Construction of invariance proof methods for parallel programs with sequential consistency

Patrick Cousot

Computer Science Department
Courant Institute of Mathematical Sciences
New York University
251 Mercer Street
New York, NY 10012, USA
pcousot@cims.nyu.edu

1. History

Program proofs can be done informally \[47\] as most mathematical reasonings or as an application of a formal method. The formal method is a “recipe” to make the prove. In general, it requires the discovery of an inductive property implying the program property to be proved. The inductive property is in general stronger/more precise than the one to be proved. The program property to be proved is given but the inductive property has to be discovered (preferably automatically, but this is an undecidable problem). It must be shown to satisfy verification conditions that imply that it is inductive (meaning that it can be proved to hold by recurrence on a well-founded relation based on the program computation steps \[11, 11, 27, 15\], the program structure \[30\], or data manipulated by the program \[12, 17\]).

Although initially applied to sequential imperative programs, such program proof methods were rapidly extended to parallel programs \[44, 43, 45, 35\] with sequential consistency hypothesis “… the result of any execution is the same as if the operations of all the processors were executed in some sequential order, and the operations of each individual processor appear in this sequence in the order specified by its program.” \[36\].

2. Empiricism

The verification conditions can be postulated out of thin air (e.g. \[27, 25\]) and claimed to define the semantics of the programming language (i.e. to “assign meanings to programs” \[24\]). The problem is that the design of the verification conditions is by trial and errors and they can be unsound and/or incomplete, without any way to prove these facts.

For example the conjunction and disjunction rules of Hoare logic are not sound for all assertion languages \[11\], Sect. 5\]. Predicate transformers \[25\] were incomplete for unbounded nondeterminism and had to be later generalized \[26\]. For similar reasons of lack of expressiveness of the assertion language, the Owicki and Gries proof method for parallel programs with shared variables \[44\] and without auxiliary variables is incomplete. The Owicki and Gries proof method for parallel programs with resources \[15\] as well \[37\]. The separation logic \[12\] as well, etc. Moreover, this empirical approach is completely language-dependent and basic principles have to be rediscovered whenever the language or its semantics is changed \[45\].
3. Constructionism

An alternative to empiricism is to derive the program proof method from an operational semantics by calculational design [3, Ch. 3], [20], [22]. The general idea is that a program property is an abstraction of the most general program property induced by the operational semantics of the programming language.

For example the operational semantics of a program can be defined by a transition system $(S, t)$ where $S$ is a set of states and $t \in \wp(S \times S)$ a transition relation [3, Ch. 3]. The set of reachable states from a set of initial states $S_0 \in \wp(S)$ is $r = \{ s' \in S \mid \exists s \in S_0 : (s, s') \in t^* \}$ where $t^*$ is the reflexive transitive closure of $t$. An invariant $P \in \wp(S)$ is an over-approximation of the reachable states $r \subseteq P$. It is invariant in that during any execution, a reachable state will always satisfy $P$ (i.e., belong to $P$). This can be formulated as a fixpoint problem in that $r = \text{lfp} F$ where $F(X) = S_0 \cup \{ s' \mid \exists s \in X : (s, s') \in t \}$ [3, Ch. 3] so that we have to prove that $\text{lfp} F \subseteq P$.

4. Fixpoint Induction

Since this strongest/most general property (reachable states $r$ in the above example) can be expressed as a fixpoint $\text{lfp} F$, and its abstractions as well, the verification conditions directly derive from fixpoint induction, a direct immediate consequence of Tarski’s theorem $\text{lfp} F = \cap \{ X \in L \mid F(X) \subseteq X \}$ (where $F \in L \mapsto L$ on the complete lattice $(L, \sqsubseteq)$) [10]. We have $\text{lfp} F \subseteq P$ if and only if $\exists I : F(I) \subseteq I \cap P$ (where $\text{lfp} F$ is the strongest program property, $\subseteq$ is logical implication, $P$ is the property to be proved, $I$ is a stronger inductive property, and $F$ (or $F(I) \subseteq I$) is the verification condition.

For the above example, this is $\exists I : S_0 \subseteq I \wedge \forall s \in I : \forall s' : (s, s') \in t \Rightarrow s' \in I$ i.e., the invariant must be true for all initial states and if it is true for one state $s$, it must be true for all its possible successors $s'$ by one more transition, if any.

Fixpoint induction is a universal induction principle for proving invariance (with numerous variants [13], for example backward methods consist in inverting the transition relation [24]).

5. Soundness and Completeness

Soundness consists in proving that the proof method is correct. Soundness follows from $I \in \{ X \in L \mid F(X) \subseteq X \}$ so $\text{lfp} F = \cap \{ X \in L \mid F(X) \subseteq X \} \subseteq I$ and $I \subseteq P$ proving $\text{lfp} F \subseteq P$ by transitivity. Completeness consists in proving that if the property to be proved does hold (i.e., $\text{lfp} F \subseteq P$) then the proof method is always applicable.

Completeness follows from the fact that it is always possible to choose $I = \text{lfp} F$ so $F(I) = I \subseteq I \cap P$. Although $I = \text{lfp} F$ can always be calculated iteratively, the iterates might have to be transfinite [10], and so this does not provide an effective algorithm to compute the inductive property. This is why in deductive methods the inductive property must be provided by the end-user, the proof system just generating the verification conditions $F(I) \subseteq I$ and $I \subseteq P$, and a theorem prover or SMT-solver is used to check the implication $\subseteq$.

To continue our example, the operational semantics of a parallel program $[P_1] \ldots [P_n]$ with shared variables $x_1, \ldots, x_m$ and sequential consistency can be formalized by the transition relation $t = \bigcup_{i=1}^n t_i$ where $t_i$ is the transition relation for process $P_i$, $i = 1, \ldots, n$. The states are of the form $(c_1, \ldots, c_n, x_1, \ldots, x_m) \in S$ where $c_i \in C_i$ is the value of the program counter of process $P_i$, $i = 1, \ldots, n$ and $x_1, \ldots, x_m$ are the values of the shared variables. So a transition $t_i$ by process $P_i$ changes the program counter $c_i$ of that process and potentially the values $x_1, \ldots, x_m$ of the shared variables $x_1, \ldots, x_m$. The program execution is the interleaving of the process executions. In the Lamport proof method [42] for parallel programs $[P_1] \ldots [P_n]$ with shared variables $x_1, \ldots, x_m$ an
assertion on \(c_1, \ldots, c_{i-1}, c_{i+1}, \ldots, c_n, x_1, \ldots, x_m\) is attached to each program point \(c_i\) of each process \(P_i, i = 1, \ldots, n\). The method is sound and complete. The verification condition states that for all processes \(i, i = 1, \ldots, n\), the local assertions attached to program points \(c_i \in C_i\) should be invariant for transitions \(t_i\) (so-called sequential proof) as well as transitions \(t_j, j \in [1, n] \setminus \{i\}\) (so-called absence of interference proof). In the Owicki and Gries proof method \([14]\) for parallel programs \([P_1][\ldots[P_n]\], the assertion attached to each program point \(c_i\) of each process \(P_i, i = 1, \ldots, n\) is on the values \(x_1, \ldots, x_m\) of the shared variables \(x_1, \ldots, x_n\) only. This is sound but incomplete (since it is not possible to specify when process \(P_i\) is at some point \(c_i\) where the other processes should be, e.g. out of a critical section). Adding auxiliary variables as proposed by \([13]\) makes the method complete (since, as shown in \([20]\), Th. 10.0.4 applied to Ex. 10.0.5 and Ex. 10.0.6, it is always possible to extend to transition system with auxiliary variables simulating the program counters of the processes). Rely-guarantee methods \([32, 3]\) are a different way of expressing the strongest invariant as a fixpoint \([35]\) based on asynchronous iterations with memory \([3]\).

6. Galois connections

More generally, proof methods do not directly refer to the semantics but to an abstraction of the semantics \([1, 11]\). Most often the abstraction can be be formalized by a Galois connection \(\langle \varphi(S), \subseteq \rangle \cong (A, \sqsubseteq)\) such that \(\alpha(P) \subseteq Q \iff P \subseteq \gamma(Q)\) where \(P \in \varphi(S)\) is a concrete property, and \(Q \in A\) is an abstract property, \(\alpha(P)\) is the best abstraction of the concrete property \(P\), and \(\gamma(Q)\) is the concretization of the abstract property \(Q\) \([11, 11]\).

For the Owicki-Gries example, \(A = \prod_{i=1}^n \prod_{c_i \in C_i} V_m\) where \(V\) is the domain of values \(x_1, \ldots, x_m\) of the shared variables. The abstraction is \(\alpha(P) \triangleq \prod_{i=1}^n \prod_{c_i \in C_i} \{\langle x_1, \ldots, x_m \rangle \mid \exists c_1, \ldots, c_{i-1}, c_{i+1}, \ldots, c_n : \langle c_1, \ldots, c_i, \ldots, c_n, x_1, \ldots, x_m \rangle \in P\}\). The concretization is \(\gamma(\prod_{i=1}^n \prod_{c_i \in C_i} Q_{i,c_i}) \triangleq \{\langle c_1, \ldots, c_{i-1}, c_i, c_{i+1}, \ldots, c_n, x_1, \ldots, x_m \rangle \mid \exists c_1, \ldots, c_{i-1}, c_{i+1}, \ldots, c_n : \langle x_1, \ldots, x_m \rangle \in Q_{i,c_i}\}\). The local assertions \(Q_{i,c_i}\) ignore the program counters of other processes (which can be any \(c_1, \ldots, c_{i-1}, c_{i+1}, \ldots, c_n\)). This is the source of the incompleteness. The partial order \(\subseteq\) is pointwise set-theoretic implication (i.e. inclusion \(\subseteq\) for sets) \(\prod_{i=1}^n \prod_{c_i \in C_i} Q_{i,c_i} \subseteq \prod_{i=1}^n \prod_{c_i \in C_i} Q'_{i,c_i}\) if and only if \(\forall i \in [1, n] : \forall c_i \in C_i : Q_{i,c_i} \subseteq Q'_{i,c_i}\).

Then for an abstract property \(Q \in A\), the program verification consists in proving is \(\alpha(\text{lfp } F) \subseteq Q\).

If \(\alpha \circ F = F \circ \gamma\) then this is equivalent to \(\text{lfp } \overline{F} \subseteq Q\) where \(\overline{F} = \alpha \circ F \circ \gamma \in A \rightarrow A\) defines the abstract verification conditions \([11\text{ theorem 7.1.0.4}(3)]\), \([22\text{ lemma 4.3}]\), \([11\text{ fact 2.3}]\). Then, as before but this time in the abstract, \(\alpha(\text{lfp } F) \subseteq Q \iff \exists I \in A : F(I) \subseteq I \wedge I \subseteq Q\).

If \(\alpha \circ F \subseteq \overline{F} \circ \gamma\) then the method is sound (i.e. \(\alpha(\text{lfp } F) \subseteq Q \iff \exists I : F(I) \subseteq I \wedge I \subseteq Q\)) but in general incomplete (\(\not\exists\)). This is almost always the case in static analysis \([11, 11]\).

7. Abstraction

Not all abstractions can be formalized by Galois connections. This is the case when there is no best/most precise abstraction.

An example is the use of logics like first-order logic, which is an abstraction. In general the strongest/most general property is not expressible in the logic which makes it incomplete. In that case only relative completeness is provable (e.g. \([11\text{ for Hoare logic }\]) i.e. under the assumption that the strongest/most general set-theoretic property is expressible in the logic. This is why we represent properties by the set of all objects which have this property rather than by a logical formula. In that case only the concretization function \(\gamma\) is used \([10]\). This is the case for formal logics or types
which may not way to have a best way to express a property (e.g. a first-order logic with addition only cannot express a multiplication while a program with addition only can expression multiplication with a loop. Adding multiplication to the logic, is not enough for exponentiation, etc.).

8. The Hierarchy of Proof Methods

Finally by choosing different abstractions, one obtained a hierarchy of famous (as well as completely forgotten) proof methods for parallel programs with sequential consistency [3, 41, 13, 7, 37, 38].

9. Methodology

In summary, given a language, an operational semantics, a fixpoint definition of runtime properties, an abstraction into properties of interest, a further abstraction to express such properties of interest locally, there is a mathematical methodology, checkable by a proof checker [33], to construct a proof method which is guaranteed to be sound, and depending on the appropriate choice of the abstractions, complete. Beyond proof methods, this methodology extends to all program semantics [8, 28].

10. Applications

Using approximate abstractions, this contructive methodology leads to the design of sound program enumerative checking methods (aiming at checking a given property in the abstract e.g. [18]) and analysis methods (aiming at automatically infer the invariant property from the program text only e.g. [8, 11]). For calculability, these abstractions have to be incomplete (sometimes even for sets of states \( S \) which are finite [18, 29]). In particular for parallel programs, approximation is indispensable for scalability [12, 13], from prototype [31, 40] to production-quality implementations [38, 39].

References


Construction of invariance proof methods for parallel programs with sequential consistency


P. Cousot


Construction of invariance proof methods for parallel programs with sequential consistency


