Given a CNF formula $E = C_1 \land C_2 \ldots \land C_k$
with each $C_i = \bigwedge_{j=1}^{k_i} l_{ij}$,
we construct a 3-CNF formula $F$ as follows.

For each clause $C_i$ of $E$

1) if $C_i = (l)$ (i.e., a single literal)
   introduce two new variables $m_i, \bar{m}_i$, and replace $C_i$ by 4 clauses
   \[
   (l + m_i + \bar{m}_i), \quad (l + m_i + \bar{m}_i), \quad (l + m_i + \bar{m}_i), \quad (l + m_i + \bar{m}_i)
   \]

   Since $m_i, \bar{m}_i$ appear in all 4 combinations, the 4 clauses can be satisfied only if $l$ is true.

2) if $C_i = (l_1 + l_2)$
   introduce a new variable $z_i$, and replace $C_i$ by 2 clauses
   \[
   (l_1 + l_2 + z_i), \quad (l_1 + l_2 + \bar{z}_i)
   \]

3) if $C_i = (l_1 + l_2 + l_3)$, just leave it.

4) if $C_i = (l_1 + l_2 + \ldots + l_m)$ with $m \geq 4$
   introduce $y_1, y_2, \ldots, y_{m-3}$ and replace $C_i$ by
   \[
   (l_1 + l_2 + y_0), (l_1 + l_2 + y_0), (l_1 + l_2 + y_0), (l_1 + l_2 + y_0), \ldots
   \]
   \[
   + (l_{m-2} + \overline{y_{m-4}} + \overline{y_{m-3}}), (l_{m-2} + \overline{y_{m-4}} + \overline{y_{m-3}})
   \]
- An assignment $T$ satisfies if $E$ makes at least one literal of $C_i$ true. Let it be $t_j$.
  Then, by making $y_{j-1}, \ldots, y_{j-2}$ true and $y_{j-1}, \ldots, y_{j-3}$ false,
  we satisfy all clauses replacing $C_i$.
  Thus we can extend $T$ to satisfy $F$.

- Conversely, if $T$ makes all $t_j$ of $C_i$ false, then not all new clauses can be satisfied.
  Why? Each $y_j$ can make at most 1 clause true, but there are $m-2$ clauses and $m-3$ $y_j$'s.

The 3CNF formula $F$ is linear in $E$ and can be constructed in linear time.

We get: CSAT $\leq_{poly} 3$-SAT

$\Rightarrow$ from CSAT NP-hard, we get 3-SAT NP-hard.

We also know that 3SAT is NP, since SAT is NP.

$\Rightarrow$ 3-SAT is NP-complete.
Exercise (10.3.2): The problem 4TA-SAT is defined as follows:

Given a propositional formula \( E \), does \( E \) have at least 4 satisfying truth assignments?

Show that 4TA-SAT is NP-complete.

Proof:

1) 4TA-SAT is in NP

We devise a non-deterministic poly-time algorithm:

1. Guess 4 truth assignments \( T_1, T_2, T_3, T_4 \).
2. Check that \( T_1, T_2, T_3, T_4 \) all satisfy \( E \).

Note that both steps require time polynomial in the size of \( E \).

2) 4TA-SAT is NP-hard.

We show this by reducing SAT to 4TA-SAT.

Let \( E \) be a propositional formula, and let \( x_1, \ldots, x_n \) be all variables in \( E \).

We construct a new formula \( E' \) such that

\[ E \in \text{SAT} \iff E' \in \text{4TA-SAT} \]

Let \( y_1, y_2 \) be two new variables. Then

\[ E' = E \lor ((x_1 \lor \neg x_2 \lor \ldots \lor x_n) \lor (y_1 \lor \neg y_2) \lor (\neg y_1 \land y_2)) \]
Consider the truth assignments for $x_1, x_2, y_1, y_2$.

<table>
<thead>
<tr>
<th>$x_1, x_2, x_n$</th>
<th>$y_1, y_2$</th>
<th>Case 1</th>
<th>Case 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>$T$</td>
<td>$E$</td>
<td>$E' $</td>
</tr>
<tr>
<td>$T$</td>
<td>$T$</td>
<td>$F$</td>
<td>$T$</td>
</tr>
<tr>
<td>$T$</td>
<td>$F$</td>
<td>$F$</td>
<td>$T$</td>
</tr>
<tr>
<td>$F$</td>
<td>$F$</td>
<td>$F$</td>
<td>$T$</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td>$F$</td>
<td>$F$</td>
<td>$F$</td>
<td>$T$</td>
</tr>
</tbody>
</table>

$2^{n+2}$

$= 3$

$\geq 4$

Alternative solution:

$E' = E \land (y_1 \lor y_2 \lor y_3)$

- If $E$ is unsatisfiable, then $E'$ is unsatisfiable, and hence $E' \not\in 47A - SAT$

- If $E$ is satisfiable, then $E'$ has at least 7 satisfying truth assignments; these are obtained by combining
  - a TA for $x_1, x_2, x_n$ satisfying $E$ with
  - the 7 TAs for $y_1, y_2, y_3$ satisfying $y_1 \lor y_2 \lor y_3$