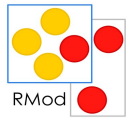
Scriptable debugging, execution querying and other advanced debugging techniques

Maximilian Ignacio Willembrinck Santander

PharoDays 2022



Get the code here: https://github.com/maxwills/PharoDays2022

Presentation Format

A mix of hands-on, demonstration and explanations.

Follow the examples:

Code here: **https://github.com/maxwills/PharoDays2022**

Let's go!

Presentation Agenda

Context	Introduction and what to expect
	Motivation
Obtaining Execution Data	Code insertion (+ Conditions)
	Instrumentation (+ Reflection)
Scriptable-Debuggers	Sindarin
	Seeker

Next

Get the code here: https://github.com/maxwills/PharoDays2022

What to expect (Spoilers)

We will explore a program execution, trying to answer a few debugging questions.

We will go from commonly-used tools and techniques to not so commonly-used:

Halts, Logging, Breakpoints, MetaLinks, MethodProxies

We will show usage of Scriptable Debuggers in Pharo.

Get the code here: https://github.com/maxwills/PharoDays2022

“The debugging activity does not take place only in the Debugger”.

Presentation Agenda

Context
- Introduction and what to expect
- Motivation

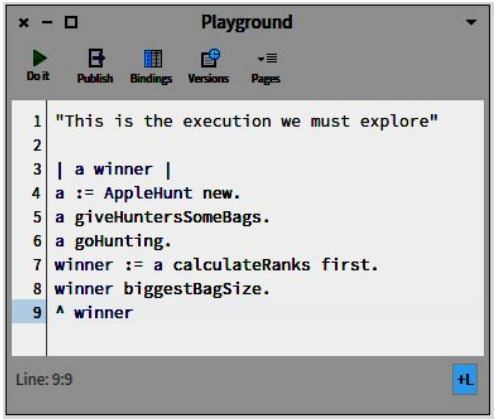
Next

Obtaining Execution Data
- Code insertion (+ Conditions)
- Instrumentation (+ Reflection)

Scriptable-Debuggers
- Sindarin
- Seeker

Motivation

We will explore a program
using several debugging tools and techniques



Our program tells a story ...

Motivation

Getting data from the execution:

Q1. How many times the method **`OrderedCollection>>add:`** is called? (and with an Apple as argument?)

Q2. How many times any method with selector **`add:`** is called? What is the actual method in every case?

Presentation Agenda

Context	Introduction and what to expect
	Motivation
Obtaining Execution Data	Code insertion (+ Conditions)
	Instrumentation (+ Reflection)
Scriptable-Debuggers	Sindarin
	Seeker

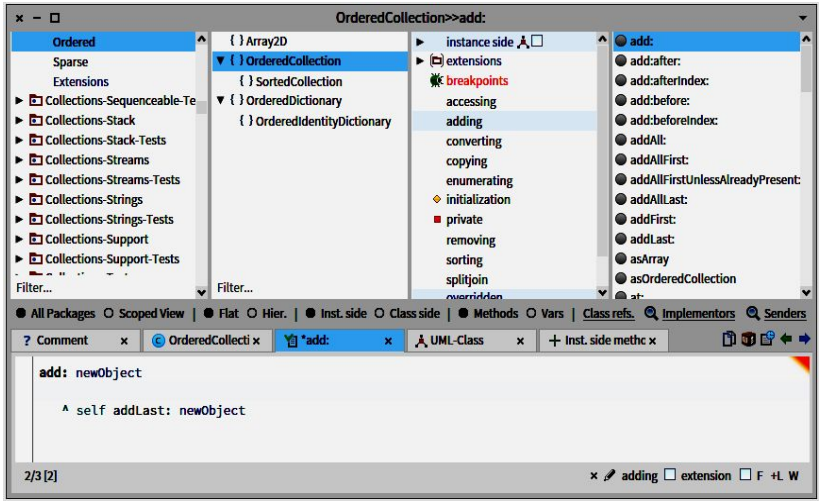
Next

(To the code!)

Obtaining Execution Data

Code Insertion

We change the code of the execution to include our inquisitive actions.



Original OrderedCollection>>add: method
(comments removed)

```
add: newObject
    self halt.
    ^ self addLast: newObject
```

```
add: newObject
    MyLoggerUtil log: 'OrderedCollection >> add: called'.
    ^ self addLast: newObject
```

Note: Images for illustrative purpose only.
Don't put an Halt in OrderedCollection>>add: :(

Obtaining Execution Data

Code Insertion

We change the code of the execution to include our inquisitive actions.

- **Logging (printing)**
- **Halt**

Getting data from the execution:

Q1. How many times the method `OrderedCollection>>add:` is called? (and with an Apple as argument?)

Obtaining Execution Data

Code Insertion

Getting data from the execution:

~~Q1. How many times the method `OrderedCollection>>add:` is called? (and with an Apple as argument?)~~

Q2. How many times any method with selector add: is called? What is the actual method in every case? Consider only cases when adding Apple, or Hunter objects.

Problem:

There are several possible methods with the #add: selector. What to do?

Presentation Agenda

Context	Introduction and what to expect
	Motivation
Obtaining Execution Data	Code insertion (+ Conditions)
	Instrumentation (+ Reflection)
Scriptable-Debuggers	Sindarin
	Seeker

Next

Obtaining Execution Data

Instrumentation

This time, to include our inquisitive actions, we change the execution without altering it's code.

```
appleHunt

    "This is the execution we must explore"

    | a winner |
    a := AppleHunt new.
    a giveHuntersSomeBags.
    a goHunting.
    winner := a calculateRanks first.
    winner biggestBagSize.
    ^ winner
```

```
appleHunt

    "This is the execution we must explore"

    | a winner |
    a := AppleHunt new.
    a giveHuntersSomeBags.
    a goHunting.
    winner := a calculateRanks first.
    winner biggestBagSize.
    ^ winner
```

A Breakpoint

A Conditional Breakpoint

\+ MetaLinks MethodProxies

Obtaining Execution Data

Instrumentation

This time, to include our inquisitive actions, we change the execution without altering it's code.

- **Breakpoints**
- **MetaLinks**
- **MethodProxies**

Instrumentation with MetaLinks

Adds a extra instructions to our execution without modifying its code.

The code:

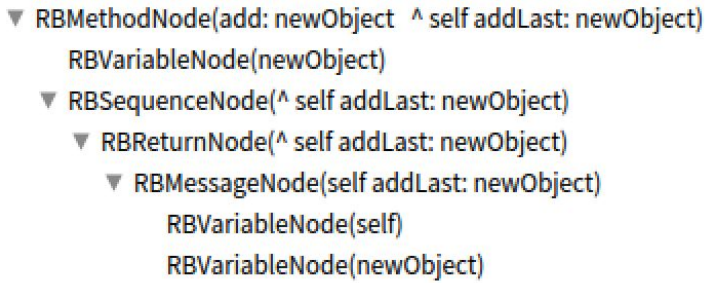
```
add: newObject

    ^ self addLast: newObject
```

Instrumentation with MetaLinks

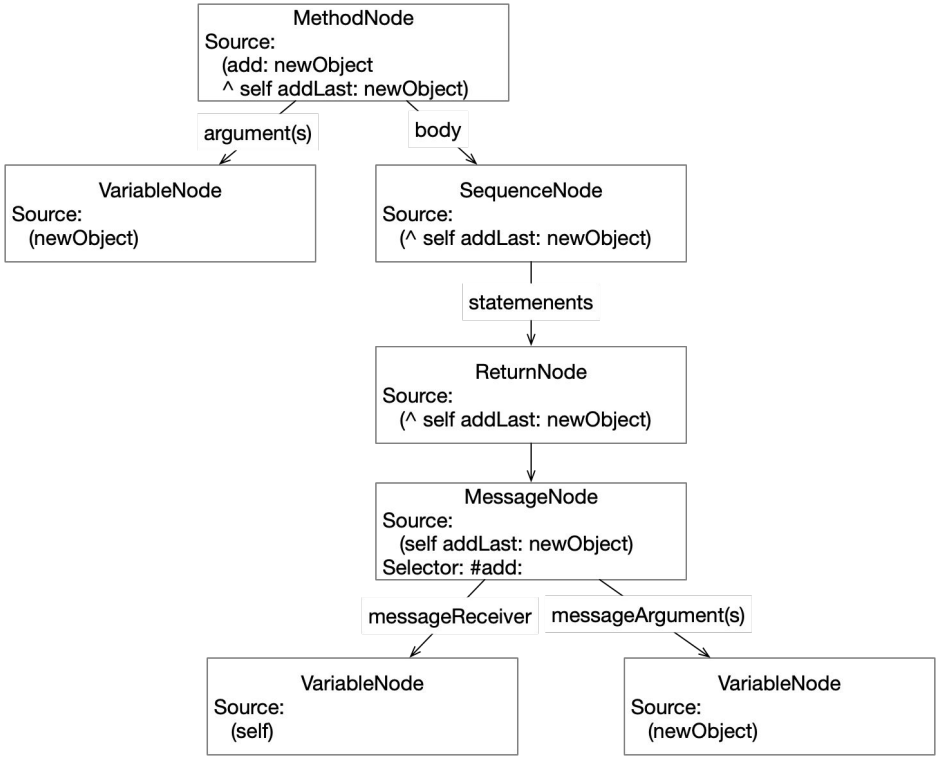
When a method is compiled:

```
add: newObject

    ^ self addLast: newObject
```

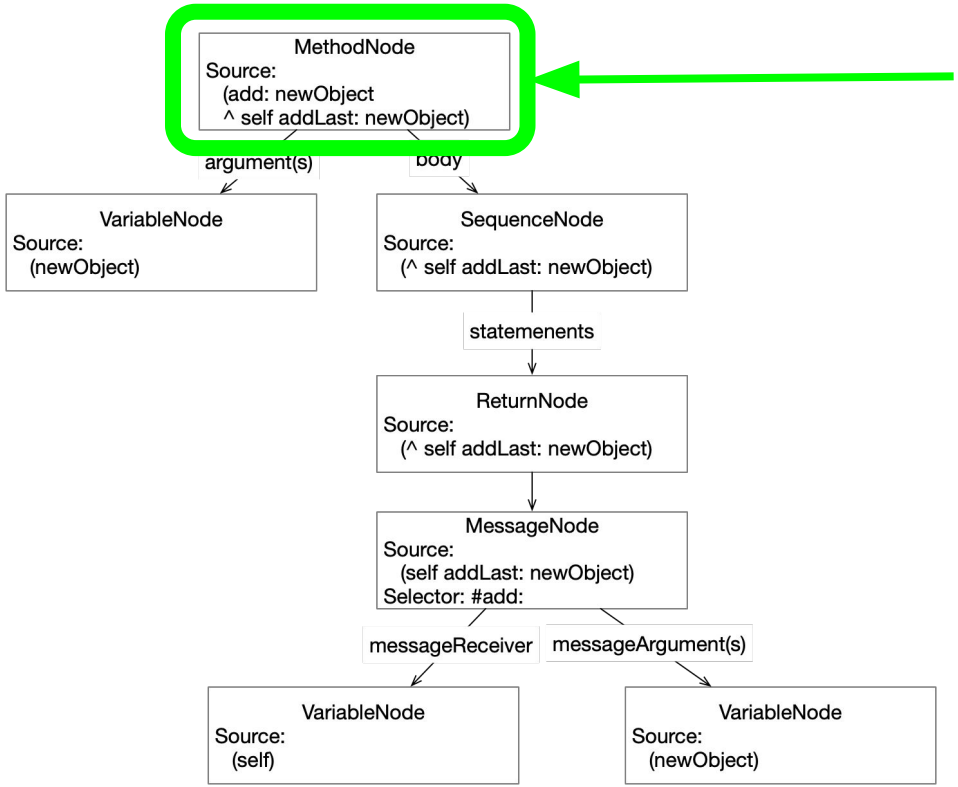
An AST is produced:



Instrumentation with MetaLinks



Instrumentation with MetaLinks



"When **that** node is executed, execute **THIS** other code!"

```
Transcript show: '#add: called'
```

Showing the code in Pharo

Instrumentation through MetaLinks Examples
+ Reflection

MethodProxies

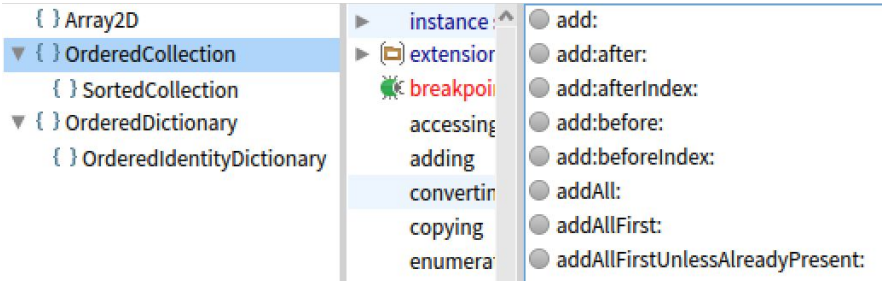
Some facts:

- Not included in Pharo.
- Get it here: https://github.com/pharo-contributions/MethodProxies

  The package is developed and maintained by S. Ducasse, G. Polito and P. Tesone, but feel free to give a hand.

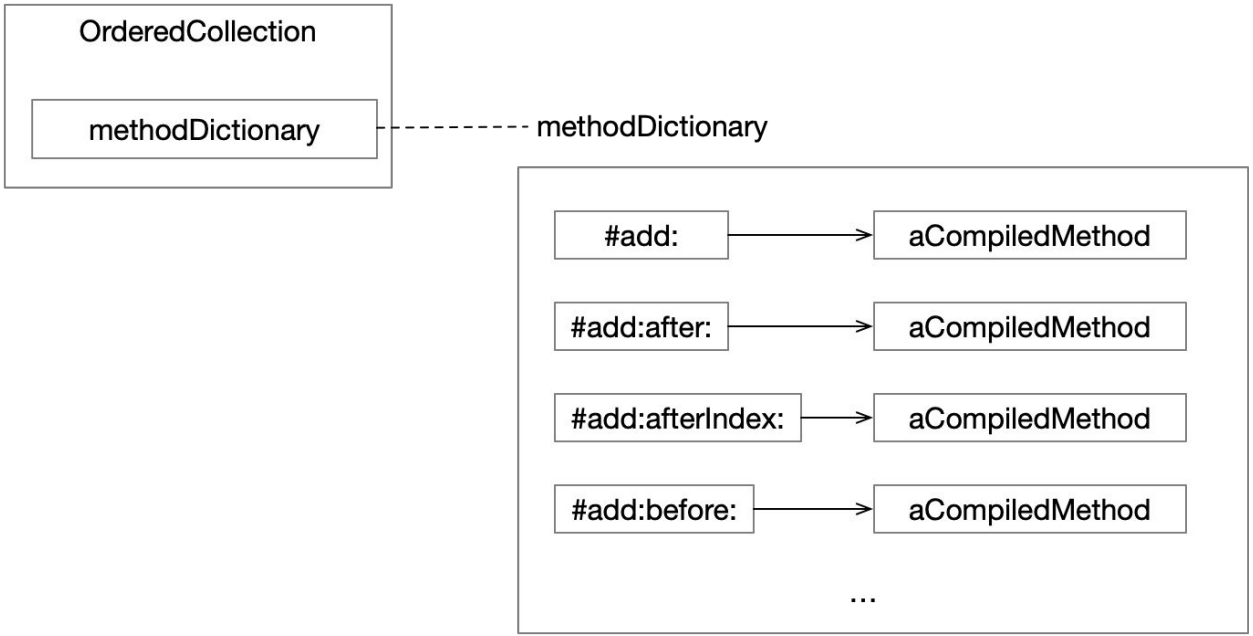
Instrumentation with MethodProxies

Adds a extra instructions to our execution without modifying its code, by “proxying” its method(s).



Instrumentation with MethodProxies

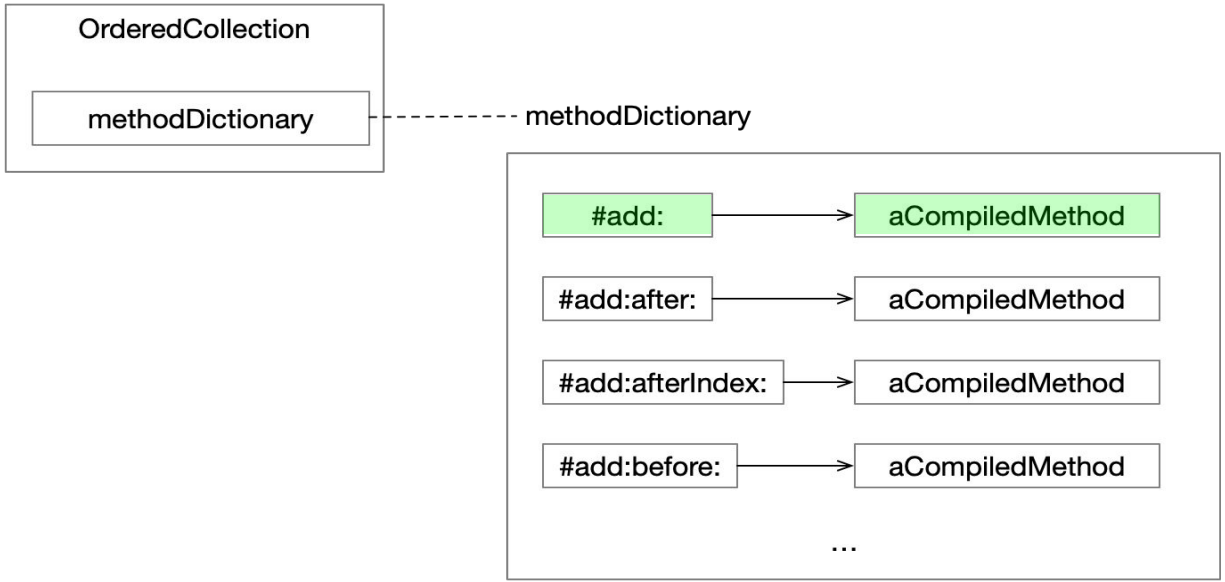
When a method is compiled, the created CompiledMethod object is stored in the methodDictionary of the class.



Instrumentation with MethodProxies

When sending a message to an object of our class, Pharo will get the CompiledMethod object of the dictionary, and will execute it.

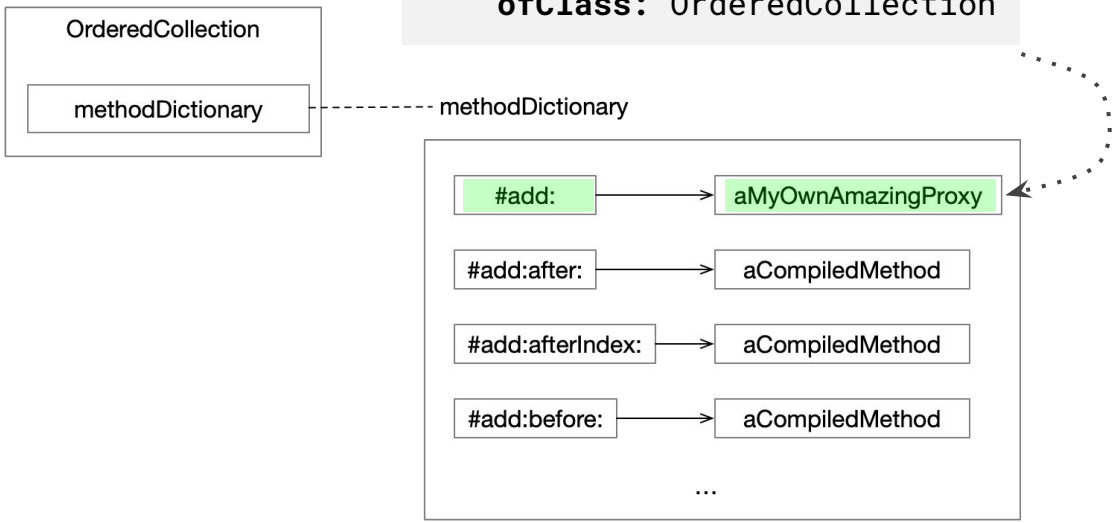
```
myCollection add: 1
```



Instrumentation with MethodProxies

To instrument our method, we can replace the CompiledMethod with a Proxy

```
MyOwnAmazingProxy
    installInMethod: #add:
    ofClass: OrderedCollection
```



(To the code!)

Obtaining Execution Data

Comparison of approaches

Code Insertion

- Halt
- Logging

Instrumentation

- Breakpoints
- MetaLinks
- MethodProxies

	Code Insertion	Instrumentation
“Simple” to understand.		
“Simple” to debug.		
Don’t modify the debugged program code.		
Persistent (recompilation)		
Good Scalability		

By instrumenting a program, we alter its execution.
(Even if we don't modify it's code)

Can we extract execution data without altering the execution at all?

Presentation Agenda

Context	Introduction and what to expect
	Motivation
Obtaining Execution Data	Code insertion (+ Conditions)
	Instrumentation (+ Reflection)
Scriptable-Debuggers	Sindarin
	Seeker

Next

Scriptable Debuggers

- Allow developers to automate debugging tasks.
- Expose an API to:
  - Manipulate the debugger and debugged execution.
  - Obtain information about the debugged execution.

Less tedious manual debugging work for the developer

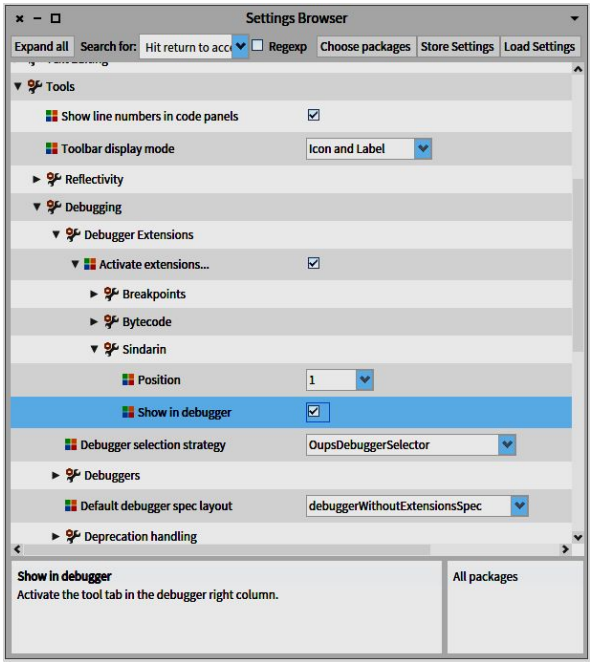
Sindarin

Included in Pharo

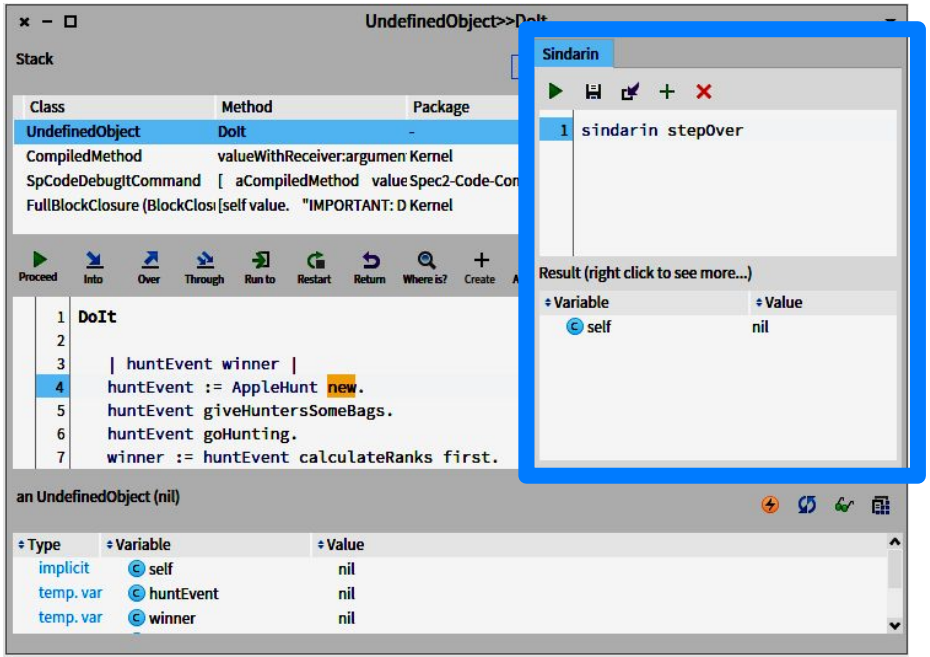
- Two flavors:
  - With UI: Extension of the StDebugger.
    - Has to activated from the Pharo settings.
  - Headless: The SindarinDebugger Object.
    - Already included in Pharo 9.0 and Pharo 10

Sindarin

Enabling the UI version



Enable it in the Pharo Settings



StDebugger with Sindarin UI Extension

(To the code)

Presentation Agenda

Context	Introduction and what to expect
	Motivation
Obtaining Execution Data	Code insertion (+ Conditions)
	Instrumentation (+ Reflection)
Scriptable-Debuggers	Sindarin
	Seeker

Next

Seeker

A prototype scriptable and queryable debugger

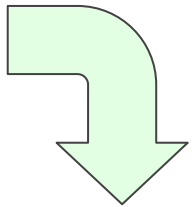
- UI version only (as an extension of the StDebugger).
- Not included in Pharo.
- Query-based debugging.
- Time-Traveling mechanics.

Get the code here: https://github.com/maxwills/PharoDays2022

(To the code)

The Query Notation
(From scripting to querying)

```
    | callsToAdd |

    callsToAdd := OrderedCollection new.

       seeker restartAtBytecodeOne.
       [ seeker canStep ] whileTrue: [
          seeker step.
          (seeker currentState isMessageSend and: [ seeker currentState node selector = #add: ])
ifTrue: [
             callsToAdd add: seeker currentState methodAboutToExecute ] ]  .
    ^ callsToAdd
```



```
^ (Query from: seeker newProgramStates
     select: [ :state | state isMessageSend and: [ state node selector = #add: ] ]
     collect: [ :state | state methodAboutToExecute ]) asOrderedCollection
```

Standard Query Notation

Equivalent queries in other languages

Pharo (Prototype implementation for Pharo 9.0/Pharo10)

```
^ (Query from: seeker newProgramStates
      select: [ :state | state isMessageSend and: [ state node selector = #add: ] ]
      collect: [ :state | state methodAboutToExecute ]) asOrderedCollection
```

SQL Query

```
SELECT state.methodAboutToExecute
FROM ProgramStates
WHERE state.isMessageSend AND
       state.node.selector = 'add:'
```

C# (Linq)

```
var results=(
  FROM state in seeker.newProgramStates
  WHERE state.isMessageSend && state.node.selector == "add:"
  SELECT state.methodAboutToExecute
).ToList();
```

Python (List comprehension + properties)

```
results = [
   state.methodAboutToExecute
   for state in seeker.newProgramStates
   if state.isMessageSend and state.node.selector == "add:"
]
```

(To the code!)

Queries In Seeker

- Not any kind of queries, but Time-Traveling Queries (TTQs).
- A set of ready-to-use TTQs are provided.
- Developers can write their own Queries and TTQs.

Summary

- **The debugger is not your only tool for debugging.**
- **Inserting extra behavior to study your program execution for your debugging sessions.**
  - **Code Insertion, Instrumentation, and dangers.**
- **Reflection is a powerful mechanism to obtain execution data.**
- **Scriptable debuggers: Sindarin.**
- **Scriptable Time-Traveling Queryable Debugger prototype: Seeker.**
  - **Query Notation and Time-Traveling Queries**

Have a good day!

Get the code here: https://github.com/maxwills/PharoDays2022

Presentation is finished.
Extra slides next.

Context

What to expect

We will explore an execution, to find answer to debugging questions.

Everything goes!

> ***The standard debugger, breakpoints, logging, proxies, reflection, scripting, speculating, etc.***

I will break some Pharo Images, and show how you can avoid that (*).

*(*In some cases. There are no guarantees, so don't sue me)*

Get the code here: https://github.com/maxwills/PharoDays2022

Context
What to expect

By the end of the session you will:

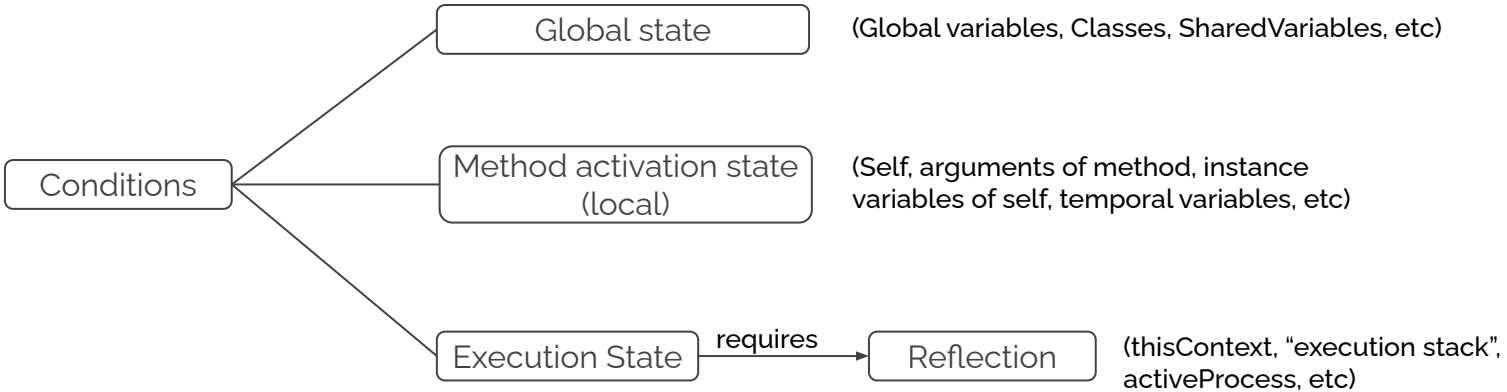
- **Be aware of currently available yet not-so-commonly used tools in Pharo.**
- **Have the knowledge on what is to come in some aspects of debugging.**

And (hopefully)

- **Your debugging techniques repertory is expanded.**
- **You had a blast!**

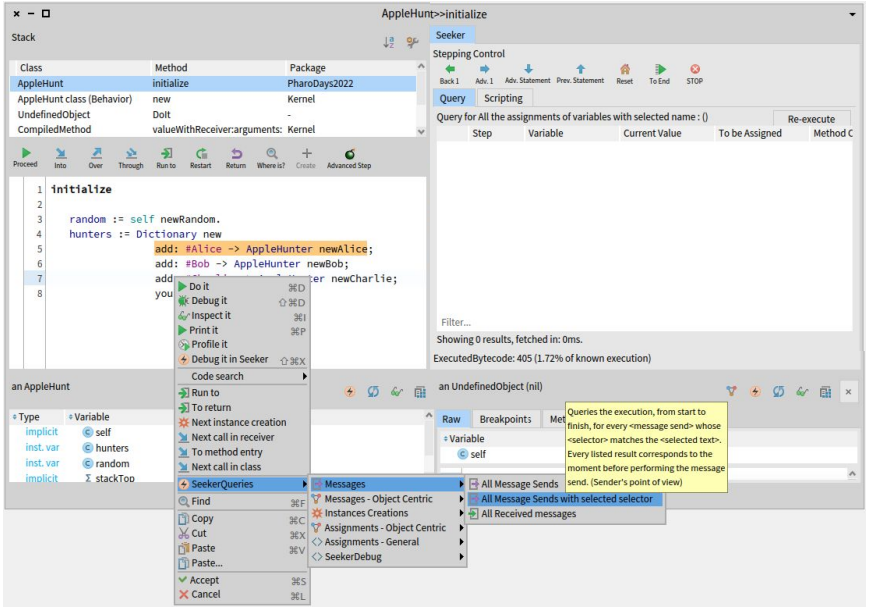
Get the code here: https://github.com/maxwills/PharoDays2022

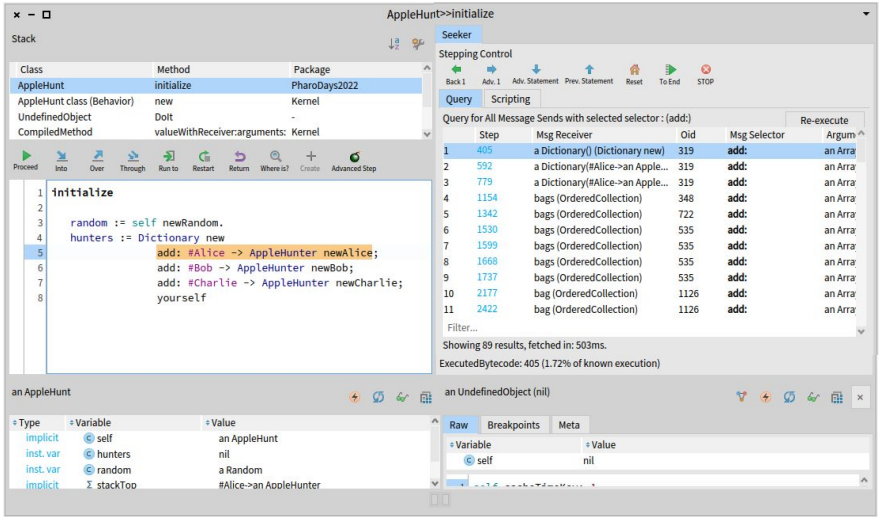
More on Conditions …



Querying the execution

What is happening?



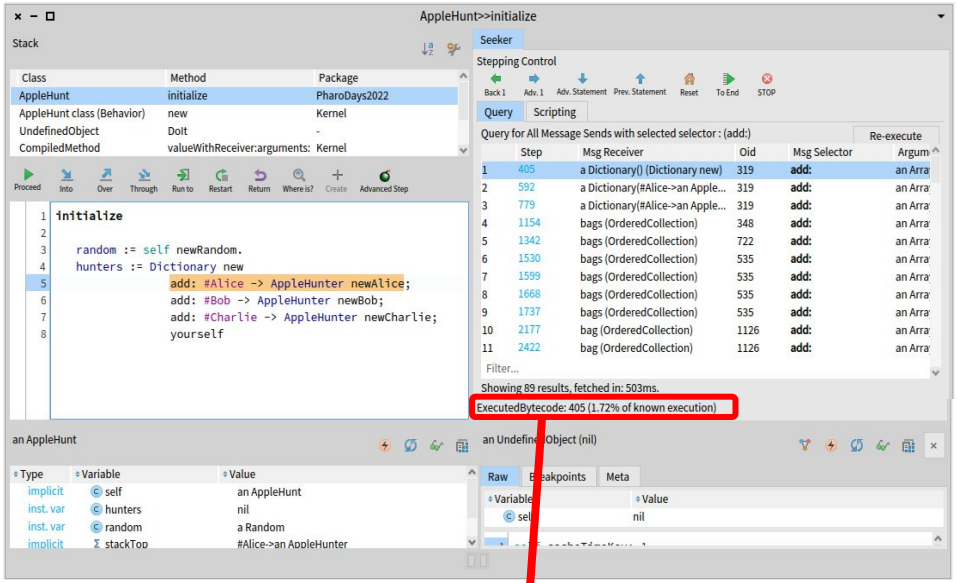


Querying during a debugging session

Results are display

Querying the execution

What is happening?



The developer is currently observing the [execution state 405]

```
^ UserTTQ from: seeker newProgramStates
    select: [ :state| state isMessageSend and: [ state messageSelector = #add: ] ]
    collect: [ :state| MessagesTTQResultItem new
      bytecodeIndex: state bytecodeIndex;
      messageArguments: state messageArguments;
      messageReceiver: state messageReceiver;
      " ... "
      messageSelector: state messageSelector;
      yourself ]
```

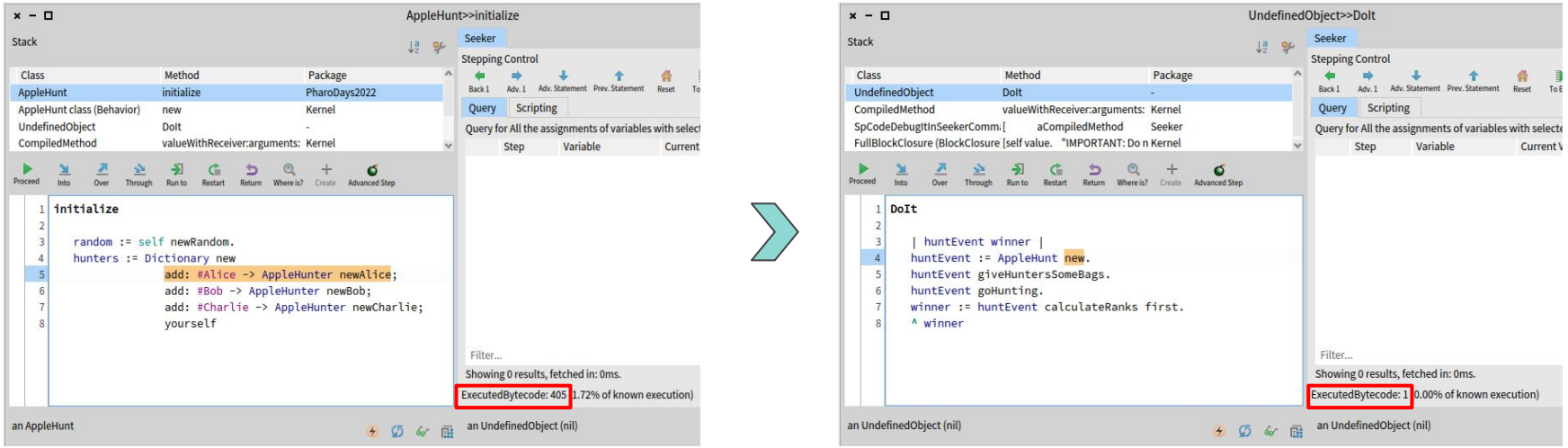
The TTQ to be executed looks like this

Querying the execution

What is happening?

On query activation:

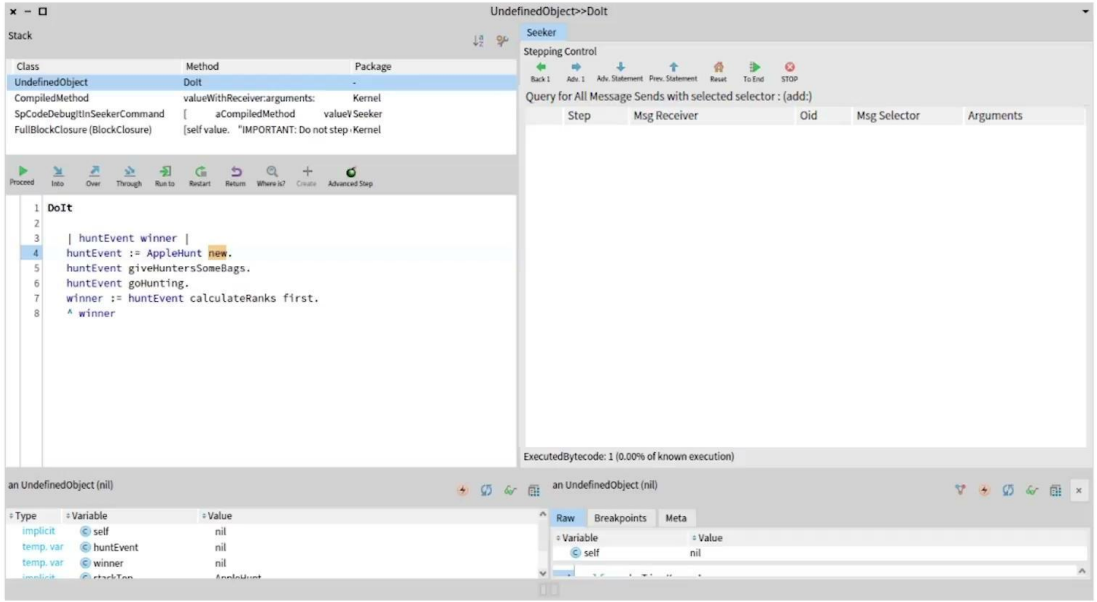
1. The debugger traversing mechanism goes back to [execution state 1].



Querying the execution

What is happening?

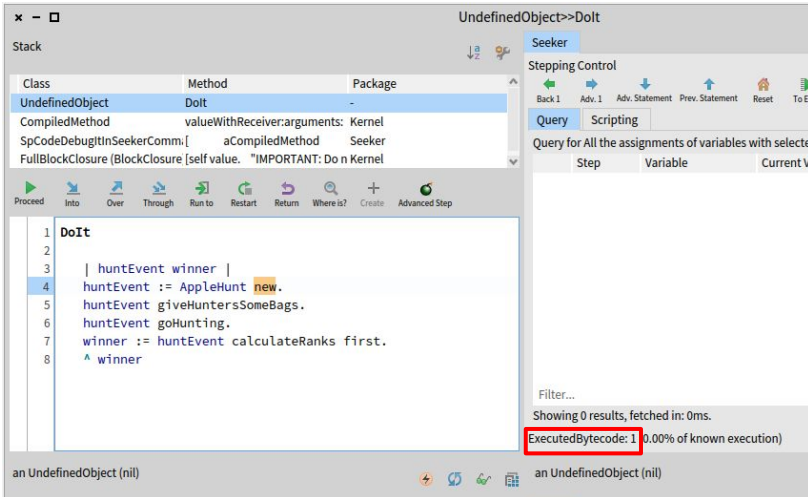
2. The debugger traversing logic is executed, while selecting and collecting relevant data, until the end of the execution.

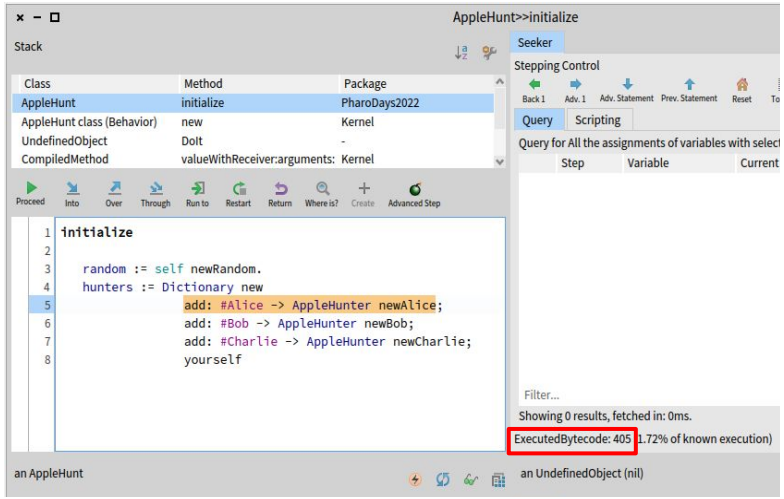


Querying the execution

What is happening?

3. The debugger goes back to the state the developer was observing (Execution state 405).





Querying the execution
What is happening?

(Remember: the developer was observing [execution state 405])

On query activation:

1. The debugger traversing mechanism goes back to [execution state 1].

2. The debugger traversing logic is executed, while selecting and collecting relevant data, until the end of the execution.

3. The debugger goes back to the state the developer was observing.

   All this, happens "*behind doors".*

   The developer doesn't see any stepping.

Installing the code used in the presentation

The code is here:
https://github.com/maxwills/PharoDays2022

In a Pharo10 image, run the following code:

Baseline in the repository.

But first

An anecdote...

(To the code)

Get the code here: https://github.com/maxwills/PharoDays2022

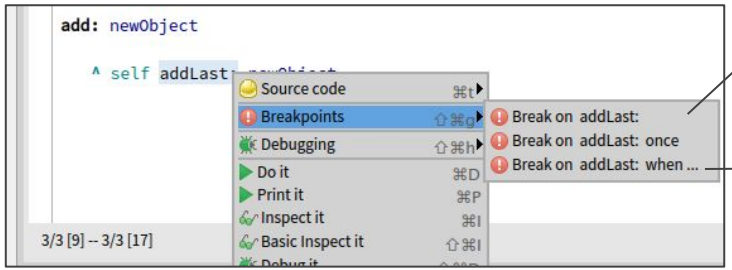
Obtaining Execution Data

Instrumentation

Example: the previous "solutions" but with breakpoints.

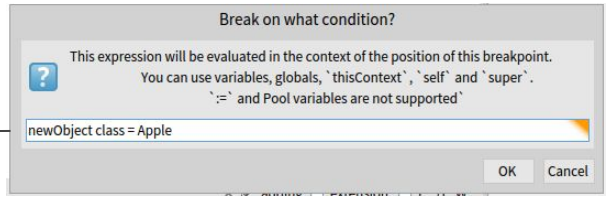
```
add: newObject

    ^ self addLast: newObject
```

Original code



Equivalent to a halt. (Don't put a breakpoint there)

Conditional



Towards a Queryable Debugger

Dissecting The Collection of Execution Data

```
    | callsToAdd |

    callsToAdd := OrderedCollection new.
    seeker doAndUpdateSessionAfter: [
        seeker restartAtBytecodeOne.
        [ seeker canStep ] whileTrue: [
            seeker step.
            (seeker currentState isMessageSend and: [ seeker currentState node selector = #add: ])
ifTrue: [ "This time, instead of logging, we store the data in a collection"
                callsToAdd add: seeker currentState methodAboutToExecute ] ] ].
    ^ callsToAdd
```

Towards a Queryable Debugger

Dissecting The Collection of Execution Data

Cleaning up the code

```
    | callsToAdd |

    callsToAdd := OrderedCollection new.
    seeker doAndUpdateSessionAfter: [
        seeker restartAtBytecodeOne.
        [ seeker canStep ] whileTrue: [
            seeker step.
            (seeker currentState isMessageSend and: [ seeker currentState node selector = #add: ])
ifTrue: [ "This time, instead of logging, we store the data in a collection"
                callsToAdd add: seeker currentState methodAboutToExecute ] ] ].
    ^ callsToAdd
```

We will mask irrelevant code and comments

Towards a Queryable Debugger

Dissecting The Collection of Execution Data

Prepare the storage of the collected results.

```
    | callsToAdd |

    callsToAdd := OrderedCollection new.

        seeker restartAtBytecodeOne.
        [ seeker canStep ] whileTrue: [
            seeker step.
            (seeker currentState isMessageSend and: [ seeker currentState node selector = #add: ])
ifTrue: [
                callsToAdd add: seeker currentState methodAboutToExecute ] ]  .
    ^ callsToAdd
```

Towards a Queryable Debugger

Dissecting The Collection of Execution Data

"Execution Traversing" Logic.

```
    | callsToAdd |

    callsToAdd := OrderedCollection new.

        seeker restartAtBytecodeOne.
        [ seeker canStep ] whileTrue: [
            seeker step.
            (seeker currentState isMessageSend and: [ seeker currentState node selector = #add: ])
ifTrue: [
                callsToAdd add: seeker currentState methodAboutToExecute ] ] .
    ^ callsToAdd
```

"Traversing the execution" logic:

1. Go to the beginning of the execution (restart).
2. Stepping the execution in a loop, until it finishes.

Towards a Queryable Debugger

Dissecting The Collection of Execution Data

"Selecting the interesting states" Logic.

```
    | callsToAdd |

    callsToAdd := OrderedCollection new.

        seeker restartAtBytecodeOne.
        [ seeker canStep ] whileTrue: [
            seeker step.
            (seeker currentState isMessageSend and: [ seeker currentState node selector = #add: ])
ifTrue: [
                callsToAdd add: seeker currentState methodAboutToExecute ] ]  .
    ^ callsToAdd
```

The "selection condition" code evaluates to true or false on each execution state.

Towards a Queryable Debugger

Dissecting The Collection of Execution Data

Addition of some execution data into the result.
The “Collecting” Logic.

```
    | callsToAdd |

    callsToAdd := OrderedCollection new.

        seeker restartAtBytecodeOne.
        [ seeker canStep ] whileTrue: [
            seeker step.
            (seeker currentState isMessageSend and: [ seeker currentState node selector = #add: ])
ifTrue: [
                callsToAdd add: seeker currentState methodAboutToExecute ] ] .
    ^ callsToAdd
```

Towards a Queryable Debugger

Dissecting The Collection of Execution Data

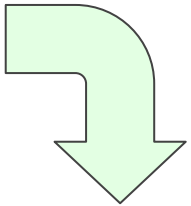
Return the collected results.

```
    | callsToAdd |

    callsToAdd := OrderedCollection new.

        seeker restartAtBytecodeOne.
        [ seeker canStep ] whileTrue: [
            seeker step.
            (seeker currentState isMessageSend and: [ seeker currentState node selector = #add: ])
ifTrue: [
                callsToAdd add: seeker currentState methodAboutToExecute ] ]  .
    ^ callsToAdd
```

From scripts to query notation

(Components mapping)

```
| callsToAdd |

callsToAdd := OrderedCollection new.

   seeker restartAtBytecodeOne.
   [ seeker canStep ] whileTrue: [
      seeker step.
      (seeker currentState isMessageSend and: [ seeker currentState node selector = #add: ])
ifTrue: [
         callsToAdd add: seeker currentState methodAboutToExecute ] ] .
  ^ callsToAdd
```



```
^ (Query from: seeker newProgramStates
      select: [ :state | state isMessageSend and: [ state node selector = #add: ] ]
      collect: [ :state | state methodAboutToExecute ]) asOrderedCollection
```