Ensuring the Correctness of Lightweight Tactics for JavaCard Dynamic Logic

Richard Bubel\textsuperscript{1} Andreas Roth\textsuperscript{2} Philipp Rümmer\textsuperscript{3}

\textit{Institut für Logik, Komplexität und Deduktionssysteme}
\textit{Universität Karlsruhe, Germany}

Abstract
The interactive theorem prover developed in the KeY project, which implements a sequent calculus for JavaCard Dynamic Logic (JavaCardDL) is based on taclets. Taclets are lightweight tactics with easy to master syntax and semantics. Adding new taclets to the calculus is quite simple, but poses correctness problems. We present an approach how derived (non-axiomatic) taclets for JavaCardDL can be proven sound in JavaCardDL itself. Together with proof management facilities, our concept allows the safe introduction of new derived taclets while preserving the soundness of the calculus.

\textit{Key words:} taclets, lightweight tactics, dynamic logic, theorem proving

1 Introduction

Background
Taclets are a new approach for constructing powerful interactive theorem provers [7]. First introduced as \textit{Schematic Theory Specific Rules} [10], they are an efficient and convenient framework for lightweight tactics. Their most important advantages are the restricted and, thus, easy to master syntax and semantics compared to an approach based on meta languages like ML, and their seamless integration with graphical user interfaces of theorem provers which they can be efficiently compiled [7] into.

Taclets contain three kinds of information, the logical content of the rule to be applied, information about side-conditions on their applicability, and pragmatic information for interactive and automatic use. Due to their easy

\textsuperscript{1} Email:bubel@ira.uka.de
\textsuperscript{2} Email:aroth@ira.uka.de
\textsuperscript{3} Email:ruemmer1@ira.uka.de

This is a preliminary version. The final version will be published in \textit{Electronic Notes in Theoretical Computer Science}
URL: \texttt{www.elsevier.nl/locate/entcs}
syntax and intuitive operational semantics, a person with some familiarity in formal methods should be able to write own taclets after a short time of study.

The interactive theorem prover developed in the KeYproject [?] is based on taclets implementing a sequent calculus for JavaCard Dynamic Logic (JavaCardDL) [?]. JavaCard is a subset of Java lacking multi-threading, garbage collection and graphical user interfaces, but with additional features like transactions.

JavaCardDL has around three hundred axiomatic rules, this means taclets that capture the JavaCard semantics. Correctness of rules is crucial since new taclets can be introduced quite easily. The work presented here ensures the correctness of derived taclets for JavaCardDL by providing means to prove them correct relatively to the core set of JavaCardDL axioms (possibly enriched with further axioms for certain theories). The soundness of taclets is proven in the calculus itself and without escaping to higher order logics. By our work, we extend an approach already described in [?] for classical first-order logic to JavaCardDL.

Related Work

Related to our approach are other projects for program verification like Bali [?,?], where consistence and correctness of rules that cover the Java semantics are ensured using Isabelle, or the LOOP project [?] where PVS is used as foundation, and the calculus rules are thus obtained as higher order logic theorems. Complementary to the presented approach—ensuring correctness for derived taclets—further work has been carried out in the KeY project in order to cross-validate a few selected axiomatic rules against the Java axiomatisation of Bali [?].

Structure of this Paper

In Sect. 1.1 we repeat the most important concepts of classical dynamic logics and JavaCardDL. A formal description of taclets and a definition of the basic vocabulary used throughout the paper is given in Sect. 2. The different steps to be performed in order to prove the correctness of derived taclets are described in Sect. 3–5. In Sect. 6 we give a justification of the complete procedure as main theorem. Finally, in Sect. 7 we discuss the current and future work to be done.

1.1 Dynamic Logic

Classical Dynamic Logics

The family of dynamic logics (DL) [?] belongs to the class of multi modal logics. As programs are first-class citizens of DL formulas, they are well-suited for program analysis and reasoning purposes. For the sake of simplicity and as a consequence of using a non-concurrent and real world programming language, we will only consider deterministic programs.
Let $p$ be an arbitrary program and $\phi$ a first-order or dynamic logic formula, then

- $\langle p \rangle \phi$ ("diamond $p \phi$")$: p$ terminates and after the execution of $p$ formula $\phi$ holds
- $[p] \phi$ ("box $p \phi$")$: if $p$ terminates then after the execution of $p$ formula $\phi$ holds

are typical representatives of DL formulas. Deterministic propositional dynamic logic (DPDL) is defined over a signature $\Sigma = (At_0, Prg_0, Op)$, where $At_0, Prg_0$ are enumerable sets of propositional variables and atomic programs respectively. Besides the classical propositional operators $\neg, \rightarrow$, the operator set $Op$ contains box $[p]$ and diamond $\langle p \rangle$ modalities for each program $p$. The set of formulas is the smallest set defined inductively over $At_0$ and $Prg_0$:

- all classical propositional formulas are formulas in DPDL
- if $\phi, \psi$ are DPDL formulas then $\phi \rightarrow \psi$ and $\neg \phi$ are DPDL formulas
- if $p \in Prg$ is a program and $\phi$ a formula in PDL then $\langle p \rangle \phi$ and $[p] \phi$ are DPDL formulas
- the set $Prg$ of programs is the smallest set satisfying
  (i) $Prg_0 \subseteq Prg$
  (ii) if $p, q \in Prg$ and $\psi$ $\in$ DPDL then
   ‘$p;q$’ (concatenation), ‘if ($\psi$) $\{p\}$ else $\{q\}$’ and ‘while ($\psi$) $\{p\}$’
   are programs.

The semantics can be defined in terms of Kripke frames $(S, (\rho_p)_{p \in Prg})$ with a set $S$ of states, and transition relations $\rho_p : S \rightarrow S$ which define the semantics of each program $p \in Prg$. The relations $\rho_p$ have to adhere certain conditions w.r.t. the program constructors $(;,$ if-else, etc.) from the definition above, for example, program composition $\rho_{p;q} = \rho_q \circ \rho_p$.

An excerpt from an axiom system for DPDL in terms of sequent calculus rules is given in Table 1.  

DPDL is useful to reason about program properties induced by the program constructors. However, as a consequence of constructing programs from atomic (anonymous) programs without any fixed semantics, they lack possibilities to talk about individual programs and, thus, about functional properties.

Like the step from propositional to first-order logic, one extends DPDL to deterministic quantified dynamic logics (DQDL). DQDL extends the propositional part to full first-order logic (with equality and universe $D$), and on the program side it replaces the anonymous atomic programs with assignments of the form $v=t$, where $v$ is a variable and $t$ an arbitrary term. In general, each program state $s \in S$ is assigned a first order structure $(D, I)$ and a variable valuation $\sigma : Var \rightarrow D$ respecting $\rho_{x=t}(s) = s'$ with $\sigma' = \sigma^{(D, I), \sigma}_x$.  

Again a relatively complete calculus can be given, the corresponding

---

4 Usually DQDL is interpreted in an arithmetic structure.
assignment rule is shown in Table 1.

**Example 1.1** For the universe $D = \mathbb{N}$ of natural numbers, the DQDL formula

$$\langle x=3; y=x \rangle y = x$$

can be proven valid with the rules of Table 1 as the proof on the right shows.

**JavaCardDL**

The step from academic languages as described in the previous paragraphs to real world programming languages like JavaCard [? ,?] leads to several complications. In the next few paragraphs, we introduce some features of JavaCardDL [?]. First some preliminaries:

- Formulas must not occur in JavaCardDL programs, instead Java expressions of type `boolean` are used as guards.
- The set of variables $Var = PVar \uplus LVar$ is the disjoint union of program variables $PVar$ and logical variables $LVar$. In contrast to logical variables, program variables can occur in programs as well as in formulas, but cannot be bound by quantifiers. For instance, let $x \in LVar$ and $o, u \in PVar$, then $\forall x. \langle o = u; \rangle x = u$ is a well-formed JavaCardDL formula, whereas $\forall x. \langle o = u; \rangle x = u$ is not.
- All states have the same universe $D$, and predicates and logical variables are assumed to be rigid.

A sequent calculus covering JavaCard has to cope with aliasing, side-effects, abrupt termination as result of thrown exceptions, breaks, continues or returns and more. The KeY approach is led by the symbolic execution paradigm, thus a majority of the calculus rules realises a JavaCard interpreter reducing expressions and statements stepwise to side-effect free assignments.
Example 1.2 An easy-to-use decomposition rule similar to (3) is not available in JavaCardDL due to abrupt termination. For example

\[ \vdash \langle 1:\{ \text{if} (v == 0) \{ \text{break} l; \} \text{else} \{ v = 0; \} v = 3; \} \rangle v = 3 \]

cannot be decomposed to

\[ \vdash \langle 1:\{ \text{if} (v == 0) \{ \text{break} l; \} \text{else} \{ v = 0; \} \} \rangle \langle v = 3 \rangle v = 3, \]

as this is obviously not equivalent for \( v = 0 \).

Decomposition was essential for DPDL and DQDL in order to reduce the complexity of programs stepwise to atomic programs or assignments, which can be handled by rules of the calculus without having a dedicated rule for each program.

JavaCardDL therefore introduces the notion of a first active statement to which a rule applies, and a program context ‘.. \( \circ_1 \) ..’ whose inactive prefix ‘..’ matches on all preceding labels, opening braces or try blocks. Consider the following rule:

\[
\begin{align*}
\#b \doteq \text{true} & \vdash \langle .. \{ \#st1 \} .. \rangle \phi \\
\#b \doteq \text{false} & \vdash \langle .. \{ \#st2 \} .. \rangle \phi \\
& \vdash \langle .. \text{if} (\#b) \{ \#st1 \} \text{else} \{ \#st2 \} .. \rangle \phi
\end{align*}
\]

(5)

where \( \#b \) is a side-effect free boolean expression and \( \#st1, \#st2 \) are arbitrary JavaCard statements.

Example 1.3 (Example 1.2 continued) Applying rule (5) to

\[ \vdash \langle 1:\{ \text{if} (v == 0) \{ \text{break} l; \} \text{else} \{ v = 0; \} v = 3; \} \rangle v = 3 \]

where \( \circ_1 \) corresponds to the program between ‘1:’ (inactive program prefix) and ‘\( v = 3; \)’ (suffix of the program context) now yields the two sequents

(i) \( (v==0) \doteq \text{true} \vdash \langle 1:\{ \text{break} l; \} v = 3; \} \rangle v = 3 \) and

(ii) \( (v==0) \doteq \text{false} \vdash \langle 1:\{ v = 0; \} v = 3; \} \rangle v = 3 \)

2 Taclets

Taclets are lightweight, stand-alone tactics with simple syntax and semantics. Their introduction was motivated by the observation that only few basic actions in proof construction are sufficient to implement most rules for first-order modal logic. These are:

- to recognise sequents as an axiom, and to close the according proof branch,
- to modify at most one formula per rule application,
- to add a finite (and fixed) number of formulas to a sequent,
- to let a proof goal split in a fixed number of branches,
to restrict the applicability according to context information. These are the only actions which taclet constructs are provided for. This restriction turns out to reduce the complexity for users of a proof system significantly [1].

**Taclets by Example**

Taclets describe rule schemas in a concise and easily readable way. A very simple example rewrites terms 1 + 1 with 2. In taclet notation such a rule schema is written:

\[ \text{find}(1 + 1) \text{replacewith}(2) \]

In a taclet—in addition to the logical content of the described rule—an operational meaning is encoded: If a user of a taclet-based prover selects the term of the \textit{find}-part (i.e. 1 + 1) of a taclet and chooses the taclet for application, the \textit{find}-part is replaced with (an instantiation of) the \textit{replacewith}-term (i.e. 2).

In this simple form, the rule schemas described by taclets are not expressive enough for practical use; \textit{schema variables} and more constructs besides \textit{find} and \textit{replacewith} make them powerful enough to fulfil the requirements posed above.

**Schema Variables and Instantiations**

Expressions\(^5\) in taclets may contain elements from a set \(SV\) of \textit{schema variables}. An \textit{instantiation} \(\iota(v)\) of a schema variable \(v \in SV\) is a concrete expression that must fulfil certain conditions depending on the kind of the schema variable (see below). We may, e.g., define a schema variable \(i\) such that \(\iota(i)\) must be a term of an integer sort.

Expressions \(e\) in taclets containing schema variables from \(SV\) are called \textit{schematic expressions} over \(SV\). The instantiation map \(\iota\) can be canonically continued on schematic expressions:

\[
\iota(op(e_1, \ldots, e_n)) = \begin{cases} 
\iota(op) & \text{if } n = 0 \text{ and } op \in SV \\
\text{op}(\iota(e_1), \ldots, \iota(e_n)) & \text{otherwise}
\end{cases}
\]

(6)

Thus \(e\) describes a set \(\{\iota(e) \mid \iota\text{ is an instantiation map for every } v \in SV\}\) of concrete expressions. For instance, a taclet \texttt{find}(i + i) \texttt{replacewith}(2 \ast i) contains schematic terms over \{\(i\)\}. Applied on a sequent containing the term 3 + 3, \(i\) is instantiated with \(\iota(i) = 3\) and the taclet replaces \(\iota(i + i) = 3 + 3\) with \(\iota(2 \ast i) = 2 \ast 3\) in the new goal.

---

5 By \textit{expression} we denote syntactical elements like terms or formulas, but in the context of JavaCardDL also Java programs.
Taclet Syntax

replacewith(2) is an example of a goal template, this means the description of how a goal changes by applying the taclet. More than one goal template may be part of a taclet, separated by semicolons, which describes that a goal is split by the taclet. If there is no goal template in a taclet, applications close the proof branch. Additionally, goal templates may contain the following clauses:

- While in the example taclets above the find- and replacewith-parts consisted of terms, they can also be sequents. All find- and replacewith-parts of a taclet must either be terms or sequents. These sequents indicate that the described expression must be a top-level formula in either the antecedent or succedent, e.g. a taclet find(\( \vdash \phi \rightarrow \psi \)) replacewith(\( \phi \vdash \psi \)) (over the schema variables \( \{\phi, \psi\} \)) is applicable only to top-level formulas in the succedent. A sequent in the find-part must have either an empty antecedent or succedent.

- Taclet applications can add formulas to the antecedent or succedent. This is denoted by the keyword add followed by a schematic sequent (similarly to replacewith).

- Taclets support the dynamic enlargement of the taclet rule base by adding new taclets described behind the key word addrules. Though a theoretical treatment in accordance with the concepts of this paper is possible \[?]\ we omit this feature in the sequel.

Often more information on the sequent a taclet is applicable to is needed. Such side conditions are described by the following optional taclet constituents:

- A taclet that contains an if followed by a schematic sequent context is only directly applicable if context is a “sub-sequent” of the sequent the taclet is applied to. If this is not the case, the taclet is however still applicable but, by an automatic cut, it is required to show the if-condition.

- Predefined clauses in a varcond-part describe conditions on the instantiations of schema variables. The most important ones are:
  - v not free in s, which disallows logical variables \( \iota(v) \) to occur unbound in \( \iota(s) \).
  - v new depending on s, which introduces a new skolem symbol \( \iota(v) \) (possibly depending on free “meta variables” occurring in \( \iota(s) \)).

The complete syntax of taclets is reiterated here as an overview:

\[
\begin{align*}
\text{[if (context)] [find (f)] [varcond (c_1, \ldots, c_k)]} \\
\text{[replacewith (rw_1)] [add (add_1)];} \\
\vdots \\
\text{[replacewith (rw_n)] [add (add_n)]}
\end{align*}
\]
For $i = 1 \ldots n$, context and add$_i$ stand for a schematic sequent, $f$ and rw$_i$ for a schematic term, formula, or sequent but all of the same kind. $c_1, \ldots, c_k$ are variable conditions.

Additionally—though out of scope of this paper—taclets can be assigned to one or several rule sets, which makes them available to be automatically executed by strategies. For a homogenous treatment in this paper $f$ and rw$_i$ are declared to be never empty: we assume that skipping \texttt{replacewith} is a shorthand for rw$_i = f$, a skipped \texttt{find} means $f = \vdash false$, and \texttt{false} always occurs in succedents of sequents. A skipped if- or add-part means context $= \vdash$ or add$_i = \vdash$ (resp.).

**Schema Variable Types**

While the above definitions have been general enough to be applied to every first-order modal logic, we are now focusing on special schema variables for JavaCardDL. Let $SV_{\text{tac}}$ denote the schema variables contained in a taclet tac. Schema variables $v \in SV_{\text{tac}}$ are assigned to one out of a predefined list of types, each having special properties concerning admissible instantiations $\iota(v)$. An instrument to define these properties is to introduce prefix sets (denoted by $\Pi_l(v)$, $\Pi_{pv}(v)$, and $\Pi_{jmp}(v)$) for schema variables $v$. A selection of the most relevant schema variable types is given below. If $v$ is of type

- **Variable**, then $v$ is assigned a sort, $\iota(v)$ must be of that sort. Moreover, $\iota(v)$ must be a logical variable. For $v \neq v' \in SV$: if $v'$ is a Variable schema variable then $\iota(v) \neq \iota(v')$. $\iota(v)$ must not occur bound in $\iota(v'')$ for all $v'' \in SV$.

- **Term**, then $v$ is assigned a sort, $\iota(v)$ must be of that sort.

  - $v$ is assigned a set $\Pi_l(v) \subseteq SV$ of schema variables. $\Pi_l(v)$ is defined to be the smallest set with, for all constituents of tac, if $v$ occurs in the scope of a Variable schema variable $v' \in SV$ then $v' \in \Pi_l(v)$ except there is a variable condition $v'$ not free in $v$ declared in $t$. $v$ is assigned a set $\Pi_{pv}(v)$ which is the smallest set of program variables that occur but are not declared in tac or are declared above every occurrence of $v$.

  - We require from instantiations $\iota$: If, for some $v' \in SV_{\text{tac}}$, $\iota(v')$ is a logical variable that occurs unbound in $\iota(v)$ then $v' \in \Pi_l(v)$; if $\iota(v')$ is a program variable that occurs undeclared in $\iota(v)$ then $v' \in \Pi_{pv}(v)$.

- **Formula**, then as for Term, $v$ is assigned $\Pi_l(v)$ and $\Pi_{pv}(v)$. $v$ must fulfil the same conditions concerning these sets.

- **Statement**, then $\iota(v)$ is a JavaCard statement. Again, $v$ is assigned $\Pi_{pv}(v)$ and it must satisfy the same conditions as above concerning this set.

  - $v$ is assigned a set $\Pi_{jmp}(v)$ consisting of JavaCard statements \texttt{break}, \texttt{continue}, \texttt{break} l, \texttt{continue} l for all labels $l$, if $v$ is enclosed with a suitable jump target. If $jst$ is a \texttt{break} or \texttt{continue} statement of $\iota(v)$ with a target

---

6 If we consider tac as abstract syntax tree.
not in \( \iota(v) \) then \( jst \in \Pi_{jmp}(v) \). Usually \textbf{Statement} schema variables have names starting with \# to distinguish them from regular \textit{Java} elements.

- **ProgramVariable**, then \( \iota(v) \) is a local program variable or class attribute of \textit{Java}. \( v \) is assigned a \textit{Java} type and \( \iota(v) \) must be of that type. Again, names of this kind of variable start with \#.

- **ProgramContext**, then \( \iota(v) \) is a program transformation\(^8\) \( pt \) that takes a \textit{Java} program element \( \alpha \) and delivers a new program element \( pt(\alpha) \), such that \( pt(\alpha) \) is a sequence of statements of which the first one contains \( \alpha \) and has only opening braces, opening \texttt{try} blocks, etc., in front. For this case, the continuation of the instantiation map (6) is then modified to

\[
\iota(op(e_1, \ldots, e_n)) := pt(\iota(e_1))
\]

if \( n = 1 \) and \( op \) is a \textbf{ProgramContext} schema variable.

Usually, \( v \) is denoted by \( \ldots e_1 \ldots \) containing the schematic \textit{Java} program \( e_1 \), as introduced in Sect. 1.1.

\textbf{Example 2.1} The following taclet performs a cut with the condition that the focused term \( t \) equals 0 and replaces it in the respective goal by 0. We declare \( t \) as \textbf{Term} schema variable of an integer sort.

\begin{align*}
\text{find}(t) &\quad \text{replacewith}(0) \quad \text{add}(t \equiv 0 \vdash ) ; \\
\text{replacewith}(t) &\quad \text{add}( \vdash t \equiv 0 ) \quad \tag{8}
\end{align*}

As an example that represents a rule of \textit{JavaCardDL}, we take a taclet that replaces the postfix increment operator applied to a program variable \( x \) behind a statement \( (#sta) \) with an equivalent statement using assignment and the + operator, and leaves the formula \( (\phi) \) behind the diamond unchanged. \( #sta \) is a \textbf{Statement} schema variable and \( \phi \) a \textbf{Formula} schema variable.

\begin{align*}
\text{find}(\langle #sta x++;\phi \rangle) &\quad \text{replacewith}(\langle #sta x=x+1;\phi \rangle) \quad \tag{9}
\end{align*}

Another taclet splits a proof for an \texttt{if} statement with the condition \( x==0 \) (where \( x \) is a concrete local variable) and produces goals, reducing the formula to the statements of the appropriate branch and the if condition put to the correct side of the sequent. \( #sta1 \) and \( #sta2 \) are \textbf{Statement} schema variables and \( \phi \) is a \textbf{Formula} schema variable.

\begin{align*}
\text{find}(l: \texttt{if } (x==0) \#sta1 \texttt{ else } \#sta2)\phi) \\
\text{replacewith}(l: \#sta1)\phi) &\quad \text{add}(x \equiv 0 \vdash ) ; \\
\text{replacewith}(l: \#sta2)\phi) &\quad \text{add}( \vdash x \equiv 0 ) \quad \tag{10}
\end{align*}

\( ^7 \) For a complete treatment of \textit{JavaCardDL} it is furthermore necessary to consider \textbf{return}-statements, which are left out in this paper

\( ^8 \) Thus being an exception from the statement above that \( \iota(v) \) must be an expression.
Semantics

Taclets have a precise operational semantics, which is described in detail in [?], and which we have sketched informally above. For the purposes of this paper it is sufficient to fix the logical meaning of a taclet in the traditional rule schema notation.

We denote the union of two sequents and the subset relation between two sequents as follows:

\[
(\Gamma_1 \vdash \Delta_1) \cup (\Gamma_2 \vdash \Delta_2) := \Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2
\]

\[
(\Gamma_1 \vdash \Delta_1) \subseteq (\Gamma_2 \vdash \Delta_2) \text{ iff } \Gamma_1 \subseteq \Gamma_2 \text{ and } \Delta_1 \subseteq \Delta_2
\]

First, we assume that \( f \) is a schematic sequent, i.e. taclet \( tac \) can only be applied to top-level formulas. By the operational semantics of taclets [?], \( tac \) represents the rule schema:

\[
\frac{rw_1 \cup add_1 \cup (\Gamma \vdash \Delta) \ldots rw_k \cup add_k \cup (\Gamma \vdash \Delta)}{f \cup (\Gamma \vdash \Delta)} \tag{11}
\]

where \( \Gamma \vdash \Delta \) is an arbitrary sequent with context \( \subseteq f \cup (\Gamma \vdash \Delta) \).

Similarly, if \( f \) is a schematic term or formula (\( seq[e] \) denotes a sequent with an arbitrary but for a rule fixed occurrence of an expression \( e \)):

\[
\frac{seq[rw_1] \cup add_1 \cup (\Gamma \vdash \Delta) \ldots seq[rw_k] \cup add_k \cup (\Gamma \vdash \Delta)}{seq[f] \cup (\Gamma \vdash \Delta)}
\]

where \( \Gamma \vdash \Delta \) is an arbitrary sequent with context \( \subseteq seq[f] \cup (\Gamma \vdash \Delta) \).

In Sect. 4, the notion of meaning formulas is derived that makes the meaning of these rule schemas induced by taclets more precise.

Because of their simplicity and operational meaning, taclets can be schematically compiled into the GUI of taclet-based interactive theorem provers: In the KeY system a mouse click over an expression displays only those taclets whose find-part can be matched with the expression in focus. This reduces drastically the cognitive burden on the user. For an extensive account on user interaction see [?].

3 Outline of Bootstrapping Taclets

After having introduced basic notions and notations, we can focus on the task of how to ensure correctness of derived taclets. We aim to prove their soundness within the JavaCardDL calculus itself. Our approach is based on [?] which has already provided this kind of bootstrapping for classical first-order logic.

Given a taclet \( tac \), we first derive a meaning formula \( M(tac) \) (see Sect. 4), which is supposed to be valid if and only if all possible applications of \( tac \) are
correct proof steps. For example, consider the following taclet $\text{tac}_0$:

$$\text{find}(\text{true} \land \phi \vdash) \text{ replacewith} (\phi \vdash)$$

with a Formula schema variable $\phi$. The corresponding meaning formula is

$$M(\text{tac}_0) = \neg\phi \rightarrow \neg(\text{true} \land \phi) \text{ or equivalently } (\text{true} \land \phi) \rightarrow \phi$$

Intuitively, the meaning formula states that if a formula in an antecedent is replaced, the new formula must be at most as strong as the old one. If this can be proven for all instantiations of $\phi$, i.e. for all formulas, then obviously $\text{tac}_0$ is sound.

Unfortunately, meaning formulas contain schema variables (here: $\phi$) and are thus no JavaCardDL formulas. Moreover, we have to quantify somehow over all formulas. Skolemisation of schema variables (see Sect. 5) helps us, however, not having to leave our original logic and not having to employ higher order logics on the object level. Skolemisation of meaning formula $M(\text{tac}_0)$ leads to

$$M_{\text{Sk}}(\text{tac}_0) = (\text{true} \land \phi_{\text{Sk}}) \rightarrow \phi_{\text{Sk}},$$

where $\phi_{\text{Sk}}$ is a new nullary predicate. We call these formulas $M_{\text{Sk}}(\text{tac})$ taclet proof obligations. $M_{\text{Sk}}(\text{tac})$ is a JavaCardDL formula (with a slightly extended vocabulary) and can be loaded into our interactive theorem prover. If the proof obligation can be proven successfully then correctness of the taclet is ensured for all possible applications according to the definition of the meaning formula. The proof of the corresponding theorem is given in [?] and sketched in Sect. 6.

On a semantic level, this theorem can be justified by arguing that if an application of the taclet $\text{tac}$ leads to an incorrect proof, a suitable interpretation $D$ can be constructed such that the meaning formula $M(\text{tac})$ is not satisfied under $D$ (which is a direct consequence of the definition of meaning formulas) and thus $M(\text{tac})$ could not have been proven. This semantic argumentation works fine for first-order logics [?], but when JavaCardDL comes into play, the complete complex JavaCard semantics would have to be incorporated in the reasoning.

Instead, we take a syntactic approach getting the JavaCard semantics via the JavaCardDL calculus for free. The basic idea is to show that an application of a taclet $\text{tac}$ can always be replaced by a transformed proof of $M_{\text{Sk}}(\text{tac})$.

4 Meaning Formulas of Taclets

The basis for our reasoning about the correctness of taclets is a meaning formula [?] derived in this section. It is declared to be the meaning of a taclet independently from concrete taclet application mechanisms, thus providing a very flexible way to address soundness issues. In fact we define a taclet application mechanism to be correct if (and only if) taclets with valid meaning
formulas are translated into sound rules.\footnote{As a schematic formula, the meaning formula is by definition valid iff all instances of the formula are valid.} To show that a taclet is correct it is thus sufficient to prove the validity of its meaning formula.

For the whole section we define $(\Gamma \vdash \Delta)^* := \bigwedge \Gamma \rightarrow \bigvee \Delta$, in particular $(\vdash \phi)^* = \phi$ and $(\phi \vdash)^* = \neg \phi$. Furthermore, in this section by the validity of a sequent we mean the validity of $(\Gamma \vdash \Delta)^*$. We define a (sequent) calculus $C$ to be \textit{sound} if only valid sequents are derivable in $C$. We conceive rules

\[
\frac{P_1 \ldots P_n}{Q}
\]

as relations between tuples of sequents (the premisses) and single sequents (the conclusion) and define that a rule $R \in C$ is sound if for all tuples $((P_1, \ldots, P_k), Q) \in R$:

if $P_1, \ldots, P_k$ are valid, then $Q$ is valid. \hfill (12)

For the calculus $C$ we can state:

\textbf{Lemma 4.1} $C$ is sound if all rules $R \in C$ are sound.

The rules $R_{tac}$ we are interested in are defined through taclets $tac$ over a set $SV$ of schema variables in the form as defined in (7). Assuming first that the find-part is a sequent, taclets induce the rule schema (11). To apply Lem. 4.1, for each instantiation $\iota$ of $SV$, (12) must be shown for $k = n$, $P_i = \iota(\text{rw}_i \cup \text{add}_i \cup \Gamma \vdash \Delta)$ $(i = 1 \ldots n)$, and $Q = \iota(f \cup \Gamma \vdash \Delta)$. Since the formulas of $\Gamma \vdash \Delta$ which are not in context are arbitrary and not influenced by the rule application we can simply omit them and show the lemma for $P_i = \iota(\text{rw}_i \cup \text{add}_i \cup \text{context})$ $(i = 1 \ldots n)$ and $Q = \iota(f \cup \text{context})$. We assume that $tac$ does not introduce skolem functions, i.e. does not contain such a variable condition. Then by the deduction theorem, the global condition (12) can be strengthened to the local implication, namely that $P_1^* \land \ldots \land P_n^* \rightarrow Q^*$ must be valid.

Since $\iota$, as defined by (6), treats propositional junctors as a homomorphism and the operator $(\cdot)^*$ is a homomorphism regarding the union of sequents up to propositional transformations, this formula can now be simplified as follows:

\[
P_1^* \land \ldots \land P_n^* \rightarrow Q^* = \bigwedge_{i=1}^{n} \iota(\text{rw}_i \cup \text{add}_i \cup \text{context})^* \rightarrow \iota(f \cup \text{context})^* \quad (13)
\]

\[
= \iota\left(\bigwedge_{i=1}^{n} (\text{rw}_i^* \lor \text{add}_i^*) \rightarrow (f^* \lor \text{context}^*)\right). \quad (14)
\]

If (14) is proven for all instantiations $\iota$, then the rule $R_{tac}$ represented by $tac$ is sound.

In the next definition our previously made assumptions are revoked: the variable condition $sv_i \text{ new depending on } \ldots$ introduces new skolem functions.
If \( P_1, \ldots, P_n \) contain skolem symbols that do not occur in \( Q \), the interpretation of the symbols can be regarded as universally quantified in (12) by the usual definition of ‘valid’. Because of their negation in (13), they are existentially bound in the meaning formula. Moreover, taclets that have terms or formulas instead of sequents as find-part and replacewith-parts are reduced to a rule that adds an equivalence \( f \leftrightarrow rw_i \) or equation \( f = rw_i \) to the antecedent.

**Definition 4.2 (Meaning Formula)** Each taclet \( tac \), as declared in (7), is assigned an unquantified meaning formula \( tac^* \), which is defined by:

\[
tac^* := \begin{cases} 
\Lambda_{i=1}^n (rw_i^* \vee add_i^*) \rightarrow (f^* \vee \text{context}^*) & \text{if } f \text{ is a sequent} \\
\Lambda_{i=1}^n \left(f \Rightarrow rw_i \rightarrow add_i^* \right) \rightarrow \text{context}^* & \text{if } f \text{ is a term} \\
\Lambda_{i=1}^n \left(f \leftrightarrow rw_i \rightarrow add_i^* \right) \rightarrow \text{context}^* & \text{if } f \text{ is a formula}
\end{cases}
\]

Suppose \( sv_1, \ldots, sv_k \in SV_{tac} \) are all schema variables, which \( tac \) contains a variable condition \( sv_i \) new depending on... for. \( M(tac) := \exists x_1 \ldots \exists x_k . \phi \) is defined to be the meaning formula of \( tac \) where \( \phi \) is obtained from \( tac^* \) by replacing each \( sv_i \) with a new variable \( x_i \) with the same sort as \( sv_i \).

**Example 4.3 (Example 2.1 continued)** The taclets \( tac_1 \), \( tac_2 \), and \( tac_3 \) defined through (8), (9), and (10), resp., have (after applying some propositional equivalence transformations) the following meaning formulas:

\[
M(tac_1) = (t \downarrow 0 \land t \downarrow 0) \lor (t \downarrow t \land \neg(t \downarrow 0)) \tag{15}
\]

\[
M(tac_2) = \langle \#sta x++; \rangle \phi \leftrightarrow \langle \#sta x=x+1; \rangle \phi \tag{16}
\]

\[
M(tac_3) = \langle(1: \text{if } (x==0) \#sta1 \text{ else } #sta2) \phi \leftrightarrow (1: \#sta1) \phi \rangle \tag{17}
\]

\[
\wedge x \downarrow 0 \lor (\langle(1: \text{if } (x==0) \#sta1 \text{ else } #sta2) \phi \leftrightarrow (1: \#sta2) \phi \rangle \wedge \neg(x \downarrow 0))
\]

## 5 Construction of Proof Obligations

Except for trivial taclets, the meaning formula \( M(tac) \) of a taclet \( tac \) contains schema variables, which is at least inconvenient for proving \( M(tac) \). Variables of these types do however not occur bound within the formula (resp. when considering validity, they can be regarded as implicitly universally quantified), and hence it is possible to replace them in a suitable way without altering the validity of the meaning formula:

- Schema variables for logical variables or program variables can simply be replaced with new concrete variables. It has to be taken in account, however, that when instantiating a schematic expression it is possible that two
different schema variables of type ProgramVariable are instantiated with the same concrete variable (which is not possible for Variable schema variables by the definitions of Sect. 2). By the presence or absence of such collisions, the set of instances of a schematic expression is divided into (finitely many) classes, which all have to be considered to capture the meaning of the schematic expression.

- Schema variables for terms, formulas or Java statements can be replaced with suitable “skolem” symbols, which are similar to the atomic programs of DPDL for Statement schema variables. To model the notion of abrupt termination, which does not exist in DPDL, tuples of Java jump statements are attached to occurrences of symbols for statements.

- Schema variables for program contexts can be replaced with a surrogate Java block containing atomic programs.

From now on, we only consider the replacement of schema variables for logical variables, terms, formulas and statements, and we also assume that the concerned taclets only contain schema variables of these kinds. Other kinds of schema variables are treated in a similar way in [?].

5.1 Skolem symbols

We define two syntactic domains that consist of symbols for the skolemisation of schema variables:

- Symbols that are placeholders for terms and formulas, and which are similar to ordinary function and predicate symbols

- Symbols that are placeholders for Java statements, similar to the atomic programs of DPDL.

As usual, the elements of both domains are assigned signatures that determine syntactically well-formed expressions. Their shape is described in more detail as follows.

Skolem Symbols for Terms and Formulas

The sets of symbols for terms and formulas are denoted with $\text{Func}_{Sk}$ and $\text{Pred}_{Sk}$ (resp.). The signature

\[
\alpha(s_{Sk}) = \begin{cases} 
(S, S_1, \ldots, S_n, T_1, \ldots, T_k) & \text{for } s_{Sk} \in \text{Func}_{Sk} \\
(S_1, \ldots, S_n, T_1, \ldots, T_k) & \text{for } s_{Sk} \in \text{Pred}_{Sk}
\end{cases}
\]

of a symbol $s_{Sk} \in \text{Func}_{Sk} \cup \text{Pred}_{Sk}$ consists of

- a result sort $S$, if $s_{Sk} \in \text{Func}_{Sk}$,

- a finite sequence $S_1, \ldots, S_n$ of sorts that determines the number and kinds of term arguments; this sequence corresponds to the signature of ordinary predicate symbols,
• a finite sequence $T_1, \ldots, T_k$ of Java types, which are the component types of a tuple of program variables with which occurrences of $s_{Sk}$ are equipped.

Accordingly, the inductive definition of well-formed terms and formulas is extended by:

If $s_{Sk} \in \text{Func}_{Sk} \cup \text{Pred}_{Sk}$ is a symbol with the signature $\alpha(s_{Sk})$ as above, $t_1, \ldots, t_n$ are terms of the sorts $S_1, \ldots, S_n$ and $pv_1, \ldots, pv_k \in PVar$ are program variables of the types $T_1, \ldots, T_k$, then

$$s_{Sk}(t_1, \ldots, t_n; pv_1, \ldots, pv_k)$$

is a term of sort $S$ or a formula (resp.).

### Skolem Symbols for Statements

The set of skolem symbols used for statements is denoted with $\text{Statement}_{Sk}$. The signature $\alpha(st_{Sk}) = (T_1, \ldots, T_k, m)$ of a symbol $st_{Sk} \in \text{Statement}_{Sk}$ consists of

• a finite sequence $T_1, \ldots, T_k$ of Java types (analogously to the symbols for terms or formulas),
• a natural number $m$ that gives the size of the jump table; this is a tuple of Java statements that are arguments of occurrences of $st_{Sk}$ within programs.

The symbols $\text{Statement}_{Sk}$ extend the definition of well-formed Java programs, i.e. the following (informal) rule is added to the Java grammar [?]:

If $st_{Sk} \in \text{Statement}_{Sk}$ is a symbol with $\alpha(st_{Sk}) = (T_1, \ldots, T_k, m)$, $pv_1, \ldots, pv_k$ are program variables of the types $T_1, \ldots, T_k$ and $jst_1, \ldots, jst_m$ are Java statements of the following kinds:\(^{10}\)

• return-statements, with or without an argument (a plain program variable)
• break- and continue-statements, with or without a label
• throw-statements whose argument is a program variable then

$$st_{Sk}(pv_1, \ldots, pv_k; jst_1; \ldots; jst_m)$$

is a statement.

### 5.2 From Meaning Formula to Proof Obligation

From now on we suppose that a taclet $tac$ with meaning formula $M(tac)$ is fixed. Let $SV_{tac}$ be the set of schema variables $M(tac)$ contains. We define an instantiation $\iota_{Sk}$ over $SV_{tac}$ that replaces each schema variable either with a $\text{JavaCardDL}$ variable or with an appropriate skolem expression. The definition refers to the properties of schema variables as introduced in Sect. 2:

^{10}Which are exactly the reasons that can lead to an abrupt termination of a statement, see [?].
• If \( x \in SV_{tac} \) is of type \( \text{Variable} \), then \( \iota_{Sk}(x) \in LVar \) is a new logical variable that has the same sort as \( x \).

• If \( sv \in SV_{tac} \) is of type \( \text{Term, Formula or Statement} \), then let \( \{pv_1, \ldots, pv_k\} = \Pi_{pr}(sv) \) be the program variables that can occur undeclared in instantiations of \( sv \). Let \( T_1, \ldots, T_k \) be the Java types of \( pv_1, \ldots, pv_k \).

• If \( sv \in SV_{tac} \) is of type \( \text{Term} \), then

\[
\iota_{Sk}(sv) = f_{Sk}(v_1, \ldots, v_l; pv_1, \ldots, pv_k)
\]

is a term, where
- \( v_1, \ldots, v_l \) with \( v_i = \iota_{Sk}(x_i) \) are the instantiations of \( x_1, \ldots, x_l \in SV_{tac} \), which are distinct \( \text{Variable} \) schema variables determined by the prefix \( \Pi_{l}(sv) = \{x_1, \ldots, x_l\} \) of \( sv \) in \( tac \)
- and \( f_{Sk} \in \text{Func}_{Sk} \) denotes a new skolem symbol with signature

\[
\alpha(f_{Sk}) = (S, S_1, \ldots, S_l, T_1, \ldots, T_k)
\]

where \( S \) is the sort of \( sv \) and \( S_1, \ldots, S_l \) are the sorts of \( v_1, \ldots, v_l \).

• Analogously, if \( sv \in SV_{tac} \) is a schema variable of type \( \text{Formula} \), then

\[
\iota_{Sk}(sv) = p_{Sk}(v_1, \ldots, v_l; pv_1, \ldots, pv_k)
\]

is a formula containing a new skolem symbol \( p_{Sk} \in \text{Pred}_{Sk} \) for formulas.

• If \( sv \in SV_{tac} \) is a schema variable of type \( \text{Statement} \), then two additional (and new) program variables are needed: \( t_{sv} \) of Java type \( \text{Throwable} \), and \( d_{sv} \) of Java type \( \text{int} \) (the latter variable is used in Sect. 5.3). Let \( \{jst_1, \ldots, jst_m\} = \Pi_{jmp}(sv) \) be jump statements that can occur uncaught in instantiations of \( sv \). The instantiation \( \iota_{Sk}(sv) \) of \( sv \) is the statement \( 11 \)

\[
\iota_{Sk}(sv) = st_{Sk}(pv_1, \ldots, pv_k, t_{sv}, d_{sv}; jst_1; \ldots; jst_m; \text{throw} t_{sv})
\]

where \( st_{Sk} \) denotes a new skolem symbol for statements with signature \( \alpha(st_{Sk}) = (T_1, \ldots, T_k, m + 1) \).

Finally, the \textit{taclet proof obligation} of \( tac \) is defined to be the formula

\[
M_{Sk}(tac) := \iota_{Sk}(M(tac))
\]

**Example 5.1 (Example 4.3 continued)** From the meaning formulas of the taclets \( tac_1, tac_2 \) and \( tac_3 \) the following proof obligations are derived:

\footnote{We always add a \textit{throw}-statement, as instantiations of \( sv \) may always terminate abruptly through an exception regardless of \( \Pi_{jmp}(sv) \).}
\[ M_{Sk}(tac_1) = (t_{Sk} \oplus 0 \land t_{Sk} \oplus 0) \lor (t_{Sk} \oplus t_{Sk} \land \neg (t_{Sk} \oplus 0)) \] (18)

\[ M_{Sk}(tac_2) = \]

\[ (\text{sta}_{Sk}(v, \text{t}_{\#sta}; \text{throw } t_{\#sta}); v++); p_{Sk}(v) \leftrightarrow \]

\[ (\text{sta}_{Sk}(v, \text{t}_{\#sta}; \text{throw } t_{\#sta}); v=v+1); p_{Sk}(v) \] (19)

\[ M_{Sk}(tac_3) = \]

\[ (((1: \text{if } (x==0) \beta_1 \text{ else } \beta_2) p_{Sk}(x) \leftrightarrow (1: \beta_1) p_{Sk}(x) \land x \neq 0) \lor \]

\[ (((1: \text{if } (x==0) \beta_1 \text{ else } \beta_2) p_{Sk}(x) \leftrightarrow (1: \beta_2) p_{Sk}(x) \land \neg (x = 0)) \] (20)

where we use the abbreviations

\[ \beta_1 = \text{sta}_{1Sk}(x, t_{\#sta_1}, d_{\#sta_1}; \text{break l}; \text{throw } t_{\#sta_1}); \]

\[ \beta_2 = \text{sta}_{2Sk}(x, t_{\#sta_2}, d_{\#sta_2}; \text{break l}; \text{throw } t_{\#sta_2}); \]

5.3 Decomposition Rules

Calculus rules for JavaCardDL programs always modify the leading statements within a program block (see Sect. 1). Unfortunately, the addition of skolem symbols for statements would destroy the (relative) completeness of a set of rules: If a skolem symbol turns up as the first active statement of a program block, no JavaCardDL rule will be applicable to that block.

As we have stated in Sect. 1.1 that a “naive” decomposition rule for JavaCardDL cannot be posed due to abrupt termination, we define a family of decomposition rules specifically for statement skolem symbols. These rules cope with abrupt termination by applying a transformation to the statement \( \alpha = \text{sta}_{Sk}(\ldots) \). This transformation splits \( \alpha \) in two parts \( \alpha_1 = \text{sta}'_{Sk}(\ldots) \) and \( \alpha_2 \), such that the concatenation \( \alpha_1; \alpha_2 \) is equivalent to the original statement \( \alpha \). Furthermore, the first program fragment \( \alpha_1 \) is constructed in a way that prevents abrupt termination, and thus, the equivalence

\[ \langle \ldots \text{sta}_{Sk}(\ldots); \beta \ldots \rangle \phi \leftrightarrow \langle \text{sta}'_{Sk}(\ldots) \rangle \langle \ldots \alpha_2; \beta \ldots \rangle \phi \] (23)

holds. The remaining statement \( \alpha_2 \) does no longer contain any skolem symbols, i.e. it is a pure JavaCard program, and hence it is possible to handle \( \alpha_2 \) by the application of regular JavaCardDL rules.

We assume that for each statement skolem symbol \( \text{sta}_{Sk} \in \text{Statement}_{Sk} \) that occurs within \( t_{Sk} \) a second new skolem symbol \( \text{Dec}(\text{sta}_{Sk}) \) is introduced, which has the same signature as \( \text{sta}_{Sk} \) except for the jump table:

\[ \alpha(\text{sta}_{Sk}) = (T_1, \ldots, T_k, m) \implies \alpha(\text{Dec}(\text{sta}_{Sk})) = (T_1, \ldots, T_k, 0). \]

Following equivalence (23), two decomposition taclets \( D^o_{\text{sta}_{Sk}} \) and \( D^{\Box}_{\text{sta}_{Sk}} \) for diamond and box modalities (resp.) are introduced for each statement skolem
symbol $st_{Sk}$ that occurs in $\iota_{Sk}$. We only give the definition of $D^c_{st_{Sk}}$, as the taclet for boxes is obtained analogously:

$$D^c_{st_{Sk}}: \{ \text{find} \left( \langle \ldots st_{Sk}(p_1,\ldots,p_k;\#jst_1;\ldots;\#jst_m); \ldots \rangle \phi \right) \}
\text{replacewith} \left( \langle \text{Dec}(st_{Sk})(p_1,\ldots,p_k); \rangle \langle \ldots ic \ldots \rangle \phi \right) \}$$

where $p_1,\ldots,p_k$ are schema variables for program variables, $\#jst_1,\ldots,\#jst_m$ are variables for statements corresponding to the signature $\alpha(st_{Sk})$ and $\phi$ is a schema variable for formulas. Furthermore the taclet contains an if-cascade $ic$, which is denoted by $\alpha_2$ in equivalence (23):

$$\{ \text{if} \ (p_k == 1) \ #jst_1
\text{else if} \ (p_k == 2) \ \ldots
\text{else if} \ (p_k == m) \ #jst_m \}$$

In this statement at most one of the jump statements represented by the schema variables $\#jst_1,\ldots,\#jst_m$ is selected and executed, depending on the value of the last program variable argument $p_k$ (note that the type of $p_k$ is $\text{int}$ by the definitions of the last section).

**Example 5.2** An application of the decomposition rule for diamond modalities could look as follows:

$$\vdash \langle st'_{Sk}(t,d)\rangle \langle \text{try} \ \{ \text{if} \ (d == 1) \ \text{throw} \ t; \ \} \ \text{catch} \ (\text{Exception} \ e) \ \{\ldots\}\rangle \phi$$

$$\vdash \langle \text{try} \ \{ \ st_{Sk}(t,d;\ \text{throw} \ t); \ \} \ \text{catch} \ (\text{Exception} \ e) \ \{\ldots\}\rangle \phi$$

### 6 Main Result

To show that $tac$ is derivable, which is by Sect. 4 equivalent to the derivability of all instances of $M(tac)$, we assume that there is a closed proof $H$ of $M_{Sk}(tac)$ using the sequent calculus for $\text{JavaCardDL}$ (extended by the skolem symbols and the decomposition taclets of Sect. 5). It is possible to transform $H$ into a proof $H_\phi$ for each instance $\phi$ of $M(tac)$:

**Theorem 6.1 (Main Result)** Suppose that a proof $H$ of $M_{Sk}(tac)$ exists. Then for each instance $\phi = \kappa(M(tac))$ of the meaning formula $M(tac)$ there is a proof $H_\phi$.

In the following we will sketch a proof of Theorem 6.1. Due to lack of space we skip most of the details of the proof; a more detailed account can be found in [?].

The proof obligation $M_{Sk}(tac) = \iota_{Sk}(M(tac))$ differs from other instances $\phi = \kappa(M(tac))$ of the meaning formula in the instantiation of schema variables for terms, formulas and statements: In $M_{Sk}(tac)$ such variables are replaced with skolem symbols as introduced in Sect. 5.1.\footnote{Schema variables for logical variables are in both cases simply instantiated with logical variables.} Hence it is possible
to obtain a “proof” $H'$ of $\phi$ by replacing each occurrence of a skolem symbol $s_{Sk}(\ldots) = \iota_{Sk}(sv)$ in $H$ with the instantiation $\kappa(sv)$ from $\phi$. In general, the tree $H'$ cannot be expected to be a proof, as it is possible that the replacement of skolem symbols leads to invalid rule applications. But by a slightly more complex transformation, as sketched below, it is possible to obtain a legal proof:

For the replacement of skolem symbols we define an appropriate kind of substitutions: We assume that a mapping $\sigma$ of the skolem symbols

$$Sym_{Sk} := Func_{Sk} \cup Pred_{Sk} \cup Stat_{Sk}$$

to terms, formulas and Java statements “with holes” is given. Namely, we allow that for a symbol $s_{Sk}$ with signature $\alpha(s_{Sk}) = ([S,s_1,\ldots,s_n,T_1,\ldots,T_k]$ (or $\alpha(s_{Sk}) = (T_1,\ldots,T_k,m)$ for statement symbols), the value $\sigma(s_{Sk})$ contains a number of holes $\circ_i$ labelled with natural numbers $i \in \{1,\ldots,n+k\}$ (or $i \in \{1,\ldots,k+m\}$, resp.).

**Example 6.2** For a predicate skolem symbol $p_{Sk} \in Pred_{Sk}$, an example of a substitution is given by the following mapping:

$$\sigma(p_{Sk}) = r(\circ_2,a) \land q(\circ_1) \land (\circ_2=1;\phi) \quad \text{for } p_{Sk} \in Pred_{Sk}, \quad \alpha(p_{Sk}) = (S,\text{int}).$$

The mapping $\sigma$ is continued to terms, formulas, Java programs, sequents, proof trees and taclets as a morphism, and by the replacement of skolem symbols. Holes are replaced with the arguments of occurrences of skolem symbols:

$$\sigma(s_{Sk}(r_1,\ldots,r_i)) := \{\circ_1/r_1,\ldots,\circ_l/r_l\}(\sigma(s_{Sk}))$$

**Example 6.3 (Example 6.2 continued)** The mapping $\sigma$ is applied in the following way to a formula containing the symbol $p_{Sk}$:

$$\sigma(\forall x.p_{Sk}(x;i)) = \forall x.(r(i,a) \land q(x) \land (i=1;\sigma(\phi))$$

### 6.1 Treatment of Taclets

The most important observation to prove Theorem 6.1 is the following lemma:

**Lemma 6.4 (Lifting of Taclet Applications)** Suppose that $R_{\text{tac'}}$ is a rule schema that is described by a taclet $\text{tac'}$, and that $\text{tac'}$ does not contain skolem symbols (as introduced in Sect. 5.1). If an instance of $R_{\text{tac'}}$ is given by

$$\frac{P_1 \ldots P_n}{Q}$$

$\text{Extensive considerations about possible collisions are omitted in this document; see [?] for details.}$

19
and \( \sigma \) is a substitution of skolem symbols, then there is a proof tree with root sequent \( \sigma(Q) \), whose open goals are exactly the sequents \( \sigma(P_1), \ldots, \sigma(P_n) \).

**Proof.** First suppose that the considered rule application is not the application of a rewrite taclet within an argument of a skolem symbol occurrence. Then it can be shown that

\[
\frac{\sigma(P_1) \cdots \sigma(P_n)}{\sigma(Q)}
\]

is an instance of \( R_{tac'} \).

Otherwise, if a rewrite taclet is applied to a term \( t \) within an argument of a skolem symbol occurrence, it is possible that a single occurrence of \( t \) in \( Q \) produces more than one occurrence of \( \sigma(t) \) in \( \sigma(Q) \) (like in example 6.3, where a single occurrence of the program variable \( i \) in the original formula yields multiple occurrences after the application of \( \sigma \)). Provided that the cut-rule and rules treating equations are available, it is then possible to perform a cut with the equation \( \sigma(t) = \sigma(t) \) and apply \( tac' \) to one side of the equation. Afterwards the equations \( \sigma(t) = \sigma(t) \) can be used to replace all occurrences of \( \sigma(t) \) successively. This is illustrated by the following proof tree fragment, in which we use the notation \( (\Gamma \vdash \Delta) = \sigma(Q) \):

\[
\frac{\sigma(P_1) \cdots \sigma(P_n)}{\Gamma, \sigma(t) \vdash \sigma(t) \vdash \Delta_{\sigma(t)}} \quad \text{tac}'
\]

\[
\frac{\Gamma \vdash \Delta, \sigma(t) = \sigma(t)}{\Gamma \vdash \Delta}
\]

**Corollary 6.5** Suppose that the proof \( H \) of \( M_{Sk}(tac) = \iota_{Sk}(M(tac)) \) only consists of applications of taclets \( tac' \), and that the concerned taclets \( tac' \) do not contain skolem symbols. Then for each instance \( \phi = \kappa(M(tac)) \) of the meaning formula \( M(tac) \) there is a proof \( H_\phi \).

**Proof.** W.l.o.g. we may assume that \( \iota_{Sk} \) and \( \kappa \) are equal w.r.t. the instantiations of schema variables of type Variable. Each taclet application within \( H \) can then be replaced with the proof tree fragment that is obtained from Lem. 6.4, for a \( \sigma \) that substitutes skolem expressions \( s_{Sk}(\ldots) = \iota_{Sk}(sv) \) with the concrete instantiation \( \kappa(sv) \), i.e. in a way such that \( \sigma(M_{Sk}(tac)) = \phi \).

### 6.2 Treatment of Decomposition Rules

Lem. 6.4 of the last section is not directly applicable to applications of the taclets \( D_{sk}^\circ, D_{sk}^\Box \) (Sect. 5.3), as these taclets contain statement skolem symbols \( s_{Sk} \) and \( Dec(s_{Sk}) \). If these symbols are replaced with arbitrary Java statements by the application of a substitution \( \sigma \) (as introduced in the previous section), then the obtained taclet will furthermore be unsound in general.
We circumvent these problems by constructing particular substitutions $\sigma$ of the symbols $s_{sk}$ and Dec($s_{sk}$) with the property that $\sigma(D^\circ_{sk})$, $\sigma(D^\Box_{sk})$ are sound taclets, so that subsequently Lem. 6.4 can be applied for obtaining a proof tree.

**Lemma 6.6** Suppose that $\sigma$ is a substitution that replaces all skolem symbols of a formula $\psi$, and $s_{sk}$ is a skolem symbol for statements. Then there is a substitution $\sigma'$ that differs from $\sigma$ only in the symbols $s_{sk}$, Dec($s_{sk}$), such that

(i) $\sigma'(D^\circ_{sk})$, $\sigma'(D^\Box_{sk})$ are sound taclets

(ii) There is a proof tree (fragment) whose root is $\vdash \sigma(\psi)$, such that the only goal left is $\vdash \sigma'(\psi)$.

Referring to this lemma it is possible to formulate an analogue of Lem. 6.4 for decomposition taclets:

**Lemma 6.7 (Lifting of Decompositions)** Suppose that $R_D$ is a rule that is described by a decomposition taclet $D$ ($D = D^\circ_{sk}$ or $D = D^\Box_{sk}$). If an instance of $R_D$ is given by $P \quad Q$ and $\sigma'$ is a substitution of skolem symbols as in Lem. 6.6 w.r.t. $D$, then there is a proof tree of $\sigma'(Q)$, whose only goal left is the sequent $\vdash \sigma'(P)$.

**Proof.** First the application of $D$ is replaced with an application of the taclet $\sigma'(D)$, which is sound by Lem. 6.6, (i) (this substitutes certain occurrences of $s_{sk}$, Dec($s_{sk}$) within $P$ and $Q$). Subsequently Lem. 6.4 can be applied to the resulting rule application w.r.t. $\sigma'$.

**Corollary 6.8** Suppose that the proof $H$ of $M_{sk}(tac) = \iota_{sk}(M(tac))$ only consists of applications of taclets tac' that do not contain skolem symbols, and of applications of decomposition taclets. Then for each instance $\phi = \kappa(M(tac))$ of the meaning formula $M(tac)$ there is a proof $H_\phi$.

**Proof.** $\sigma$ is chosen as in the proof of Cor. 6.5. By repeated application of Lem. 6.6, (ii) it is possible to construct a proof tree with root sequent $\vdash \phi$ and a single goal $\vdash \sigma'(M_{sk}(tac))$, with a substitution $\sigma'$ that is chosen according to Lem. 6.6 for each skolem symbol $s_{sk}$ for statements.

It is then possible to construct a closed proof tree of $\vdash \sigma'(M_{sk}(tac))$ by transforming $H$: Each taclet application within $H$ is replaced with the proof tree fragment that is obtained from Lem. 6.4 or Lem. 6.7 (according to the kind of the taclet).

**7 Conclusions**

In this paper, we have outlined how to ensure correctness of derived taclets. Because of limited space, we have only sketched the basic idea and covered only some few kinds of schema variables. The presented concept is completely
integrated in the taclet-based KeY prover. Even a greater class of possible *JavaCardDL* taclets is supported.

As future work, it remains

- to generalise the concept of skolemisation of meaning formulas,
- to study quantified first-order logics with skolemised statements as ‘atomic’ programs, and
- to explore further areas of application, as for example, proofs of program transformation properties.

Taclets are a simple but powerful concept. By their syntactic and semantic simplicity, users are enabled to write new rules and add them to the system easily. We have shown that, despite this fact, the correctness of the rule base can be efficiently ensured—even for a special purpose logic like *JavaCardDL*.

**Acknowledgements**

We would like to thank Martin Giese and Steffen Schlager for useful comments on earlier versions of this paper, as well as Bernhard Beckert and P.H. Schmitt for fruitful discussions. Also we want to thank the anonymous referees and workshop organisers.