Traces Properties
Semantics and applications to verification

Xavier Rival

École Normale Supérieure

February 22, 2019
Program of this lecture

Goal of verification

Prove that $\lbrack P \rbrack \subseteq S$

(i.e., all behaviors of $P$ satisfy specification $S$)

where $\lbrack P \rbrack$ is the program semantics and $S$ the desired specification

Last week, we studied a form of $\lbrack P \rbrack$...

Today’s lecture: we look back at program’s properties

- families of properties:
  what properties can be considered “similar” ? in what sense ?

- proof techniques:
  how can those kinds of properties be established ?

- specification of properties:
  are there languages to describe properties ?
A high level overview

- In this lecture we look at trace properties
- A property is a set of traces, defining the admissible executions

Safety properties:
- something (e.g., bad) will never happen
- proof by invariance

Liveness properties:
- something (e.g., good) will eventually happen
- proof by variance

Beyond safety and liveness: hyperproperties (e.g., security...)

Xavier Rival
Traces Properties
February 22, 2019
State properties

As usual, we consider $S = (S, \rightarrow, S_I)$

First approach: properties as sets of states

- A property $P$ is a set of states $P \subseteq S$
- $P$ is satisfied if and only if all reachable states belong to $P$, i.e., $[S]_R \subseteq P$ where $[S]_R = \{ s_n \in S \mid \exists \langle s_0, \ldots, s_n \rangle \in [S]^*, s_0 \in S_I \}$

Examples:

- Absence of runtime errors:
  
  $P = S \setminus \{ \Omega \}$ where $\Omega$ is the error state

- Non termination (e.g., for an operating system):

  $P = \{ s \in S \mid \exists s' \in S, s \rightarrow s' \}$
Second approach: properties as sets of traces

- A property $\mathcal{T}$ is a set of traces $\mathcal{T} \subseteq S^\infty$
- $\mathcal{T}$ is satisfied if and only if all traces belong to $\mathcal{T}$, i.e., $[S]^\infty \subseteq \mathcal{T}$

Examples:

- Obviously, state properties are trace properties
- Functional properties:
  e.g., “program $P$ takes one integer input $x$ and returns its absolute value”
- Termination: $\mathcal{T} = S^*$ (i.e., the system should have no infinite execution)
Monotonicity

**Property 1**

Let \( P_0, P_1 \subseteq S \) be two state properties, such that \( P_0 \subseteq P_1 \).

Then \( P_0 \) is stronger than \( P_1 \), i.e. if program \( S \) satisfies \( P_0 \), then it also satisfies \( P_1 \).

**Property 2**

Let \( T_0, T_1 \subseteq S \) be two trace properties, such that \( T_0 \subseteq T_1 \).

Then \( T_0 \) is stronger than \( T_1 \), i.e. if program \( S \) satisfies \( T_0 \), then it also satisfies \( T_1 \).

**Property 3**

Let \( S_0, S_1 \) two transition systems, such that \( S_1 \) has more behaviors than \( S_0 \) (i.e., \([ S_0 ] \subseteq [ S_1 ]\)), and \( P \) be a (trace or state) property. Then, if \( S_1 \) satisfies \( P \), so does \( S_0 \).

**Proofs:** straightforward application of the definitions
Outline
Safety properties

Informal definition: safety properties
A safety property is a property which specifies that some (bad) behavior will never occur, at any time

- **Absence of runtime errors** is a safety property ("bad thing": error)
- **State properties** is a safety property ("bad thing": reaching \( S \setminus P \))
- **Non termination** is a safety property ("bad thing": reaching a blocking state)
- **“Not reaching state \( b \) after visiting state \( a \)”** is a safety property (and **not** a state property)
- **Termination** is **not** a safety property

We now intend to provide a **formal definition** of safety.
Towards a formal definition

**How to refute a safety property?**

- We assume $S$ does not satisfy safety property $\mathcal{P}$
- Thus, there exists a **counter-example trace**
  \[ \sigma = \langle s_0, \ldots, s_n, \ldots \rangle \in \mathbb{S} \setminus \mathcal{P}; \]
  it may be finite or infinite...
- The intuitive definition says this trace **eventually exhibits some bad behavior**, at some given time, corresponding to some index $i$
- Therefore, trace $\sigma' = \langle s_0, \ldots, s_i \rangle$ violates $\mathcal{P}$, i.e. $\sigma' \notin \mathcal{P}$
- We remark $\sigma'$ is finite

**A safety property that does not hold can always be refuted with a finite counter-example**
A Few Operators on Traces

Prefix: We write $\sigma[i]$ for the prefix of length $i$ of trace $\sigma$:

$$
\langle s_0, \ldots, s_n \rangle[i]_0 = \epsilon
$$

$$
\langle s_0, \ldots, s_n \rangle[i+1] = \begin{cases} 
\langle s_0, \ldots, s_i \rangle & \text{if } i < n \\
\langle s_0, \ldots, s_n \rangle & \text{otherwise}
\end{cases}
$$

$$
\langle s_0, \ldots \rangle[i+1] = \langle s_0, \ldots, s_i \rangle
$$

Suffix (or tail):

$$
\sigma[i] = \epsilon \text{ if } |\sigma| < i
$$

$$
(\langle s_0, \ldots, s_i \rangle \cdot \sigma)[i+1] =: \sigma \text{ otherwise}
$$
Upper closure operators

Definition: upper closure operator (uco)

We consider a preorder \((S, \sqsubseteq)\). Function \(\phi : S \rightarrow S\) is an upper closure operator iff:

- **monotone**
- **extensive:** \(\forall x \in S, \ x \sqsubseteq \phi(x)\)
- **idempotent:** \(\forall x \in S, \ \phi(\phi(x)) = \phi(x)\)

Dual: lower closure operator, monotone, reductive, idempotent

Examples:

- on real/decimal numbers, or on fraction: the **ceiling** operator, that returns the next integer is an upper-closure operator
Prefix closure

Definition: prefix closure

The prefix closure operator is defined by:

$$PCI : \mathcal{P}(S^\infty) \longrightarrow \mathcal{P}(S^*)$$

$$X \longmapsto \{\sigma_i \mid \sigma \in X, i \in \mathbb{N}\}$$

Example: assuming

$$S = \{\langle a, b, c \rangle, \langle a, c \rangle\}$$

then,

$$PCI(S) = \{\epsilon, \langle a \rangle, \langle a, b \rangle, \langle a, b, c \rangle, \langle a, c \rangle\}$$

Properties:

- PCI is monotone
- PCI is idempotent, i.e., $$PCI \circ PCI(X) = PCI(X)$$
- PCI is not extensive (infinite traces do not appear anymore)
Limit

Definition: limit

The **limit operator** is defined by:

\[
\text{Lim} : \mathcal{P}(S^\infty) \longrightarrow \mathcal{P}(S^\infty)
\]

\[X \longmapsto X \cup \{\sigma \in S^\infty | \forall i \in \mathbb{N}, \sigma[i] \in X\}\]

Operator **Lim** is an upper-closure operator

**Proof**: exercise!

**Example**: assuming

\[S = \{\epsilon, \langle a\rangle, \langle a, b\rangle, \langle a, b, a\rangle, \langle a, b, a, b\rangle, \ldots\}\]

then,

\[\text{Lim}(S) = S \cup \{\langle a, b, a, b, a, b, \ldots\}\}\]
Towards a formal definition for safety

**Operator Safe**

Operator *Safe* is defined by \( \text{Safe} = \text{Lim} \circ \text{PCI} \).

Operator **Safe saturates** a set of traces \( S \) with

- prefixes
- infinite traces all finite prefixes of which can be observed in \( S \)

Thus, if \( \text{Safe}(S) = S \) and \( \sigma \) is a trace, to establish that \( \sigma \) is not in \( S \), it is sufficient to discover a **finite prefix of** \( \sigma \) that cannot be observed in \( S \).

- if \( \sigma \) is finite the result is clear (consider \( \sigma \))
- otherwise, if all finite prefixes of \( \sigma \) are in \( S \), then \( \sigma \) is in the limit, thus in \( S \).

**Safety: definition**

A trace property \( \mathcal{T} \) is a **safety** property if and only if \( \text{Safe}(\mathcal{T}) = \mathcal{T} \)
Safety properties: formal definition

An upper closure operator

Operator **Safe** is an upper closure operator over $\mathcal{P}(S^\infty)$

Proof:

**Safe is monotone** since $\text{Lim}$ and $\text{PCI}$ are monotone

**Safe is extensive:**
indeed if $X \subseteq S^\infty$ and $\sigma \in X$, we can show that $\sigma \in \text{Safe}(X)$:

- if $\sigma$ is a finite trace, it is one of its prefixes, so $\sigma \in \text{PCI}(X) \subseteq \text{Lim}(\text{PCI}(X))$
- if $\sigma$ is an infinite trace, all its prefixes belong to $\text{PCI}(X)$, so $\sigma \in \text{Lim}(\text{PCI}(X))$
Safety properties: formal definition

Proof (continued):

Safe is idempotent:

- as Safe is extensive and monotone Safe ⊆ Safe ◦ Safe, so we simply need to show that Safe ◦ Safe ⊆ Safe
- let X ⊆ S^∞, σ ∈ Safe(Safe(X)); then:

\[ \sigma \in \text{Safe}(\text{Safe}(X)) \]
\[ \Rightarrow \forall i, \sigma[i] \in \text{PCI} \circ \text{Safe}(X) \quad \text{by def. of Lim} \]
\[ \Rightarrow \forall i, \exists \sigma', j, \sigma[i] = \sigma'[j] \land \sigma' \in \text{Safe}(X) \quad \text{by def. of PCI} \]
\[ \Rightarrow \forall i, \exists \sigma', j, \sigma[i] = \sigma'[j] \land \forall k, \sigma'[k] \in \text{PCI}(X) \]
\[ \quad \text{by def. of Lim and case analysis over finiteness of } \sigma' \]
\[ \Rightarrow \forall i, \exists \sigma', j, \sigma[i] = \sigma'[j] \land \sigma'[j] \in \text{PCI}(X) \quad \text{if we take } k = j \]
\[ \Rightarrow \forall i, \sigma[i] \in \text{PCI}(X) \quad \text{by simplification} \]
\[ \Rightarrow \sigma \in \text{Lim} \circ \text{PCI}(X) \quad \text{by def. of Lim} \]
\[ \Rightarrow \sigma \in \text{Safe}(X) \]
Safety properties: formal definition

Safety: definition

A trace property $\mathcal{T}$ is a safety property if and only if $\text{Safe}(\mathcal{T}) = \mathcal{T}$

Theorem

If $\mathcal{T}$ is a trace property, then $\text{Safe}(\mathcal{T})$ is a safety property

Proof:

Straightforward, by idempotence of $\text{Safe}$

Intuition:

- if $\mathcal{T}$ is a trace property (not necessarily a safety property), $\text{Safe}(\mathcal{T})$ is the strongest safety property, that is weaker than $\mathcal{T}$
- at this point, this observation is not so useful... but it will be soon!
Example

We assume that:

- $\mathcal{S} = \{a, b\}$
- $\mathcal{T}$ states that *a should not be visited after state b is visited*; elements of $\mathcal{T}$ are of the general form $\langle a, a, a, \ldots, a, b, b, b, b, \ldots \rangle$ or $\langle a, a, a, \ldots, a, a, \ldots \rangle$

Then:

- $\text{PCI}(\mathcal{T})$ elements are all finite traces which are of the above form (i.e., made of $n$ occurrences of $a$ followed by $m$ occurrences of $b$, where $n, m$ are positive integers)
- $\text{Lim}(\text{PCI}(\mathcal{T}))$ adds to this set the trace made made of infinitely many occurrences of $a$ and the infinite traces made of $n$ occurrences of $a$ followed by infinitely many occurrences of $b$
- thus, $\text{Safe}(\mathcal{T}) = \text{Lim}(\text{PCI}(\mathcal{T})) = \mathcal{T}$

Therefore $\mathcal{T}$ is indeed formally a safety property.
Theorem

Any **state property** is also a **safety property**.

Proof:

Let us consider **state property** $\mathcal{P}$.

It is equivalent to **trace property** $\mathcal{T} = \mathcal{P}^\infty$:

\[
\begin{align*}
\text{Safe}(\mathcal{T}) &= \text{Lim}(\text{PCI}(\mathcal{P}^\infty)) \\
&= \text{Lim}(\mathcal{P}^*) \\
&= \mathcal{P}^* \cup \mathcal{P}^\omega \\
&= \mathcal{P}^\infty \\
&= \mathcal{T}
\end{align*}
\]

Therefore $\mathcal{T}$ is indeed a safety property.
Intuition of the formal definition

Operator **Safe saturates** a set of traces $S$ with

- prefixes
- infinite traces all finite prefixes of which can be observed in $S$

Thus, if $\text{Safe}(S) = S$ and $\sigma$ is a trace, to establish that $\sigma$ is not in $S$, it is sufficient to discover a **finite prefix of $\sigma$** that cannot be observed in $S$.

Alternatively, if all finite prefixes of $\sigma$ belong to $S$ or can observed as a prefix of another trace in $S$, by definition of the limit operator, $\sigma$ **belongs to $S$** (even if it is infinite).

Thus, our definition **indeed captures properties that can be disproved with a finite counter-example**.
Proof by invariance

- We consider transition system $S = (S, \rightarrow, S_I)$, and safety property $T$. Finite traces semantics is the least fixpoint of $F_*$.
- We seek a way of verifying that $S$ satisfies $T$, i.e., that $[S]^{\infty} \subseteq T$.

Principle of invariance proofs

Let $\Pi$ be a set of finite traces; it is said to be an invariant if and only if:

- $\forall s \in S_I, \langle s \rangle \in \Pi$
- $F_*(\Pi) \subseteq \Pi$

It is stronger than $T$ if and only if $\Pi \subseteq T$.

The “by invariance” proof method is based on finding an invariant that is stronger than $T$. 
Safety properties

Soundness

Theorem: soundness

The invariance proof method is sound: if we can find an invariant for $S$, that is stronger than safety property $T$, then $S$ satisfies $T$.

Proof:
We assume that $I$ is an invariant of $S$ and that it is stronger than $T$, and we show that $S$ satisfies $T$:

- by induction over $n$, we can prove that $F^n_* (\{ \langle s \rangle \mid s \in S_I \}) \subseteq F^n_*(I) \subseteq I$
- therefore $[S]^* \subseteq I$
- thus, $\text{Safe}([S]^*) \subseteq \text{Safe}(I) \subseteq \text{Safe}(T)$ since $\text{Safe}$ is monotone
- we remark that $[S]^{\infty} = \text{Safe}([S]^*)$
- $T$ is a safety property so $\text{Safe}(T) = T$
- we conclude $[S]^{\infty} \subseteq T$, i.e., $S$ satisfies property $T$
Completeness

Theorem: completeness
The invariance proof method is complete: if $S$ satisfies safety property $\mathcal{T}$, then we can find an invariant $I$ for $S$, that is stronger than $\mathcal{T}$.

Proof:
We assume that $[S]^*$ satisfies $\mathcal{T}$, and show that we can exhibit an invariant.
Then, $I = [S]^*$ is an invariant of $S$ by definition of $[.]^*$, and it is stronger than $\mathcal{T}$.

Caveat:
- $[S]^\infty$ is most likely not a very easy to express invariant
- it is just a convenient completeness argument
- so, completeness does not mean the proof is easy!
Example

We consider the proof that the program below computes the sum of the elements of an array, i.e., when the exit is reached, \( s = \sum_{k=0}^{n-1} t[k] \):

\[
\begin{align*}
\text{i, s integer variables} \\
\text{t integer array of length n}
\end{align*}
\]

\[
\begin{align*}
\ell_0 : \text{(true)} \\
& \quad s = 0; \\
\ell_1 : \text{(s = 0)} \\
& \quad i = 0; \\
\ell_2 : \text{(i = 0 \& s = 0)} \\
& \quad \text{while}(i < n) \\
\ell_3 : \text{(0 \leq i < n \& s = \sum_{k=0}^{i-1} t[k])} \\
& \quad s = s + t[i]; \\
\ell_4 : \text{(0 \leq i < n \& s = \sum_{k=0}^{i} t[k])} \\
& \quad i = i + 1; \\
\ell_5 : \text{(1 \leq i \leq n \& s = \sum_{k=0}^{i-1} t[k])} \\
& \quad \} \\
\ell_6 : \text{(i = n \& s = \sum_{k=0}^{n-1} t[k])}
\end{align*}
\]

Principle of the proof:

- for each program point \( \ell \), we have a local invariant \( I_\ell \) (denoted by a logical formula instead of a set of states in the figure)
- the global invariant \( I \) is defined by:
  \[
  I = \{ \langle \ell_0, m_0 \rangle, \ldots, \langle \ell_n, m_n \rangle \mid \forall n, m_n \in I_{\ell_n} \}\]
Informal definition: liveness properties

A liveness property is a property which specifies that some (good) behavior will eventually occur.

- **Termination** is a liveness property
  “good behavior”: reaching a blocking state (no more transition available)
- “State a will eventually be reached by all execution” is a liveness property
  “good behavior”: reaching state a
- The absence of runtime errors is *not* a liveness property

As for safety properties, we intend to provide a **formal definition** of liveness.
Intuition towards a formal definition

How to refute a liveness property?

- We consider liveness property $\mathcal{T}$ (think $\mathcal{T}$ is termination)
- We assume $S$ does not satisfy liveness property $\mathcal{T}$
- Thus, there exists a counter-example trace $\sigma \in [S] \setminus \mathcal{T}$;
- Let us assume $\sigma$ is actually finite...
  the definition of liveness says some (good) behavior should eventually occur:
  - how do we know that $\sigma$ cannot be extended into a trace $\sigma \cdot \sigma'$ that will satisfy this behavior?
  - maybe that after a few more computation steps, $\sigma$ will reach a blocking state...
Intuition towards a formal definition

To refute a liveness property, we need to look at infinite traces.

**Example:** if we run a program, and do not see it return...
- should we do Ctrl+C and conclude it does not terminate?
- should we just wait a few more seconds minutes, hours, years?

Towards a formal definition:
we expect any finite trace be the prefix of a trace in $\mathcal{T}$

... since finite executions cannot be used to disprove $\mathcal{T}$

**Formal definition (incomplete)**

$$PCI(\mathcal{T}) = S^*$$
Definition

Formal definition

Operator \textbf{Live} is defined by \( \text{Live}(\mathcal{T}) = \mathcal{T} \cup (\mathcal{S}^\infty \setminus \text{Safe}(\mathcal{T})) \). Given property \( \mathcal{T} \), the following three statements are equivalent:

\begin{enumerate}[(i)]
    \item \( \text{Live}(\mathcal{T}) = \mathcal{T} \)
    \item \( \text{PCI}(\mathcal{T}) = \mathcal{S}^* \)
    \item \( \text{Lim} \circ \text{PCI}(\mathcal{T}) = \mathcal{S}^\infty \)
\end{enumerate}

When they are satisfied, \( \mathcal{T} \) is said to be a \textit{liveness property}

Example: termination

- The property is \( \mathcal{T} = \mathcal{S}^* \)
  (i.e., there should be no infinite execution)
- Clearly, it satisfies (ii): \( \text{PCI}(\mathcal{T}) = \mathcal{S}^* \)
  thus termination indeed satisfies this definition
Proof of equivalence

Proof of equivalence:

(i) implies (ii):

We assume that \( \text{Live}(\mathcal{T}) = \mathcal{T} \), i.e., \( \mathcal{T} \cup (\mathbb{S}^\infty \setminus \text{Safe}(\mathcal{T})) = \mathcal{T} \)
therefore, \( \mathbb{S}^\infty \setminus \text{Safe}(\mathcal{T}) \subseteq \mathcal{T} \).

Let \( \sigma \in \mathbb{S}^* \), and let us show that \( \sigma \in \text{PCI}(\mathcal{T}) \); clearly, \( \sigma \in \mathbb{S}^\infty \), thus:

- either \( \sigma \in \text{Safe}(\mathcal{T}) = \text{Lim}(\text{PCI}(\mathcal{T})) \), so all its prefixes are in \( \text{PCI}(\mathcal{T}) \) and \( \sigma \in \text{PCI}(\mathcal{T}) \)
- or \( \sigma \in \mathcal{T} \), which implies that \( \sigma \in \text{PCI}(\mathcal{T}) \)

(ii) implies (iii):

If \( \text{PCI}(\mathcal{T}) = \mathbb{S}^* \), then \( \text{Lim} \circ \text{PCI}(\mathcal{T}) = \mathbb{S}^\infty \)

(iii) implies (i):

If \( \text{Lim} \circ \text{PCI}(\mathcal{T}) = \mathbb{S}^\infty \), then
\[
\text{Live}(\mathcal{T}) = \mathcal{T} \cup (\mathbb{S}^\infty \setminus (\text{Lim} \circ \text{PCI}(\mathcal{T}))) = \mathcal{T} \cup (\mathbb{S}^\infty \setminus \mathbb{S}^\infty) = \mathcal{T}
\]
Example

We assume that:

- $S = \{a, b, c\}$
- $T$ states that $b$ should eventually be visited, after $a$ has been visited; elements of $T$ can be described by

$$T = S^* \cdot a \cdot S^* \cdot b \cdot S^\infty$$

Then $T$ is a liveness property:

- let $\sigma \in S^*$; then $\sigma \cdot a \cdot b \in T$, so $\sigma \in PCI(T)$
- thus, $PCI(T) = S^*$
A property of **Live**

**Theorem**

If \(\mathcal{T}\) is a trace property, then \(\text{Live}(\mathcal{T})\) is a liveness property (i.e., operator \(\text{Live}\) is idempotent).

**Proof:** we show that \(\text{PCI} \circ \text{Live}(\mathcal{T}) = \mathbb{S}^*\), by considering \(\sigma \in \mathbb{S}^*\) and proving that \(\sigma \in \text{PCI} \circ \text{Live}(\mathcal{T})\); we first note that:

\[
\text{PCI} \circ \text{Live}(\mathcal{T}) = \text{PCI}(\mathcal{T}) \cup \text{PCI}(\mathbb{S}^\infty \setminus \text{Safe}(\mathcal{T})) \\
= \text{PCI}(\mathcal{T}) \cup \text{PCI}(\mathbb{S}^\infty \setminus \text{Lim} \circ \text{PCI}(\mathcal{T}))
\]

- if \(\sigma \in \text{PCI}(\mathcal{T})\), this is obvious.
- if \(\sigma \notin \text{PCI}(\mathcal{T})\), then:
  - \(\sigma \notin \text{Lim} \circ \text{PCI}(\mathcal{T})\) by definition of the limit
  - thus, \(\sigma \in \mathbb{S}^\infty \setminus \text{Lim} \circ \text{PCI}(\mathcal{T})\)
  - \(\sigma \in \text{PCI}(\mathbb{S}^\infty \setminus \text{Lim} \circ \text{PCI}(\mathcal{T}))\) as \(\text{PCI}\) is extensive when applied to sets of finite traces, which proves the above result
Outline
Termination proof with ranking function

- We consider only termination
- We consider transition system $S = (S, \rightarrow, S_I)$, and liveness property $T$
- We seek a way of verifying that $S$ satisfies termination, i.e., that $[S]^\infty \subseteq S^*$

**Definition: ranking function**

A **ranking function** is a function $\phi : S \rightarrow E$ where:

- $(E, \sqsubseteq)$ is a well-founded ordering
- $\forall s_0, s_1 \in S, \ s_0 \rightarrow s_1 \implies \phi(s_1) \sqsubseteq \phi(s_0)$

**Theorem**

If $S$ has a ranking function $\phi$, it satisfies termination.
We consider the termination of the array sum program:

\[ i, s \text{ integer variables} \]
\[ t \text{ integer array of length } n \]
\[ l_0: \quad s = 0; \]
\[ l_1: \quad i = 0; \]
\[ l_2: \quad \textbf{while}(i < n)\{ \]
\[ l_3: \quad s = s + t[i]; \]
\[ l_4: \quad i = i + 1; \]
\[ l_5: \quad \} \]
\[ l_6: \quad \ldots \]

**Ranking function:**

\[ \phi: \quad S \longrightarrow \mathbb{N} \]
\[ (l_0, m) \quad \mapsto \quad 3 \cdot n + 6 \]
\[ (l_1, m) \quad \mapsto \quad 3 \cdot n + 5 \]
\[ (l_2, m) \quad \mapsto \quad 3 \cdot n + 4 \]
\[ (l_3, m) \quad \mapsto \quad 3 \cdot (n - m(i)) + 3 \]
\[ (l_4, m) \quad \mapsto \quad 3 \cdot (n - m(i)) + 2 \]
\[ (l_5, m) \quad \mapsto \quad 3 \cdot (n - m(i)) + 4 \]
\[ (l_6, m) \quad \mapsto \quad 0 \]
Proof by variance

- We consider transition system $S = (S, \rightarrow, S_I)$, and liveness property $T$; infinite traces semantics is the greatest fixpoint of $F_\omega$.
- We seek a way of verifying that $S$ satisfies $T$, i.e., that $[S]^\infty \subseteq T$

Principle of variance proofs

Let $(I_n)_{n \in \mathbb{N}}, I_\omega$ be elements of $S^\infty$; these are said to form a variance proof of $T$ if and only if:

- $S^\omega \subseteq I_0$
- For all $k \in \{1, 2, \ldots, \omega\}$, $\forall s \in S, \langle s \rangle \in I_k$
- For all $k \in \{1, 2, \ldots, \omega\}$, there exists $l < k$ such that $F_\omega(I_l) \subseteq I_k$
- $I_\omega \subseteq T$

Proofs of soundness and completeness: exercise, similar to the previous proof but using the definition of $[S]^\infty$ instead
Outline
Decomposition of trace properties

The decomposition theorem

**Theorem**

Let $T \subseteq S^\infty$; it can be decomposed into the conjunction of safety property $\text{Safe}(T)$ and liveness property $\text{Live}(T)$:

$$T = \text{Safe}(T) \cap \text{Live}(T)$$


- **Consequence of this result:** the proof of any trace property can be decomposed into
  - a proof of safety
  - a proof of liveness
Proof

- **Safety part:**
  Safe is idempotent, so $\text{Safe}(\mathcal{T})$ is a safety property.

- **Liveness part:**
  Live is idempotent, so $\text{Live}(\mathcal{T})$ is a liveness property.

- **Decomposition:**

  \[
  \text{Safe}(\mathcal{T}) \cap \text{Live}(\mathcal{T}) = \text{Safe}(\mathcal{T}) \cap (\mathcal{T} \cup S^\infty \setminus \text{Safe}(\mathcal{T})) \\
  = \text{Safe}(\mathcal{T}) \cap \mathcal{T} \\
  \cup \text{Safe}(\mathcal{T}) \cap (S^\infty \setminus \text{Safe}(\mathcal{T})) \\
  = \mathcal{T} \cup \emptyset \\
  = \mathcal{T}
  \]
Decomposition of trace properties

Example: verification of total correctness

\[\begin{align*}
& i, s \text{ integer variables} \\
& t \text{ integer array of length } n \\
& \ell_0 : \quad s = 0; \\
& \ell_1 : \quad i = 0; \\
& \ell_2 : \quad \textbf{while}(i < n)\{ \\
& \quad \ell_3 : \quad s = s + t[i]; \\
& \quad \ell_4 : \quad i = i + 1; \\
& \ell_5 : \quad } \\
& \ell_6 : \quad \ldots
\end{align*}\]

**Property to prove:**

- total correctness

1. the program terminates
2. and it computes the sum of the elements in the array

Application of the decomposition principle

**Conjunction of two proofs:**

1. Proved with a ranking function
2. Proved with local invariants
Safety and Liveness Decomposition Example

We consider a very simple greatest common divider code function:

\begin{verbatim}
\textbf{l}_{0}: \textbf{int} f(\textbf{int} a, \textbf{int} b)\{
\textbf{l}_{1}: \ \textbf{while}(a > 0)\{
\textbf{l}_{2}: \ \textbf{int} d = b/a;
\textbf{l}_{3}: \ \textbf{int} r = b - a * d;
\textbf{l}_{4}: \ \textbf{b} = a;
\textbf{l}_{5}: \ \textbf{a} = r;
\textbf{l}_{6}: \ \}\textbf{}}
\textbf{l}_{7}: \ \textbf{return} \ \textbf{b};
\textbf{l}_{8}: \ \}
\end{verbatim}

Specification

When applied to positive integers, function \texttt{f} should always return their GCD.
Safety and Liveness Decomposition Example

We consider a very simple greatest common divider code function:

```c
int f(int a, int b) {
    while (a > 0) {
        int d = b / a;
        int r = b - a * d;
        b = a;
        a = r;
    }
    return b;
}
```

**Specification**

When applied to positive integers, function $f$ should always return their GCD.

**Safety part**

For all trace starting with positive inputs, a conjunction of two properties:

- no runtime errors
- the value of $b$ is the GCD

**Liveness part**

Termination, on all traces starting with positive inputs
The Zoo of semantic properties: current status

Trace properties
  total correctness

Safety properties
  never reach $s_0$ before $s_1$

State properties
  absence or runtime errors
  partial correctness

Liveness properties
  termination

Safety: if wrong, can be refuted with a finite trace proof done by invariance

Liveness: if wrong, has to be refuted with an infinite trace proof done by variance
Ultimately, we would like to **verify or compute** properties.
So far, we simply describe properties with **sets of executions** or worse, with English / French / ... statements.
Ideally, we would prefer to use a **mathematical language** for that:
- to *gain in concision, avoid ambiguity*
- to *define sets of properties to consider, fix the form of inputs for verification tools...*

**Definition: specification language**

A *specification language* is a set of terms \( \mathbb{L} \) with an *interpretation function* (or *semantics*)

\[
[.] : \mathbb{L} \rightarrow \mathcal{P}(S^\infty) \quad \text{(resp., } \mathcal{P}(S))
\]

We are now going to consider specification languages for states, for traces...
A State specification language

A first example of a (simple) specification language:

A state specification language

**Syntax:** we let terms of $L_S$ be defined by:

$$p \in L_S ::= \emptyset \ell \mid x < x' \mid x < n \mid \neg p' \mid p' \land p'' \mid \Omega$$

**Semantics:** $\llbracket p \rrbracket_s \subseteq S_\Omega$ is defined by

- $\llbracket \emptyset \ell \rrbracket_s = \{ \ell \} \times M$
- $\llbracket x \leq x' \rrbracket_s = \{ (\ell, m) \in S \mid m(x) \leq m(x') \}$
- $\llbracket x \leq n \rrbracket_s = \{ (\ell, m) \in S \mid m(x) \leq n \}$
- $\llbracket \neg p \rrbracket_s = S_\Omega \setminus \llbracket p \rrbracket_s$
- $\llbracket p \land p' \rrbracket_s = \llbracket p \rrbracket_s \cap \llbracket p' \rrbracket_s$
- $\llbracket \Omega \rrbracket_s = \{ \Omega \}$

**Exercise:** add $=, \lor, \implies \ldots$
State properties: examples

Unreachability of control state $l_0$:
- specification: $\Omega \lor \neg @l_0$
- property: $[\Omega \lor \neg @l_0]_s = S_\Omega \setminus \{(l_0, m) | m \in M\}$

Absence of runtime errors:
- specification: $\neg \Omega$
- property: $[\neg \Omega]_s = S_\Omega \setminus \{\Omega\} = S$

Intermittent invariant:
- principle: attach a local invariant to each control state
- example:

\[
\begin{align*}
  l_0 & : \textbf{if}(x \geq 0)\{ \\
  l_1 & : \quad y = x; \quad \circ l_1 \implies x \geq 0 \\
  l_2 & : \textbf{else}\{ \quad \circ l_2 \implies x \geq 0 \land y \geq 0 \\
  l_3 & : \quad y = -x; \quad \circ l_3 \implies x < 0 \\
  l_4 & : \textbf{else}\{ \quad \circ l_4 \implies x < 0 \land y > 0 \\
  l_5 & : \quad y \geq 0
\end{align*}
\]
Propositional temporal logic: syntax

We now consider the **specification of trace properties**

- **Temporal logic**: specification of properties in terms of events that occur at distinct times in the execution (hence, the name “temporal”)
- There are **many** instances of temporal logic
- We study a simple one: Pnueli’s Propositional Temporal Logic

**Definition: syntax of PTL (Propositional Temporal Logic)**

Properties over traces are defined as terms of the form

$$ t(\in L_{PTL}) ::= p \quad \text{state property, i.e., } p \in L_S $$

$$ t' \lor t'' \quad \text{disjunction} $$

$$ \neg t' \quad \text{negation} $$

$$ \circ t' \quad "next" $$

$$ t' \mathbb{U} t'' \quad "until", \text{ i.e., } t' \text{ until } t'' $$
Propositional temporal logic: semantics

The semantics of a temporal property is a set of traces, and it is defined by induction over the syntax:

**Semantics of Propositional Temporal Logic formulae**

\[
[p]_t = \{ s \cdot \sigma \mid s \in [p]_s \land \sigma \in S^\infty \}
\]

\[
[t_0 \lor t_1]_t = [t_0]_t \cup [t_1]_t
\]

\[
[\neg t_0]_t = S^\infty \setminus [t_0]_t
\]

\[
[\bigcirc t_0]_t = \{ s \cdot \sigma \mid s \in S \land \sigma \in [t_0]_t \}
\]

\[
[t_0 \mathcal{U} t_1]_t = \{ \sigma \in S^\infty \mid \exists n \in \mathbb{N}, \forall i < n, \sigma_i \in [t_0]_t \land \sigma_n \in [t_1]_t \}
\]
Temporal logic operators as syntactic sugar

Many useful operators can be added:

- **Boolean constants:**

  \[
  \text{true} ::= (x < 0) \lor \neg(x < 0)  \\
  \text{false} ::= \neg\text{true}
  \]

- **Sometime:**

  \[
  \diamond t ::= \text{true} \mathbin{\text{U}} t
  \]

  **intuition:** there exists a rank \( n \) at which \( t \) holds

- **Always:**

  \[
  \Box t ::= \neg(\diamond(\neg t))
  \]

  **intuition:** there is no rank at which the negation of \( t \) holds

**Exercise:** what do \( \diamond \Box t \) and \( \Box \diamond t \) mean?
Propositional temporal logic: examples

We consider the program below:

\[ l_0: \text{x = input();} \]
\[ l_1: \text{if(x < 8){} } \]
\[ l_2: \text{x = 0;} \]
\[ l_3: \text{}} \text{else } \}
\[ l_4: \text{x = 1;} \]
\[ l_5: } \]
\[ l_6: \ldots \]

Examples of properties:

- “when \( l_4 \) is reached, \( x \) is positive”
  \[ \square (\diamond l_4 \implies x \geq 0) \]

- “if the value read at point \( l_0 \) is negative, and when \( l_6 \) is reached, \( x \) is equal to 0”
  \[ \square ((\diamond l_1 \land x < 0) \implies \square (\diamond l_6 \implies x = 0)) \]
We now consider other interesting properties of programs, and show that they do not all reduce to trace properties.

Security

- Collects many kinds of properties
- So we consider just one:
  - An unauthorized observer should not be able to guess anything about private information by looking at public information

Example: another user should not be able to guess the content of an email sent to you

- We need to formalize this property
A few definitions

Assumptions:
- We let $\mathcal{S} = (S, \rightarrow, S_I)$ be a transition system
- States are of the form $(\ell, m) \in L \times M$
- Memory states are of the form $X \rightarrow \mathbb{V}$
- We let $\ell, \ell' \in L$ (program entry and exit)
  and $x, x' \in X$ (private and public variables)

Security property we are looking at

Observing the value of $x'$ at $\ell'$
gives no information on the value of $x$ at $\ell$
A few examples

A secure program (no information flow, no way to guess x):

\[
\ell : \quad x' = 84;
\]

\[
\ell' : \quad \ldots
\]

An insecure program (explicit information flow, \(x'\) gives a lot of information about \(x\), so that we can simply recompute it):

\[
\ell : \quad x' = x - 2;
\]

\[
\ell' : \quad \ldots
\]

An insecure program (implicit information flow, through a test):

\[
\ell : \quad \text{if(} x < 0 \text{)\{}x' = 0; \}\}
\]

\[
\ell' : \quad \ldots
\]

How to characterize information flow in the semantic level?
Non-interference

We consider the transformer $\Phi$ defined by:

$$\Phi : \mathcal{M} \rightarrow \mathcal{P}(\mathcal{M})$$

$$m \mapsto \{ m' \in \mathcal{M} \mid \exists \sigma = \langle (l, m), \ldots, (l', m') \rangle \in [\mathcal{S}] \}$$

**Definition: non-interference**

There is no interference between $(l, x)$ and $(l', x')$ and we write $(l', x') \not\hookrightarrow (l, x)$ if and only if the following property holds:

$$\forall m \in \mathcal{M}, \forall \nu_0, \nu_1 \in \mathcal{V},$$

$$\{ m'(x') \mid m' \in \Phi(m[x \leftarrow \nu_0]) \} = \{ m'(x') \mid m' \in \Phi(m[x \leftarrow \nu_1]) \}$$

**Intuition:**

- if two observations at point $l$ differ only in the value of $x$, there is no difference in observation of $x'$ at $l'$
- in other words, observing $x'$ at $l'$ (even on many executions) gives no information about the value of $x$ at point $l$...
Beyond safety and liveness

Non-interference is not a trace property

- We assume $\mathbb{V} = \{0, 1\}$ and $\mathbb{X} = \{x, x'\}$ (store $m$ is defined by the pair $(m(x), m(x'))$, and denoted by it)
- We assume $\mathbb{L} = \{\ell, \ell'\}$ and consider two systems such that all transitions are of the form $(\ell, m) \rightarrow (\ell', m')$ (i.e., system $S$ is isomorphic to its transformer $\Phi[S]$)

\[
\Phi[S_0] : 
\begin{align*}
(0,0) & \mapsto M \\
(0,1) & \mapsto M \\
(1,0) & \mapsto M \\
(1,1) & \mapsto M
\end{align*}
\]

\[
\Phi[S_1] : 
\begin{align*}
(0,0) & \mapsto M \\
(0,1) & \mapsto M \\
(1,0) & \mapsto \{(1,1)\} \\
(1,1) & \mapsto \{(1,1)\}
\end{align*}
\]

- $S_1$ has fewer behaviors than $S_0$: $[S_1]^* \subset [S_0]^*$
- $S_0$ has the non-interference property, but $S_1$ does not
- If non interference was a trace property, $S_1$ should have it (monotony)

Thus, the non interference property is not a trace property
Dependence properties

Many notions of dependences
So we consider just one:
what inputs may have an impact on the observation of a given output

Applications:
- reverse engineering: understand how an input gets computed
- slicing: extract the fragment of a program that is relevant to a result

This corresponds to the negation of non-interference
Definition: interference

There is **interference** between \((l, x)\) and \((l', x')\) and we write \((l', x') \rightsquigarrow (l, x)\) if and only if the following property holds:

\[
\exists m \in M, \exists v_0, v_1 \in V, \\
\{ m'(x') | m' \in \Phi(m[x \leftarrow v_0]) \} \neq \{ m'(x') | m' \in \Phi(m[x \leftarrow v_1]) \}
\]

- This expresses that there is at least one case, where the value of \(x\) at \(l\) has an impact on that of \(x'\) at \(l'\).
- It may not hold even if the computation of \(x'\) reads \(x\):

\[
\begin{align*}
\ell &: \quad x' = 0 \times x; \\
\ell' &: \quad \ldots
\end{align*}
\]
Interference is not a trace property

- We assume $\mathbb{V} = \{0, 1\}$ and $\mathbb{X} = \{x, x'\}$ (store $m$ is defined by the pair $(m(x), m(x'))$, and denoted by it).
- We assume $\mathbb{L} = \{\ell, \ell'\}$ and consider two systems such that all transitions are of the form $(\ell, m) \rightarrow (\ell', m')$ (i.e., system $S$ is isomorphic to its transformer $\Phi[S]$).

\[
\begin{align*}
\Phi[S_0] : & \quad (0, 0) & \mapsto & \quad M \quad & \quad (0, 0) & \mapsto & \quad \{(1, 1)\} \\
& \quad (0, 1) & \mapsto & \quad M \quad & \quad (0, 1) & \mapsto & \quad \{(1, 1)\} \\
& \quad (1, 0) & \mapsto & \quad \{(1, 1)\} \quad & \quad (1, 0) & \mapsto & \quad \{(1, 1)\} \\
& \quad (1, 1) & \mapsto & \quad \{(1, 1)\} \quad & \quad (1, 1) & \mapsto & \quad \{(1, 1)\}
\end{align*}
\]

- $S_1$ has fewer behavior than $S_0$: $[S_1]^* \subset [S_0]^*$.
- $S_0$ has the interference property, but $S_1$ does not.
- If interference was a trace property, $S_1$ should have it (monotony).

Thus, the interference property is not a trace property.
Hyperproperties

Conclusion:
- The absence of interference between \((\ell, x)\) and \((\ell', x')\) is not a trace property: we cannot describe as the set of programs the semantics of which is included into a given set of traces.
- It can however be described by a set of sets of traces: we simply collect the set of program semantics that satisfy the property.

This is what we call a hyperproperty:

Hyperproperties

- **Trace hyperproperties** are described by sets of sets of executions
- **Trace properties** are described by sets of executions

**2-safety**: to disprove the absence of interference (i.e., to show there exists an interference), we simply need to exhibit **two finite traces**.
The Zoo of semantic properties

Sets of sets of executions
non-interference, dependency

Trace properties
total correctness

Safety properties
never reach $s_0$ before $s_1$

State properties
absence or runtime errors
partial correctness

Liveness properties
termination
Summary

To sum-up:

- **Trace properties** allow to express a large range of program properties
- **Safety** = absence of bad behaviors
- **Liveness** = existence of good behaviors
- Trace properties can be **decomposed** as conjunctions of safety and liveness properties, with **dedicated proof methods**
- Some interesting properties are **not trace properties**
  security properties are **sets of sets of executions**
- Notion of **specification languages** to describe program properties