CSC240 Problem Set II

Administrativia: discussed problem 1 with Ming Feng Wan and Darren Chan, no extra material consulted

1. Given three variables, **P**, **Q**, and **R**, note that there are 8 possible truth value assignments:

P	**Q**	**R**
T	T	T
T	T	F
T	F	T
T	F	F
F	T	T
F	T	F
F	F	T
F	F	F

   Suppose a propositional formula f contains at most 3 different variables. Let $[f]_{\text{min}} = [f]$ be the representation of f in the DNF.

   Note that the truth value of each minterm in $[f]$ is completely determined by some row in the table above (if some minterm contains less than three variables, it can be considered *completable* by adding the value T in conjunction until there are three terms in each minterm). Since there are 8 rows, 8 clauses in DNF, each containing at most 3 variables, will completely determine any propositional formula.

   DNF allows us to reduce the number of sufficient clauses by half, since, given the fact that 8 clauses represent all possible propositional formulas, in each disjunction only one of the minterms determines whether the disjunction is true or false (in the sense that removing one of the minterms would not change the equivalent value of the clause representing the disjunction of two minterms).

   Suppose the maxmimum number of sufficient minterms is lower than 4. Since a propositional formula g can be constructed in DNF with 4 mutually exclusive minterms (see (2)), these terms cannot be sufficient.

   Therefore, 4 clauses are necessary and sufficient to represent all possible propositional formulas with at most three variables. Hence, the total frequency of variable occurence is at most 12.

2. Consider the propositional formula P IFF Q.

P IFF Q

is logically equivalent to
((P IMPLIES Q) AND (Q IMPLIES P))
which is logically equivalent to
((NOT P) OR Q) AND ((NOT Q) OR P)
which is logically equivalent to
((NOT P) OR Q) AND (NOT Q)) OR (((NOT P) OR Q) AND P)
which is logically equivalent to
((NOT P) AND (NOT Q)) OR
(Q AND (NOT Q)) OR
((NOT P) AND P) OR (Q AND P)
which is logically equivalent to
((NOT P) AND (NOT Q)) OR (P AND Q)

Let $A = P$ IFF Q be the propositional formula in P, Q.

From the second line of the derivation above, observe that

P IFF Q

is logically equivalent to
(NOT (P AND (NOT Q)) AND NOT (Q AND (NOT P)))
which is logically equivalent to
NOT ((P AND (NOT Q)) OR (Q AND (NOT P)))

and thus

NOT $A = (P$ AND (NOT Q)) OR ((NOT P) AND Q)

Consider now A IFF R. By substitution, we obtain that it is logically equivalent to

((NOT A) AND (NOT R)) OR (A AND R)

which is logically equivalent to
(((P AND (NOT Q)) OR ((NOT P) AND Q)) AND NOT R) OR
[([(NOT P) AND (NOT Q)] OR (P AND Q)) AND R]
which is logically equivalent to
(P AND (NOT Q) AND NOT R) OR
((NOT P) AND Q AND NOT R) OR
((NOT P) AND (NOT Q) AND R) OR
(P AND Q AND R)

Since there are 4 mutually exclusive conjunctive clauses, the formula cannot be simplified further. There are 3 variables in each clause, and thus the number of occurences is 12.

3. Note that CNF can be obtained from DNF by the following procedure:

   a) construct a truth table for all the variables and the corresponding values of f

   b) write down a column for the negation of f

   c) for each row where the negation of f is true, construct its minterm representation to obtain the DNF of the negation of f

   d) negate the negation of f to obtain CNF

By De Morgan's Laws, the number of variable frequencies in each negated minterm does not change. From (1), the number of variables for f in DNF is at most 12. Therefore, CNF also contains at most 12 variables in total.

4. Suppose a propositional formula f containing b binary connectives is given.

   Consider an algorithm:

   a) parenthesise the formula so that any pair of related subformulas is either a pair of literals in parentheses, a literal related to a parenthesised multiliteral subformula, itself partenthesisesd, a pair of parenthesised subformulas, also parenthesised

   b) write down all the pairs of related subformulas, first listing parenthesised literals, then literals with multilinear subformulas, then pairs of multilinear subformulas, and assign index to the each element in the resultant list

   c) for each parenthesised pair of literals, assign a new variable X_i to the corresponding pair, where i is the index of the pair.

   d) for each pairs of literals with multilinear subformulas assign new variables X_j to the pair, where j is the index of the corresponding pair, with multilinear subformulas substitituted by the variables given in the previous step. If there are no variables left for the pairs of literals, use variables for multilinears produced in the current step

   e) for each pair of multilinears, replace multilinears by the variables given in the previous step

   Note that the procedure terminates, since there is a finite number of variables in the propositional formula given.

   Note that the total number of new variables produced by the described algorithm is equal to the number of binary connectives, which itself corresponds to the number of literal-literal and literal-multiliteral pairs in f.

   Note that all new variables and the corresponding pairs make satisfiable clauses. Making a conjuction out of those clause, we obtain a formula which is satisfiable if and only if f is satisfiable. Note that, as described in the procedure, there are at most three variables in each obtained clause.

5. By (4), for any propositional formula f containing b binary connectives there is a conjuction g of at most b propositional formulas with at most 3 different variables each such that f is satisfiable if and only if g is satisfiable.

   By (3), each clause of such a representation can be written in CNF with the frequency of variable occurence of at most 12. This shapes the whole representation into the CNF, giving $[g]$. Note that each clause of $[g]$ may contain at most 1 or 2 variables originally present in f (by the algorithm). Observe that, since there are b binary connectives in f, it contains at most $b + 1$ different variables. Similarly, the number of different variables in $[g]$ which are also in f is at most $b + 1$.

   Because each clause in $[g]$ is in one-to-one correspondence with these connectives in f, and all the different variables in f are also contained in $[g]$, while in each clause the frequency of variable occurence is at most 12, then it follows by the remark in the previous paragraph that the frequency of variable occurence in $[g]$ corresponding to the variables which are both in $[g]$ and f is at most $12(b + 1)$.

   Therefore, if any propositional formula G is satisfiable if and only if a propositional formula H in conjunctive normal form, such that each clause of H contains at most

3 different variables, is satisfiable, then the number of occurrences of variables in H is at most 12 times the number of occurrences of variables in G.