1 Review

1.1 Planning Problem and Plans

The planning problem we are considering is a 3-tuple \( \langle D, I, G \rangle \) described in the ADL language whose syntax is given in the book, where \( I \) is the initial state, \( G \) is the goal and \( D \) is the domain description.

**Plans.** We define a plan \( \pi \) as a sequence of sets of actions, i.e., \( \pi = A_0; A_1; \ldots; A_{n-1} \) where \( A_i, 0 \leq i \leq n - 1, \) is a set of actions. That is, we allow actions to be executed parallelly.

**Transition Function.** Given a planning problem \( \langle D, I, G \rangle \) with the set of actions \( A \) and set of fluents \( F \), the transition function \( \Phi \) of the planning domain that maps sets of actions and states into states is defined as follows.

\[
\Phi(A, \sigma) = \begin{cases} 
\bot & \text{if } \sigma = \bot \text{ or } \exists a \in A \text{ s.t. } a \text{ is not executable in } \sigma \text{ or } Add_A \cap Del_A \neq \emptyset \\
(\sigma \cup Add_A) \setminus Del_A & \text{otherwise}
\end{cases}
\]

(1)

where \( A \subseteq A \) is some set of actions, \( s \) is a state, \( Add_A = \bigcup_{a \in A}(Add_a) \), and \( Del_A = \bigcup_{a \in A}(Del_a) \).

It is worthwhile extending the definition of \( \Phi \) for a plan \( \pi \) as well. More precisely, the transition function \( \Phi \) for plans are defined recursively as follows:

\[
\Phi([A], \sigma) = \Phi(A, \sigma)
\]

\[
\Phi([A;p], \sigma) = \Phi(p, \Phi(A, \sigma)) \text{ if } p \neq \emptyset
\]

\( \sigma_0 = I \) is called the initial state. We say that \( G \) can be achieved from \( \sigma_0 \) after executing plan \( \pi \) if \( \Phi(\pi, \sigma_0) \neq \bot \) and \( \Phi(\pi, \sigma_0) \) satisfies \( G \).

1.2 SAT Encodings

Given a planning problem \( \langle D, I, G \rangle \) in the ADL language with the transition function given in (1), the general procedure for translating \( P \) into the corresponding propositional theory \( T \) with \( n \) time instants is sketched out as follows:

- Each fluent \( f \in F \) corresponds to the set of propositions \( \{f^i\} \) in \( T \), where \( i = 0 \ldots n \).
- Each action \( a \in A \) corresponds to the set of propositions \( \{a^i\} \), where \( i = 0 \ldots n - 1 \).
- The encoding consists of the following sentences

\[
\bigwedge_{f \in F} f^0
\]

(2)

\[
\bigwedge_{g \in G} g^n
\]

(3)

\[
a^i \Rightarrow \bigwedge_{l \in Pr_c a} l^i \quad \forall a \in A, i = 0 \ldots n - 1
\]

(4)

\[
f^{i+1} \leftrightarrow (f^i \land \left( \bigwedge_{a \in A, f \in Del_a} \neg a^i \right)) \lor \left( \bigvee_{a \in A, f \in Add_a} a^i \right) \quad \forall f \in F, i = 0 \ldots n - 1
\]

(5)

\[
\neg f^{i+1} \leftrightarrow (-f^i \land \left( \bigwedge_{a \in A, f \in Add_a} \neg a^i \right)) \lor \left( \bigvee_{a \in A, f \in Del_a} a^i \right) \quad \forall f \in F, i = 0 \ldots n - 1
\]

(6)
1.3 Correctness

We will show that the encoding $T$ from (2)-(6) is equivalent to the original planning problem $P = \langle D, I, G \rangle$.

2 Proof of the Correctness

To prove the correctness of the SAT-encoding, we need prove the following

1. [Soundness] If $T$ has a model $M$ then we can extract from $M$ a plan $\pi$ for $P$ that achieves $G$ from the initial state $I$.
2. [Completeness] If $P$ has a plan $\pi$ of length $n$ that achieves $G$ from the initial state then $T$ has at least one model $M$ from which $\pi$ can be extracted.

2.1 Soundness

Assume that $M$ is a model of $T$. Let $A_i = \{ a \mid a$ is an action such that $a^t \in M \}$, $\pi = [A_0; A_1; \ldots ; A_{n-1}]$ and $s_i = \{ f \mid f^t \in M \}$. Note that $A_i's$ might be empty. By (2) and (3), it is easy to see that $s_0$ is the initial state of $P$ and $G$ holds in $s_n$. Before proving the soundness of the encoding, let us prove the following lemmas.

Lemma 1 For every integer $i$, $0 \leq i \leq n$, a fluent literal $l$ holds in a state $s_i$ if and only if $M \models l^i$.

Proof. It is obvious from the definition of $s_i$.

Lemma 2 For every integer $i$, $0 \leq i \leq n-1$, and for every action $a$ in $A_i$, $a$ is executable in $s_i$.

Proof. Let $l$ be a fluent literal in $Pre_a$. By the definition of $A_i$, we have $M \models a^i$. On the other hand, $a^i \Rightarrow l^i$ (by (4)). Hence, $M \models l^i$. As a result, $l^i$ holds in $s_i$ (by Lemma 1). That is, $a$ is executable in $s_i$.

Lemma 3 For every integer $i$, $0 \leq i \leq n-1$, $Add_{A_i} \cap Del_{A_i} = \emptyset$.

Proof. Assume that $Add_{A_i} \cap Del_{A_i} \neq \emptyset$. Thus, there exists a fluent $f \in F$ s.t. $f \in Add_{A_i} \cap Del_{A_i}$. This implies that there exist two actions $b, c \in A_i$ s.t. $f \in Add_b$ and $f \in Del_c$.

By (5), we have $f^{i+1}$ is true in $M$. And by (6), we have $\neg f^{i+1}$ is true in $M$. This is a contradiction. Therefore, $Add_{A_i} \cap Del_{A_i} = \emptyset$.

Lemma 4 For every integer $i$, $0 \leq i \leq n-1$, we have $\Phi(A_i, s_i) = s_{i+1}$.

Proof. According to Lemma 2 and Lemma 3, every action $a$ in $A_i$ is executable in $s_i$ and $Add_{A_i} \cap Del_{A_i} = \emptyset$ for all $0 \leq i \leq n-1$. Therefore, according to (1), $\Phi(A_i, s_i) = (s_i \cup Add_{A_i}) \backslash Del_{A_i} = s$ for some $s$. We will prove $s_{i+1} = s$ by showing that, for every $f \in F$,

(1) if $f \in s_{i+1}$ then $f \in s$
(2) if $f \not\in s_{i+1}$ then $f \not\in s$

Proof of (1). Assume that $f \in s_{i+1}$.

According to Lemma 1, we have $M \models f^{i+1}$. Hence,

$$M \models (f^i \land (\bigwedge_{a \in A_i, f \not\in Del_a} \neg a^i)) \lor (\bigvee_{a \in A_i, f \in Add_a} a^i)$$

There are two possibilities:

a) $M \models (f^i \land (\bigwedge_{a \in A_i, f \not\in Del_a} \neg a^i))$

It is easy to see that $M \models f^i$. Thus, $f \in s_i$.

b) $M \models \bigvee_{a \in A_i, f \in Add_a} a^i$. That is, there must exist an action $b$ such that $M \models b^i$ and $f \in Add_b$. Since $M \models b^i$, we have $b \in A_i$. Hence, $f \in Add_b \subseteq Add_{A_i}$. In addition, $Add_{A_i} \cap Del_{A_i} = \emptyset$ (Lemma 4), we have $f \not\in Del_{A_i}$. As a result $f \in (s_i \cup Add_{A_i}) \backslash Del_{A_i} = s$.

Therefore, in both cases, we have $f \in s$. **
Proof of (2). Assume that \( f \notin s_{i+1} \).

According to Lemma 1, \( f^{i+1} \) is not in \( M \), i.e., \( M \models \neg f^{i+1} \). By (6), we have two possibilities:

a) \( M \models (\neg f^i \land (\bigwedge_{a \in A_i} \neg a^i)) \).

On the other hand, since \( M \models (\neg f^i \land (\bigwedge_{a \in A_i} \neg a^i)) \), we have \( M \models (\bigwedge_{a \in A_i} \neg a^i) \). Thus, \( f \notin \text{Add}_{A_i} \). Consequently, \( f \notin (s_i \cup \text{Add}_{A_i}) \setminus \text{Del}_{A_i} = s_{i+1} \).

b) \( M \models (\bigvee_{a \in A_i} \neg a^i) \).

That is, there must exist an action \( b \) such that \( f \in \text{Del}_{b} \) and \( b^i \) is true in \( M \).

Thus, \( f \notin \text{Del}_{A_i} \) and thus \( f \notin s_{i+1} \) \( \setminus \text{Del}_{A_i} = s \).

So, we can conclude that if \( f \notin s_{i+1} \) then \( f \notin s \).

The lemma is therefore proved.

**Theorem 1** The encoding is sound.

**Proof.** Let \( \sigma_i = \Phi(A_0 \ldots A_{i-1}, \sigma_0) \) for \( 1 \leq i \leq n \). As we mentioned above, \( \sigma_0 = s_0 \). By Lemma 4, using induction, we can easily prove that \( s_i = \sigma_i \) for all \( 0 \leq i \leq n \). Hence, \( \Phi(\pi, \sigma_0) = \sigma_n = s_n \). On the other hand, \( G \) holds \( s_n \). Thus, \( G \) holds in \( \Phi(\pi, \sigma_0) \). That is, \( G \) can be achieved from \( \sigma_0 \) after executing \( \pi \).

The encoding is therefore proved to be sound.

### 2.2 Completeness

Assume that \( P \) has a plan \( \pi = A_0; A_1; \ldots; A_{n-1} \), where \( A_i \), \( 0 \leq i \leq n-1 \), is a set of actions, that achieves goal \( G \) from \( \sigma_0 \). We denote \( \Phi(A_0; A_1; \ldots; A_{j-1}, \sigma_0) \) by \( \sigma_j \). Since \( \pi \) is a plan that achieves goal \( G \), we have the following properties:

1. \( \sigma_0 = I \).
2. For every integer, \( i, 0 \leq i \leq n - 1 \), and for every action \( a \in A_i \), \( a \) is executable in \( \sigma_i \).
3. \( \sigma_i \neq \bot \) and \( \text{Del}_{A_i} \cap \text{Add}_{A_i} = \emptyset \) for every \( i, 0 \leq i \leq n \).
4. \( \sigma_{i+1} = (\sigma_i \cup \text{Add}_{A_i}) \setminus \text{Del}_{A_i} \).
5. For every literal \( l \in G \), \( l \) holds in \( \sigma_n \).

Let \( T \) be the corresponding propositional theory for \( P \) with \( n \) time instants and for every \( i, 0 \leq i \leq n \), let \( \delta_i \) denote the set of propositions \( \{f^i \mid f \in \sigma_i\} \). We construct an interpretation \( M \) of \( T \) as follows:

\[
M = ( \bigcup_{0 \leq i \leq n} \delta_i ) \cup ( \bigcup_{0 \leq i \leq n-1} \{a^i \mid a \in A_i\})
\]

First of all, we will prove the following lemmas and then use them to prove the completeness of the encoding.

**Lemma 5** For every integer \( i, 0 \leq i \leq n \), a fluent literal \( l \) holds in \( \sigma_i \) iff \( M \models l^i \).

**Proof.** It is easy to see that \( M \models l^i \) if \( l^i \) is contained in \( \delta_i \), that is, \( l \) holds in \( \sigma_i \).

**Lemma 6** For every integer \( i, 0 \leq i \leq n - 1 \), we have

\[
M \models a^i \quad \forall a \in A_i
\]

\[
M \models \neg a^i \quad \forall a \notin A_i
\]

**Proof.** It is obvious from the construction of \( M \).

**Theorem 2** The encoding is complete.

**Proof.** To prove that the encoding is complete we need to prove

1. \( M \) is a model of \( T \) and,
2. \( M \) can construct the plan \( \pi \).
Proving (2) is trivial. We will show (1) by checking all the propositional sentences in $T$.

By Lemma 5, the sentence (2) and (3) are apparently true in $M$.

For every integer $i, 0 \leq i \leq n - 1$, and an action $a \in A_i$, since $a$ is executable in $\sigma_i$, every fluent literal $l \in Pre_a$ must hold in $\sigma_i$. By Lemma 5, we have $M \models l$. Thus, the sentence (4) is true in $M$.

From Lemma 6, we have:

$$\bigwedge_{a \in A, f \in Del_a} \neg a^i \iff \bigwedge_{a \in A, f \in Del_a} \neg a^i$$

$$\bigwedge_{a \in A, f \in Add_a} \neg a^i \iff \bigwedge_{a \in A, f \in Add_a} \neg a^i$$

$$\bigvee_{a \in A, f \in Add_a} a^i \iff \bigvee_{a \in A, f \in Add_a} a^i$$

$$\bigvee_{a \in A, f \in Del_a} a^i \iff \bigvee_{a \in A, f \in Del_a} a^i$$

So, to prove the sentences (5) and (6) are true in $M$, we only need to prove the following sentences are true in $M$.

$$f^{i+1} \iff (f^i \land (\bigwedge_{a \in A_i, f \in Del_a} \neg a^i)) \lor (\bigvee_{a \in A_i, f \in Add_a} a^i)$$

(7)

$$\neg f^{i+1} \iff (-f^i \land (\bigwedge_{a \in A_i, f \in Add_a} \neg a^i)) \lor (\bigvee_{a \in A_i, f \in Del_a} a^i)$$

(8)

For each fluent $f \in F$, and an integer $i, 0 \leq i \leq n - 1$, consider the following four cases:

**Case 1.** $f$ holds in both $\sigma_i$ and $\sigma_{i+1}$.

According to Lemma 5, we have $M \models f^i$ and $M \models f^{i+1}$.

Since $\sigma_{i+1} = (\sigma_i \cup Add_{A_i}) \setminus Del_{A_i}$, $f \not\in Del_{A_i}$. Hence, we can reduce (7) and (8) to:

$$f^{i+1} \iff f^i \lor (\bigvee_{a \in A_i, f \in Add_a} a^i)$$

(9)

$$\neg f^{i+1} \iff (-f^i \land (\bigwedge_{a \in A_i, f \in Add_a} \neg a^i))$$

(10)

(9) is true in $M$ as both $f^i$ and $f^{i+1}$ are true in $M$. (9) is also true in $M$ as both the left hand side and the right hand side are false in $M$.

**Case 2.** $f$ holds in $\sigma_i$ but does not hold in $\sigma_{i+1}$.

By Lemma 5, we have $M \models f^i$ and $M \not\models f^{i+1}$.

Since $\sigma_{i+1} = (\sigma_i \cup Add_{A_i}) \setminus Del_{A_i}$, $f \not\in Del_{A_i}$. So, there exists $a \in A_i$ such that $f \in Del_a$. Thus, $(\bigvee_{a \in A_i, f \in Del_a} -a^i)$ is false in $M$ and $(\bigwedge_{a \in A_i, f \in Add_a} a^i)$ are true in $M$. On the other hand, since $Del_{A_i} \cap Add_{A_i} = \emptyset$, we have $f \not\in Add_{A_i}$. As a result, both sides of (7) are false, whereas both sides of (8) are true. That is, (7) and (8) are true in $M$.

**Case 3.** $f$ does not hold in $\sigma_i$ but holds in $\sigma_{i+1}$.

In this case, we have $f^{i+1}$ is true in $M$ but $f^i$ is not. Since, $\sigma_{i+1} = (\sigma_i \cup Add_{A_i}) \setminus Del_{A_i}$, we have $f \in Add_{A_i}$. Thus $f \not\in Del_{A_i}$. Similarly to Case 2, (7) is true in $M$ as both sides are true and (8) is also true because its both sides are false.

**Case 4.** $f$ does not hold in both $\sigma_i$ and $\sigma_{i+1}$.

In this case, we have $M \models \neg f^i$ and $M \models \neg f^{i+1}$. Since $\sigma_{i+1} = (\sigma_i \cup Add_{A_i}) \setminus Del_{A_i}$, we have either $f \in Del_{A_i}$ or $f \not\in Add_{A_i}$. However, $f \in Del_{A_i}$ also implies that $f \not\in Add_{A_i}$. Hence, in both cases, we have $f \not\in Add_{A_i}$. 

4
(7) and (8) are therefore equivalent to

\[ f^{i+1} \leftrightarrow f^i \land (\bigwedge_{a \in A_i, f \in D_{e_{a}}} \neg a^i) \]  \hspace{1cm} (11)

\[ \neg f^{i+1} \leftrightarrow \neg f^i \lor (\bigvee_{a \in A_i, f \in D_{e_{a}}} a^i) \]  \hspace{1cm} (12)

It is easy to see that both of them are true in \( M \).

\( M \) is therefore a model of \( T \). So, the encoding is complete.