Proof of Theorem (continued)

- Clearly, if $A_\phi > 0$, the protocol convinces Bob of this.
- We next show that if $A_\phi = 0$, then Bob will be cheated with only negligible probability.

Lemma 90 *Suppose $A_\phi = 0$ and Alice claims a nonzero value $\boldsymbol{a}$. Then with probability $\geq (1 - \frac{2n}{2^n})^{i-1}$, the value of $\boldsymbol{a}$ claimed at the ith stage is wrong.*

©2006 Prof. Yuh-Dauh Lyuu, National Taiwan University

Proof of Lemma 90 (continued)

- The first $\boldsymbol{a}$ claimed by Alice is nonzero, which is certainly wrong.
- The lemma therefore holds for $i = 1$.
- By induction, for $i > 1$, the $(i-1)$st value was wrong with probability $\geq (1 - \frac{2n}{2^n})^{i-2}$.
- Suppose it is indeed wrong.
- The polynomial $A'(x)$ produced by Alice in the ith stage must be such that $A'(0) \cdot A'(1)$ or $A'(0) + A'(1)$ equals the wrong value $\boldsymbol{a}$.

Proof of Lemma 90 (continued)

- Alice must therefore supply a wrong polynomial $A'(x)$, different from the true polynomial $C(x)$.
  - Recall that Bob uses $A'(x)$ not $C(x)$.
- $C(x) - A'(x)$ is a polynomial of degree $2n$.
- Hence it has at most $2n$ roots.
- The random number between 0 and $p - 1$ picked by Bob will be one of these roots with probability at most $2n/p$.

Proof of Lemma 90 (concluded)

- The probability that $\boldsymbol{a}$ at the ith stage is *correct* is

$$
\begin{aligned}
&\leq \left[1-\left(1-\frac{2n}{2^n}\right)^{i-2}\right]\left(1-\frac{2n}{p}\right)\\
&\leq 1-\left(1-\frac{2n}{2^n}\right)^{i-2}\left(1-\frac{2n}{p}\right)\\
&\leq 1-\left(1-\frac{2n}{2^n}\right)^{i-1}.
\end{aligned}
$$

  - Recall that $p \geq 2^n$.

Proof of Theorem (concluded)

- In the last round, Bob will catch Alice's deception with probability $(1 - \frac{2n}{2^n})^n \to 1$.
- To achieve the confidence level of $1 - 2^{-n}$ required by the definition of IP, simply repeat the protocol.

©2006 Prof. Yuh-Dauh Lyuu, National Taiwan University

The Algorithm

```
Alice and Bob both arithmetize $\phi$ to obtain $\Phi$;
Alice picks a prime $p$ and sends it to Bob;
Bob rejects if $p$ does not satisfy the desired conditions;
Alice claims $A_\phi = \boldsymbol{a} \bmod p$ to Bob;
Bob set $A = A_\phi$;
repeat
    Alice sends $A'(x)$ to Bob;
    Bob rejects if $\boldsymbol{a} \neq A'(0) \cdot A'(1) \bmod p$ when $A = \prod_x \cdots$ or
    $\boldsymbol{a} \neq A'(0) + A'(1) \bmod p$ when $A = \sum_x \cdots$;
    Bob picks a random number $r$ and sends it to Alice;
    Bob calculates $\boldsymbol{a} = A'(r)$;
    Alice and Bob both set $A = A'(r)$; {Some details left out.}
until there no $\prod$ or $\sum$ left in $A$
Bob accepts iff $A'(x)$ is as claimed in the last stage;
```

©2006 Prof. Yuh-Dauh Lyuu, National Taiwan University

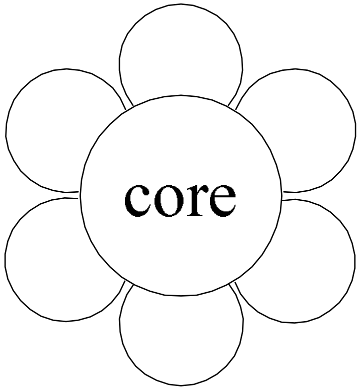
Exponential Circuit Complexity

- Almost all boolean functions require $\frac{2^n}{2n}$ gates to compute (generalized Theorem 14 on p. 153).
- Progress of using circuit complexity to prove exponential lower bounds for NP-complete problems has been slow.
  - As of January 2006, the best lower bound is $5n - o(n)$.[a]
- We next establish exponential lower bounds for depth-3 circuits.

[a]Iwama and Morizumi (2002).

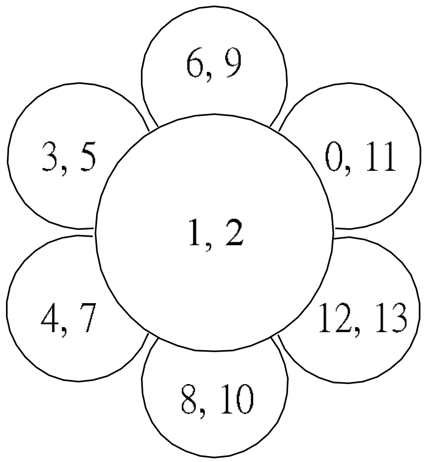
Sunflowers

- Fix $p \in \mathbb{Z}^+$ and $\ell \in \mathbb{Z}^+$.
- A **sunflower** is a family of p sets $\{P_1, P_2, \ldots, P_p\}$, called **petals**, each of cardinality at most ℓ.
- All pairs of sets in the family must have the same intersection (called the **core** of the sunflower).



A Sample Sunflower

$$\{\{1, 2, 3, 5\}, \{1, 2, 6, 9\}, \{0, 1, 2, 11\},$$
$$\{1, 2, 12, 13\}, \{1, 2, 8, 10\}, \{1, 2, 4, 7\}\}$$



The Erdős-Rado Lemma

Lemma 91 *Let $\mathcal{Z}$ be a family of more than $M = (p-1)^{\ell}\ell!$ nonempty sets, each of cardinality ℓ or less. Then $\mathcal{Z}$ must contain a sunflower (of size p).*

- Induction on ℓ.
- For $\ell = 1$, p different singletons form a sunflower (with an empty core).
- Suppose $\ell > 1$.
- Consider a *maximal* subset $\mathcal{D} \subseteq \mathcal{Z}$ of *disjoint* sets.
  - Every set in $\mathcal{Z} - \mathcal{D}$ intersects some set in $\mathcal{D}$.

The Proof of the Erdős-Rado Lemma (continued)

- Suppose $\mathcal{D}$ contains at least p sets.
  - $\mathcal{D}$ constitutes a sunflower with an empty core.
- Suppose $\mathcal{D}$ contains fewer than p sets.
  - Let D be the union of all sets in $\mathcal{D}$.
  - $|D| \leq (p-1)\ell$ and D intersects every set in $\mathcal{Z}$.
  - There is a $d \in D$ that intersects more than $\frac{M}{(p-1)\ell} = (p-1)^{\ell-1}(\ell-1)!$ sets in $\mathcal{Z}$.
  - Consider $\mathcal{Z}' = \{Z - \{d\} : Z \in \mathcal{Z}, d \in Z\}$.
  - $\mathcal{Z}'$ has more than $M' = (p-1)^{\ell-1}(\ell-1)!$ sets.
  - M' is just M with ℓ decreased by one.

The Proof of the Erdős-Rado Lemma (concluded)

- (continued)
  - $\mathcal{Z}'$ contains a sunflower by induction, say

$$\{P_1, P_2, \ldots, P_p\}.$$

  - Now,

$$\{P_1 \cup \{d\}, P_2 \cup \{d\}, \ldots, P_p \cup \{d\}\}$$

    is a sunflower in $\mathcal{Z}$.

Comments on the Erdős-Rado Lemma

- A family of more than M sets must contain a sunflower.
- **Plucking** a sunflower entails replacing the sets in the sunflower by its core.
- By repeatedly finding a sunflower and plucking it, we can reduce a family with more than M sets to a family with at most M sets.
- If $\mathcal{Z}$ is a family of sets, the above result is denoted by pluck($\mathcal{Z}$).

An Example of Plucking

- Recall the sunflower on p. 733:

$$\begin{aligned}\mathcal{Z} &= \{\{1,2,3,5\},\{1,2,6,9\},\{0,1,2,11\},\\ &\quad \{1,2,12,13\},\{1,2,8,10\},\{1,2,4,7\}\}\end{aligned}$$

- Then

$$\text{pluck}(\mathcal{Z}) = \{\{1,2\}\}.$$

Exponential Circuit Complexity for NP-Complete Problems

- We shall prove exponential lower bounds for NP-complete problems using *monotone* circuits.
  - Monotone circuits are circuits without $\neg$ gates.
- Note that this does not settle the P vs. NP problem or any of the conjectures on p. 489.

The Power of Monotone Circuits

- Monotone circuits can only compute monotone boolean functions.
- They are powerful enough to solve a P-complete problem, MONOTONE CIRCUIT VALUE (p. 241).
- There are NP-complete problems that are not monotone; they cannot be computed by monotone circuits at all.
- There are NP-complete problems that are monotone; they can be computed by monotone circuits.
  - HAMILTONIAN PATH and CLIQUE.

©2006 Prof. Yuh-Dauh Lyuu, National Taiwan University

$\text{CLIQUE}_{n,k}$

- $\text{CLIQUE}_{n,k}$ is the boolean function deciding whether a graph $G = (V, E)$ with n nodes has a clique of size k.
- The input gates are the $\binom{n}{2}$ entries of the adjacency matrix of G.
  - Gate g_{ij} is set to true if the associated undirected edge $\{ i, j \}$ exists.
- $\text{CLIQUE}_{n,k}$ is a monotone function.
- Thus it can be computed by a monotone circuit.
- This does not rule out that nonmonotone circuits for $\text{CLIQUE}_{n,k}$ may use fewer gates.

Crude Circuits

- One possible circuit for $\text{CLIQUE}_{n,k}$ does the following.
  1. For each $S \subseteq V$ with $|S| = k$, there is a subcircuit with $O(k^2)$ $\wedge$-gates testing whether S forms a clique.
  2. We then take an OR of the outcomes of all the $\binom{n}{k}$ subsets $S_1, S_2, \ldots, S_{\binom{n}{k}}$.
- This is a monotone circuit with $O(k^2 \binom{n}{k})$ gates, which is exponentially large unless k or $n-k$ is a constant.
- A **crude circuit** $\text{CC}(X_1, X_2, \ldots, X_m)$ tests if *any* of $X_i \subseteq V$ forms a clique.
  - The above-mentioned circuit is $\text{CC}(S_1, S_2, \ldots, S_{\binom{n}{k}})$.

Razborov's Theorem

Theorem 92 (Razborov (1985)) *There is a constant c such that for large enough n, all monotone circuits for* $\text{CLIQUE}_{n,k}$ *with $k = n^{1/4}$ have size at least $n^{cn^{1/8}}$.*

- We shall approximate any monotone circuit for $\text{CLIQUE}_{n,k}$ by a restricted kind of crude circuit.
- The approximation will proceed in steps: one step for each gate of the monotone circuit.
- Each step introduces few errors (false positives and false negatives).
- But the resulting crude circuit has exponentially many errors.

©2006 Prof. Yuh-Dauh Lyuu, National Taiwan University

The Proof

- Fix $k = n^{1/4}$.
- Fix $\ell = n^{1/8}$.
- Note that
$$2\binom{\ell}{2} \leq k.$$
- p will be fixed later to be $n^{1/8} \log n$.
- Fix $M = (p-1)^{\ell}\ell!$.
  - Recall the Erdős-Rado lemma (p. 734).

The Proof (continued)

- Each crude circuit used in the approximation process is of the form $\mathrm{CC}(X_1, X_2, \ldots, X_m)$, where:
  - $X_i \subseteq V$.
  - $|X_i| \leq \ell$.
  - $m \leq M$.
- We shall show how to approximate any circuit for $\textsc{clique}_{n,k}$ by such a crude circuit, inductively.
- The induction basis is straightforward:
  - Input gate g_{ij} is the crude circuit $\mathrm{CC}(\{i, j\})$.

The Proof (continued)

- Any monotone circuit can be considered the OR or AND of two subcircuits.
- We shall show how to build approximators of the overall circuit from the approximators of the two subcircuits.
  - We are given two crude circuits CC($\mathcal{X}$) and CC($\mathcal{Y}$).
  - $\mathcal{X}$ and $\mathcal{Y}$ are two families of at most M sets of nodes, each set containing at most ℓ nodes.
  - We construct the approximate OR and the approximate AND of these subcircuits.
  - Then show both approximations introduce few errors.

©2006 Prof. Yuh-Dauh Lyuu, National Taiwan University

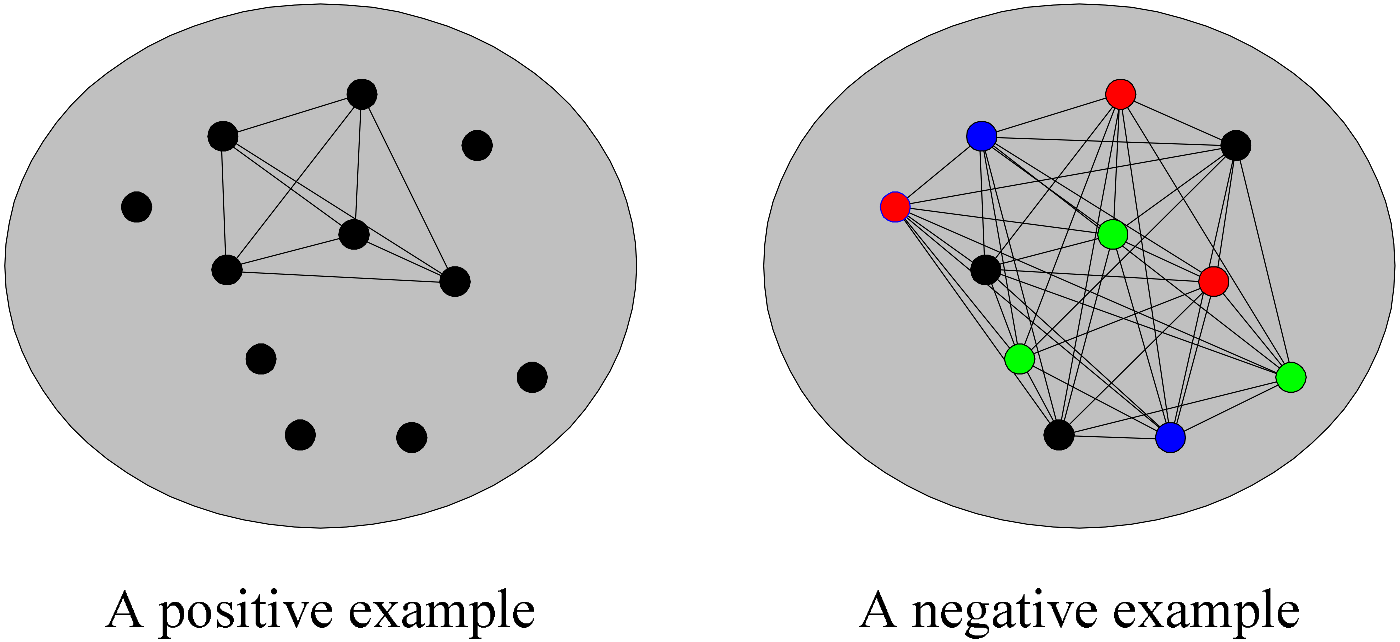
The Proof: Positive Examples

- Error analysis will be applied to only **positive examples** and **negative examples**.
- A positive example is a graph that has $\binom{k}{2}$ edges connecting k nodes in all possible ways.
- There are $\binom{n}{k}$ such graphs.
- They all should elicit a true output from $\text{CLIQUE}_{n,k}$.

The Proof: Negative Examples

- Color the nodes with $k - 1$ different colors and join by an edge any two nodes that are colored differently.
- There are $(k - 1)^n$ such graphs.
- They all should elicit a false output from $\text{CLIQUE}_{n,k}$.

Positive and Negative Examples with $k = 5$



The Proof: OR

- $\mathrm{CC}(\mathcal{X} \cup \mathcal{Y})$ is *equivalent to* the OR of $\mathrm{CC}(\mathcal{X})$ and $\mathrm{CC}(\mathcal{Y})$.
- Violations occur when $|\mathcal{X} \cup \mathcal{Y}| > M$.
- Such violations can be eliminated by using
$$\mathrm{CC}(\mathrm{pluck}(\mathcal{X} \cup \mathcal{Y}))$$
as the approximate OR of $\mathrm{CC}(\mathcal{X})$ and $\mathrm{CC}(\mathcal{Y})$.
- We now count the numbers of errors this approximate OR makes on the positive and negative examples.

©2006 Prof. Yuh-Dauh Lyuu, National Taiwan University

The Proof: OR (concluded)

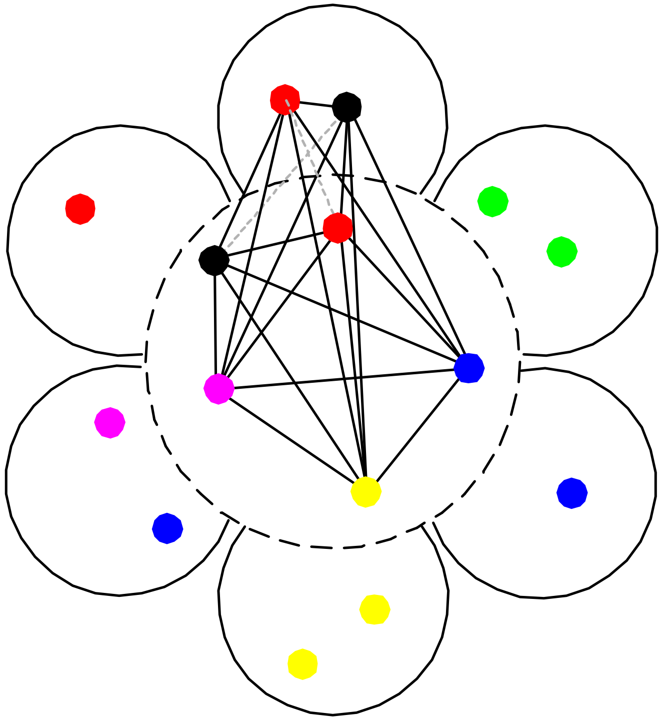
- CC(pluck($\mathcal{X} \cup \mathcal{Y}$)) *introduces* a **false positive** if a negative example makes both CC($\mathcal{X}$) and CC($\mathcal{Y}$) return false but makes CC(pluck($\mathcal{X} \cup \mathcal{Y}$)) return true.
- CC(pluck($\mathcal{X} \cup \mathcal{Y}$)) *introduces* a **false negative** if a positive example makes either CC($\mathcal{X}$) or CC($\mathcal{Y}$) return true but makes CC(pluck($\mathcal{X} \cup \mathcal{Y}$)) return false.
- How many false positives and false negatives are introduced by CC(pluck($\mathcal{X} \cup \mathcal{Y}$))?

The Number of False Positives

Lemma 93 CC(pluck($\mathcal{X} \cup \mathcal{Y}$)) *introduces at most* $\frac{M}{p-1}\, 2^{-p}(k-1)^n$ *false positives.*

- Assume a plucking replaces the sunflower $\{Z_1, Z_2, \ldots, Z_p\}$ with its core Z.
- A false positive is *necessarily* a coloring such that:
  - There is a pair of identically colored nodes in each petal Z_i (and so both crude circuits return false).
  - But the core contains distinctly colored nodes.
    * This implies at least one node from each same-color pair was plucked away.
- We now count the number of such colorings.

Proof of Lemma 93 (continued)



Proof of Lemma 93 (continued)

- Color nodes V at random with $k-1$ colors and let $R(X)$ denote the event that there are repeated colors in set X.
- Now $\text{prob}[R(Z_1) \wedge \cdots \wedge R(Z_p) \wedge \neg R(Z)]$ is at most

$$\begin{aligned} & \text{prob}[R(Z_1) \wedge \cdots \wedge R(Z_p) | \neg R(Z)] \\ = \quad & \prod_{i=1}^{p} \text{prob}[R(Z_i) | \neg R(Z)] \leq \prod_{i=1}^{p} \text{prob}[R(Z_i)]. \quad (16) \end{aligned}$$

  - First equality holds because $R(Z_i)$ are independent given $\neg R(Z)$ as Z contains their only common nodes.
  - Last inequality holds as the likelihood of repetitions in Z_i decreases given no repetitions in $Z \subseteq Z_i$.

Proof of Lemma 93 (continued)

- Consider two nodes in Z_i.
- The probability that they have identical color is $\frac{1}{k-1}$.
- Now $\text{prob}[\, R(Z_i) \,] \le \frac{\binom{|Z_i|}{2}}{k-1} \le \frac{\binom{\ell}{2}}{k-1} \le \frac{1}{2}$.
- So the probability[a] that a random coloring is a new false positive is at most 2^{-p} by inequality (16).
- As there are $(k-1)^n$ different colorings, each plucking introduces at most $2^{-p}(k-1)^n$ false positives.

[a] Proportion, i.e.

Proof of Lemma 93 (concluded)

- Recall that $|\,\mathcal{X} \cup \mathcal{Y}\,| \leq 2M$.
- Each plucking reduces the number of sets by $p - 1$.
- Hence at most $\frac{M}{p-1}$ pluckings occur in pluck$(\mathcal{X} \cup \mathcal{Y})$.
- At most
$$\frac{M}{p-1}\, 2^{-p}(k-1)^n$$
false positives are introduced.

The Number of False Negatives

Lemma 94 $\mathrm{CC}(\mathrm{pluck}(\mathcal{X} \cup \mathcal{Y}))$ *introduces no false negatives.*

- Each plucking replaces a set in a crude circuit by a subset.
- This makes the test less stringent.
  - For each $Y \in \mathcal{X} \cup \mathcal{Y}$, there must exist at least one $X \in \mathrm{pluck}(\mathcal{X} \cup \mathcal{Y})$ such that $X \subseteq Y$.
  - So if $Y \in \mathcal{X} \cup \mathcal{Y}$ is a clique, then $\mathrm{pluck}(\mathcal{X} \cup \mathcal{Y})$ also contains a clique, in X.
- So plucking can only increase the number of accepted graphs.

The Proof: AND

- The approximate AND of crude circuits $\mathrm{CC}(\mathcal{X})$ and $\mathrm{CC}(\mathcal{Y})$ is

$$\mathrm{CC}(\mathrm{pluck}(\{X_i \cup Y_j : X_i \in \mathcal{X}, Y_j \in \mathcal{Y}, |X_i \cup Y_j| \leq \ell\})).$$

- We now count the numbers of errors this approximate AND makes on the positive and negative examples.

©2006 Prof. Yuh-Dauh Lyuu, National Taiwan University

The Proof: AND (concluded)

- The approximate AND *introduces* a **false positive** if a negative example makes either CC($\mathcal{X}$) or CC($\mathcal{Y}$) return false but makes the approximate AND return true.
- The approximate AND *introduces* a **false negative** if a positive example makes both CC($\mathcal{X}$) and CC($\mathcal{Y}$) return true but makes the approximate AND return false.
- How many false positives and false negatives are introduced by the approximate AND?

The Number of False Positives

Lemma 95 *The approximate* AND *introduces at most $M^2 2^{-p}(k-1)^n$ false positives.*

- $\mathrm{CC}(\{X_i \cup Y_j : X_i \in \mathcal{X}, Y_j \in \mathcal{Y}\})$ introduces no false positives.
  - If $X_i \cup Y_j$ is a clique, both X_i and Y_j must be cliques, making both $\mathrm{CC}(\mathcal{X})$ and $\mathrm{CC}(\mathcal{Y})$ return true.
- $\mathrm{CC}(\{X_i \cup Y_j : X_i \in \mathcal{X}, Y_j \in \mathcal{Y}, |X_i \cup Y_j| \leq \ell\})$ introduces no false positives for the same reason as above.

Proof of Lemma 95 (concluded)

- $|\{X_i \cup Y_j : X_i \in \mathcal{X}, Y_j \in \mathcal{Y}, |X_i \cup Y_j| \leq \ell\}| \leq M^2$.
- Each plucking reduces the number of sets by $p - 1$.
- So pluck$(\{X_i \cup Y_j : X_i \in \mathcal{X}, Y_j \in \mathcal{Y}, |X_i \cup Y_j| \leq \ell\})$ involves $\leq M^2/(p-1)$ pluckings.
- Each plucking introduces at most $2^{-p}(k-1)^n$ false positives by the proof of Lemma 93 (p. 752).
- The desired upper bound is

$$[\,M^2/(p-1)\,]\,2^{-p}(k-1)^n \leq M^2 2^{-p}(k-1)^n.$$

©2006 Prof. Yuh-Dauh Lyuu, National Taiwan University

The Number of False Negatives

Lemma 96 *The approximate* AND *introduces at most* $M^2\binom{n-\ell-1}{k-\ell-1}$ *false negatives.*

- We follow the same three-step proof as before.
- $\text{CC}(\{X_i \cup Y_j : X_i \in \mathcal{X}, Y_j \in \mathcal{Y}\})$ introduces no false negatives.
  - Suppose both $\text{CC}(\mathcal{X})$ and $\text{CC}(\mathcal{Y})$ accept a positive example with a clique of size k.
  - This clique must contain an $X_i \in \mathcal{X}$ and a $Y_j \in \mathcal{Y}$.
    - This is why both $\text{CC}(\mathcal{X})$ and $\text{CC}(\mathcal{Y})$ return true.
  - As the clique contains $X_i \cup Y_j$, the new circuit returns true.

©2006 Prof. Yuh-Dauh Lyuu, National Taiwan University

Proof of Lemma 96 (concluded)

- $\mathrm{CC}(\{X_i \cup Y_j : X_i \in \mathcal{X}, Y_j \in \mathcal{Y}, |X_i \cup Y_j| \leq \ell\})$ introduces $\leq M^2 \binom{n-\ell-1}{k-\ell-1}$ false negatives.
  - Deletion of set $Z = X_i \cup Y_j$ larger than ℓ introduces false negatives which are cliques containing Z.
  - There are $\binom{n-|Z|}{k-|Z|}$ such cliques.
    * It is the number of positive examples whose clique contains Z.
  - $\binom{n-|Z|}{k-|Z|} \leq \binom{n-\ell-1}{k-\ell-1}$ as $|Z| > \ell$.
  - There are at most M^2 such Zs.
- Plucking introduces no false negatives.

Two Summarizing Lemmas

From Lemmas 93 (p. 752) and 95 (p. 760), we have:

Lemma 97 *Each approximation step introduces at most* $M^2 2^{-p}(k-1)^n$ *false positives.*

From Lemmas 94 (p. 757) and 96 (p. 762), we have:

Lemma 98 *Each approximation step introduces at most* $M^2\binom{n-\ell-1}{k-\ell-1}$ *false negatives.*

The Proof (continued)

- The above two lemmas show that each approximation step introduce “few” false positives and false negatives.
- We next show that the resulting crude circuit has “a lot” of false positives or false negatives.

The Final Crude Circuit

Lemma 99 *Every final crude circuit either is identically false—thus wrong on all positive examples—or outputs true on at least half of the negative examples.*

- Suppose it is not identically false.
- By construction, it accepts at least those graphs that have a clique on some set X of nodes, with $|X| \leq \ell$, which at $n^{1/8}$ is less than $k = n^{1/4}$.
- The proof of Lemma 93 (p. 752ff) shows that at least half of the colorings assign different colors to nodes in X.
- So half of the negative examples have a clique in X and are accepted.

The Proof (continued)

- Recall the constants on p. 744: $k = n^{1/4}$, $\ell = n^{1/8}$, $p = n^{1/8} \log n$, $M = (p-1)^{\ell} \ell! < n^{(1/3) n^{1/8}}$ for large n.
- Suppose the final crude circuit is identically false.
  - By Lemma 98 (p. 764), each approximation step introduces at most $M^2 \binom{n-\ell-1}{k-\ell-1}$ false negatives.
  - There are $\binom{n}{k}$ positive examples.
  - The original crude circuit for $\text{CLIQUE}_{n,k}$ has at least

$$\frac{\binom{n}{k}}{M^2 \binom{n-\ell-1}{k-\ell-1}} \geq \frac{1}{M^2} \left(\frac{n-\ell}{k} \right)^{\ell} \geq n^{(1/12) n^{1/8}}$$

gates for large n.

The Proof (concluded)

- Suppose the final crude circuit is not identically false.
  - Lemma 99 (p. 766) says that there are at least $(k-1)^n/2$ false positives.
  - By Lemma 97 (p. 764), each approximation step introduces at most $M^2 2^{-p}(k-1)^n$ false positives.
  - The original crude circuit for $\text{CLIQUE}_{n,k}$ has at least

$$\frac{(k-1)^n/2}{M^2 2^{-p}(k-1)^n} = \frac{2^{p-1}}{M^2} \geq n^{(1/3)n^{1/8}}$$

gates.

©2006 Prof. Yuh-Dauh Lyuu, National Taiwan University

$P \neq NP$ Proved?

- Razborov's theorem says that there is a monotone language in NP that has no polynomial monotone circuits.
- If we can prove that all monotone languages in P have polynomial monotone circuits, then $P \neq NP$.
- But Razborov proved in 1985 that some monotone languages in P have no polynomial monotone circuits!

Finis