Hardness of Approximating Graph Transformation Problem*

Chih-Long Lin

Department of Computer Science
State University of New York at Stony Brook
Stony Brook, NY 11794-4400, U.S.A.
e-mail: longlin@sbcs.sunysb.edu

1 Introduction

Given two graphs G and H, the *Graph Transformation* problem is to find the minimum integer k such that there exists a graph G' which is *isomorphic* to H, and can be obtained from G by *relocating* k edges of G, that is, removing k existing edges from G and adding k new edges to G. Given G, H and k, the decision problem GT of determining whether G can be made isomorphic to H by relocating k edges is NP-complete, even if both G and H are restricted to planar planar. Notice that if we fix $k = 0$, then the problem becomes the *graph isomorphism* problem, which is solvable in polynomial time on planar graphs [HW74]. Furthermore, it is interesting to note the following results on GT_k, the version of GT restricted to specific integer $k(m)$ (m being the size of G): (1) If $k(m) = O(1)$ then GT_k cannot be NP-complete unless $\Sigma_2^P = \Pi_2^P$ (cf. [Sch87]); and (2) if $k(m) = \Omega(m^\epsilon)$ for any constant ϵ, then GT_k is NP-complete, even restricted to planar graphs (details are omitted due to page limit).

In this paper, we consider the hardness of approximating the graph transformation problem. Formally speaking, for any two graphs $G = (V, E)$ and H, we write $G \xrightarrow{k} H$ if G can be made isomorphic to H by relocating k edges; let $\Delta(G, H)$ be the function computing the minimum integer k such that $G \xrightarrow{k} H$. Our main result states that if $\mathsf{P} \neq \mathsf{NP}$, then there exists a constant $c > 1$ such that no polynomial-time algorithm can approximate $\Delta(G, H)$ within the factor c.

Our result continues the recent trend of developing new complexity results on the approximation of intractable optimization problems. Based on a new characterization of NP using the concept of Probabilistic Checkable Proof systems, Arora et al. have shown that, if $\mathsf{P} \neq \mathsf{NP}$, then the problem MAX-CLIQUE of finding a maximum-sized clique in a given graph does not have a polynomial-time approximation of the optimum solution to within a factor of n^ϵ for some constant $\epsilon > 0$ [ALM+92]. Following this breakthrough, many new non-approxiability results have been developed based on reasonable assumptions on the relations among complexity classes such as P, NP, PSPACE, and $\mathsf{DTIME}(2^{poly(\log n)})$ (see, for example, [CFLS93, LY93b, ABSS93]).

* Research supported in part by NSF grant CCR 9121472.

Our problem is also closely related to a group of intractable optimization problems called the *maximum induced subgraph* problems using *node-deletion* or *edge-deletion* operations [LY80, AH82]. More precisely, for a given graph property Π, the node-deletion (or, edge-deletion) problem for Π is to find an induced subgraph of a given graph G with property Π by removing a *minimum* number of nodes (or, respectively, edges) from G. Let MAX-NODE-DELETION (and MAX-EDGE-DELETION) denote the problem of finding the maximum induced subgraph using node-deletion (and, respectively, edge-deletion) operations. It was shown in [LY93a] that the MAX-NODE-DELETION problem for a wide class of graph properties cannot have any polynomial-time constant-ratio approximations if $\mathsf{NP} \not\subseteq \mathsf{DTIME}(2^{poly(\log n)})$. There is, however, no similar result known for the MAX-EDGE-DELETION problems.

The rest of the paper can be briefly summarized as follows. In section 2, we present a general concept of reductions from decision problems to approximation problems; this concept generalizes the notion of *linear reduction*, or *L-reduction*, introduced by [PY91] to capture the idea of reductions that preserve approximation ratios up to a constant factor. The main result is proved in section 3. Modifying the NP-completeness proof for 3DM, the three-dimensional matching problem (see [GJ79, pages 50–53]), we reduce to our problem the MAX-3SAT-B problem, which is of finding the maximum number of simultaneously satisfiable three-literal clauses among a given set of such clauses, with the number of occurrences for each literal *bounded* by some constant B [PY91]. The non-approximability of the MAX-3SAT-B problem is a direct corollary of [ALM+92] and [PY91].

2 Preliminaries

In this section, we define the notion of *G-reductions*. Let Q^+ be positive rationals and R^+ be positive reals.

Definition 1. Let $f, g : \Sigma^* \to \mathsf{Q}^+$ and $c : \mathsf{N} \to \mathsf{R}^+$, $c(n) > 1$ for all n, be given. We say that g *approximates* f to *within a factor of* c (c-approximates f in short) if for all x, $x \in \Sigma^*$, we have $f(x)/c(|x|) < g(x) < c(|x|) \cdot f(x)$. The *c-approximation problem* of f is to compute a function g that c-approximates f.

To characterize the hardness of approximating *function-evaluation* problems in contrast to standard decision problems, we first define the notation of a *pair* of hard problems.

Definition 2. Let $A, B \subseteq \Sigma^*$, $A \cap B = \emptyset$, and $\mathcal{C}$ be a decision problem class. We say $\langle A, B\rangle \in \mathcal{C} \times co\text{-}\mathcal{C}$ if $A \in \mathcal{C}$ and $B \in co\text{-}\mathcal{C}$. Given two pairs $\langle A, B\rangle$ and $\langle A', B'\rangle$ in $\mathcal{C} \times co\text{-}\mathcal{C}$, we say that $\langle A, B\rangle$ *G-reduces* to $\langle A', B'\rangle$ if there is a polynomial time computable function f such that $f(A) \subseteq A'$ and $f(B) \subseteq B'$. We say that $\langle A, B\rangle$ is $\mathcal{C}$-hard if there exists a $\mathcal{C}$-hard set C such that $\langle C, \overline{C}\rangle$ G-reduces to $\langle A, B\rangle$.