Formalisation of component based systems

Lars-Åke Fredlund

Explore methods for component specification



- Explore methods for component specification
- Methods will cover at least specification of input/output types and partial descriptions of functional behaviour



- Explore methods for component specification
- Methods will cover at least specification of input/output types and partial descriptions of functional behaviour
- We will mention extensions to specify concurrent behaviour, timing behaviour, and so on (check transparencies for more details)



- Explore methods for component specification
- Methods will cover at least specification of input/output types and partial descriptions of functional behaviour
- We will mention extensions to specify concurrent behaviour, timing behaviour, and so on (check transparencies for more details)
- Methods will be based on formal methods



Formalisations

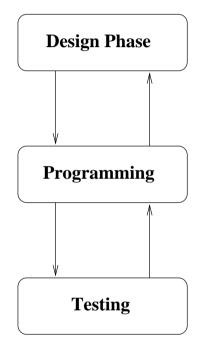
What do we mean by formalisations?

- Components (or systems) specified using formal methods
- Formal Methods are based in *sound mathematics* and provide the ability to *reason* rigourously about programs and specifications
- Reasoning may be automatic or manual



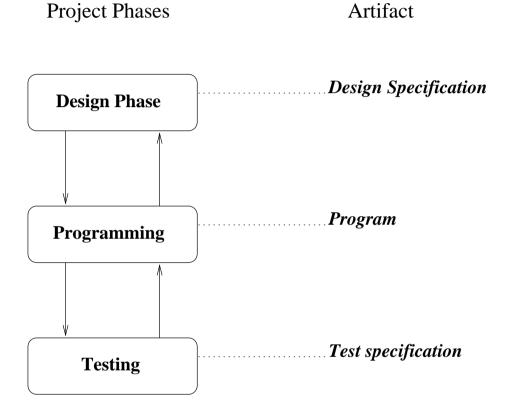
Formal Methods in a Project Design Cycle

Project Phases



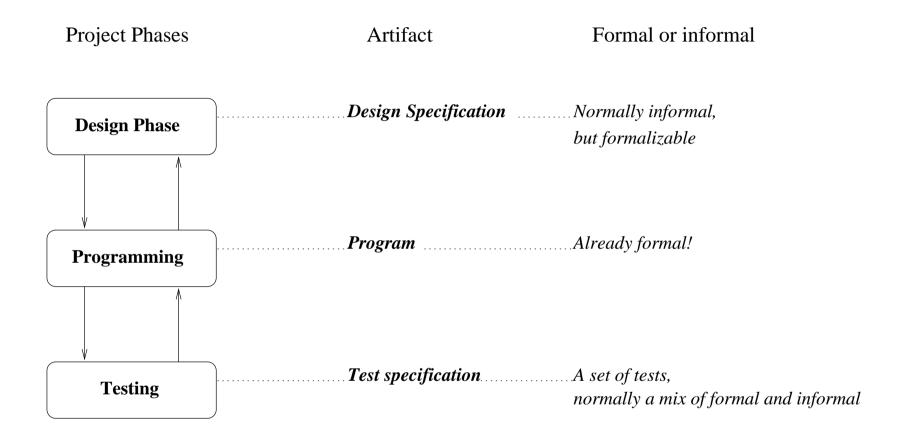


Formal Methods in a Project Design Cycle





Formal Methods in a Project Design Cycle





Example: formalized Test Specifications

Since formal specifications have a precise meaning, we can use them for **analysis**

Test Specification: a formalization of which tests to run, using a formal language



Example: formalized Test Specifications

Since formal specifications have a precise meaning, we can use them for analysis

- Test Specification: a formalization of which tests to run, using a formal language
- Analysis possible measures of how good the testing process is – how big part of the program covered by tests:
 - how large percentage of the *lines* of the program are tested
 - or: how large percentage of the *paths* through the program tested
 - or: how many *states* of the program covered by tests



• We construct *models* of systems



- We construct *models* of systems
- A model *abstracts* from aspects of the real system



- We construct models of systems
- A model *abstracts* from aspects of the real system
- But can be used to *predict* some properties of the real system (e.g., for testing, ...)



- We construct models of systems
- A model *abstracts* from aspects of the real system
- But can be used to *predict* some properties of the real system (e.g., for testing, ...)
- Problem: how to keep the model "synchronised" with the system under development



- We construct models of systems
- A model *abstracts* from aspects of the real system
- But can be used to *predict* some properties of the real system (e.g., for testing, ...)
- Problem: how to keep the model "synchronised" with the system under development
- Many initiatives: (Model-drive architecture (MDA) from OMG)



Non-software models

Building models is standard practice in "normal engineering":



Non-software models

Building models is standard practice in "normal engineering":

A real airplane (Saab Safir):





Non-software models

Building models is standard practice in "normal engineering":

A real airplane (Saab Safir):



assembled by an experienced tester. Reve thentic Kits HOBBY KITS 1960

A 1:72 scale plastic model kit:





Benefits of formalized Design Specification

- Suppose we have a formalized design specification (a system model)
- We can then derive system tests from the formal design specification
- Test metrics become:
 - how big part of the design specification tested
 - That is, which percentage of paths through the design specification tested against the program, and which percentage of design specification states have been tested
- Other uses:
 - Generating a program (implementation) from the design specification
 - Validating/simulating high-level specifications: do they make sense?

Why Study Formalisation (and analysis)?

Formal methods have a bad reputation, but are becoming more and more used

- To document design decisions (UML)
- Concise specifications of tests (QuickCheck)
- In hardware (to verify Intel CPUs)
- In software:
 - Microsoft, Linux: security analysis, device driver safety, ...
 - Static analysis in compilers to detect runtime errors at compile time
- Often works well in limited domain settings (with special rules for how to write software)
- In general very useful techniques for verification of safety critical systems or in situations where failures are very costly
- Important: formal methods are debugging techniques. They cannot guarantee correctness, but can make errors less likely

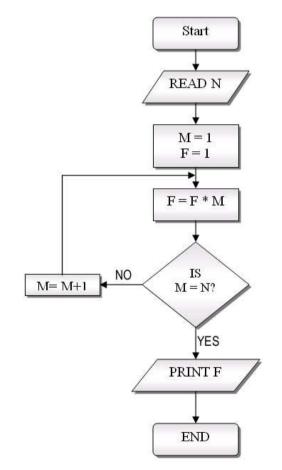
Formalisation

- In general formalisation is a big area we will only introduce the topic here
- Today: *formalisation* examples, following lecture: *analysis* examples



Formalisations& Verification – History

- Turing verified programs
- Floyd: program flowcharts





Central Concepts

Pre- and post-condition

If a pre-condition *pre* holds before a statement, the post-condition *post* holds after



Central Concepts

Pre- and post-condition

If a pre-condition *pre* holds before a statement, the post-condition *post* holds after

Invariant

A property which holds always during an execution Loop invariants are used to characterise loop behaviour



Central Concepts

Pre- and post-condition

If a pre-condition *pre* holds before a statement, the post-condition *post* holds after

Invariant

A property which holds always during an execution Loop invariants are used to characterise loop behaviour

Termination condition

Specifies under which condition a computation terminates Usually proved by providing a *measure* – something which decreases during a computation, but cannot go on decreasing forever



Hoare logic: putting pre-and-post conditions in syntactic form: {*Pre*} *Command* {*Post*}



- Hoare logic: putting pre-and-post conditions in syntactic form: {*Pre*} *Command* {*Post*}
- Example proof rules:



- Hoare logic: putting pre-and-post conditions in syntactic form: {*Pre*} *Command* {*Post*}
- Example proof rules:



- Hoare logic: putting pre-and-post conditions in syntactic form: {*Pre*} *Command* {*Post*}
- Example proof rules:

An example proof:

 $\{i = N \land j = 0\}$ while i>0 do i:=i-1;j:=j+1 $\{j = N\}$



- Hoare logic: putting pre-and-post conditions in syntactic form: {*Pre*} *Command* {*Post*}
- Example proof rules:

$$\{j = N - i\}$$
while i>0 do i:=i-1;j:=j+1 $\{i = 0 \land j = N - i\}$
$$i = N \land j = 0 \supset j = N - i \qquad i = 0 \land j = N - i \supset j = N$$
$$\{i = N \land j = 0\}$$
while i>0 do i:=i-1;j:=j+1 $\{j = N\}$



- Hoare logic: putting pre-and-post conditions in syntactic form: {*Pre*} *Command* {*Post*}
- Example proof rules:

$$\{i > 0 \land j = N - i\} \ i := i - 1; \ j := j + 1 \ \{j = N - i\}$$
$$\{j = N - i\} \text{ while } i > 0 \ \text{do } i := i - 1; \ j := j + 1 \ \{i = 0 \land j = N - i\}$$
$$i = N \land j = 0 \supset j = N - i \qquad i = 0 \land j = N - i \supset j = N$$
$$\{i = N \land j = 0\} \text{ while } i > 0 \ \text{do } i := i - 1; \ j := j + 1 \ \{j = N\}$$



- Hoare logic: putting pre-and-post conditions in syntactic form: {*Pre*} *Command* {*Post*}
- Example proof rules:

$$\models (N - i) + 1 = N - (i - 1)$$

$$\{i > 0 \land j = N - i\} \ i := i - 1; \ j := j + 1 \ \{j = N - i\}$$

$$\{j = N - i\} \text{ while } i > 0 \ \text{do } i := i - 1; \ j := j + 1 \ \{i = 0 \land j = N - i\}$$

$$i = N \land j = 0 \supset j = N - i \qquad i = 0 \land j = N - i \supset j = N$$

$$\{i = N \land j = 0\} \text{ while } i > 0 \ \text{do } i := i - 1; \ j := j + 1 \ \{j = N\}$$



- Dijkstra: weakest pre condition: wp(C)
- Owicki and Gries: extensions to concurrency
- Lamport: proving an invariant *I*:
 - I holds in the initial state of the program
 - For all program statements S prove {I} S {I}
 (if I holds before the statement S, it should afterwards also)

Correctness claims – classical

For classical terminating programs (or functions) a number of properties needs to be proved

- *Partial correctness*: if the program halts it satisfies a property
- *Termination*: the program halts
- **Total correctness:** both Partial correctness and Termination



Correctness claims – reactive systems

- The previous correctness claims are sufficient for terminating programs
- But reactive systems keep on running
- A reactive systems is a system that responds to stimuli (input) and responds with actions (output) – a process
- To formulate correctness properties about reactive systems people started experimenting with *temporal logics*, *program equivalences*, and so on...



Temporal logic

- Pneuli 1977: added discrete and linear time operators to propositional logic, to be able to specify properties of reactive systems
- Program meaning (semantics):
 - a *program state s* maps the program variables to values
 - a *run* of the program is an infinite sequence of program states (s₀, s₁, s₂,...) from an initial state s₀
 - for a terminating system simply add a self-loop in the terminating state to yield an infinite run
 - the semantics of a program p is its set of runs, ||p||
 - If the program is nondeterministic (or accepts input) there will be more than one run of the program



Runs of concurrent programs: examples

Consider the following simple shared variable program:

```
if x>0 then x:=x-1 || if x<3 then x:=x+1
```

where S1 || S2 runs the atomic statements S1 and S2 in parallel



Runs of concurrent programs: examples

Consider the following simple shared variable program:

if x>0 then x:=x-1 || if x<3 then x:=x+1

where S1 || S2 runs the atomic statements S1 and S2 in parallel

Its runs starting from the state x=0 is the **infinite** set:

$$\begin{cases} \langle \mathbf{x} = 0 \rangle \cdot \langle \mathbf{x} = 1 \rangle \cdot \langle \mathbf{x} = 0 \rangle \cdot \dots \\ \langle \mathbf{x} = 0 \rangle \cdot \langle \mathbf{x} = 1 \rangle \cdot \langle \mathbf{x} = 2 \rangle \cdot \langle \mathbf{x} = 1 \rangle \cdot \dots \\ \langle \mathbf{x} = 0 \rangle \cdot \langle \mathbf{x} = 1 \rangle \cdot \langle \mathbf{x} = 2 \rangle \cdot \langle \mathbf{x} = 3 \rangle \cdot \langle \mathbf{x} = 2 \rangle \dots \\ \dots \end{cases}$$

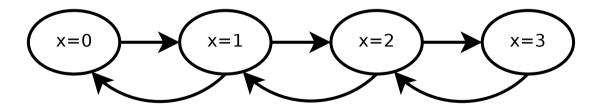


Program runs

We can also depict the runs

$$\left\{ \begin{array}{l} \langle \mathbf{x} = 0 \rangle \cdot \langle \mathbf{x} = 1 \rangle \cdot \langle \mathbf{x} = 0 \rangle \cdot \dots \\ \langle \mathbf{x} = 0 \rangle \cdot \langle \mathbf{x} = 1 \rangle \cdot \langle \mathbf{x} = 2 \rangle \cdot \langle \mathbf{x} = 1 \rangle \cdot \dots \\ \langle \mathbf{x} = 0 \rangle \cdot \langle \mathbf{x} = 1 \rangle \cdot \langle \mathbf{x} = 2 \rangle \cdot \langle \mathbf{x} = 3 \rangle \cdot \langle \mathbf{x} = 2 \rangle \dots \end{array} \right\}$$

as a state graph:





Consider the parallel program

```
if x>0 then x:=x-1 || if x>0 then x:=x-1
```

with the starting state $\langle x = 3 \rangle$



Consider the parallel program

```
if x>0 then x:=x-1 || if x>0 then x:=x-1
```

with the starting state $\langle x = 3 \rangle$

• If statements are **atomic** the program has the single run $\langle x = 3 \rangle \langle x = 2 \rangle \langle x = 1 \rangle \langle x = 0 \rangle$



Consider the parallel program

```
if x>0 then x:=x-1 || if x>0 then x:=x-1
```

with the starting state $\langle x = 3 \rangle$

- If statements are atomic the program has the single run $\langle x = 3 \rangle \langle x = 2 \rangle \langle x = 1 \rangle \langle x = 0 \rangle$
- If statements are not atomic the program has the two runs

$$\left\{ \begin{array}{l} \langle \mathbf{x} = 3 \rangle \cdot \langle \mathbf{x} = 2 \rangle \cdot \langle \mathbf{x} = 1 \rangle \cdot \langle \mathbf{x} = 0 \rangle \\ \langle \mathbf{x} = 3 \rangle \cdot \langle \mathbf{x} = 2 \rangle \cdot \langle \mathbf{x} = 1 \rangle \cdot \langle \mathbf{x} = 0 \rangle \cdot \langle \mathbf{x} = -1 \rangle \end{array} \right\}$$



Consider the parallel program

if x>0 then x:=x-1 || if x>0 then x:=x-1

with the starting state $\langle x = 3 \rangle$

- If statements are atomic the program has the single run $\langle x = 3 \rangle \langle x = 2 \rangle \langle x = 1 \rangle \langle x = 0 \rangle$
- If statements are not atomic the program has the two runs

$$\left\{ \begin{array}{l} \langle \mathbf{x} = 3 \rangle \cdot \langle \mathbf{x} = 2 \rangle \cdot \langle \mathbf{x} = 1 \rangle \cdot \langle \mathbf{x} = 0 \rangle \\ \langle \mathbf{x} = 3 \rangle \cdot \langle \mathbf{x} = 2 \rangle \cdot \langle \mathbf{x} = 1 \rangle \cdot \langle \mathbf{x} = 0 \rangle \cdot \langle \mathbf{x} = -1 \rangle \end{array} \right\}$$

Conclusion: programming concurrent programs is hard



Consider the parallel program

if x>0 then x:=x-1 || if x>0 then x:=x-1

with the starting state $\langle x = 3 \rangle$

- If statements are atomic the program has the single run $\langle x = 3 \rangle \langle x = 2 \rangle \langle x = 1 \rangle \langle x = 0 \rangle$
- If statements are not atomic the program has the two runs

$$\left\{ \begin{array}{l} \langle \mathbf{x} = 3 \rangle \cdot \langle \mathbf{x} = 2 \rangle \cdot \langle \mathbf{x} = 1 \rangle \cdot \langle \mathbf{x} = 0 \rangle \\ \langle \mathbf{x} = 3 \rangle \cdot \langle \mathbf{x} = 2 \rangle \cdot \langle \mathbf{x} = 1 \rangle \cdot \langle \mathbf{x} = 0 \rangle \cdot \langle \mathbf{x} = -1 \rangle \end{array} \right\}$$

- Conclusion: programming concurrent programs is hard
- The Erlang programming language we shall hear about soon has good concurrency primitives

Temporal logic – operators

Classical linear temporal operators (defined over runs):

• Always ϕ

 ϕ holds in all future states of the run

• Eventually ϕ

 ϕ holds in some future state of the run

Next φ

. . .

 ϕ holds in the next state

• ϕ_1 Until ϕ_2

 ϕ_1 holds in all states until ϕ_2 holds

- And the normal ones: negation $\neg \phi$, conjunction $\phi_1 \land \phi_2$, implication $\phi_1 \supset \phi_2, \ldots$
- And propositional operators (over variables): x < y, even(z),



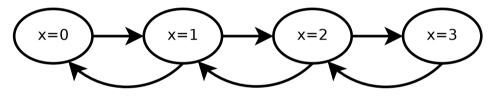
Temporal logic – meaning

- A program p satisfies a formula d when all the runs of the program are satisfied by the formula
- The logic is linear because it doesn't talk about the branching structure of the state graph of the program (what is set of the possible next states of the program)
- So called *branching time* logics do consider the branching structure of the state graph of the program



Consider the atomic parallel program

if
$$x>0$$
 then $x:=x-1 \mid \mid x<3$ then $x:=x+1$

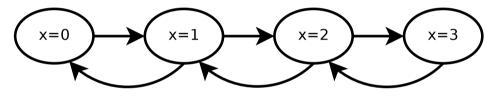




Consider the atomic parallel program

if
$$x>0$$
 then $x:=x-1 \mid \mid x<3$ then $x:=x+1$

with the starting state $\langle x = 3 \rangle$ and the state graph

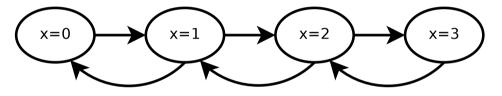


• Does Always $x \ge 0$ hold?



Consider the atomic parallel program

if
$$x>0$$
 then $x:=x-1 \mid \mid x<3$ then $x:=x+1$

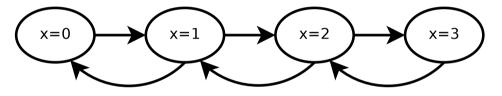


- Does Always $x \ge 0$ hold?
- **Yes**; if x=0 then the guard prevents further decrease



Consider the atomic parallel program

if
$$x>0$$
 then $x:=x-1 \mid \mid x<3$ then $x:=x+1$

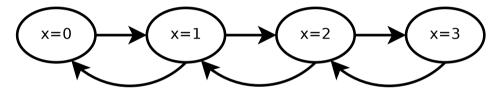


- Does Always $x \ge 0$ hold?
- **Yes**; if x=0 then the guard prevents further decrease
- Does Always ($x = 3 \supset$ Eventually x = 0) hold?



Consider the atomic parallel program

if
$$x>0$$
 then $x:=x-1 \mid \mid x<3$ then $x:=x+1$



- Does Always $x \ge 0$ hold?
- **Yes**; if x=0 then the guard prevents further decrease
- Does Always ($x = 3 \supset$ Eventually x = 0) hold?
- **No**; there is a run $\langle x = 3 \rangle \cdot \langle x = 2 \rangle \cdot \langle x = 3 \rangle \cdot \langle x = 2 \rangle \cdot \ldots$



• Eventually $\phi \equiv \neg$ Always $(\neg \phi)$



- Eventually $\phi \equiv \neg$ Always $(\neg \phi)$
- A safety property expresses that something bad φ never happens: Always ¬ φ



- Eventually $\phi \equiv \neg$ Always $(\neg \phi)$
- A safety property expresses that something bad φ never happens: Always ¬ φ
- A *liveness property* expresses that something good ϕ eventually happens: *Eventually* ϕ



- Eventually $\phi \equiv \neg$ Always $(\neg \phi)$
- A safety property expresses that something bad φ never happens: Always ¬ φ
- A *liveness property* expresses that something good ϕ eventually happens: *Eventually* ϕ
- Often one uses fairness assumptions to rule out bad program behaviours

 ϕ eventually holds under the assumption that ψ doesn't always hold:

 $(\neg A l ways \psi) \supset (E ventually \phi)$



Temporal logic

- Many variants exists:
 - Continuous time, time interval models
 - Additional operators (e.g., talking about past time)
 - Branching time logics: $\mathcal{A} \phi \phi$ holds on all program states that are reachable from the current one
- In general temporal logics are good for expressing program correctness properties
- But it is difficult to give *complete* descriptions of what a correct program behaviour is
- More on algorithms and tools for checking programs against properties later on...



Specification Languages

- A specification is an abstract definition of what the correct behaviour of a program is
- A specification abstracts away from irrelevant details
- Verification techniques (later) permits to check whether a program satisfies its specification
- Two large families of specification languages:
 - abstract model based (Z, VDM, Slam-SL)
 - and *algebraic* specifications (CCS, π calculus)

An model based language: Z

- Based on mathematical logic (first-order predicate logic) and set theory
- Objects are typed
- Schemas permits to specify in a modular fashion both static (data) and dynamic (behaviour) properties of a system
- The acceptable data states of a program are specified using predefined mathematical concepts (sequences, sets, ...)
- Operations upon states are characterised using pre- and post-conditions



The Z specification language

Different Schema Types:

- State schemas characterises reachable states of a system using a state invariant; defines what must be respected by operations
- Operation schemas describe how operations with input parameters cause state changes, and return parameters
- Operation schemas are given using pre- and post-conditions

Z schema example: static part

Result ::= ok | nospace







Z schema example, dynamic part

 $\begin{array}{c} \textit{InsertOk} \\ \Delta \textit{Queue} \\ \textit{insert?} : \mathbb{N} \\ \textit{result!} : \textit{Result} \\ \#\textit{queue} < 10 \\ \textit{queue'} = \textit{queue} \land \langle \textit{insert?} \rangle \\ \textit{result!} = \textit{ok} \\ \end{array} \\ \hline \begin{array}{c} \textit{InsertWithError} \\ \Xi \textit{Queue} \\ \textit{insert?} : \mathbb{N} \end{array} \end{array}$

result! : Result

#queue ≥ 10 result! = nospace

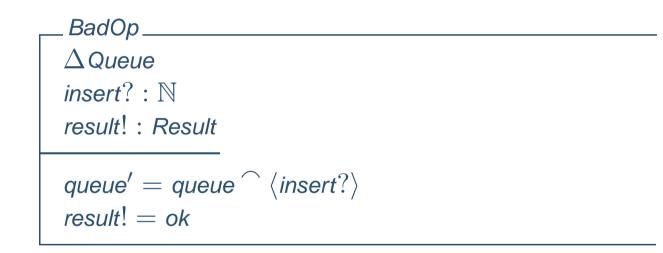
Insert = *InsertOk* ∨ *InsertWithError*



Abstract models – Z

Proof challenges (for a theorem prover):

- An initial state exists
- The pre-condition of each operation guarantees that the resulting state exists (is a proper state). Consider:



Violates the state invariant



Z limitations and extensions

- Serious limitations with respect to:
 - Lack of modularisation and Object Orientation
 - Specifying reactive systems, realtime, concurrency
- Has been used in quite a few industrial projects (in the UK)
- A mathematical clean notation, expressive, but rather abstract (can be difficult to implement)
- To implement a Z spec one can use program derivation and program refinement techniques (in Z itself)
- Object-oriented extensions exists: Object-Z and Z++
- Nowadays the B method receives more attention



Axiomatic Specifications

- A system is specified as a set of abstract data types
- Operations on data are characterised as axioms of equality
- Rewrite rules define transitions between system states (instances of the data types)
- Specifications are executable
- Examples: Maude/rewriting logic, OBJ, FOOPS



Maude

- Home page: http://maude.cs.uiuc.edu/
- Origin at Stanford, USA (Messeguer and others, big Spanish community)
- Rewriting of equations modulo commutativity, associativity and idem-potency
- Equations are evaluated nondeterministically, in parallel
- Permits specification of reactive (and concurrent) systems
- Reflexive language (Maude can be represented in Maude) and object-oriented (inheritance, polymorphism)
- Uses reflection to control which equations and rewrite rules to apply (rewriting strategies)



Maud example

```
fmod NatQueue is
```

```
sorts NatQueue .
```

```
protecting NAT .
```

```
op empty : -> NatQueue [ctor] .
op enq : Nat NatQueue -> NatQueue [ctor] .
op head : NatQueue -> Nat .
op deq : NatQueue -> NatQueue .
```

```
vars N N1 : Nat . vars Q : NatQueue .
```

```
eq head(enq(N,empty)) = N .
eq head(enq(N,enq(N1,Q))) = head(enq(N1,Q)) .
```

```
eq deq(enq(N,empty)) = empty .
eq deq(enq(N,enq(N1,Q))) = enq(N,deq(enq(N1,Q))) .
endfm
```

Usage examples

What is the head of the queue after inserting 3 and then 2? red head(enc(2,enc(3,empty))).



Usage examples

What is the head of the queue after inserting 3 and then 2? red head(enc(2,enc(3,empty))).

Answer: result NzNat: 3



Maud example: with rewrite rules

Rewrite rules express transitions between states:

```
mod CommChannel is
  protecting NatOueue .
  vars N : Nat .
  vars O : NatOueue .
  rl [receive] :
       enq(N,Q) => deq(enq(N,Q)).
  rl [loose_msq] :
       enq(N, O) => O.
  rl [duplicate_msg] :
       enq(N,Q) => enq(N,enq(N,Q)).
endm
```



Usage examples

• What are the possible queues resulting after at most two transitions from the state $2 \cdot 1$?

```
search [,2] enq(2,enq(1,empty)) =>+ q:NatQueue .
```

Answer: 11 states:

 $\begin{array}{cccccccc} \epsilon & & & & 1 \\ 2 & 1 & & & & \\ 2 \cdot 2 & & & 1 \cdot 1 & & 2 \cdot 1 \\ 2 \cdot 2 \cdot 1 & & & 2 \cdot 1 \cdot 1 & & \\ 2 \cdot 2 \cdot 2 \cdot 1 & & & 2 \cdot 2 \cdot 1 \cdot 1 & & 2 \cdot 1 \cdot 1 \cdot 1 \end{array}$

• Path to $2 \cdot 1 \cdot 1 \cdot 1$:

Maude> show path 10 .
state 0, NatQueue: enq(2, enq(1, empty))
===[rl enq(N, Q) => enq(N, enq(N, Q)) [label duplicate_msg] .]===>
state 4, NatQueue: enq(2, enq(1, enq(1, empty)))
===[rl enq(N, Q) => enq(N, enq(N, Q)) [label duplicate_msg] .]===>
state 10, NatQueue: enq(2, enq(1, enq(1, enq(1, empty))))

Maude conclusions

- Useful for developing executable prototypes rapidly
- Yields reasonable efficient prototypes
- But specifications become pretty algorithmic (not abstract)

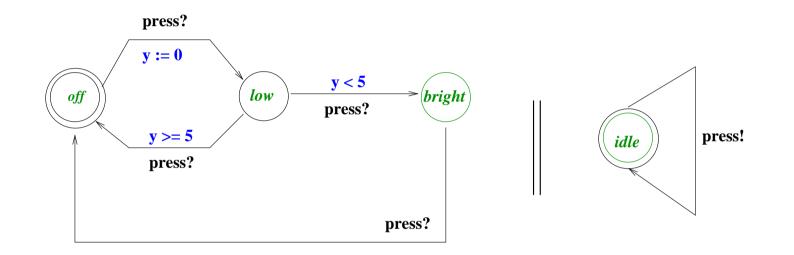


Specifying real-time and hybrid systems

- Specification languages generally tailored to the task of constructing verifyable models
- Typical system: network of clocked automata
- Languages and verification systems: UPPAAL (http://www.uppaal.com)
- Alternatives: modelling using real-time UML variants

A Lamp Example in UPPAAL

A lamp that has two intensities (low and bright), and a user:

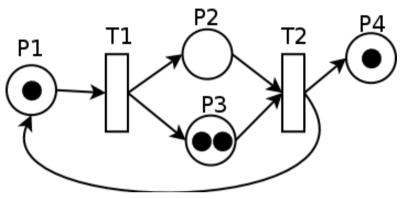


If the user presses the button (press!) twice within 5 time units the intensity of the lamp is set to bright



Formalisation of Concurrency and Distribution

Petri-Nets: a mainly graphical notation for transition systems:



Nondeterministic, highly concurrent Intuitive with a good notation, but do they scale?

State machines that communicate by exchanging messages: SDL, Promela, I/O-automatons, cleanly written Erlang, ...

Often serious specification languages useful for checking complex systems

Process algebras: CCS, CSP, π-calculus, ambient calculus Mathematically clean formalisms, often less suited for specifying larger systems

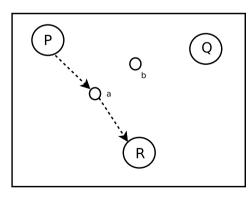
Process Algebras

- A large variant of calculi for providing a mathematically elegant theory of concurrency
 - For basic concerns: CCS, CSP
 - With extensions to mobility: π -calculus
 - With extensions to distribution: mobile ambients
 - With abstract data types: LOTOS, μ CRL
 - With (many) extensions to real-time, to stochastic behaviours, to ...
- As a summary: popular as a (mathematical) tool for reasoning about concurrency; for real programs?
- Due to their popularity one should have a basic knowledge of the field



Process Algebras: CCS

- CCS is a very basic process algebra (due to Milner 1980)
- Basic entities are processes and ports (used for binary communication between two processes)
- We let *P*, *Q*, *R*, ... stand for processes and *a*, *b*, *c*, ... for ports
- A process/port graph:





CCS syntax

Communication by synchronisation: (a is a port) output action a(v) input action a(x) or internal action r

If there is no value being sent or received we omit the action parameter: $\overline{a}(v)$ becomes \overline{a} and a(x) becomes a

- Sequential composition: α.P where α is an action (input or output action, or internal τ)
 Behaviour: after performing a it behaves as P
- Choice: P + Q can behave as P or as Q
- Parallel behaviour: the agent P | Q behaves as P running in parallel with Q
- The agent which can do nothing: 0

A simple example: a coffee/tea machine

A simple (one-use) coffee machine:

 $coin.\left(\overline{coffee}.0+\overline{tea}.0\right)$



A simple example: a coffee/tea machine

A simple (one-use) coffee machine:

 $coin. \left(\overline{coffee}.0 + \overline{tea}.0\right)$

• A user that always wants tea:

coin.tea.0



A simple example: a coffee/tea machine

A simple (one-use) coffee machine:

 $coin. \left(\overline{coffee}.0 + \overline{tea}.0\right)$

• A user that always wants tea:

coin.tea.0

The combination of a user and the coffee machine:

 $coin. \left(\overline{coffee}.0 + \overline{tea}.0\right) \mid \overline{coin}.tea.0$



Process Algebras: CCS operators part II

- Restriction: the agent $P \setminus \{a\}$ cannot communicate with the environment using the port *a*
- Relabelling: in communications with its environment the agent P[f] relabels all channel names using the relabelling function f
- Recursive agents can be defined: $P \stackrel{\text{def}}{=} a.(P \mid b.0)$
- And simple test on boolean conditions: if b then P else Q



A simple example: a coffee/tea machine continue

A simple coffee machine:

$$CM1 \stackrel{\text{def}}{=} coin. \left(\overline{coffee}.CM1 + \overline{tea}.CM1\right)$$

A user that always wants tea:

User $\stackrel{\text{def}}{=}$ $\overline{coin}.tea.0$

The combination of a user and the coffee machine: $(User \mid CM1) \setminus \{coffee, tea, coin\}$



Defining the Behaviour of CCS Agents

- A transition rule based semantics, defined using the syntactic shape of terms, is often called a structured operational semantics (abbreviated SOS)
- Such semantics are used to define the meaning of many programming languages and systems
- One should at least have a basic grasp of how to read such semantic definitions
- We use an operational semantics to define the behaviour of CCS agents



CCS, Operational behaviour

Semantics defined by transition rules:

• prefix
$$\frac{}{\alpha.P \xrightarrow{\alpha} P'}$$
• choice $\frac{P \xrightarrow{\alpha} P'}{P+Q \xrightarrow{\alpha} P'}$, $\frac{P \xrightarrow{\alpha} P'}{Q+P \xrightarrow{\alpha} P'}$
• interleaving $\frac{P \xrightarrow{\alpha} P'}{P \mid Q \xrightarrow{\alpha} P' \mid Q}$, $\frac{P \xrightarrow{\alpha} P'}{Q \mid P \xrightarrow{\alpha} Q \mid P'}$
• synchronisation, $\frac{P \xrightarrow{a(x)} P' \quad Q \xrightarrow{\overline{a}(v)} Q'}{P \mid Q \xrightarrow{\tau} P'[v/x] \mid Q'}$
• synchronisation, $\frac{Q \xrightarrow{a(x)} Q' \quad P \xrightarrow{\overline{a}(v)} P'}{P \mid Q \xrightarrow{\tau} P'[v/x] \mid Q'}$



CCS, Operational behaviour part II

• restriction
$$\frac{P \xrightarrow{\alpha} P' \quad a \notin fn(\alpha)}{P \setminus \{a\} \xrightarrow{\alpha} P' \setminus \{a\}}$$
where
$$\begin{bmatrix} fn(\alpha) \equiv \begin{cases} \{a\} & \text{if } \alpha = \overline{a}(v) \\ \{a\} & \text{if } \alpha = a(x) \\ \{\} & \text{if } \alpha = \tau \end{cases}$$
• relabelling
$$\frac{P \xrightarrow{\alpha} P'}{P[f] \xrightarrow{f(\alpha)} P'[f]}$$
• recursion
$$\frac{P[v/x] \xrightarrow{\alpha} P' \quad N(x) \stackrel{\text{def}}{=} P}{N(v) \xrightarrow{\alpha} P'}$$
• eval
$$\frac{P \xrightarrow{\alpha} P' \quad b \ true}{if \ b \ then \ P \ else \ Q \xrightarrow{\alpha} P'}, \quad \frac{P \xrightarrow{\alpha} P' \quad b \ false}{if \ b \ then \ Q \ else \ P \xrightarrow{\alpha} P'}$$



• Coffee machine 1 always work: $CM1 \stackrel{\text{def}}{=} coin. \left(\overline{coffee}.CM1 + \overline{tea}.CM1\right)$



- Coffee machine 1 always work: $CM1 \stackrel{\text{def}}{=} coin. \left(\overline{coffee}.CM1 + \overline{tea}.CM1\right)$
- Coffee machine 2 sometimes (unspecified when) only allows the coffee choice:
 - $CM2 \stackrel{\text{def}}{=} coin. \left(\overline{coffee}.CM2 + \overline{tea}.CM2 + \tau.\overline{coffee}.CM2\right)$



- Coffee machine 1 always work: $CM1 \stackrel{\text{def}}{=} coin. \left(\overline{coffee}.CM1 + \overline{tea}.CM1\right)$
- Coffee machine 2 sometimes (unspecified when) only allows the coffee choice:

 $CM2 \stackrel{\text{def}}{=} coin. \left(\overline{coffee}.CM2 + \overline{tea}.CM2 + \tau.\overline{coffee}.CM2\right)$

A typical user that always wants tea:
 User ^{def} = coin.tea.0



- Coffee machine 1 always work: $CM1 \stackrel{\text{def}}{=} coin. \left(\overline{coffee}.CM1 + \overline{tea}.CM1\right)$
- Coffee machine 2 sometimes (unspecified when) only allows the coffee choice:

 $CM2 \stackrel{\text{def}}{=} coin. \left(\overline{coffee}.CM2 + \overline{tea}.CM2 + \tau.\overline{coffee}.CM2\right)$

- A typical user that always wants tea: $User \stackrel{\text{def}}{=} \overline{coin}.tea.0$
- The combination of a user and machine 1 (let $S \equiv \{coffee, tea, coin\}$):

 $(\textit{User} \mid \textit{CM1}) \setminus \texttt{S} \xrightarrow{\tau} \xrightarrow{\tau} (0 \mid \textit{CM1}) \setminus \texttt{S}$



- Coffee machine 1 always work: $CM1 \stackrel{\text{def}}{=} coin. \left(\overline{coffee}.CM1 + \overline{tea}.CM1\right)$
- Coffee machine 2 sometimes (unspecified when) only allows the coffee choice:

 $CM2 \stackrel{\text{def}}{=} coin. \left(\overline{coffee}.CM2 + \overline{tea}.CM2 + \tau.\overline{coffee}.CM2\right)$

- A typical user that always wants tea:
 User ^{def} = coin.tea.0
- The combination of a user and machine 1 (let S ≡ {coffee, tea, coin}):

 $(\textit{User} \mid \textit{CM1}) \setminus \texttt{S} \xrightarrow{\tau} \xrightarrow{\tau} (0 \mid \textit{CM1}) \setminus \texttt{S}$

The combination of a user and machine 2 may deadlock after two machine steps:

 $(User \mid CM2) \setminus S \xrightarrow{\tau} \xrightarrow{\tau} (tea.0 \mid \overline{coffee}.CM2) \setminus S$

Specifications for Process Algebras

Suppose that we have written a complex agent *P*, and we want to develop a simpler specification for that agent. What can we do?

Two options:

- Write the specification as a temporal logic formula φ, and show that P : φ (P satisfies φ)
- Write the specification as another CCS agent S, and show that P = S, with regards to some notion of equality "="



Process algebra – process equality

Crucial question: when do two processes *P* and *Q* exhibit the same behaviour?

- First question: what does it mean for P and Q to have the same behaviour?
- Do we require that they have (almost) the same set of traces? In practise this is often too weak, e.g., CM1 and (a slight variant of) CM2 have the same set of traces but have very different behaviour
- Or do we need a stronger notion of equivalence? There are many options out there: strong equivalence, observation equivalence, ...



Proving processes equal

- Most algebras have an axiomatic theory, e.g., a set of equations of the type P + Q = Q + P, $P \mid 0 = P$ and so on...
- Hence two processes P and Q are equal if we can prove P = Q using the axioms
- A more behavioural alternative is to find a *bisimulation relation* relating *P* and Q
- Two process P and Q are (strong bisimulation) equivalent if we can find a bisimulation relation S containing the pair (P, Q)

A pair $(P, Q) \in S$ if and only if

- If P → P' for some α and P' then there exists a Q' such that Q → Q', and (P', Q') ∈ S
- If Q → Q' for some α and Q' then there exists a P' such that P → P', and (P', Q') ∈ S

Often it is far easier to find a bisimulation relation than to use equational reasoning



π calculus

- CCS is a fairly static calculus what if we allow names (channels) to be communicated?
- The result is the π calculus (Milner, Walker and Parrow 1989)
- A process that receives a new name can later communicate using it (new communication capabilities arise during the execution)
- The distinction between channels and data is removed
- A very basic calculus (but expressive!) for experimenting with one form of mobility
- Nowadays very popular inspiration for some standards proposals for composition languages of web services



π calculus operators

- Most operators come from CCS: 0 the inactive process, choice P + Q, parallelism P | Q
- Communication primitives are different:
 - output: x
 y.P: the name y is sent over the name x; then behaves as P
 - input: x(w).P: a name y is received on the channel x; then behaves as P[y/w] (P with y substituted for w)
- Matching: [x = y]P behaves as *P* if *x* is the same name as *y*, otherwise as 0
- Private names: (x)P creates a new name x that is private to P
- Replication: !P is equivalent to !P | P (an infinite number of copies of P in parallel)

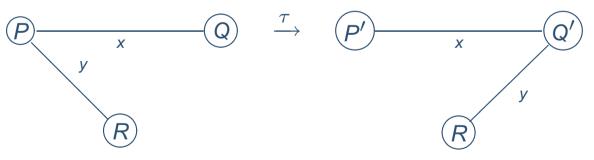


π calculus: name mobility

Mobility example:

Receive a new name at *a* and use the new name to send *z*: a(x). $\overline{x} z$. *P*

- Evolution of communication capabilities, let: $P(x,y) \stackrel{\text{def}}{=} \overline{x} y \cdot P'(x) \text{ and } Q(x) \stackrel{\text{def}}{=} x(z) \cdot Q'(x,z)$
- The following action is enabled $P(x, y) \mid Q(x) \mid R(y) \xrightarrow{\tau} P'(x) \mid Q'(x, z)[y/z] \mid R(y)$
- Evolution of communication capabilities depicted graphically:





An encoding of True and False:

False(x) $\stackrel{\text{def}}{=}$!(query, false, true) \overline{x} query. \overline{query} false. \overline{query} true. $\overline{false.0}$ True(x) $\stackrel{\text{def}}{=}$!(query, false, true) \overline{x} query. \overline{query} false. \overline{query} true. $\overline{true.0}$



An encoding of True and False:

 $False(x) \stackrel{\text{def}}{=} !(query, false, true) \overline{x} query.\overline{query} false.\overline{query} true.\overline{false.0}$ $True(x) \stackrel{\text{def}}{=} !(query, false, true) \overline{x} query.\overline{query} false.\overline{query} true.\overline{true.0}$

Lets define a process P(x) that behaves as P₁ if its argument x represents true and P₂ if it represents false:

 $P(x) \stackrel{\text{def}}{=} x(query).query(false).query(true).(true.P_1 + false.P_2)$



An encoding of True and False:

False(x) $\stackrel{\text{def}}{=}$!(query, false, true) \overline{x} query. \overline{query} false. \overline{query} true. $\overline{false.0}$ True(x) $\stackrel{\text{def}}{=}$!(query, false, true) \overline{x} query. \overline{query} false. \overline{query} true. $\overline{true.0}$

Lets define a process P(x) that behaves as P₁ if its argument x represents true and P₂ if it represents false:

 $P(x) \stackrel{\text{def}}{=} x(query).query(false).query(true).(true.P_1 + false.P_2)$

If we execute P(x) | True(x) we will eventually end up in the new state $P_1 | 0 | True(x) = P_1 | True(x)$



An encoding of True and False:

 $False(x) \stackrel{\text{def}}{=} !(query, false, true) \overline{x} query.\overline{query} false.\overline{query} true.\overline{false}.0$ $True(x) \stackrel{\text{def}}{=} !(query, false, true) \overline{x} query.\overline{query} false.\overline{query} true.\overline{true}.0$

Lets define a process P(x) that behaves as P₁ if its argument x represents true and P₂ if it represents false:

 $P(x) \stackrel{\text{def}}{=} x(query).query(false).query(true).(true.P_1 + false.P_2)$

- If we execute P(x) | True(x) we will eventually end up in the new state $P_1 | 0 | True(x) = P_1 | True(x)$
- The syntax is ugly; it is better in the polyadic π -calculus:

 $False(x) \stackrel{\text{def}}{=} !(false, true) \overline{x} \langle false, true \rangle . \overline{false}.0$ $P(x) \stackrel{\text{def}}{=} x(false, true). (true.P_1 + false.P_2)$

$\pi\text{-calculus transition rules}$

act
$$\frac{P \xrightarrow{\alpha} P'}{\alpha \cdot P \xrightarrow{\alpha} P}$$
 sum $\frac{P \xrightarrow{\alpha} P'}{P + Q \xrightarrow{\alpha} P'}$
par $\frac{P \xrightarrow{\alpha} P'}{P \mid Q \xrightarrow{\alpha} P' \mid Q}$ repl $\frac{P \mid P \xrightarrow{\alpha} P'}{P \mid Q \xrightarrow{\alpha} P' \mid Q}$
equiv $\frac{Q \xrightarrow{\alpha} Q' \quad P \equiv Q \quad P' \equiv Q'}{P \xrightarrow{\alpha} P'}$

The rules uses the congruence \equiv which is defined:

- $\blacksquare P \mid Q == Q \mid P$
- $\blacksquare P + Q \equiv Q + P$
- $\blacksquare [x = x]P \equiv P$
- if $A(x) \stackrel{\text{def}}{=} P'$ then $A(y) \equiv P'[y/x]$
- $P \equiv Q$ if *P* and *Q* are α -equivalent, i.e., only bound variables are different, e.g., $(x)\overline{y}x.0 \equiv (z)\overline{y}z.0$

Transition rules, part II

$$l\text{-com } \underbrace{ \begin{array}{c} P \xrightarrow{\overline{x} \, y} P' & Q \xrightarrow{x(z)} Q' \\ \hline P \mid Q \xrightarrow{\tau} P' \mid Q'[y/z] \end{array} }_{P \mid Q \xrightarrow{\tau} P' \mid Q'[y/z]}$$

close
$$\xrightarrow{P \xrightarrow{\overline{x}(y)} P'} Q \xrightarrow{x(y)} Q'$$

 $\xrightarrow{P \mid Q \xrightarrow{\alpha} (y)(P' \mid Q')}$

res
$$\xrightarrow{P \xrightarrow{\alpha} P' \quad y \notin n(\alpha)}{(y)P \xrightarrow{\alpha} (y)P'}$$

open
$$\xrightarrow{P \xrightarrow{\overline{x}y} P' \quad y \neq x}{(y)P \xrightarrow{x(y)} P'}$$



$\pi\text{-calculus:}$ variants and implementations

- Asynchronous π-calculus
 Only 0 allowed after an output prefix
 x̄ u.0 is ok, x̄ u.x(z).0 is not!
- Higher-order π-calculus: communicating of processes as well as names
- spi-calculus A variant of the π-calculus for reasoning about security
- Ambient calculus: a process algebra for reasoning about distribution
- Pict: a programming language based on the asynchronous π -calculus
- WS-CDL: a web choreography language inspired by the π-calculus



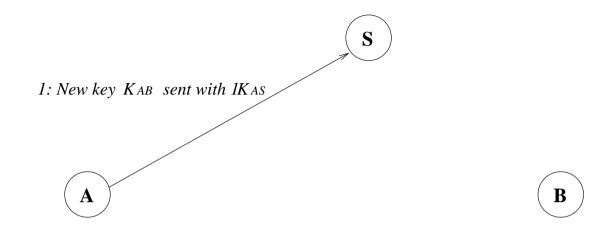
$\pi\text{-calculus}$ variant: the spi-calculus

- spi-calculus: a variant of the π-calculus for reasoning about security
- Specially suited for reasoning about shared key cryptography
- Extends the normal π -calculus with a few new primitives:
 - {*M*}*N* represents the ciphertext obtained by encrypting *M* under the key *N*
 - case L of {x}N in P attempts to decrypt the term L with the key N. If L is a ciphertext of the form {M}N, then the process behaves as P[M/x]. Otherwise, the process is stuck.
- The normal operational semantics of the π -calculus is extended, i.e., we get a lot of reasoning power for free



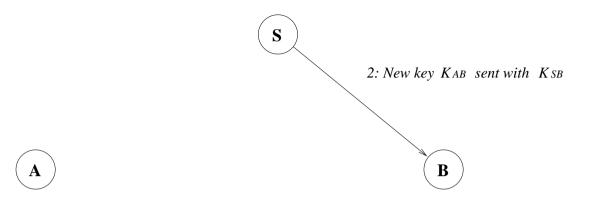
. . .

Spi Example: Wide Mouthed Frog protocol



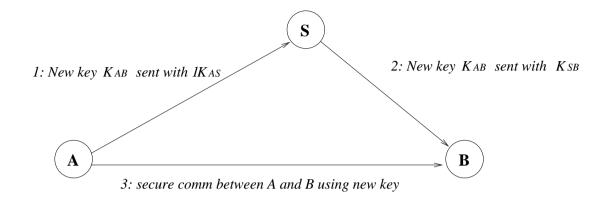


Spi Example: Wide Mouthed Frog protocol



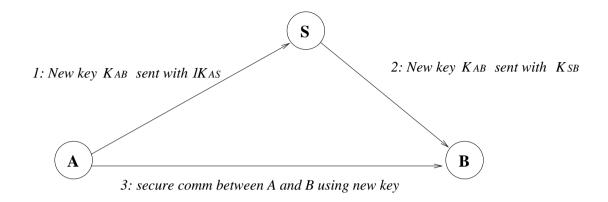


Spi Example: Wide Mouthed Frog protocol





Spi Example: Wide Mouthed Frog protocol



In the spi-calculus:

$$\begin{array}{lll} \mathcal{A}(\mathcal{M}) &\equiv & \nu(\mathcal{K}_{AB})(\overline{c_{AS}}\langle\{\mathcal{K}_{AB}\}\mathcal{K}_{AS}\rangle.\overline{c_{AB}}\langle\{\mathcal{M}\}\mathcal{K}_{AB}\rangle) \\ & \mathcal{S} &\equiv & c_{AS}(x).\textit{case } x \textit{ of } \{y\}\mathcal{K}_{AS} \textit{ in } \overline{c_{SB}}\langle\{y\}\mathcal{K}_{SB}\rangle \\ & \mathcal{B} &\equiv & c_{SB}(x).\textit{case } x \textit{ of } \{y\}\mathcal{K}_{SB} \textit{ in } c_{AB}(z).\textit{case } z \textit{ of } \{w\}y \textit{ in } F(w) \end{array}$$



Part I: The Ambient Calculus

- Due to Cardelli and Walker
- For modelling mobile computation mobile code that moves between locations (ambients)
- Used to reason about administrative domains when does a program have the right to migrate from one computing location to another computing location and start computing there?
- An ambient is a bounded place where computation takes place
- An ambient can be nested in other ambients
- Each ambient has a name used to control access, and a set of local agents (processes) that control the actions of the ambient
- A name is something that can be created, communicated and from which capabilities can be extracted

Ambient Calculus: operators

- As in the π-calculus: 0 (the process which can do nothing),
 P | Q parallel composition, replication !P and creation of a new name n in (n)P.
- An ambient is written n[P] where n is the name and P is the process running inside the ambient
- If $P \to P'$ then $n[P] \to n[P']$
- The general shape of an ambient is $n[P_1 | \dots | P_n | m_1[\dots] | \dots | m_k[\dots]]$ where P_i is a non-ambient process and $m_i[\dots]$ is a subambient of n

In graphical notation:



Ambient Calculus: operators

- An action prefix is written M.P, where M enters, exits or opens an ambient
- Entry capability: in m.P instructs the ambient surrounding in m.P to enter a sibling named m

$$\begin{bmatrix} n \\ in \ m.P \mid Q \end{bmatrix} \mid \begin{bmatrix} m \\ R \end{bmatrix} \rightarrow \begin{bmatrix} m \\ P \mid Q \end{bmatrix} \mid R$$

Exit capability: out m.P instructs the ambient surrounding out m.P to exit its parent named m





Open capability and communication

Open capability: open m.P which provides a way of dissolving the boundary of an ambient m located at the same level as openm.P

$$openm.P \mid \begin{array}{c} m \\ \hline Q \end{array} \rightarrow P \mid Q$$

- The full calculus adds two items:
 - Capabilities can be paths M.M'
 - A communication rule between processes in the same ambient: (x)P | ⟨M⟩ → P[M/x] where (x)P is an input prefix and ⟨M⟩ an output

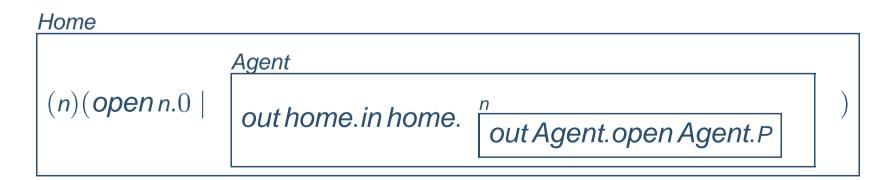


Ambient example

Motivating example: an agent *P* leaves its home ambient and later comes back (with authentification)



or graphically:





Home[(*n*) (*open n*.0 | *Agent*[*out home.in home.n*[*out Agent.open Agent.P*]])]



Home[(n) (open n.0 | Agent[out home.in home.n[out Agent.open Agent.P]])]

 $= (n) Home[open n.0 | \\ Agent[out home.in home.n[out Agent.open Agent.P]]]$



Home[(n) (open n.0 | Agent[out home.in home.n[out Agent.open Agent.P]])]

 $\equiv (n) Home[open n.0 |$ Agent[out home.in home.n[out Agent.open Agent.P]]]



Home[(n) (open n.0 | Agent[out home.in home.n[out Agent.open Agent.P]])]

 $\equiv (n) Home[open n.0 |$ Agent[out home.in home.n[out Agent.open Agent.P]]]

↓ in home
(n) (Home[open n.0 | Agent[n[out Agent.open Agent.P]]])



Home[(n) (open n.0 | Agent[out home.in home.n[out Agent.open Agent.P]])]

 $\equiv (n) Home[open n.0 | \\Agent[out home.in home.n[out Agent.open Agent.P]]]$

in home
(n) (Home[open n.0 | Agent[n[out Agent.open Agent.P]]])

↓ out Agent (n) (Home[open n.0 | n[open Agent.P] | Agent[]])



Home[(n) (open n.0 | Agent[out home.in home.n[out Agent.open Agent.P]])]

 $\equiv (n) Home[open n.0 | \\Agent[out home.in home.n[out Agent.open Agent.P]]]$

in home
(n) (Home[open n.0 | Agent[n[out Agent.open Agent.P]]])

↓ out Agent (n) (Home[open n.0 | n[open Agent.P] | Agent[]])

↓ open n (n) (Home[0 | open Agent.P | Agent[]])



Home[(n) (open n.0 | Agent[out home.in home.n[out Agent.open Agent.P]])]

= (n) Home[open n.0 |
Agent[out home.in home.n[out Agent.open Agent.P]]]

in home
(n) (Home[open n.0 | Agent[n[out Agent.open Agent.P]]])

↓ out Agent (n) (Home[open n.0 | n[open Agent.P] | Agent[]])

↓ open n (n) (Home[0 | open Agent.P | Agent[]])

 $\downarrow open Agent$ (n) (Home[0 | P | 0]) \equiv Home[P]

Tools for CCS and the $\pi\text{-calculus}$

There exists a set of tools for CCS and the π -calculus that permits to:

simulate specifications,

to check whether two specifications are equivalent,

to check a specification against a property in temporal logic, and so on...

Tools:

For CCS: Concurrency workbench, Concurrency Workbench of the New Century (CWB-NC), ...

For π -calculus: Mobility workbench, ...

