

Behavioral Program Logic^{*}

Eduard Kamburjan

Department of Computer Science, Technische Universität Darmstadt, Germany
kamburjan@cs.tu-darmstadt.de

Abstract. We present Behavioral Program Logic (BPL), a dynamic logic for trace properties that incorporates concepts from behavioral types and allows reasoning about non-functional properties within a sequent calculus. BPL uses *behavioral modalities* $[s \Vdash \tau]$, to verify statements s against *behavioral specifications* τ . Behavioral specifications generalize postconditions and behavioral types. They can be used to specify other static analyses, e.g., data flow analyses. This enables deductive reasoning about the results of multiple analyses on the same program, potentially implemented in different formalisms. Our calculus for BPL verifies the behavioral specification gradually, as common for behavioral types. This vastly simplifies specification, calculus and composition of local results. We present a sequent calculus for object-oriented actors with futures that integrates a pointer analysis and bridges the gap between behavioral types and deductive verification.

1 Introduction

When reasoning about concurrent programs, the intermediate states of an execution are of more relevance than when reasoning about sequential programs. In an object-oriented setting, it does not suffice to specify pre- and postcondition of some method m . Instead, the *traces* generated by m must be specified.

Recently, dynamic logics for trace properties have been developed [3, 7, 12] to leverage well-established verification techniques from dynamic logic [1] to a concurrent setting. The application of these approaches to real world models of distributed systems [13, 25] revealed two shortcomings: (1) the composition of method-local verification results to a guarantee for the whole system is not automatic and (2) the specification of trace properties is too complex. Thus, the current approaches are deemed as not practical for serious verification efforts.

Another group of verification techniques, *behavioral types*, aim “*to describe properties associated with the behavior of programs and in this way also describe how a computation proceeds.*” [20]. For object-oriented languages, behavioral types can also be seen as specifications of traces of methods. Behavioral types, especially session types [19], are restricted in their expressive power to easily compose their local results to global guarantees, and are natural specifications for protocols. However, they lack precision when handling state [5] or require

^{*} This work is supported by the **FormbaR** project, part of AG Signalling/DB RailLab.

additional static analyses [23]. For Active Objects [9] (object-oriented actors with futures), a translation from session types to a trace logic has been given [23].

We introduce *Behavioral Program Logic* (BPL) to combine precise state reasoning from program logics with the relative simplicity of behavioral types and enable the integration of static analyses into deductive reasoning. The main difference to previous approaches in dynamic logic for trace properties is the *behavioral modality* $[s \Vdash \tau]$, which expresses that all traces of statement s satisfy specification τ . The specification τ is not a formula, as the postcondition of modalities in classical dynamic logic, but is a specification translated into a monadic second order formula over traces. Similarly to behavioral types, τ may contain syntactic elements and allows to syntactically match with s . Sequent calculi for BPL may reduce s and τ in one rule. Contrary to previous dynamic logics for traces, behavioral specifications are more succinct and easier to compose and decompose by, e.g., using the projection mechanism of session types.

We distinguish between behavioral types, that have a sequent calculus of the above kind, and *behavioral specifications*, which do not. Behavioral specifications interface with external properties, such as a data-flow points-to analysis. Beyond integrating external analyses into the sequent calculus, this modularizes the sequent calculus by expressing different properties with different behavioral specifications. Behavioral specifications are clear interfaces that allow to close proofs once more context is known and generalize proof repositories [6].

Our main contributions are (1) BPL, a trace program logic that integrates deductive reasoning with static analyses (2) *method types*, a behavioral type in BPL that generalizes method contracts, object invariants and local types for Active Objects. Due to space constraints, we do not give (de-)compositions and full semantics and refer to [23] and our technical report [22] for full details. We introduce our programming language in Sec. 2 and BPL in Sec. 3. In Sec. 4 we introduce method types. Sec. 5 summarizes previous approaches and concludes.

2 Preliminaries: An Actor Language with Futures

We introduce Behavioral Program Logic using a Core Active Object language [9] (*CAO*) with futures, *CAO* uses strong encapsulation (i.e., all fields are object-private) and cooperative scheduling. *CAO* is based on ABS [21] and we use a locally abstract, globally concrete (LAGC) semantics [11]. An LAGC semantics consists of two layers: A locally abstract (LA) layer for statements and methods, and a globally concrete (GC) layer for objects and systems. The LA layer is a denotational semantics that abstractly describes the behavior of a method in every possible context, while the GC layer is an operational semantics that concretizes the LA semantics of processes in a concrete context. LA semantics enable one to analyze a method in isolation and Active Objects allow us to demonstrate that BPL is suited for complex concurrency models.

Definition 1 (Syntax). Let \sim range over $\&\&, ||, +, -, *, /, >=, >, <, <=, \nu$ over variables, \mathbf{f} over fields, \mathbf{C} over class names, \mathbf{m} over method names, i and \mathbf{n} over \mathbb{N} . The syntax of *CAO* is defined in Fig. 1, where $\vec{\tau}$ denotes (possibly empty) lists.

$$\begin{aligned}
\text{Prgm} &::= \overrightarrow{\text{Class}} \text{ Main} \quad \text{Main} ::= \mathbf{main}\{\text{si}\} \quad \text{Class} ::= \mathbf{class} \text{ C } (\overrightarrow{\text{C f}}) \{\overrightarrow{\text{Field}} \overrightarrow{\text{Meth}}\} \quad \text{Field} ::= \text{D f} = \text{e}; \\
\text{Meth} &::= \text{D m}(\overrightarrow{\text{D v}})\{\text{s}; \mathbf{return} \text{e};\} \quad \text{D} ::= \mathbf{Int} \mid \mathbf{Bool} \mid \mathbf{Fut}\langle \text{D} \rangle \quad \text{si} ::= \text{C v} = \text{C}(\overrightarrow{\text{v}}); \text{si} \mid \text{v}!\text{m}(\overrightarrow{\text{e}}) \\
\text{s} &::= [\text{D}] \text{l} = \text{e} \mid [\text{D}] \text{v} = \text{e}.\mathbf{get}_i \mid [\text{D}] \text{v} = \text{f}!\text{m}(\overrightarrow{\text{e}}) \mid \mathbf{skip} \mid \mathbf{while}(\text{e})\{\text{s}\} \mid \mathbf{if}(\text{e})\{\text{s}\}\mathbf{else}\{\text{s}\} \mid \text{s}; \text{s} \\
\text{e} &::= \text{l} \mid \text{n} \mid \mathbf{True} \mid \mathbf{False} \mid \text{e} \sim \text{e} \mid !\text{e} \mid -\text{e} \quad \text{l} ::= \mathbf{this.f} \mid \text{v}
\end{aligned}$$

Fig. 1. Syntax of CAO.

A program consists of a set of classes and a main block. The main block contains object instantiations and a method call to initialize the communication. All objects are created at once, not in the order of their instantiations. A class contains (1) parameter fields which reference other objects, (2) fields for data, initialized upon creation and (3) methods. Multiple instances can share their parameters. Parameters cannot be reassigned. As data types we use integers, booleans and parametric futures. Each class has a `run` method that is started upon creation. We omit `run` if it is empty.

```

1 class T(Comp S, Log L){
2   Int test(Int i){
3     Fut<Int> f = S!cmp(i);
4     Int r = f.get0;
5     if(r < 0){
6       r = -r;
7       f = L!log(i);
8     }
9     return r;
10  }
11 }

```

Fig. 2. An example method

Example 1. Fig. 2 shows a simple method that passes its input to `Comp.cmp` and reads the result. If the result is negative, its sign is inverted and the original input data is logged by `Log.log`. The possibly inverted result is returned.

The semantics of a method is a set of symbolic traces, to describe the behavior of the method in every possible context, i.e., for every possible heap, call parameters and accessed futures. Additionally to semantic values (semantic values are, e.g., object identifiers, rationals, futures etc.), symbolic traces contain symbolic expressions. Structurally, symbolic expressions mirror syntactic expressions, but do not contain variables or fields. Instead they contain symbolic values and symbolic fields. Symbolic values have no operations defined on them and act as placeholders. They are replaced by semantic values once the method is running and the context is known. Symbolic fields are special symbolic values that contain the name of the field they are abstracting.

Definition 2 (Symbolic Expressions). *Let v range over semantic values, \underline{v} over symbolic values, i over \mathbb{N} and $\mathbf{this.f}_i$ over symbolic fields. Symbolic expres-*

Statement $\text{v} = \text{f}!\text{m}(\overrightarrow{\text{e}})$ calls method `m` asynchronously on the object `f` with parameters $\overrightarrow{\text{e}}$. A fresh future is generated and stored in `v`. This future identifies the called process. We say that the called process will *resolve* the future by executing `return e` and storing the value of `e` in the future. The synchronizing statement $\text{v} = \text{e}.\mathbf{get}_i$ reads from the future in `e` into `v`. Until the future is resolved, the reading process blocks its object. The identifier `i` is used to distinguish multiple synchronization points. The other statements, expressions and methods are standard.

sions \underline{e} are defined below. We highlight symbolic elements by underlining.

$$\underline{e} ::= \underline{e} \sim \underline{e} \mid !\underline{e} \mid -\underline{e} \mid \underline{v} \mid v \mid \underline{\text{this}}.\underline{f}_i$$

To model the points where processes, objects and futures interact, traces contain events as markers for visible communication.

Definition 3 (Events). Events are defined by the following grammar.

$$\text{ev} ::= \text{invEv}(\underline{x}, \underline{x}', \underline{f}, \underline{m}, \overrightarrow{\underline{e}}) \mid \text{invREv}(\underline{x}, \underline{f}, \underline{m}, \overrightarrow{\underline{e}}) \mid \text{futEv}(\underline{x}, \underline{f}, \underline{m}, \underline{e}) \mid \text{futREv}(\underline{x}, \underline{f}, \underline{m}, \underline{e}, i) \mid \text{noEv}$$

Event $\text{invEv}(\underline{x}, \underline{x}', \underline{f}, \underline{m}, \overrightarrow{\underline{e}})$ models a call from \underline{x} to \underline{x}' on method \underline{m} with future \underline{f} and call parameters $\overrightarrow{\underline{e}}$. The future and the callee may be symbolic: locally it is not possible to know the used future and the called object. Event $\text{invREv}(\underline{x}, \underline{f}, \underline{m}, \overrightarrow{\underline{e}})$ is the callee view on a call. The object \underline{x} here is the callee, the caller is not visible to the callee. Event $\text{futEv}(\underline{x}, \underline{f}, \underline{m}, \underline{e})$ models the termination of a process for future \underline{f} , computing method \underline{m} in object \underline{x} and returning \underline{e} . Event $\text{futREv}(\underline{x}, \underline{f}, \underline{m}, \underline{e}, i)$ models a **get**_{*i*} statement in object \underline{x} on the future \underline{f} , which was computed by \underline{m} and returned \underline{e} . Finally, **noEv** models an internal step.

Local traces consist of a *selection condition*, a set of symbolic expressions that express when a trace executes and a *history*, a sequence of events and states.

Definition 4 (Local Semantics and Traces). A heap ρ maps from fields to symbolic expressions and a local state σ maps variables to symbolic expressions. Pairs of local states and heaps are object states and we write (σ) . The evaluation function $\llbracket \underline{e} \rrbracket_{(\sigma)}$ maps a syntactic expression to a symbolic expression.

A local trace θ has the form $\text{sc} \triangleright \text{hs}$, where sc is a set of symbolic expressions, called selection condition, and hs is a non-empty sequence, called history, such that every odd-indexed element is an object state and every even-indexed element an event. The semantics of methods and statements is defined by a function $\llbracket \cdot \rrbracket_{\underline{x}, \underline{f}, \underline{m}, (\sigma)}$ where \underline{x} is the object name, \underline{f} the future the method is resolving, \underline{m} the method name and (σ) the current object state. Future, object name and state may be symbolic. The semantics of a method \underline{m} with body \underline{s} is, for a symbolic (σ) :

$$\llbracket \underline{m} \rrbracket_{\underline{x}, \underline{f}, \underline{m}, (\sigma)} = \left\{ \emptyset \triangleright \left\langle \left(\sigma \right), \text{invREv}(\underline{x}, \underline{f}, \underline{m}, \overrightarrow{\underline{e}}) \right\rangle \circ \theta \mid \theta \in \llbracket \underline{s} \rrbracket_{\underline{x}, \underline{f}, \underline{m}, (\sigma)} \right\}$$

where $\overrightarrow{\underline{e}}$ is extracted from the parameter names in σ . E.g., for a method **Int** $\underline{m}(\text{Int } \underline{a}, \text{Rat } \underline{b})$ we set $\overrightarrow{\underline{e}} = (\sigma(\underline{a}), \sigma(\underline{b}))$. Fig. 3 shows selected rules. All variables are initialized, futures with **no**, a special future that is never resolved.

The rules for assignment both update the local state or the heap, and add a **noEv** event. The rule for branching evaluates both branches and adds the corresponding guard evaluation to the selection condition. The rule for **get** is similar to the variable assignment, but receives a fresh symbolic value and stores it in the local state. As the event, a resolving reaction event is added, which stores the accessed future and the fresh symbolic value. The rule for method

$$\begin{aligned}
\llbracket v = e \rrbracket_{X,f,m,(\rho)} &= \left\{ \emptyset \triangleright \left\langle \left(\frac{\sigma}{\rho} \right), \text{noEv}, \left(\sigma[v \mapsto \llbracket e \rrbracket_{(\rho)}] \right) \right\rangle \right\} & \llbracket \text{this.f} \rrbracket_{(\rho)} &= \rho(\mathbf{f}) \\
& & \llbracket v \rrbracket_{(\rho)} &= \sigma(v) \\
\llbracket \text{this.f} = e \rrbracket_{X,f,m,(\rho)} &= \left\{ \emptyset \triangleright \left\langle \left(\frac{\sigma}{\rho} \right), \text{noEv}, \left(\rho[\mathbf{f} \mapsto \llbracket e \rrbracket_{(\rho)}] \right) \right\rangle \right\} & \llbracket \text{!}e \rrbracket_{(\rho)} &= \begin{cases} - \llbracket e \rrbracket_{(\rho)} & \text{if } \llbracket e \rrbracket_{(\rho)} \text{ symbolic} \\ - \llbracket \llbracket e \rrbracket_{(\rho)} \rrbracket & \text{otherwise} \end{cases} \\
\llbracket \text{return } e \rrbracket_{X,f,m,(\rho)} &= \left\{ \emptyset \triangleright \left\langle \left(\frac{\sigma}{\rho} \right), \text{futEv}(X, f, m, \llbracket e \rrbracket_{(\rho)}), \left(\frac{\sigma}{\rho} \right) \right\rangle \right\} & \llbracket \text{skip} \rrbracket_{X,f,m,(\rho)} &= \left\{ \emptyset \triangleright \left\langle \left(\frac{\sigma}{\rho} \right) \right\rangle \right\} \\
\llbracket \text{if}(e) \{s\} \text{else} \{s'\} \rrbracket_{X,f,m,(\rho)} &= \left\{ \text{sc} \cup \{ \llbracket e \rrbracket_{(\rho)} \} \triangleright \text{hs} \mid \text{sc} \triangleright \text{hs} \in \llbracket s \rrbracket_{X,f,m,(\rho)} \right\} \\
& \quad \cup \left\{ \text{sc} \cup \{ \llbracket \text{!}e \rrbracket_{(\rho)} \} \triangleright \text{hs} \mid \text{sc} \triangleright \text{hs} \in \llbracket s' \rrbracket_{X,f,m,(\rho)} \right\} \\
\llbracket v = e.\text{get}_i \rrbracket_{X,f,m,(\rho)} &= \left\{ \emptyset \triangleright \left\langle \left(\frac{\sigma}{\rho} \right), \text{futREv}(X, \llbracket e \rrbracket_{(\rho)}, \underline{m}, v, i), \left(\sigma[v \mapsto \underline{v}] \right) \right\rangle \right\} \text{ where } \underline{v} \text{ is fresh} \\
\llbracket v = \mathbf{f}!\mathbf{n}(\vec{e}) \rrbracket_{X,f,m,(\rho)} &= \left\{ \emptyset \triangleright \left\langle \left(\frac{\sigma}{\rho} \right), \text{invEv}(X, \llbracket \mathbf{f} \rrbracket_{(\rho)}, \underline{v}, \mathbf{n}, \overline{\llbracket e \rrbracket_{(\rho)}}, \left(\sigma[v \mapsto \underline{v}] \right) \right\rangle \right\} \text{ where } \underline{v} \text{ is fresh} \\
\llbracket \text{while}(e) \{s\} \rrbracket_{X,f,m,(\rho)} &= \llbracket \text{if}(e) \{s; \text{while}(e) \{s\}; \text{skip}\} \text{else} \{\text{skip}\} \rrbracket_{X,f,m,(\rho)} \\
\llbracket s; s' \rrbracket_{X,f,m,(\rho)} &= \left\{ \text{sc} \cup \text{sc}' \triangleright \text{hs} \circ \text{hs}' \mid \text{sc} \triangleright \text{hs} \in \llbracket s \rrbracket_{X,f,m,(\rho)}, \text{sc}' \triangleright \text{hs}' \in \llbracket s' \rrbracket_{X,f,m,(\rho)} \right\}
\end{aligned}$$

Fig. 3. Selected rules of the LA semantics of statements and expression. Evaluation $\llbracket e \rrbracket$ of semantic values has its natural definition.

$$\begin{array}{l|l}
\{ \underline{v} < 0 \} \triangleright & \\
\langle (\sigma \cup \{ \mathbf{f} \mapsto \mathbf{no}, \mathbf{r} \mapsto 0 \}, \rho), \text{invREv}(\underline{X}, \underline{f}, \mathbf{T}, \text{test}, \langle \underline{i} \rangle) \rangle \circ & \{ \underline{v} \geq 0 \} \triangleright \\
\langle (\sigma \cup \{ \mathbf{f} \mapsto \mathbf{no}, \mathbf{r} \mapsto 0 \}, \rho), \text{invEv}(\underline{X}, \underline{S}, \underline{f}', \text{Comp. cmp}, \langle \underline{i} \rangle) \rangle \circ & \langle (\sigma \cup \{ \mathbf{f} \mapsto \mathbf{no}, \mathbf{r} \mapsto 0 \}, \rho), \text{invREv}(\underline{X}, \underline{f}, \mathbf{T}, \text{test}, \langle \underline{i} \rangle) \rangle \circ \\
\langle (\sigma \cup \{ \mathbf{f} \mapsto \underline{f}', \mathbf{r} \mapsto 0 \}, \rho), \text{futREv}(\underline{X}, \underline{f}', \text{Comp. cmp}, \underline{v}, 0) \rangle \circ & \langle (\sigma \cup \{ \mathbf{f} \mapsto \mathbf{no}, \mathbf{r} \mapsto 0 \}, \rho), \text{invEv}(\underline{X}, \underline{S}, \underline{f}', \text{Comp. cmp}, \langle \underline{i} \rangle) \rangle \circ \\
\langle (\sigma \cup \{ \mathbf{f} \mapsto \underline{f}', \mathbf{r} \mapsto \underline{v} \}, \rho), \text{noEv} \rangle \circ & \langle (\sigma \cup \{ \mathbf{f} \mapsto \underline{f}', \mathbf{r} \mapsto 0 \}, \rho), \text{futREv}(\underline{X}, \underline{f}', \text{Comp. cmp}, \underline{v}, 0) \rangle \circ \\
\langle (\sigma \cup \{ \mathbf{f} \mapsto \underline{f}', \mathbf{r} \mapsto -\underline{v} \}, \rho), \text{invEv}(\underline{X}, \underline{L}, \underline{f}'', \text{Log. log}, \langle \underline{i} \rangle) \rangle \circ & \langle (\sigma \cup \{ \mathbf{f} \mapsto \underline{f}', \mathbf{r} \mapsto \underline{v} \}, \rho), \text{futEv}(\underline{X}, \underline{f}, \mathbf{T}, \text{test}, \underline{v}) \rangle \circ \\
\langle (\sigma \cup \{ \mathbf{f} \mapsto \underline{f}'', \mathbf{r} \mapsto -\underline{v} \}, \rho), \text{futEv}(\underline{X}, \underline{f}, \mathbf{T}, \text{test}, -\underline{v}) \rangle \circ & \langle (\sigma \cup \{ \mathbf{f} \mapsto \underline{f}', \mathbf{r} \mapsto \underline{v} \}, \rho) \rangle \\
\langle (\sigma \cup \{ \mathbf{f} \mapsto \underline{f}'', \mathbf{r} \mapsto -\underline{v} \}, \rho) \rangle &
\end{array}$$

Fig. 4. LA semantics of $\mathbf{T.test}$, with $\sigma = \{ \mathbf{i} \mapsto \underline{i} \}, \rho = \{ \mathbf{S} \mapsto \underline{S}, \mathbf{L} \mapsto \underline{L} \}$.

calls is analogous, but uses a fresh future for the call instead of a fresh read value. The added event is an invocation event with the evaluated parameters. Fig. 4 shows the two traces in the semantics of Ex. 1.

Symbolic traces represent a set of concrete traces, which contain only semantic values and correspond to possible behaviors of the statement. The set of concrete traces represented by a symbolic trace is a vast overapproximation and we only consider *selected* traces: concrete traces used in some terminating run of a given program. For a formal definition and the GC semantics, we refer to [22].

Definition 5 (Selected Traces). *A trace θ continues trace θ' , written $\theta \preceq \theta'$, if its history is a suffix of the history of θ' , with all symbolic elements replaced by concrete values, such that this substitution evaluates all expressions in the selection condition to true. A trace θ is selected in a program Prgm , if it is used during some run of Prgm . Let \mathbf{m} be the method containing \mathbf{s} .*

$$\llbracket \mathbf{s} \rrbracket_{\mathbf{x},f,m,(\sigma)}^{\text{Prgm}} = \{ \theta \in \llbracket \mathbf{s} \rrbracket_{\mathbf{x},f,m,(\sigma)} \mid \exists \theta' \in \llbracket \mathbf{m} \rrbracket_{\mathbf{x},f,m,(\sigma)} \cdot \theta \preceq \theta' \wedge \theta \text{ used in a terminating run of Prgm} \}$$

We use a first-order state (FOS) logic to express properties of states and a monadic second-order (MSO) logic to express properties of traces. The MSO logic embeds the FOS by using FOS formulas similar to predicates on states. Similarly, it uses terms that allow to specify events.

Definition 6 (FOS Syntax). *Let p range over predicate symbols, f over function symbols, x over logical variable names and S over sorts. As sorts we take all data types \mathbf{D} , all class names and additionally \mathbf{N} and **Heap**. The logical heaps are functions from field names to semantic values. Formulas φ and terms \mathbf{t} are defined by the following grammar, where \mathbf{v} are program variables, consisting of local variables and the special variables **heap** and **result**, and \mathbf{f} are all field names.*

$$\varphi ::= p(\vec{\mathbf{t}}) \mid \mathbf{t} \doteq \mathbf{t} \mid \varphi \vee \varphi \mid \neg \varphi \mid \exists x \in S. \varphi \quad \mathbf{t} ::= x \mid \mathbf{v} \mid \mathbf{f} \mid f(\vec{\mathbf{t}})$$

We demand the usual constants, (e.g., 0, **True**) and that each operator defined in syntactic expressions \mathbf{e} is a function symbol, so one can directly translate a syntactic expression into a FOS term. We additionally assume the following function symbols to handle heaps: **select**(\mathbf{t}, \mathbf{t}) \mid **store**($\mathbf{t}, \mathbf{t}, \mathbf{t}$), where **select**(h, \mathbf{f}) reads field \mathbf{f} from heap h and **store**($h, \mathbf{f}, \mathbf{t}$) stores the value of \mathbf{t} in field \mathbf{f} of heap h . As only one object is considered, we do not require an object parameter.

Definition 7 (FOS Semantics). *Interpretation I maps function names to functions and predicate names to predicates. Assignment β maps logical variable to semantic values of the resp. sort. Evaluation of terms in state (σ) is defined as a function $\llbracket \mathbf{t} \rrbracket_{(\sigma),I,\beta}$ and satisfiability of formulas by a relation $(\sigma), \beta, I \models \varphi$.*

For the special variable **heap** we set $\llbracket \mathbf{heap} \rrbracket_{\mathbf{x},f,m,(\sigma)} = \rho$ and for the heap functions we follow JavaDL [1] and demand, e.g., the following connection axiom for all heaps h , all fields \mathbf{f} and terms \mathbf{t} : $I(\mathbf{select})(I(\mathbf{store})(h, \mathbf{f}, \mathbf{t}), \mathbf{f}) = \mathbf{t}$.

The models for the MSO logic are local traces and the whole semantic domain. This allows to quantify over method names etc. – it is not a logic over finite sequences. In addition to standard MSO constructs, we use $[\mathbf{t}_{\text{tr}}] \doteq \mathbf{t}$ to say that the event at position \mathbf{t}_{tr} of the trace is equal to the term \mathbf{t} . Similarly, $[\mathbf{t}_{\text{tr}}] \vdash \varphi$ expresses that the state at position \mathbf{t}_{tr} is a model for the FOS formula φ .

Definition 8 (MSO Syntax). *Let p, f, x range over the same sets as before, S over sorts. As sorts we take all data types \mathbf{D} and additionally \mathbf{I} , the set of trace indices, \mathbf{O} , the set of all object names, **Fut**, the set of all futures of all types, the supertype **Any**, the set of all well-typed expressions and \mathbf{M} , the set of all method names. Formulas ψ are defined as follows. Terms \mathbf{t}_{tr} are standard.*

$$\psi ::= p(\vec{\mathbf{t}}_{\text{tr}}) \mid \psi \vee \psi \mid \neg \psi \mid \mathbf{t}_{\text{tr}} \subseteq \mathbf{t}_{\text{tr}} \mid \exists x \in S. \psi \mid \exists X \subseteq S. \psi \mid [\mathbf{t}_{\text{tr}}] \doteq \mathbf{t}_{\text{tr}} \mid [\mathbf{t}_{\text{tr}}] \vdash \varphi$$

The predicate **isEvent**(i) that holds iff $\theta[i]$ is an event. For each type of event, there is a function symbol that maps its parameters to an event of its type and a predicate that holds iff the given position is an event of that kind, e.g.,

$$\text{isfutEv}(i) \iff \exists f \in \mathbf{Fut}. \exists o \in \mathbf{O}. \exists m \in \mathbf{M}. \exists v \in \mathbf{Any}. [i] \doteq \text{futEv}(o, f, m, v)$$

Definition 9 (MSO Semantics). *The semantics of terms and event terms is defined by a function $\llbracket \cdot \rrbracket_{I,\beta}$. The satisfiability of MSO-formulas is defined by a relation $\theta, I, \beta \models \psi$. The semantics of our extensions follows.*

$$\begin{aligned} \theta, I, \beta \models [\mathbf{tr}_1] \doteq \mathbf{tr}_2 &\iff 1 \leq \llbracket \mathbf{tr}_1 \rrbracket_{I,\beta} \leq |\theta| \wedge \theta[\llbracket \mathbf{tr}_1 \rrbracket_{I,\beta}] = \llbracket \mathbf{tr}_2 \rrbracket_{I,\beta} \\ \theta, I, \beta \models [\mathbf{tr}] \vdash \varphi &\iff 1 \leq \llbracket \mathbf{tr} \rrbracket_{I,\beta} \leq |\theta| \wedge \theta[\llbracket \mathbf{tr} \rrbracket_{I,\beta}] \text{ is a state} \wedge \theta[\llbracket \mathbf{tr} \rrbracket_{I,\beta}], I, \beta \models \varphi \end{aligned}$$

Example 2. Let $\mathbf{r} = \mathbf{f.get}_0$ be the statement from Ex. 1. The following MSO-formula expresses that if all values read from futures of `cmp` is positive, and every future read at point 0 is from `cmp`, then after the read the value of \mathbf{r} is positive.

$$\begin{aligned} &(\forall i \in \mathbf{I}. (\forall v \in \mathbf{Int}. [i] \doteq \text{futREv}(-, -, \mathbf{cmp}, v, -) \rightarrow v > 0) \wedge \\ &\forall i \in \mathbf{I}. (\forall m \in \mathbf{M}. [i] \doteq \text{futREv}(-, -, m, -, 0) \rightarrow m \doteq \mathbf{cmp})) \\ &\rightarrow \forall i \in \mathbf{I}. ([i] \doteq \text{futREv}(-, -, -, -, 0) \rightarrow [i+1] \vdash \mathbf{r} > 0) \end{aligned}$$

Relativization [17], an established technique in abstract model theory [15], syntactically restricts a formula ψ on a substructure defined by another formula ψ' . It is denoted $\psi[x \in S \setminus \psi']$, where x is a free variable in ψ' of S sort. Each quantifier of S sort is restricted to elements that fulfill ψ' .

Example 3. Formula φ expresses that every trace-element is either an event, or a state with $\mathbf{r} > 0$. The relativization with ψ expresses that φ holds for every index above 9. Both traces of Fig. 2 satisfy $\varphi[j \in \mathbf{I} \setminus \psi]$, neither satisfies φ .

$$\begin{aligned} \varphi &= \forall i \in \mathbf{I}. \text{isEvent}(i) \vee [i] \vdash \mathbf{r} > 0 & \psi &= j \geq 9 \\ \varphi[j \in \mathbf{I} \setminus \psi] &= \forall i \in \mathbf{I}. i \geq 9 \rightarrow (\text{isEvent}(i) \vee [i] \vdash \mathbf{r} > 0) \end{aligned}$$

We use common abbreviations, e.g., $\forall x \in S. \varphi$ for $\neg \exists x \in S. \neg \varphi$ and `true` and shorten comparisons of `Bool` terms by writing, e.g., $i > j$ instead of $i > j \doteq \mathbf{True}$.

3 Behavioral Program Logic

Behavioral Program Logic (BPL) is an extension of FOS with *behavioral modalities* $[s \Vdash^\alpha \tau]$ that contain a statement \mathbf{s} and a behavioral specification (τ, α) . A behavioral specification consists of (1) a syntactic component (the type τ) and (2) a translation α of the type into an MSO formula that has to hold for all traces generated by the statement. Behavioral specifications can be seen as representations of a certain class of MSO formulas, which are deemed useful for verification of distributed systems. For the rest of this section, we assume fixed parameters `Prgm`, \mathbf{x} , f , \mathbf{m} for evaluation.

Definition 10 (Behavioral Program Logic). *A behavioral specification \mathbb{T} is a pair $(\tau_{\mathbb{T}}, \alpha_{\mathbb{T}})$, where $\alpha_{\mathbb{T}}$ maps elements of $\tau_{\mathbb{T}}$ to MSO formulas.*

BPL-formulas φ , terms t and updates U are defined by the following grammar, which extends Def. 6. The meta variables range as in Def. 6. Additionally let \mathbf{s} range over statements and $(\tau_{\mathbb{T}}, \alpha_{\mathbb{T}})$ over behavioral specifications.

$$\varphi ::= \dots \mid [s \Vdash^{\alpha_{\mathbb{T}}} \tau_{\mathbb{T}}] \mid \{U\}\varphi \quad \mathbf{t} ::= \dots \mid \{U\}\mathbf{t} \quad U ::= \epsilon \mid U \parallel U \mid \{U\}U \mid \mathbf{v} := \mathbf{t}$$

$$\begin{aligned}
\llbracket \{U\}t \rrbracket_{(\sigma), I, \beta} &= \llbracket t \rrbracket_{\llbracket U \rrbracket_{(\sigma), I, \beta}, I, \beta} \quad \llbracket \epsilon \rrbracket_{(\sigma), I, \beta}(x) = x \quad \left(\begin{array}{c} \sigma \\ \rho \end{array} \right), I, \beta \models \{U\}\varphi \Leftrightarrow \llbracket U \rrbracket_{(\sigma), I, \beta}, I, \beta \models \varphi \\
\llbracket \mathbf{v} := \mathbf{t} \rrbracket_{(\sigma), I, \beta} \left(\left(\begin{array}{c} \sigma' \\ \rho' \end{array} \right) \right) &= \begin{cases} \left(\begin{array}{c} \sigma'' \\ \rho'' \end{array} \right) & \text{if } \mathbf{v} = \mathbf{heap}, \rho'' = \llbracket \mathbf{t} \rrbracket_{(\sigma), I, \beta} \\ \left(\begin{array}{c} \sigma'' \\ \rho'' \end{array} \right) & \text{otherwise, } \sigma'' = \sigma'[\mathbf{v} \mapsto \llbracket \mathbf{t} \rrbracket_{(\sigma), I, \beta}] \end{cases} \\
\llbracket U \parallel U' \rrbracket_{(\sigma), I, \beta}(x) &= \llbracket U' \rrbracket_{(\sigma), I, \beta}(\llbracket U \rrbracket_{(\sigma), I, \beta}(x)) \quad \llbracket \{U\}U' \rrbracket_{(\sigma), I, \beta} = \llbracket U' \rrbracket_{\llbracket U \rrbracket_{(\sigma), I, \beta}, I, \beta} \\
\left(\begin{array}{c} \sigma \\ \rho \end{array} \right), I, \beta \models [\mathbf{s} \Vdash^{\alpha_{\mathbb{T}}} \tau_{\mathbb{T}}] &\Leftrightarrow \forall \theta \in \llbracket \mathbf{s} \rrbracket_{\mathbf{X}, \mathbf{f}, \mathbf{m}, (\sigma)}^{\text{Prgm}} \cdot \theta, I, \beta \models \alpha_{\mathbb{T}}(\tau_{\mathbb{T}})
\end{aligned}$$

Fig. 5. Semantics of BPL. The satisfiability relation on the right of the semantics of behavioral modalities is the one of MSO.

The semantics of a behavioral modality $[\mathbf{s} \Vdash^{\alpha_{\mathbb{T}}} \tau_{\mathbb{T}}]$ is that all traces generated by \mathbf{s} *selected within* Prgm are models for $\alpha_{\mathbb{T}}(\tau_{\mathbb{T}})$. We use updates $[2, 1]$ to keep track of state changes, their semantics is a state transition. Update $\mathbf{v} := \mathbf{t}$ changes the state by updating \mathbf{v} to \mathbf{t} . The parallel update $U \parallel U'$ applies U and U' in parallel, with U' winning in case of clashes. ϵ is the empty update and application $\{U\}$ evaluates the term (resp. formula) in the state after applying U .

Definition 11 (Semantics of BPL). *The semantical extension of FOS to BPL is given in Fig. 5. The interpretation I has the properties described above. A formula φ is valid if every (σ) and every β make it true.*

Object, program, method name, resolved future and type of **result** are implicitly known, but we omit them for readability's sake. We use a sequent calculus to reason about BPL (resp. FOS).

Definition 12 (Sequents and Rules). *Let Δ, Γ be sets of BPL-formulas. A sequent $\Gamma \Rightarrow \Delta$ has the semantics of $\bigwedge \Gamma \rightarrow \bigvee \Delta$. Γ is called the antecedent and Δ the succedent. Let C, P_i be sequents. A rule has the form*

$$(\text{name}) \frac{P_1 \quad \dots \quad P_n}{C} \text{ cond}$$

Where C is called the conclusion and P_i the premise, while *cond* is a side-condition. Side-conditions are always decidable. For readability's sake, we apply side conditions containing equalities directly in the premises.

Rules may contain, in addition to expressions, schematic variables. Their handling is standard [1]. We assume the usual FO rules for the FOS part of BPL handling all FO operators such as quantifiers.

Definition 13 (Soundness). *A rule is sound if validity of all premisses implies validity of the conclusion.*

Soundness implicitly refers to a program Prgm , as behavioral modalities are defined over Prgm -selectable traces. Rewrite rules $\tau_1 \rightsquigarrow \tau_2$ syntactically replace one type τ_1 by another, τ_2 (and vice versa) and are sound if $\alpha(\tau_1) \equiv \alpha(\tau_2)$.

Discussion. Before we introduce method types, a particular behavioral specification, we illustrate BPL with further examples. To reason about postconditions, as standard modal logics, we define a behavioral specification that only uses the last state of a trace (denoted by the function symbol $last$) for its semantics.

Example 4. The specification for postconditions is the pair of the set of all FOS sentences and the function pst , defined below. \top is the type of `result`. The first case accesses the return value stored in the `futEv` when `result` is used.

$$pst(\varphi) = \begin{cases} \exists v \in \top. [last - 1] \doteq futEv(_, _, _, v) \wedge [last] \vdash \varphi[result \setminus v] & \text{if } \varphi \text{ contains } \mathbf{result} \\ [last] \vdash \varphi & \text{otherwise} \end{cases}$$

A Hoare triple $\{\varphi\} \mathbf{s} \{\psi\}$ has the same semantics as the formula $\varphi \rightarrow [\mathbf{s} \Vdash^{pst} \psi]$. A standard dynamic logic modality $[\mathbf{s}] \psi$ has the same semantics as the behavioral modality $[\mathbf{s} \Vdash^{pst} \psi]$ ¹. Behavioral modalities generalize these systems and can be used to express any (MSO) trace property, independent of the form of its verification system. The following defines a points-to analysis for futures [14] (for the next statement), normally implemented in a data-flow framework.

Example 5 (Points-To). The behavioral specification of a *points-to analysis* specifies that the next statement reads a future resolved by a method from set M .

$\mathbb{T}_{p2} = (\mathcal{P}(M), p2)$ with

$$p2(M) = \exists x \in \mathcal{O}. \exists f \in \mathbf{Fut}. \exists m \in M. \exists v \in \mathbf{Any}. \exists i \in \mathbb{N}. [1] \doteq futREv(x, f, m, v, i) \wedge \bigvee_{m' \in M} m \doteq m'$$

The following formula expresses that the `get` statement reads a positive number, if the future is resolved by `Comp.cmp`. This is the case if `Comp.cmp` always returns positive values. The identifier connects the two modalities semantically.

$$\varphi_p = [r = \mathbf{f}. \mathbf{get}_0 \Vdash^{p2} \{\mathbf{Comp.cmp}\}] \rightarrow [r = \mathbf{f}. \mathbf{get}_0 \Vdash^{pst} r > 0]$$

It is not necessary to include postcondition reasoning. Rule $(\mathbf{ex1})$ in Fig. 6 expresses that if the next read from \mathbf{s} is from some set E' and it is required to show that the next read is from E , it suffices to check whether E is a subset of E' . Rule $(\mathbf{ex}\Vdash)$ connects two analyses and generalizes Ex. 2: one may assume some formula ψ for a read value, if this synchronization always reads from method `Comp.cmp` and that the method body of `Comp.cmp` establishes ψ .

The above example illustrates the difference between modalities and typing judgments. Modalities are formulas and can be used for deductive reasoning about a type judgment (which, in our case, is encoded into \Vdash). While a calculus for pst is easily carried over from other sequent calculi, this is not possible for all behavioral specifications. The proof can still be closed in two ways.

- There may be some rules, such as $(\mathbf{ex1})$ above, that enable to reason about the analysis without reducing the statement at all.

¹ This justifies our use of the term “modality”. Contrary to standard modalities, behavioral modalities are not formulas that express modal statements about *formulas*, but formulas that express a modal statement about *more general specifications*.

$$\begin{array}{c}
\text{(ex1)} \frac{\Gamma, \psi(v) \Rightarrow \{v := v\} \{s \Vdash^{\text{pst}} \varphi\}, \Delta}{\Gamma, [s \Vdash^{\text{p2}} E'] \Rightarrow [s \Vdash^{\text{p2}} E], \Delta} E \subseteq E' \\
\text{(ex-||-)} \frac{\Gamma \Rightarrow [v = \mathbf{f.get}_0 \{ \text{Comp.comp} \}] \wedge [s \text{Comp.comp} \Vdash^{\text{pst}} \psi] \quad v \text{ fresh}}{\Gamma \Rightarrow [v = \mathbf{f.get}_0; s \Vdash^{\text{pst}} \varphi], \Delta}
\end{array}$$

Fig. 6. Two example rules for behavioral specifications. $\psi(v)$ replaces **result** by v and we assume that ψ contains no fields.

- If the proof contains only open branches containing behavioral specification, one may run a static analysis to evaluate them to true or false directly. E.g., if for the formula φ_p above the pointer analysis returns that the synchronization point 0 reads from `L.log`, the first behavioral modality evaluates to false and the whole formula to true.

Using external analyses increases modularity: (1) the BPL-calculus is simpler because it does not need to encode the implementation and (2) one may verify functional correctness of a method *up to its context*. Open branches are then a description of the context which the method requires. This may be verified once more context is known, thus extending proof repositories [6] to external analyses.

4 A Sequent Calculus for BPL: Behavioral Types

In this section we characterize behavioral types as behavioral specifications with a set of sequent calculus rules and a constraint on the proof obligations of the methods within a program. Before we formalize this in general, we introduce method types [23, 24], a behavioral type for Active Objects that suffices to generalize method contracts and object invariants by integrating the behavioral specifications for postcondition reasoning and points-to analysis. The method type of a method describes the local view of a method on a protocol.

Definition 14. *The local protocol \mathbf{L} and method type \mathbf{L} of a method are defined by the grammar below. The behavioral specification for method types is $\mathbb{T}_{\text{met}} = (\mathbf{L}, \alpha_{\text{met}})$. Let $\mathbf{x}_0, \dots, \mathbf{x}_n$ be roles, and $\mathbf{f}_{\mathbf{x}_0}, \dots, \mathbf{f}_{\mathbf{x}_n}$ fields of fitting type. $\alpha_{\text{met}}(\mathbf{L})$ is defined as $\exists \mathbf{x}_0, \dots, \mathbf{x}_n \in \mathbf{O}. \bigwedge_{i \leq n} \mathbf{x}_i \doteq \mathbf{f}_{\mathbf{x}_i} \wedge \alpha'_{\text{met}}(\mathbf{L})$. The first part models the (generated [24]) assignment of roles (as function symbols) to fields.*

$$\mathbf{L} ::= ?_{\mathbf{m}}(\varphi). \mathbf{L} \quad \mathbf{L} ::= \mathbf{x}!_{\mathbf{m}}(\varphi) \mid \downarrow(\varphi) \mid \text{skip} \mid \mathbf{L}.\mathbf{L} \mid \mathbf{L}^* \mid \oplus \{\mathbf{L}_i\}_{i \in I} \mid \&(\overrightarrow{\mathbf{m}}, \varphi)\{\mathbf{L}, \mathbf{L}\}$$

The local protocol of a method contains the receiving action $?_{\mathbf{m}}(\varphi)$, which models that the parameters satisfy the predicate φ . The method body is checked against the method type – there is no statement corresponding to receiving. Roles keep track of an object through the protocol. We stress that statements and method types share syntactic elements – it is possible to pattern match on statements/expressions on one side and a method type on the other side in rules.

Calls are specified with the call action $\mathbf{x}!_{\mathbf{m}}(\varphi)$, where $\mathbf{x.m}$ is the receiver and the predicate φ has to hold. Here, φ does not only specify the sent data but

also local variables and fields. It can express properties such as “the sent data is larger than some field”. The termination action $\downarrow(\varphi)$ models termination in a state satisfying φ (which again may include `result`). The empty action `skip` models no visible actions and $L_1.L_2$ to sequential composition: all interactions in L_1 must happen before L_2 . Repetition L^* corresponds to the Kleene star (and loops) and models zero or more repetitions of the interactions in L .

There are two choice operators: $\oplus\{L_i\}_{i \in I}$ is the active choice, the method must select one branch L_i . It is not necessary to implement all branches, the method may choose to never select some branches. The index set I must not be empty. $\&(\vec{m}, \varphi)\{L_1, L_2\}$ is the passive choice: some other method made a choice and this method has to follow the protocol according to this choice. The choice is communicated via a future which has to be resolved by one of the methods in \vec{m} . If the choice condition φ , which may only include the program variable `result`, is fulfilled by the read data, L_1 has to be followed, otherwise L_2 has to be followed. Both branches have to be implemented.

The semantics of the call and termination actions specify a trace with at least three elements with the correct event on second position and a state fulfilling the given predicate on the third position. Every other event is `noEv`. The semantics of the empty action and active choice are straightforward. Sequential composition uses relativization: some position i is chosen, such that the left translation holds before i and the right translation afterwards. Note that i is included in both relativization, to uphold the invariant that a trace always starts and ends with a state. The semantics of repetition are the only point where we require second order quantifiers: set I is a set of indices, such that the first and last position are included and for every consecutive pair k, l of elements of I , the translation of the repeated type holds in the relativization between k and l . Passive choice specifies that the first event is a read on a correct future (i.e., resolved by the correct method) and the suffix afterwards follows the communicated choice correctly.

Example 6. The following formalizes the behavior described informally in Ex. 2:

$$?T.test(true).S!Comp.comp(data \dot{=} i).\&(\{Comp.comp\}, result < 0) \left\{ \begin{array}{l} L!Log.log(data \dot{=} i), \\ skip \end{array} \right\} . \downarrow(result \geq 0)$$

The `result` variable in the guard of the passive choice is referring to the result of the read value, not the specified method.

We define behavioral types from a program logic perspective² by a type system, which is a set of sequent calculus rules that match on behavioral modalities and an obligation scheme, that maps every method to a proof obligation

Definition 15 (Behavioral Types). A behavioral type \mathbb{T} is a behavioral specification $(\tau_{\mathbb{T}}, \alpha_{\mathbb{T}})$ extended with $(\gamma_{\mathbb{T}}, \nu_{\mathbb{T}})$.

The obligation scheme $\nu_{\mathbb{T}}$ maps method names m to proof obligations, sequents of the form $\varphi_m \Rightarrow [s_m \Vdash^{\alpha_{\mathbb{T}}} \tau_m]$, which have to be proven. s_m is the method

² Behavioral types are sometimes (informally) distinguished from data types by having a subject reduction theorem where the typing relation is preserved, but not the type itself [10]. In BPL this would correspond to the property that one of the rules has a premise where the type in the behavioral modality is different than in the conclusion.

Fig. 7. Semantics for \mathbb{T}_{met} . Unbound variables are implicitly existentially quantified.

$$\begin{aligned}
\alpha'_{\text{met}}(\mathbf{X!m}(\varphi)) &= \forall i \in \mathbf{I}. \text{isEvent}(i) \wedge [i] \neq \text{noEv} \rightarrow [i] \doteq \text{invEv}(x, \mathbf{X}, f, \mathbf{m}, \vec{\mathbf{e}}) \wedge [i-1] \vdash \varphi(\vec{\mathbf{e}}) \\
&\quad \wedge \exists i \in \mathbf{I}. [i] \neq \text{noEv} \wedge \text{isEvent}(i) \\
\alpha'_{\text{met}}(\downarrow(\varphi)) &= \forall i \in \mathbf{I}. \text{isEvent}(i) \wedge [i] \neq \text{noEv} \rightarrow [i] \doteq \text{futEv}(x, f, \mathbf{m}, \mathbf{e}) \wedge [i-1] \vdash \varphi[\mathbf{result} \setminus \mathbf{e}] \\
&\quad \wedge \exists i \in \mathbf{I}. [i] \neq \text{noEv} \wedge \text{isEvent}(i) \\
&\quad \text{where } \varphi(\vec{\mathbf{e}}) \text{ replaces its free variables by } \vec{\mathbf{e}}. \varphi[\mathbf{result} \setminus \mathbf{e}] \text{ replaces } \mathbf{result} \text{ by } \mathbf{e}. \\
\alpha'_{\text{met}}(\text{skip}) &= \forall l \in \mathbf{I}. [l] \doteq \text{noEv} \vee [l] \vdash \mathbf{true} \quad \alpha'_{\text{met}}(\oplus\{\mathbf{L}_i\}_{i \in I}) = \bigvee_{i \in I} \alpha'_{\text{met}}(\mathbf{L}_i) \\
\alpha'_{\text{met}}(\mathbf{L}_1.\mathbf{L}_2) &= \exists i \in \mathbf{I}. \alpha'_{\text{met}}(\mathbf{L}_1)[n \in \mathbf{I} \setminus n \leq i] \wedge \alpha'_{\text{met}}(\mathbf{L}_2)[n \in \mathbf{I} \setminus n \geq i] \\
\alpha'_{\text{met}}(\mathbf{L}^*) &= \exists I \subseteq \mathbf{I}. \exists a, b \in I. a < b \wedge \\
&\quad \forall k \in \mathbf{I}. ((k < a \wedge \text{isEvent}(i) \rightarrow [i] \neq \text{noEv}) \vee (a \leq k \wedge k \leq b)) \wedge \\
&\quad \forall i_1, i_2 \in I. ((\forall l \in I. l \leq i_1 \wedge i_2 \leq l) \rightarrow \alpha'_{\text{met}}(\mathbf{L})[n \in \mathbf{I} \setminus i_1 \leq n \wedge n \leq i_2]) \\
\alpha'_{\text{met}}(\&\{\mathbf{m}_l\}_{l \in I}, \varphi)\{\mathbf{L}_1, \mathbf{L}_2\} &= \exists i, j, k \in \mathbf{I}. i < j \wedge j < k \wedge \\
&\quad (\forall l \in \mathbf{I}. l \doteq j \vee l \geq k \vee (l \leq i \wedge ([l] \doteq \text{noEv} \vee [l] \vdash \mathbf{true})) \wedge [j] \doteq \text{futREv}(x, \mathbf{m}, f, \mathbf{e}, n) \wedge \\
&\quad \bigvee_{l \in I} \mathbf{m} \doteq \mathbf{m}_l \wedge ([k] \vdash \varphi \rightarrow \alpha'_{\text{met}}(\mathbf{L}_1)[n \in \mathbf{I} \setminus n \geq k]) \wedge ([k] \not\vdash \varphi \rightarrow \alpha'_{\text{met}}(\mathbf{L}_2)[n \in \mathbf{I} \setminus n \geq k])
\end{aligned}$$

body of \mathbf{m} . The type system $\gamma_{\mathbb{T}}$ is a set of rewrite rules for $\tau_{\mathbb{T}}$ and sequent calculus rules with conclusions matching the sequent $\Gamma \Rightarrow \{\mathbf{U}\}[\mathbf{s} \Vdash^{\alpha_{\mathbb{T}}} \tau_{\mathbb{T}}], \Delta$.

We demand that obligation schemes are consistent, i.e., proof obligations do not contradict each other. This would be the case if, for example a method is called and its precondition φ is checked caller-side, then φ must truly be used as a precondition by the proof obligation for the called method.

Definition 16. Let $\mathbf{L}_m = ?_m(\varphi_m).\mathbf{L}_m$ be the local protocols in Prgm . We require that all \mathbf{L}_m are consistent: If \mathbf{m} is called in the method type of any other method \mathbf{m}' , then the call condition implies φ_m . Furthermore, $\varphi_{\mathbf{X}.run} = \mathbf{true}$.

The extension of the behavioral specification \mathbb{T}_{met} of method types to a behavioral type is given by the calculus in Fig. 8 and $\iota_{\text{met}}(\mathbf{m}) = \varphi_m \wedge \Phi \Rightarrow [\mathbf{s}_m \Vdash^{\alpha_{\text{met}}} \mathbf{L}_m]$. Formula $\Phi = \bigwedge_{\mathbf{X}} \mathbf{X} \doteq \text{select}(\mathbf{heap}, \mathbf{f}_{\mathbf{X}})$ encodes the assignment of roles to fields.

The call condition may contain fields of the other objects, but this is not an issue when checking consistency, as the precondition only contains fields of the own object and the fields are simply uninterpreted function symbols. The method in Fig. 2 can be typed with the type in Ex. 6.

Rule **(met-v)** translates a variable-assignment into an update and **(met-f)** is analogous for fields. Rule **(met-get)** has three premises: one premise checks via \mathbb{T}_{p2} that the correct methods are synchronized with. The two others use a fresh constant v for the read value and assign it to the target variable. The two premises differ in the branch that is checked afterwards, depending on whether or not the choice condition holds. Rule **(met-while)** is a standard loop invariant rule. An invariant I holds before the first iteration and is preserved by the loop to remove all other information afterwards. The loop body is checked against the repeated type and the continuation against the continuation of the type. Method types have no special action for the end of a statement, so \mathbb{T}_{pst} is used

$$\begin{array}{c}
\text{(met-V)} \frac{\Gamma \Rightarrow \{U\}\{v := e\}[\mathbf{s} \Vdash^{\alpha_{\text{met}}} L], \Delta}{\Gamma \Rightarrow \{U\}[v = e; \mathbf{s} \Vdash^{\alpha_{\text{met}}} L], \Delta} \quad \text{(met-F)} \frac{\Gamma \Rightarrow \{U\}\{\text{heap} := \text{store}(\text{heap}, \mathbf{f}, \mathbf{e})\}[\mathbf{s} \Vdash^{\alpha_{\text{met}}} L], \Delta}{\Gamma \Rightarrow \{U\}[\mathbf{this.f} = \mathbf{e}; \mathbf{s} \Vdash^{\alpha_{\text{met}}} L], \Delta} \\
\\
\Gamma \Rightarrow \{U\}\{v := v\}(\varphi(v) \rightarrow [\mathbf{s} \Vdash^{\alpha_{\text{met}}} L_1]), \Delta \\
\text{(met-get)} \frac{\Gamma \Rightarrow \{U\}\{v := v\}(\neg\varphi(v) \rightarrow [\mathbf{s} \Vdash^{\alpha_{\text{met}}} L_2]), \Delta \quad \Rightarrow [v = \mathbf{e.get}_i; \mathbf{s} \Vdash^{\text{p2}}\{\vec{m}\}]}{\Gamma \Rightarrow \{U\}[v = \mathbf{e.get}_i; \mathbf{s} \Vdash^{\alpha_{\text{met}}} \&(\vec{m}, \varphi)\{L_1, L_2\}], \Delta} \quad v \text{ fresh} \\
\\
\text{(met-while)} \frac{\Gamma \Rightarrow \{U\}I, \Delta \quad I, \mathbf{e} \Rightarrow [\mathbf{s} \Vdash^{\text{pst}} I] \quad I, \mathbf{e} \Rightarrow [\mathbf{s} \Vdash^{\alpha_{\text{met}}} L] \quad I, \neg\mathbf{e} \Rightarrow [\mathbf{s}' \Vdash^{\alpha_{\text{met}}} L'], \Delta}{\Gamma \Rightarrow \{U\}[\mathbf{while} \ \mathbf{e} \ \mathbf{do} \ \mathbf{s} \ \mathbf{od} \ \mathbf{s}' \Vdash^{\alpha_{\text{met}}} L^*.L'], \Delta} \\
\\
\Gamma \Rightarrow \{U\}(\mathbf{e} \rightarrow [\mathbf{s}; \mathbf{s}''] \Vdash^{\alpha_{\text{met}}} \oplus\{L_i\}_{i \in I_1}), \Delta \\
\text{(met-if)} \frac{\Gamma \Rightarrow \{U\}(\neg\mathbf{e} \rightarrow [\mathbf{s}'; \mathbf{s}'''] \Vdash^{\alpha_{\text{met}}} \oplus\{L_i\}_{i \in I_2}), \Delta}{\Gamma \Rightarrow \{U\}[\mathbf{if} \ \mathbf{e} \ \mathbf{then} \ \mathbf{s} \ \mathbf{else} \ \mathbf{s}' \ \mathbf{fi} \ \mathbf{s}'''] \Vdash^{\alpha_{\text{met}}} \oplus\{L_i\}_{i \in I}], \Delta} \quad I_1 \cup I_2 \subseteq I \\
\\
\text{(met-call)} \frac{\Gamma \Rightarrow \{U\}(\varphi(\mathbf{e}) \wedge \text{select}(\text{heap}, \mathbf{f}) \doteq X), \Delta \quad \Gamma \Rightarrow \{U\}\{v := f\}[\mathbf{s} \Vdash^{\alpha_{\text{met}}} L], \Delta}{\Gamma \Rightarrow \{U\}[v = \mathbf{f!m}(\mathbf{e}); \mathbf{s} \Vdash^{\alpha_{\text{met}}} X!\mathbf{m}(\varphi).L], \Delta} \quad f \text{ fresh} \\
\\
\text{(met-return)} \frac{\Gamma \Rightarrow \{U\}\{\mathbf{result} := \mathbf{e}\}\varphi, \Delta}{\Gamma \Rightarrow \{U\}[\mathbf{return} \ \mathbf{e} \Vdash^{\alpha_{\text{met}}} \downarrow(\varphi)], \Delta} \quad \text{(met-skip)} \frac{}{\Gamma \Rightarrow \{U\}[\mathbf{skip} \Vdash^{\alpha_{\text{met}}} \text{skip}], \Delta} \\
L \rightsquigarrow \oplus\{L\} \quad \text{skip.L} \rightsquigarrow L \quad L.\text{skip} \rightsquigarrow L
\end{array}$$

Fig. 8. Rules for \mathbb{T}_{met} . We remind that the sets I_1, I_2 are defined as non-empty. For simplicity, we assume that every branch and every loop body implicitly ends in **skip**.

for checking that the loop preserves its invariant. Rule **(met-if)** splits the set of possible choices into two and checks each branch against one of these sets. These sets may overlap and do not need to cover all original choices, but may not be empty. Rule **(met-call)** checks the annotated condition of the called method and the correct target explicitly and that the correct method is called by matching call type and call statement. We remind that references are not reassigned, so call targets can be verified locally. The other rules are straightforward.

Contracts and Invariants. Method types generalize method contracts and object invariants as follows. An object invariant is encoded by adding it to the formula in the receiving and terminating actions of all method in an object – except the constructor **run**, where it is only added to the terminating action. A method contract (consisting of a precondition on the parameters and a postcondition) is encoded analogously by adding the precondition to the receiving and the postcondition to the terminating actions. However, one additional step is required: Method types are generated by projection of global types [23], so to use them for object invariants or method contracts requires to infer a method type first. This is done by mapping every call to a call action, every branching

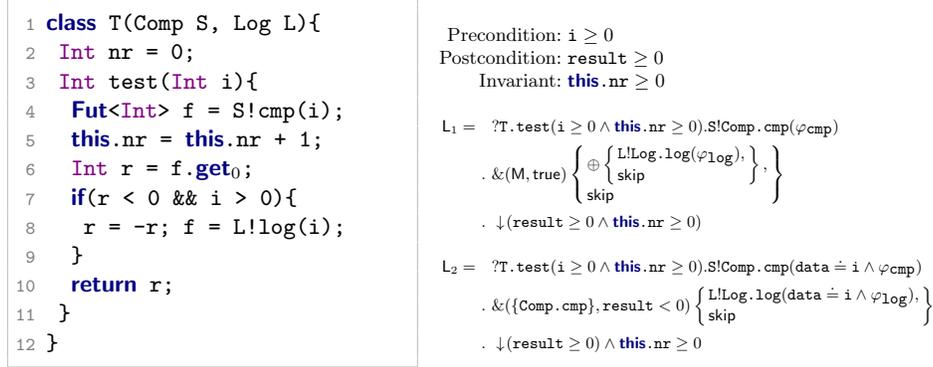


Fig. 9. An example method and two method types for method contracts and invariants.

to an active choice, every loop to a repetition, termination to a terminating action and using `true` at every position where a formula is required, before adding precondition, postcondition or object invariant. The most complex construct is synchronization. Each such read is mapped to a passive choice with all methods as the method set and `true` as the choice condition. The following code is added in the first branch. The second branch is `skip`. Invariants require fields in the precondition and a fitting notion of consistency, which was developed in [23].

Example 7. Consider the code in Fig. 9, a variation of our running example. It tracks the number of calls to `T.test` and inverts the result if the input is positive. It adheres to the contract with precondition $i \geq 0$ and postcondition $\text{result} \geq 0$ and the invariant $\text{this.nr} \geq 0$. The algorithm above derives the following type:

$$?T.\text{test}(\text{true}) \cdot S!Comp.\text{cmp}(\text{true}) \cdot \&(M, \text{true}) \left\{ \oplus \left\{ \begin{array}{l} L!Log.\text{log}(\text{true}), \\ \text{skip} \end{array} \right\}, \right\} \cdot \downarrow(\text{true})$$

Let φ_{cmp} and φ_{log} be the preconditions of the called methods. The final specification, after adding the contract and the invariant, is shown on the right in Fig. 9 as L_1 . The inferred type is not the one we gave in Ex. 6: For one, it differs in its shape (two choice operators). For another, it neither keeps track of the passed data, nor specifies the relation between the return value of `Comp.cmp` and the taken branch. These properties are typical for protocol specifications and require a global view, contrary to the local view of method contracts and object invariants. However, one can add the pre- and postcondition and the object invariant also to the type given in Ex. 6 and combine local and global specification. The result is shown as L_2 in in Fig. 9. L_2 expresses that the method follows the protocol and adheres to contract and object invariant.

Theorem 1. \mathbb{T}_{met} is sound for every program.

The proof is standard [22]. Consistency of the obligation scheme is required to establish that all selected traces are models for the type of their method. The first two elements are not described by the method type and, thus, removed.

Corollary 1. *If (1) for every method m with type $?_m(\varphi).L_m$ the formula $\iota_{\text{met}}(m)$ is valid and (2) the obligation scheme is consistent, then for every selected trace θ of any method m , the trace after the invocation reaction event follows its type:*

$$\theta[2..|\theta|], I, \emptyset \models \alpha_{\text{met}}(L_m)$$

5 Conclusion and Related Work

This work presents BPL, a program logic for object-oriented distributed programs that enables deductive reasoning about the results of static analyses and integrates concepts from behavioral types by pattern-matching statement and specification. The *method type* behavioral type generalizes method contracts, session types and object invariants. In the following, we discuss related work.

Dynamic Logics. Beckert and Bruns [3] use LTL formulas in dynamic logic modalities in their Dynamic Trace Logic (DTL) for Java. Given an LTL formula φ , the DTL-formula $[s]\varphi$ expresses that φ describes all traces of s . DTL uses a restricted form of pattern matching: its three loop invariant rules depend on the outermost operator of φ and other rules may consume a “next” operator. DTL does not use events and specifies patterns of state changes, not of interactions.

The Abstract Behavior Specification Dynamic Logic (ABSDDL) of Din and Owe [12] is for the ABS language [21]. In ABSDDL, a formula $[s]\varphi$, where φ is a first-order formula over the program state, has the standard meaning that φ holds after s is executed. ABSDDL uses a special program variable to keep track of the visible events. Its rules are tightly coupled with object-invariant reasoning. This makes it impossible to specify the state at arbitrary interactions.

Bubel et al. [7] define dynamic logic with coinductive traces (DLCT). In DLCT, a formula $[s]\varphi$, where φ is a trace modality formula, containing symbolic trace formulas, has the meaning that every trace of s is a model for φ . Contrary to ABSDDL, DLCT keeps track of the whole trace, not just the events. DLCT is not able to specify the property that between two states, some form of event does *not* occur, as symbolic trace formulas are not closed under negation.

Behavioral Types. A number of behavioral types deals with assertions [4, 5] or Actors [16, 18, 26]. Stateful Behavioral Types for Active Objects (STAO) [23] uses both and defines the judgment $\varphi, s' \vdash s : \tau$, that expresses that all traces of s are models for the translation of τ . φ and s' keep track of the chosen path so far. STAO is not able to reason about multiple judgments, but relies on external analyses for precision. Reasoning about these results happens on a meta-level.

Finally, Propositions-as-Types theorems (PaT) have been established [8, 27] between session types for the π -calculus and intuitionistic linear logic. They are specific to this setting and do not characterize general behavioral types. To our best knowledge, Def. 15 is the first formal characterization of behavioral types.

Future Work. An implementation of BPL for full ABS is ongoing and as future work, we plan to investigate further types and concurrency models, in particular systems with shared memory and effect type systems.

References

1. AHRENDT, W., BECKERT, B., BUBEL, R., HÄHNLE, R., SCHMITT, P. H., AND ULBRICH, M., Eds. *Deductive Software Verification - The KeY Book - From Theory to Practice*, vol. 10001 of *Lecture Notes in Computer Science*. Springer, 2016.
2. BECKERT, B. A dynamic logic for the formal verification of java card programs. In *Java on Smart Cards: Programming and Security* (Berlin, Heidelberg, 2001), I. Attali and T. Jensen, Eds., Springer Berlin Heidelberg, pp. 6–24.
3. BECKERT, B., AND BRUNS, D. Dynamic logic with trace semantics. In *Automated Deduction - CADE 2013. Proceedings* (2013), M. P. Bonacina, Ed., vol. 7898 of *Lecture Notes in Computer Science*, Springer, pp. 315–329.
4. BERGER, M., HONDA, K., AND YOSHIDA, N. Completeness and logical full abstraction in modal logics for typed mobile processes. In *Automata, Languages and Programming, 35th International Colloquium, ICALP 2008, Proceedings, Part II* (2008), L. Aceto, I. Damgård, L. A. Goldberg, M. M. Halldórsson, A. Ingólfssdóttir, and I. Walukiewicz, Eds., vol. 5126 of *Lecture Notes in Computer Science*, Springer, pp. 99–111.
5. BOCCHI, L., LANGE, J., AND TUOSTO, E. Three algorithms and a methodology for amending contracts for choreographies. *Sci. Ann. Comp. Sci.* *22*, 1 (2012), 61–104.
6. BUBEL, R., DAMIANI, F., HÄHNLE, R., JOHNSEN, E. B., OWE, O., SCHAEFER, I., AND YU, I. C. Proof repositories for compositional verification of evolving software systems - managing change when proving software correct. *Trans. Found. Mastering Chang.* *1* (2016), 130–156.
7. BUBEL, R., DIN, C. C., HÄHNLE, R., AND NAKATA, K. A dynamic logic with traces and coinduction. In *Automated Reasoning with Analytic Tableaux and Related Methods TABLEAUX 2015. Proceedings* (2015), H. de Nivelle, Ed., vol. 9323 of *Lecture Notes in Computer Science*, Springer, pp. 307–322.
8. CAIRES, L., AND PFENNING, F. Session types as intuitionistic linear propositions. In *Concurrency Theory, CONCUR 2010. Proceedings* (2010), P. Gastin and F. Laroussinie, Eds., vol. 6269 of *Lecture Notes in Computer Science*, Springer, pp. 222–236.
9. DE BOER, F. S., SERBANESCU, V., HÄHNLE, R., HENRIO, L., ROCHAS, J., DIN, C. C., JOHNSEN, E. B., SIRJANI, M., KHAMESPANAH, E., FERNANDEZ-REYES, K., AND YANG, A. M. A survey of active object languages. *ACM Comput. Surv.* *50*, 5 (2017), 76:1–76:39.
10. DEZANI-CIANCAGLINI, M. personal communication, 19.10.2018.
11. DIN, C. C., HÄHNLE, R., JOHNSEN, E. B., PUN, K. I., AND TAPIA TARIFA, S. L. Locally abstract, globally concrete semantics of concurrent programming languages. In *Automated Reasoning with Analytic Tableaux and Related Methods - 26th International Conference, TABLEAUX 2017, Proceedings* (2017), R. A. Schmidt and C. Nalon, Eds., vol. 10501 of *Lecture Notes in Computer Science*, Springer, pp. 22–43.
12. DIN, C. C., AND OWE, O. A sound and complete reasoning system for asynchronous communication with shared futures. *J. Log. Algebr. Meth. Program.* *83*, 5-6 (2014), 360–383.
13. DIN, C. C., TAPIA TARIFA, S. L., HÄHNLE, R., AND JOHNSEN, E. B. History-based specification and verification of scalable concurrent and distributed systems. In *International Conference on Formal Engineering Methods, ICFEM 2015. Proceedings* (2015), M. J. Butler, S. Conchon, and F. Zaidi, Eds., vol. 9407 of *Lecture Notes in Computer Science*, Springer, pp. 217–233.

14. FLORES-MONTOYA, A. E., ALBERT, E., AND GENAIM, S. May-happen-in-parallel based deadlock analysis for concurrent objects. In *Formal Techniques for Distributed Systems* (Berlin, Heidelberg, 2013), D. Beyer and M. Boreale, Eds., Springer Berlin Heidelberg, pp. 273–288.
15. GARCÍA-MATOS, M., AND VÄÄNÄNEN, J. Abstract model theory as a framework for universal logic. In *Logica Universalis* (Basel, 2005), J.-Y. Beziau, Ed., Birkhäuser Basel, pp. 19–33.
16. GIACHINO, E., JOHNSEN, E. B., LANEVE, C., AND PUN, K. I. Time complexity of concurrent programs - A technique based on behavioural types -. In *Formal Aspects of Component Software - 12th International Conference, FACS 2015, Revised Selected Papers* (2015), C. Braga and P. C. Ölveczky, Eds., vol. 9539 of *Lecture Notes in Computer Science*, Springer, pp. 199–216.
17. HENKIN, L. Relativization with respect to formulas and its use in proofs of independence. *Compositio Mathematica* 20 (1968), 88–106.
18. HENRIO, L., LANEVE, C., AND MASTANDREA, V. Analysis of synchronisations in stateful active objects. In *Integrated Formal Methods - 13th International Conference, iFM 2017. Proceedings* (2017), N. Polikarpova and S. Schneider, Eds., vol. 10510 of *Lecture Notes in Computer Science*, Springer, pp. 195–210.
19. HONDA, K., YOSHIDA, N., AND CARBONE, M. Multiparty asynchronous session types. vol. 63, pp. 9:1–9:67.
20. HÜTTEL, H., LANESE, I., VASCONCELOS, V. T., CAIRES, L., CARBONE, M., DENIÉLOU, P.-M., MOSTROUS, D., PADOVANI, L., RAVARA, A., TUOSTO, E., VIEIRA, H. T., AND ZAVATTARO, G. Foundations of session types and behavioural contracts. *ACM Comput. Surv.* 49, 1 (Apr. 2016), 3:1–3:36.
21. JOHNSEN, E. B., HÄHNLE, R., SCHÄFER, J., SCHLATTE, R., AND STEFFEN, M. ABS: A core language for abstract behavioral specification. In *Formal Methods for Components and Objects - 9th International Symposium, FMCO 2010. Revised Papers* (2010), B. K. Aichernig, F. S. de Boer, and M. M. Bonsangue, Eds., vol. 6957 of *Lecture Notes in Computer Science*, Springer, pp. 142–164.
22. KAMBURJAN, E. Behavioral program logic and LAGC semantics without continuations (technical report). *CoRR abs/1904.13338* (2019).
23. KAMBURJAN, E., AND CHEN, T. Stateful behavioral types for active objects. In *Integrated Formal Methods - 14th International Conference, iFM 2018. Proceedings* (2018), C. A. Furia and K. Winter, Eds., vol. 11023 of *Lecture Notes in Computer Science*, Springer, pp. 214–235.
24. KAMBURJAN, E., DIN, C. C., AND CHEN, T. Session-based compositional analysis for actor-based languages using futures. In *Formal Methods and Software Engineering - 18th International Conference on Formal Engineering Methods, ICFEM 2016. Proceedings* (2016), K. Ogata, M. Lawford, and S. Liu, Eds., vol. 10009 of *Lecture Notes in Computer Science*, pp. 296–312.
25. KAMBURJAN, E., AND HÄHNLE, R. Deductive verification of railway operations. In *Reliability, Safety, and Security of Railway Systems. Modelling, Analysis, Verification, and Certification - Second International Conference, RSSRail 2017. Proceedings* (2017), A. Fantechi, T. Lecomte, and A. B. Romanovsky, Eds., vol. 10598 of *Lecture Notes in Computer Science*, Springer, pp. 131–147.
26. NEYKOVA, R., AND YOSHIDA, N. Multiparty session actors. *Logical Methods in Computer Science* 13, 1 (2017).
27. WADLER, P. Propositions as types. *Commun. ACM* 58, 12 (2015), 75–84.