Motivation

Consider the goal formula

\[ AonB \land BonC \]

regressed with operator

\[ \langle AonC \land Aclear \land Bclear, AonB \land \neg Bclear \land Cclear \rangle \]

resulting in the new goal

\[ AonC \land Aclear \land Bclear \land BonC. \]

It is intuitively clear that no state satisfying this formula is reachable by any plan from a legal blocks world state.

Goal formulae and formulae obtained by regressing them often represent some states that are not reachable from the initial state.

If none of the states is reachable from the initial state, there are no plans reaching the formula.

We would like to have reachable states only, if possible.

The same problem shows up in satisfiability planning: partial valuations considered by satisfiability algorithms may represent unreachable states, and this may result in unnecessary search.
Invariants

**Goal:** Restriction to states that are reachable.

**Problem:** Testing reachability is computationally as complex as testing whether a plan exists.

**Solution:** Use an approximate notion of reachability.

**Implementation:** Compute in polynomial time formulae that characterize a superset of the reachable states.

Invariants: definition

**Definition**
A formula $\phi$ is an invariant of $\langle A, I, O, G \rangle$ if $s \models \phi$ for every state $s$ reachable from $I$.

**Example**
The formula $\neg(AonB \land AonC)$ is an invariant in a blocks world task.

**Remark**
Invariants are usually proved inductively:
- Prove that $\phi$ is true in the initial state.
- Prove that operator application preserves $\phi$.

Invariants: the strongest invariant

**Definition**
An invariant $\phi$ is the strongest invariant of $\langle A, I, O, G \rangle$ if for any invariant $\psi$, $\phi \models \psi$.

The strongest invariant exactly characterizes the set of all states that are reachable from the initial state:
- For all states $s$, $s \models \phi$ if and only if $s$ is reachable.

**Remark**
There are infinitely many strongest invariants for any given planning task, but they are all logically equivalent. (If $\phi$ is a strongest invariant, then so is $\phi \lor \psi$. . .)

Invariants: the strongest invariant for blocks world

The strongest invariant for the blocks world

Let $X$ be the set of blocks, for example $X = \{A, B, C, D\}$.

The conjunction of the following formulae is the strongest invariant for the set of all states for the blocks $X$.

For all $x \in X$: clear$(x) \leftrightarrow \bigwedge_{y \in X} \neg on(y, x)$

For all $x \in X$: ontable$(x) \leftrightarrow \bigwedge_{y \in X} \neg on(x, y)$

For all $x, y, z \in X$ with $y \neq z$: $\neg on(x, y) \lor \neg on(x, z)$

For all $x, y, z \in X$ with $y \neq z$: $\neg on(y, x) \lor \neg on(z, x)$

For all $n \geq 1$ and $x_1, \ldots, x_n \in X$:

$$\neg (on(x_1, x_2) \land on(x_2, x_3) \land \cdots \land on(x_{n-1}, x_n) \land on(x_n, x_1))$$
Invariants: connection to plan existence

Theorem

Let \( \phi \) be the strongest invariant for \( \langle A, I, O, G \rangle \). Then \( \langle A, I, O, G \rangle \) has a plan if and only if \( G \land \phi \) is satisfiable.

Proof.

Very easy!

Theorem

Computing the strongest invariant \( \phi \) is PSPACE-hard. Even deciding whether or not \( \top \) is the strongest invariant is already PSPACE-hard.

Invariants: connection to plan existence

Proof continues . . .

(\( \Rightarrow \)) If a plan exists for \( T \), then the same plan is applicable in \( T' \). We can thus reach a state satisfying \( G \) in \( T' \).

From this state, we can reach any state \( s \) by first applying \( \langle G, a' \land \bigwedge_{a \in A} a \rangle \) and then applying the operators \( \langle a', \neg a \rangle \) for each variable \( a \) with \( s(a) = 0 \). (If \( s(a') = 0 \), the corresponding operator must be applied last.)

If all states are reachable in \( T' \), then \( \top \) is the strongest invariant for \( T' \).

(\( \Leftarrow \)) (by contraposition): If \( T \) is not solvable, then no state satisfying \( G \) is reachable in \( T \). In that case, no state satisfying \( G \) is reachable in \( T' \), and thus \( a' \) cannot be made true in \( T' \). Thus, \( \neg a' \) is an invariant in \( T' \) which is stronger than \( \top \), so \( \top \) is not the strongest invariant in \( T' \).
Computation of invariants: informally

1. Start with all 1-literal clauses that are true in the initial state.
2. Repeatedly test every operator vs. every clause to check whether the clause can be shown to be true after applying the operator:
   2.1 One of the literals in the clause is necessarily true: **retain**.
   2.2 Otherwise, if the clause is too long: **forget it**.
   2.3 Otherwise, replace the clause by **new clauses** obtained by adding literals that are now true.
3. When all clauses are retained, stop: they are invariants.

Example

Let $C_0 = \{ A\text{clear}, \neg B\text{clear}, A\text{on}B, \neg A\text{on}T, B\text{on}T \}$ and $o = \langle A\text{clear} \land A\text{on}B, B\text{clear} \land \neg A\text{on}B \land A\text{on}T \rangle$.

1. $C_0 \cup \{ A\text{clear} \land A\text{on}B \}$ is satisfiable: $o$ is applicable.
2. The 1-literal clauses $\neg B\text{clear}, A\text{on}B$ and $\neg A\text{on}T$ become false when $o$ is applied.
3. They are not thrown away, though: they are replaced by **weaker** clauses.
4. Literals true after applying $o$ in state $s$ such that $s \models C_0$: $A\text{clear}, B\text{clear}, \neg A\text{on}B, \neg B\text{on}A, A\text{on}T, B\text{on}T$.
5. 2-literal clauses that are **weaker** than $\neg B\text{clear}$ and **now true** are $\neg B\text{clear} \lor A\text{clear}, \neg B\text{clear} \lor B\text{clear}, \neg B\text{clear} \lor \neg A\text{on}B, \neg B\text{clear} \lor \neg B\text{on}A, B\text{clear} \lor A\text{on}T$, and $\neg B\text{clear} \lor B\text{on}T$.

Example (continues...)  

6. Similar 2-literal clauses are obtained from $A\text{on}B$ and from $\neg A\text{on}T$.
7. By eliminating logically equivalent ones, tautologies, and clauses that follow from those in $C_0$ not falsified we get $C_1 = \{ A\text{clear}, \neg B\text{on}A, B\text{on}T, \neg B\text{clear} \lor \neg A\text{on}B, \neg B\text{clear} \lor A\text{on}T, A\text{on}B \lor B\text{clear}, A\text{on}B \lor A\text{on}T, \neg A\text{on}T \lor B\text{clear}, \neg A\text{on}T \lor \neg A\text{on}B \}$ for distance 1 states.
8. Some clauses in $C_1$ can be refined further by checking other operators whose preconditions are consistent with $C_1$. With a bit more computation, $C_i$ settles to a set containing all invariants for two blocks.

Computation of invariants

Example

Let $C_i = \{ \neg A\text{in}Rome \lor \neg A\text{in}Paris, \neg A\text{in}Rome \lor \neg A\text{in}NYC, \neg A\text{in}Paris \lor \neg A\text{in}NYC \}$, $o = \langle A\text{in}Rome, A\text{in}Paris \land \neg A\text{in}Rome \rangle$.

1. Does $o$ preserve truth of $\neg A\text{in}Paris \lor \neg A\text{in}NYC$?
2. Because $o$ makes $\neg A\text{in}Paris$ false, we must show that $\neg A\text{in}NYC$ is true after applying $o$.
3. But $\neg A\text{in}NYC$ is not even mentioned in $o$!
4. However, since $A\text{in}Rome$ is the precondition of $o$ and $\neg A\text{in}Rome \lor \neg A\text{in}NYC$ was true before applying $o$, we can infer that $\neg A\text{in}NYC$ was true before applying $o$.
5. Since $o$ does not make $\neg A\text{in}NYC$ false, it is true also after applying $o$, and then so is $\neg A\text{in}Paris \lor \neg A\text{in}NYC$. 

Example

Let $C_i = \{ \neg A\text{in}Rome \lor \neg A\text{in}Paris, \neg A\text{in}Rome \lor \neg A\text{in}NYC, \neg A\text{in}Paris \lor \neg A\text{in}NYC \}$, $o = \langle A\text{in}Rome, A\text{in}Paris \land \neg A\text{in}Rome \rangle$.

1. Does $o$ preserve truth of $\neg A\text{in}Paris \lor \neg A\text{in}NYC$?
2. Because $o$ makes $\neg A\text{in}Paris$ false, we must show that $\neg A\text{in}NYC$ is true after applying $o$.
3. But $\neg A\text{in}NYC$ is not even mentioned in $o$!
4. However, since $A\text{in}Rome$ is the precondition of $o$ and $\neg A\text{in}Rome \lor \neg A\text{in}NYC$ was true before applying $o$, we can infer that $\neg A\text{in}NYC$ was true before applying $o$.
5. Since $o$ does not make $\neg A\text{in}NYC$ false, it is true also after applying $o$, and then so is $\neg A\text{in}Paris \lor \neg A\text{in}NYC$. 

M. Helmert, B. Nebel (Universität Freiburg) — AI Planning — December 1st, 2006
Computation of invariants: function preserved

Test whether a clause remains true when operator is applied

Test if an operator preserves a clause

```python
def preserved(l_1 ∨ ... ∨ l_n, C, o):
    for each l ∈ {l_1, ..., l_n}:
        if not preserved-literal(C, o, {l_1, ..., l_n} \ {l}, l):
            return False
    return True
```

Test if an operator preserves a literal

```python
def preserved-literal(C, o, L', l):
    ⟨c, e⟩ := o
    C_Γ := C ∪ {c} ∪ {EPC_Γ(e)}
    return C_Γ is unsatisfiable
    or C_Γ |= EPC_Γ(e) for some l' ∈ L'
    or C_Γ |= l' ∧ ¬EPC_Γ(e) for some l'' ∈ L'
```

Correctness

Lemma

Let \( C \) be a set of clauses, \( \phi = l_1 \lor ... \lor l_n \) a clause, and \( o \) an operator. If preserved(\( \phi \), \( C \), \( o \)) returns true, then app_o(s) |= \( \phi \) for every state \( s \) such that \( s |= C \cup \{ \phi \} \) and app_o(s) is defined.

Computation of invariants: function preserved

Why is preserved incomplete?

Example (incompleteness)

Let \( o = \langle a, \neg b \land (c \triangleright d) \land (\neg c \triangleright e) \rangle \).

\( \text{preserved}(b \lor d \lor e, \emptyset, o) \) returns false because the preserved-literal check for \( l = b \) fails:

- Operator \( o \) can make \( b \) false.
- It is not guaranteed that \( d \) is true in the resulting state.
- It is not guaranteed that \( e \) is true in the resulting state.

However, \( d \lor e \) is true after applying \( o \), and hence \( b \lor d \lor e \) will be true as well.
Computation of invariants: the main procedure

Outline

1. \( C = \text{the set of 1-literals clauses that are true in the initial state.} \)
2. For each operator \( o \) and clause \( \phi \in C \), test if \( \phi \) remains true when \( o \) is applied.
3. If not, remove \( \phi \), and if the number of literals in \( \phi \) is less than \( n \), add clauses \( \phi \lor l \) for each literal \( l \) which is guaranteed to be true after applying \( o \).
4. Remove all dominated invariants.
5. Repeat from step 2 if \( C \) has changed in the previous two steps.
6. Otherwise every clause in \( C \) is an invariant.

For any fixed limit \( n \) on the size of the clauses, the number of iterations is \( O(n^m) \) (where \( m = |A| \) is the number of state variables) and hence polynomial.

Correctness

Theorem

The procedure invariants\((A, I, O, n)\) returns a set \( C \) of clauses with at most \( n \) literals such that for any applicable operator sequence \( o_1, \ldots, o_m \in O \): \( \text{app}_{o_1;\ldots;o_m}(I) \models C \).

Proof.

A \( I \models C \):

- The initial state satisfies the initial set of 1-literal clauses.
- All modifications to the clause set only make it logically weaker (i.e., \( C' \models C \) after each iteration of the main loop.)
- Thus the initial state satisfies the resulting clause set \( C \) by induction over the number of iterations.

\( \blacksquare \)
Why is the strongest invariant not always found?

1. Practical implementations of the algorithm use polynomial time approximations of the tests for satisfiability and $|=\$.
2. The function preserved is incomplete for operators in general (but complete for STRIPS operators.) Making it complete makes it NP-hard.
3. The strongest invariant may require arbitrarily long clauses, so the restriction to clauses of any fixed length makes it impossible to represent it.

Example
The acyclicity of the on relation in the blocks world needs clauses of length $n$ when there are $n$ blocks.

Computation of invariants

Example

Initial state: $I |= a \land \neg b \land \neg c$

Operators: $o_1 = \langle a, \neg a \land b \rangle$, $o_2 = \langle b, \neg b \land c \rangle$, $o_3 = \langle c, \neg c \land a \rangle$

Computation: Find invariants with at most 2 literals:

$C_0 = \{ a, \neg b, \neg c \}$
$C_1 = \{ \neg c, a \lor b, \neg b \lor \neg a \}$
$C_2 = \{ \neg b \lor \neg a, \neg c \lor \neg a, \neg c \lor \neg b \}$
$C_3 = \{ \neg b \lor \neg a, \neg c \lor \neg a, \neg c \lor \neg b \}$
$C_j = C_2$ for all $j \geq 2$

Invariants in satisfiability planning

Invariants in satisfiability planning

For every invariant $I_1 \lor \cdots \lor I_n$, add the clauses

$I_1^t \lor \cdots \lor I_n^t$

for all time points $t$.

Notice that the above formulae are logical consequences of $\Phi_{\text{seq}}^i$ and $\Phi_{\text{par}}^i$, so the invariants do not change the set of valuations of these formulae.

Invariants are critical for the efficiency of satisfiability planning on many types of problems.
Invariants in backward search

Motivating example

Problem: Regression produces sets $T$ of states such that
1. some states in $T$ are not reachable from $I$, or
2. none of the states in $T$ are reachable from $I$.
The first is not always a serious problem (but may worsen the quality of distance estimates, for example.)

Solution: Use invariants to avoid formulae that do not represent any reachable states.
1. Compute invariant $\phi$.
2. Do only regression steps such that $\text{regr}_\phi(\psi) \land \phi$ is satisfiable.

Invariants for problem reformulation

Mutexes

Binary clause invariants are called mutexes because they state that certain variable assignments cannot be simultaneously true and are hence mutually exclusive.

Example
The invariant $\neg A_{onB} \lor \neg A_{onC}$ states that $A_{onB}$ and $A_{onC}$ are mutex. Often, a larger set of literals is mutually exclusive because every pair of them forms a mutex.

Example
In blocks world, $B_{onA}$, $C_{onA}$, $D_{onA}$ and $A_{clear}$ are mutex.

Summary

Invariants are needed for making backward search and satisfiability planning more efficient and (in the case of mutexes) can be used for problem reformulation.

We gave an algorithm for computing a class of invariants.
1. Start with 1-literal clauses true in the initial state.
2. Repeatedly weaken clauses that could not be shown to be invariants.
3. Stop when all clauses are guaranteed to be invariants.

The algorithm runs in polynomial time if the satisfiability and logical consequence tests are approximated by a polynomial time algorithm and the size of the invariant clauses is bounded by a constant.