_3^{(2)}, W)$ is a constant function (of x and y).

We can also ask the dual question. Call a graphon *finitely forcible from the left*, if it is determined by finite number of conditions $\mathcal{E}(W, H_i) = a_i$.

Problem 29 (V.T. Sós) Which graphons W are finitely forcible from the left?

As an example, quasirandom graph sequences (G_n) with edge density p can be characterized by the property that every set $S \subseteq V(G_n)$ with $|S| = |V(G_n)|/2$ contains asymptotically $p\binom{|S|}{2}$ edges. This means that if H_a denotes the complete looped graph with 2 nodes of weight 1/2, one of the loops having weight a, and all other edges having weights 0, then

$$\mathcal{E}(W, H_1) = p, \qquad \mathcal{E}(W, H_{-1}) = -p$$

determine the function $W \equiv p$.

2.7 Colored graphs, hypergraphs

Problem 30 Describe convergence and limit objects for multigraphs, digraphs, colored graphs etc.

This can be done along the same lines as for simple graphs, but there are some surprises. For example, it seems that limits of multigraphs with edge-multiplicities bounded by d are not real valued functions, but 2-variable functions with values in $\mathbb{R}^d$. Unbounded multiplicities represent further difficulties of analytic nature.

Formulating regularity lemmas and constructing limits of sequences of r-uniform hypergraphs, where r is fixed, is a highly nontrivial task, but it essentially solved now, thanks to the work of Rödl and Skokan, Gowers, Tao, Elek and Szegedy [77, 46, 86, 35].

Problem 31 How to define the distance of hypergraphs?

Problem 32 Extend these results to nonuniform hypergraphs, with unbounded edge-size.

The semidefiniteness conditions for homomomorphism functions can be extended to hypergraphs (see e.g. [61]), and it is natural to ask if these can be useful in extremal hypergraph theory, especially since extremal problems for hypergraphs tend to be much harder than for graphs, and even basic questions are unsolved.

Problem 33 Derive any nontrivial hypergraph extremal result from the semidefiniteness of the hypergraph connection matrix.

Problem 34 (G. Kalai) Is the sequence of 3-graphs with maximum number of edges not containing a complete 3-graph on 4 nodes convergent?

2.8 Dual extremal problems

Since homomorphism functions, fixing the first variable, are dual reflection positive, one expects that the above extremal graph results problems concerning densities of various subgraphs have analogues for problems concerning numbers of homomorphisms into various graphs (e.g., number of colorings with different numbers of colors).

3 Graphs with bounded degree

3.1 Regularity Lemma

Problem 35 Is there a Regularity Lemma for graphs with bounded degree?

Extensions of the regularity lemma for non-dense graphs are known, but they are void for graphs that are so sparse. Is there a Regularity Lemma for graphs with bounded degree? There are great difficulties here, but three results justify cautious optimism.

An observation of Alon (unpublished) implies that a weak analogue of the Regularity Lemma holds: *for every $d, r \geq 1$ and $\varepsilon > 0$ there is a $k = k(d, r, \varepsilon)$ such that for every graph G with degrees bounded by d there is a graph H with degrees bounded by d and $|V(H)| \leq k$, such that the distributions r-neighborhoods in G and H are closer than ε in variation distance.* (Unfortunately, no effective bound on k follows from the proof.)

Problem 36 Give any explicit bound on the function $k = k(d, r, \varepsilon)$. Give an algorithm to construct H from G.

It was proved recently by Elek [32], and independently by Angel and Szegedy (unpublished) that every graph with degrees bounded by d can be decomposed by deleting εn edges, into highly homogeneous parts, where the number of these parts is bounded by a function of d and ε. Unfortunately, the highly homogeneous parts can still have a complex structure, but I feel that this is a first important step in the direction of finding an analogue of the Regularity Lemma.

A third idea of decomposition is related to Følner sequences in the theory of amenable groups, and is called *hyperfiniteness* for general graph sequences [34, 78]. It is likely that large real-life networks can often be approximated by hyperfinite graphings; on the other hand, hyperfinite sequences and graphings seem to be much better behaved, and much of the theory of dense graph sequences can be extended at least to this case.

The problem of the Regularity Lemma is related to the following.

Conjecture 7 (Aldous–Lyons) Every unimodular distribution on rooted countable graphs with bouded degree is the limit of a bounded degree graph sequence.

Problem 37 Is there a good notion of "distance" for graphs with bounded degree?

Problem 38 Is there a notion of convergence for graphs with bounded degree that is stronger than Benjamini–Schramm? (For example, one could read off from the limit that the graphs are expanders.)

To illustrate this by a simple example, let (G_n) be a sequence of 3-regular bipartite expander graphs with their girth tending to infinity. Let H_i consist of two disjoint copies of G_i. The Benjamini–Schramm limit of both sequences is a distribution concentrated on a single 3-regular rooted tree. In the Elek description, we get a dynamical system (Ω, T_1, T_2, T_3), where T_1, T_2 and T_3 generate a free group which acts on Ω without fixed points. In the case of the limit of the sequence (G_n), this action is ergodic, while in the case of the H_n, Ω splits into two invariant subsets of measure 1/2. So it appears that in the limit object, the underlying σ-algebra also carries combinatorial information. This is in stark contrast with the dense case [21].

We can consider our graphs "on a different scale", and study them as metric spaces with the usual graph distance as metric, normalized by the diameter. We can then consider the limit of these metric spaces in the sense of Gromov [48]. For example, the limit of a sequence of larger and larger square grids is a (full) square. This global structure is not revealed by the Benjamini–Schramm limit. It is easy to construct examples where the interesting structure of the graphs appears on an intermediate scale.

Problem 39 Describe and possibly unify limit objects belonging to different scales.

Problem 40 Can we understand different limit objects using ultraproducts, similarly to the work of Elek and Szegedy in th dense case?

3.2 Property testing for sparse graphs

Problem 41 Characterize the closures of testable properties.

For the dense case, this was done by Lovász and Szegedy [69].

Problem 42 Suppose that instead of exploring the neighborhood of a single random node, we could select two random nodes and test simple quantities associated with them, like distance, maximum flow, electrical resistance. What information can be gained by such tests? Is there a "complete" set of tests that would give enough information to determine the global structure of the graph to a reasonable accuracy?

4 Edge coloring models

Problem 43 Characterize edge coloring functions whose value may be infinite on the nodeless loop.

Problem 44 Define limit objects for the edge-coloring models.

5 Misc

Problem 45 Is $\mathcal{E}(W, H)$ testable as a parameter of H, for a fixed W?