Disjunctive Normal Form and Two Facts in Boolean Logic

The following material is used for proving 3SAT $\in$ NPC.

A *conjunct* is $l_1 \wedge \cdots \wedge l_m$ where the l_i are literals. A *disjunctive normal form* (*dnf*) is a disjunction of conjuncts, i.e., $C_1 \vee \cdots \vee C_n$ where the C_i are conjuncts. A *uniform* disjunctive normal form is a dnf where all the conjuncts C_i have the same number of literals. So every uniform dnf is in the form

$$\bigvee_{1 \le j \le n} \bigwedge_{1 \le i \le m} l_i^j = (l_1^1 \wedge \cdots \wedge l_m^1) \vee \cdots \vee (l_1^n \wedge \cdots \wedge l_m^n)$$

An example with m = 2, n = 3 is $\bigvee_{1 \le j \le 3} \bigwedge_{1 \le i \le 2} l_i^j = (l_1^1 \wedge l_2^1) \vee (l_1^2 \wedge l_2^2) \vee (l_1^3 \wedge l_2^3)$.

Fact 1 (General Distributive Law): Every uniform dnf has the following equivalent cnf:

$$\bigvee_{1 \le j \le n} \bigwedge_{1 \le i \le m} l_i^j \iff \bigwedge_{1 \le i_1,\ldots,i_n \le m} (l_{i_1}^1 \vee \cdots \vee l_{i_n}^n)$$

For example,

$$(l_1^1 \wedge l_2^1) \vee (l_1^2 \wedge l_2^2) \vee (l_1^3 \wedge l_2^3) \iff$$
$$(l_1^1 \vee l_1^2 \vee l_1^3) \wedge (l_1^1 \vee l_1^2 \vee l_2^3) \wedge (l_1^1 \vee l_2^2 \vee l_1^3) \wedge (l_1^1 \vee l_2^2 \vee l_2^3) \wedge$$
$$(l_2^1 \vee l_1^2 \vee l_1^3) \wedge (l_2^1 \vee l_1^2 \vee l_2^3) \wedge (l_2^1 \vee l_2^2 \vee l_1^3) \wedge (l_2^1 \vee l_2^2 \vee l_2^3)$$

Note that the left side has $m \times n$ literals while the right side has $m^n \times n$ literals.

Fact 2 Every cnf ϕ can be converted to a 3-cnf ϕ' such that ϕ is satisfiable iff ϕ' is satisfiable (but not necessarily equivalent).

Let ϕ be $C_1 \wedge \cdots \wedge C_n$ where C_i is $l_1 \vee \cdots \vee l_m$. Convert each C_i as follows:

If m = 1, convert to $l_1 \vee l_1 \vee l_1$
If m = 2, convert to $l_1 \vee l_2 \vee l_1$
If m = 3, convert to itself
If m > 3, convert to $(l_1 \vee l_2 \vee z_1) \wedge (\neg z_1 \vee l_3 \vee z_2) \wedge (\neg z_2 \vee l_4 \vee z_3) \wedge \cdots \wedge (\neg z_{m-3} \vee l_{m-1} \vee l_m)$ where the z_i, $1 \le i \le m-3$, are new variables.

For example, $l_1 \vee l_2 \vee l_3 \vee l_4 \vee l_5 \vee l_6$ is converted to $(l_1 \vee l_2 \vee z_1) \wedge (\neg z_1 \vee l_3 \vee z_2) \wedge (\neg z_2 \vee l_4 \vee z_3) \wedge (\neg z_3 \vee l_5 \vee l_6)$.

If m = 1, 2, or 3, the converted C_i has 3 literals. If m > 3, the converted C_i has $m + 2(m-3) = 3m - 6$ literals. Hence the converted C_i has $\max(3, 3m-6)$ literals. Hence ϕ' has $\Sigma_{1 \le i \le n} \max(3, 3m_i-6)$ literals where m_i is the # of literals in C_i.