18.405J/6.841J Advanced Complexity Theory	April 16, 1998
Lecture 18	
Lecturer: Daniel A. Spielman	*Scribe: Originally E.Lehman, edited by A.Lysyanskaya*

In the last lecture, we constructed an interactive proof of graph non-isomorphism. This construction relied on the verifier having "private coins", inaccessible to the prover. In this lecture, we will construct an interactive proof for (a special case of) graph non-isomorphism that uses only "public coins", visible to both the prover and verifier. In a subsequent lecture, we will see that private coins can *always* be replaced by public coins. Formally:

Theorem 1 $IP(f(n)) \subseteq AM(f(n)+2)$

That is, for any language with an interactive proof, there is an Arthur-Merlin game with two more rounds.

1 Graph Automorphisms

In this lecture we prove the preceding theorem only for a special case of graph-nonisomorphism. In this special case, we prove the non-isomorphism of two connected graphs, A and B, each with only the trivial automorphism[1]. Hereafter, n will denote the number of nodes in both A and B. If they were different, the proof of non-isomorphism would be trivial.

If G is a graph, then let $[G]$ denote the set of graphs obtained by permuting the vertices of G. All of the graphs in $[G]$ have the same set of vertices, and all are isomorphic. However, a pair of vertices might be adjacent in one graph in $[G]$, but not in another.

For a graph G, the size of $[G]$ is $n!/|Aut(G)|$ where n is the number of vertices in G and $|Aut(G)|$ is the number of automorphisms of G. Since the graphs A and B have only one, trivial automorphism, $[A] = [B] = n!$.

Define $A \cup B$ to be the graph obtained by unioning the vertex and edge sets of A and B. (Intuitively, a picture of $A \cup B$ is obtained by drawing A beside B.) Assume that the vertex sets of A and B are disjoint, so that $A \cup B$ has $2n$ nodes.

If A and B are not isomorphic, then $A \cup B$ has only the trivial automorphism. If A and B are isomorphic, then $A \cup B$ has the trivial automorphism and one more automorphism obtained by swapping corresponding vertices of A and B. As consequences, we obtain two fact that we will use extensively:

- If A and B are not isomorphic, then $|[A \cup B]| = (2n)!$.
- If A and B are isomorphic, then $|[A \cup B]| = (2n)!/2$.

2 The Main Idea

The overall idea is that the prover convinces the verifier that $|[A \cup B]|$ is big— more like $(2n)!$ than $(2n)!/2$. By the preceding facts, this implies that A and B are not isomorphic.

This approach is not unlike the one used to show that BPP $\subseteq \Sigma_2\text{P} \cap \Pi_2\text{P}$. But there is a new complication. Informally, the earlier the problem was to determine whether a set was a "big

[1] Recall that an automorphism of a graph is a permutation of the vertices that leaves the edge set unchanged. For example, a cycle on n nodes has $2n$ automorphisms. Every graph has a trivial automorphism defined by the identity permutation.

fraction" or "small fraction" of a universe. The present problem is to determine whether a set is "a tiny fraction" or "half a tiny fraction" of a universe. In this case, the universe U is defined to be the set of all graphs with the same nodes as $A \cup B$. It is easy to see that $[A \cup B]$ is at most "a tiny fraction" of this universe U.

To see why this is a problem, consider the following procedure for proving graph non-isomorphism. The verifier uses public coins to generate a random graph $G \in U$. If $G \in [A \cup B]$, then the prover returns a proof of this fact in the form of a correspondence between the vertices of G and the vertices of $A \cup B$. The verifier looks at how often the prover returns a valid proof. If A and B are not isomorphic, then $[A \cup B]$ is twice as big as if they were isomorphic. This means that the prover can return a valid proof twice as often. The verifier detects this and concludes that A and B are not isomorphic.

The flaw in this procedure is that a random graph $G \in U$ is so rarely in $[A \cup B]$— whether A and B are isomorphic or not— that the number of rounds must be exponential for the verifier to reach any conclusion.

The solution is to use hashing. The set $[A \cup B]$ is a tiny fraction of U. But now we cleverly hash U down to a small set S. It will turn out that if A and B are not isomorphic, then the image of $[A \cup B]$ under the hash function will be a big fraction of S. If A and B are isomorphic, then the image of $[A \cup B]$ will be only a small fraction of S.

In this way, hashing from U to S converts an "tiny fraction" vs. "half a tiny fraction" problem in U into a "big fraction" vs. "small fraction" problem in S. Now a procedure similar to the one described above will work properly.

3 Universal Hashing

A family of hash functions $H : U \mapsto S$ is *universal* if two conditions hold. Probabilities are taken over the choice of hash function $h \in H$.

1. For all $s \in S$ and $x \in U$, $\Pr[h(x) = s] = \frac{1}{|S|}$.

2. For all $s_1, s_2 \in S$ and distinct $x_1, x_2 \in U$, $\Pr[h(x_1) = s_1 \wedge h(x_2) = s_2] = \frac{1}{|S|^2}$.

The family of all functions from U to S is universal, but is so large that even specifying a particular element requires excessive space.

Fortunately, more manageable universal hash families exist. For example, suppose that $U = \mathrm{GF}(2)^a$ and $S = \mathrm{GF}(2)^b$. Elements of U and S can be regarded as $0-1$ column vectors with sizes a and b. Then all functions of the form $x \mapsto Mx + N$ form a universal hash family, where M is a $0-1$ matrix and N is a $0-1$ column vector.

The following lemma is the tool we use to convert a "tiny fraction" vs. "half a tiny fraction" problem in U into a "big fraction" vs. "small fraction" problem in S.

Lemma 2 *Let H be a universal family of hash functions from U to S. Let T be a subset of U and define $\alpha = |T|/|S|$. Then*

$$\alpha - \frac{\alpha^2}{2} \leq \frac{\mathit{Exp}[|h(T)|]}{|S|} \leq \alpha$$

where the expectation is taken over the choice of hash function $h \in H$.

Proof The right inequality is easy. The image of T can never be larger than T, so $|h(T)|/|S| \leq |T|/|S| = \alpha$.

The left inequality is more tricky. The quantity $\mathrm{Exp}[|h(T)|]/|S|$ can be regarded as the probability that s is an element of $h(T)$ taken over a random choice of $h \in H$ and $s \in S$. We can bound this probability as follows:

$$
\begin{aligned}
\Pr[s \in h(T)] &\geq \sum_{x \in T} \Pr[s = h(x)] - \sum_{\{x,y\} \in T} \Pr[s = h(x) = h(y)] \\
&= \frac{|T|}{|S|} - \frac{\binom{|T|}{2}}{|S|^2} \\
&> \alpha - \frac{\alpha^2}{2}
\end{aligned}
$$

The first line uses an inclusion-exclusion expansion truncated after two terms. The second line follows from the definition of universal hashing. The final line follows from the definition of α.

■

Note that as α grows greater than 1, the lower bound $\alpha - \alpha^2$ begins to drop from 1/2. However, since $\text{Exp}[|h(T)|]/|S|$ must be non-decreasing as the set T grows, 1/2 is still a valid lower bound. We will use this fact.

4 Last Trick

Recall that if graphs A and B are not isomorphic, then the size of $[A \cup B]$ is $(2n)!$. If A and B are isomorphic, then $[A \cup B]$ is only half as large. For technical reasons we will need this ratio to be just a little greater than two-to-one even before we use hashing.

The trick is to look at a product:

$$\underbrace{[A \cup B] \times [A \cup B] \times \ldots \times [A \cup B]}_{k\ times} = [A \cup B]^k$$

If $[A \cup B]$ changes size by a factor of two, then the product changes size by a factor of 2^k. In fact, $k = 4$ will be sufficient for the procedure given below

5 Procedure

Now we can assemble the tools of the preceding sections into an interactive proof with public coins for graph-nonisorphism. Call the verifier Arthur and the prover Merlin.

Let U be the set of all k-tuples of graphs with the same nodes as $A \cup B$, encoded as $0-1$ column vectors. Let T be the subset of U corresponding to k-tuples of graphs in $[A \cup B]$. Let S be the set of 0-1 column vectors of size b, where b is defined by the inequality:

$$2^b \leq (2n)!^k < 2^{b+1}$$

Here k is a constant, the size of the product described in the preceding section. The Arthur-Merlin protocol for graph non-isomorphism is as follows:

Merlin presents a hash function h from the universal family defined previously.

Arthur chooses two random elements s_1 and s_2 from S and sends them to Merlin.

Merlin finds t_1 and t_2 in T such that $h(t_1) = s_1$ and $h(t_2) = s_2$ if such values exist. Merlin returns these values to Arthur along with proofs that they are elements of T.

Recall that elements of T correspond to k-tuples of graphs isomorphic to $A \cup B$. Therefore, Merlin's proofs indicate how vertices of $A \cup B$ are associated with vertices in each of the graphs in the k-tuples to give isomorphisms.

Arthur checks that $h(t_1) = s_1$ and $h(t_2) = s_2$ and checks Merlin's proofs that $t_1, t_2 \in T$. If either t_1 or t_2 passes both checks, then Arthur accepts.

All that remains is to verify the correctness of the procedure. We must show that if A and B are isomorphic, then Arthur accepts with probability at most $1/4$ and that if A and B are not isomorphic, then Arthur accepts with probability at least $3/4$.

First, suppose that A and B are isomorphic. Then $|[A \cup B]| = (2n)!/2$, $|T| = (2n)!^k/2^k$, and $|S| > (2n)!^k/2$. The probability that a random element of S is contained in $h(T)$ is at most $|h(T)|/|S| \le |T|/|S| \le 1/2^{k-1}$. The probability that either s_1 or s_2 is an element of $h(T)$ is at most $1/2^{k-2}$ or at most $1/4$ for $k = 4$. Note that we used no special properties of the hash function here; no matter how Merlin chooses the hash function h, Arthur accepts with probability at most $1/4$.

Now suppose that A and B are not isomorphic. Then $|[A \cup B]| = (2n)!$, $|T| = (2n)!^k$, $|S| \le (2n)!^k$, and $\alpha = |T|/|S| \ge 1$. Lemma 2 says that $\text{Exp}\,[|h(T)|]/|S| \ge \alpha - \alpha^2/2 \ge 1/2$. This is a probabilistic proof that there exists some hash function h in the universal family such that $|h(T)|/|S| \ge 1/2$. If Merlin chooses this hash function, then the probability that either s_1 or s_2 is and element of $h(T)$ is at least $3/4$. Therefore Arthur accepts with probability at least $3/4$.