Lecture 9: Complete Languages (Concluded)

1 Introduction

We are on the subject of "natural complete languages" for the various canonical classes.

An interesting theme in this area is that we can define families of closely related problems, where each member of the family is complete for some member of the canonical list. An example comes from theorem proving. Cook show that SAT is NP-complete. By restrictions of this problem, we can show construct problems that are complete for NL and P. This lecture shows:

(1) 2UNSAT is NL-complete.

(2) 3UNIT is P-complete.

On the other hand, if we generalize SAT (to allow quantification, we can get a problem that is complete for $PSPACE$.

2 Background in Theorem Proving

Some concepts: boolean variable, literal, formula, assignment, satisfaction, CNF, DNF, clauses. If we are interested in satisfiablity, then we may restrict our formulas to CNF.

3 2UNSAT

We already saw that GAP is NL-complete. We now show that the problem 2UNSAT is NL-complete:

2UNSAT:

Given: 2CNF formula F

Property: F is unsatisfiable

Curiosity: we have "unsatisfiablity" not "satisfiability" in this problem.

Fix a 2CNF formula F. Write $u \to v$ if $\{\overline{u}, v\} \in F$. Suppose I is a satisfying assignment for F.

(i) If $u \to v$ and $I(u) = 1$ then $I(v) = 1$.

(ii) $\stackrel{*}{\to}$ is reflexive and transitive.

(iii) $u \stackrel{*}{\to} v$ if and only if $\bar{v} \stackrel{*}{\to} \bar{u}$.

Thus if $u \stackrel{*}{\to} \bar{u}$ and I satisfies F then $I(u) = 0$.

Define $cl(X) = \{v : (\exists u \in X) u \stackrel{*}{\to} v\}$.

THEOREM: F is unsatisfiable iff there is a literal u such that $u \stackrel{*}{\to} \bar{u}$ and $\bar{u} \stackrel{*}{\to} u$.

Using this lemma, we see that 2UNSAT is in $NLOG$.

To show that 2UNSAT is $NLOG$-hard, let us reduce GAP to it: suppose $\langle n, G\rangle$ is input to GAP. We construct the following CNF formula:

$$1 \wedge \bar{n} \wedge \bigwedge_{\langle i,j\rangle} (\bar{i} \vee j)$$

where $\langle i, j\rangle$ range over edges in G. Note that this is close to a 2CNF.

Interpret: variable i means "node i is reachable from node 1".

(i) Suppose there is a path from 1 to n. Then this formula is not satisfiable.

(ii) Suppose there is no path. Then assign all nodes that can reach n to be false. Assign the remaining nodes to be true.

4 3UNIT

Unit Resolution: if C, C' are clauses, then they are RESOLVABLE if there is some u such that $u \in C$ and $\overline{u} \in C'$. Their resolution is $C \cup C' \setminus \{u, \overline{u}\}$. A clause is UNIT if it has only one literal. The resolution step is UNIT if C or C' is unit. A DEDUCTION of C_m from F is a sequence

$$C_1, \ldots, C_m = \Box$$

where C_i is either in F or resolved from two previous clauses. It is a UNIT DEDUCTION if each resolution step is UNIT.

RESOLUTION THEOREM: F is unsatisfiable iff there is a deduction of $\Box$ from F.

Here is the problem of 3UNIT:

GIVEN: a 3CNF formula F.

PROPERTY: $\Box$ can be deduced from F using unit deduction.

FACT: 3UNIT is in P.

THEOREM: 3UNIT is P-hard.

Proof: Similar to Cook's proof.

Let M be a TM that accepts in deterministic time $< n^k$ for some k. MAY ASSUME: M is simple, never moves left of its initial tape cell, and accepts by writing the symbol a_0 on cell 1 after entering the accept state q_a.

Let $C_1, \ldots, C_m$ be an accepting configuration $m = n^k$.

Represent C_i as the string $w_1[q, a]w_2$ where $w_1 a w_2 \in \Sigma^*$ and $q \in Q$. We will pad this string with blanks to have length exactly n^k. Let $P_{i,t}^a$ say that "symbol a is in the ith cell of C_t". Here $a \in \Sigma \cup (\Sigma \times Q)$.

For any input $x = a_1, \ldots, a_n$, let $f(x) = F_0 \wedge \cdots F_{m+1}$.

- F_0 says that C_0 is the initial configuration.
- F_{m+1} is the unit clause $\neg P_{1,m}^{[a_0, q_a]}$.
- F_i says that C_i is obtained from C_{i-1} for $i = 1, \ldots, m$.

DETAILS: let $\partial(a, b, c) = b'$ say that if the neighbors of symbol b are a, c then b is next replaced by b'. Then

$$F_i = \bigwedge_i \left(\bigwedge_{a,b,c} (\neg P_{i-1,t-1}^a \vee \neg P_{i,t-1}^b \vee \neg P_{i+1,t-1}^c \vee P_{i,t}^{\partial(a,b,c)}) \right).$$

CLAIM: (a) $P_{i,t}^a$ is unit deducible from $f(x)$ iff a is the i symbol in C_t. (Use induction on t. (b) No other other unit clauses can be unit deduced.

Now we show: x is accepted by M iff $\Box$ is unit deducible from $f(x)$.

Q.E.D.