# **Concurrency, Intuition and Formal** Verification: Yes, We Can!

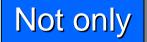
Matt Pedersen<sup>a</sup> and Peter Welch<sup>b</sup> <sup>a</sup>School of Computer Science, UNLV, USA <sup>b</sup>School of Computing, University of Kent, UK

> phw@kent.ac.uk matt@cs.unlv.edu

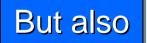
#### Curricula for Concurrency and Parallelism SPLASH 2010, 17<sup>th</sup>. Oct

12-Oct-10

Copyleft (GPL) P.H.Welch and J.B.Pedersen



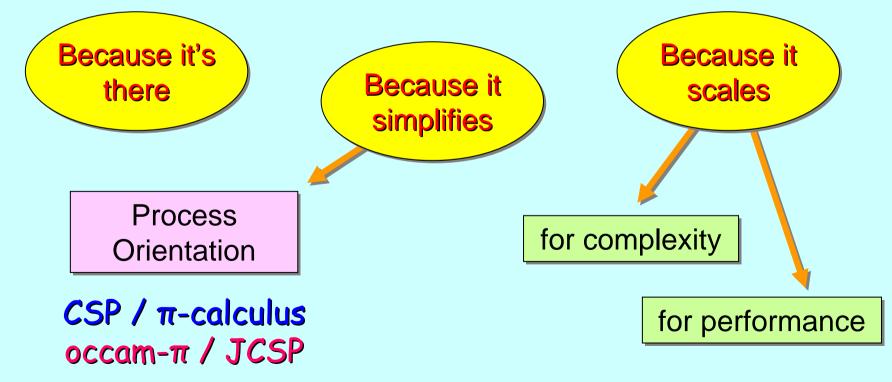
*can* we (and *should* we) teach concurrency at the start of the undergraduate CS curriculum ...



we **can** (and we **should**) teach formal analysis and verification of this concurrency at the same time ...

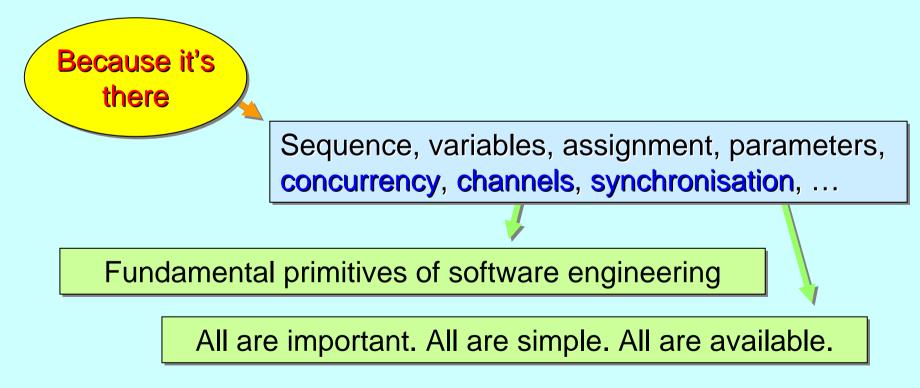
#### Not only

can we (and should we) teach concurrency at the start of the undergraduate CS curriculum ...





*can* we (and *should* we) teach concurrency at the start of the undergraduate CS curriculum ...



Complex and high-performance systems cannot avoid concurrent design, implementation *and reasoning*.

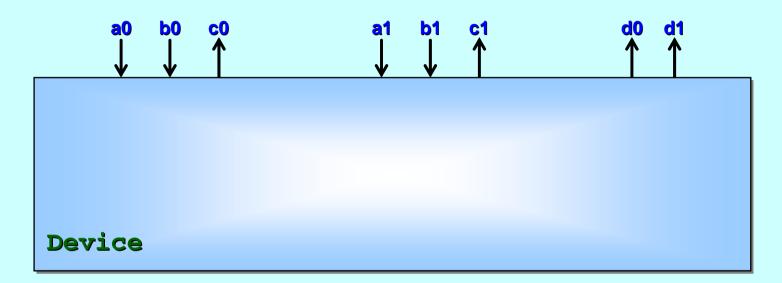
Common concurrency bugs are intermittent – not repeatable on demand. *Untestable in practice*.

We stand on the shoulders of giants (who made the theory and model checkers). We verify programs just by writing programs ... it becomes everyday practice.

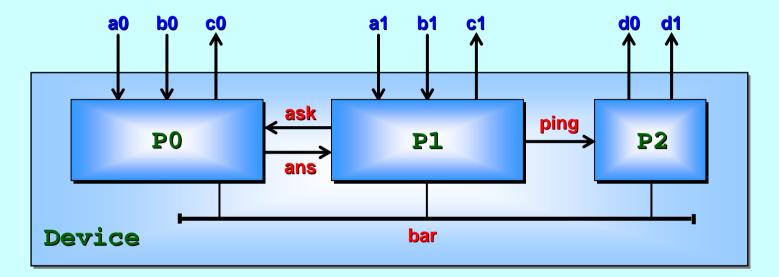
#### But also

we **can** (and we **should**) teach formal analysis and verification of this concurrency at the same time ...

The following example has been developed from one first worked through in a single lesson of a graduate class in concurrency at UNLV in the spring of 2010.



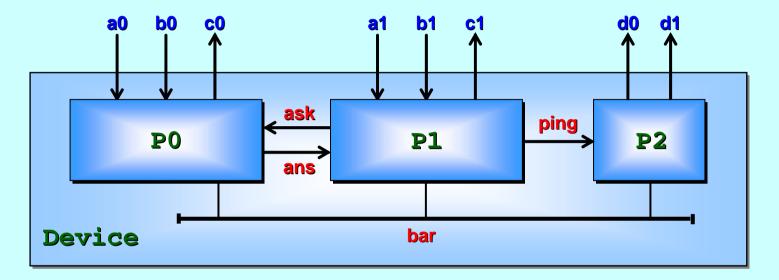
**Device** : real-time controller for 8 channels (4 input, 4 output).



**Device** : real-time controller for 8 channels (4 input, 4 output).

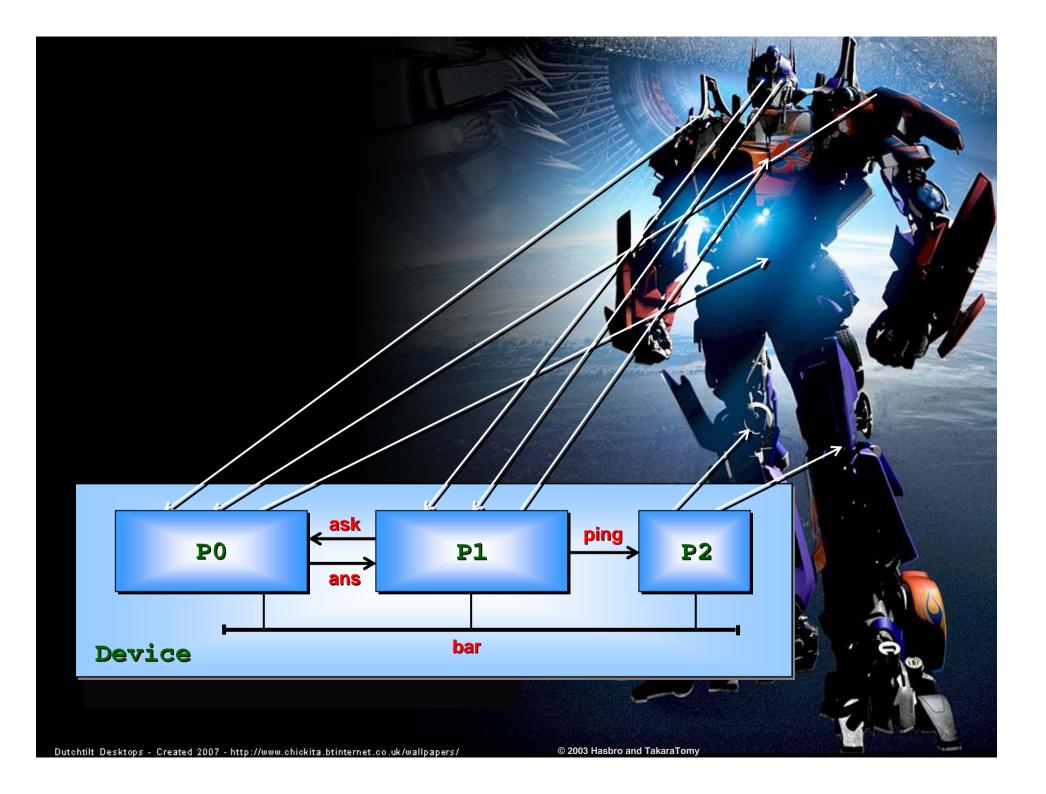
There are 3 sub-components: **P0** (weapons systems), **P1** (vision processing) and **P2** (motion stabilizer).

They exchange information over internal channels (**ask**, **ans**, **ping**) and all coordinate actions with an internal barrier (**bar**).

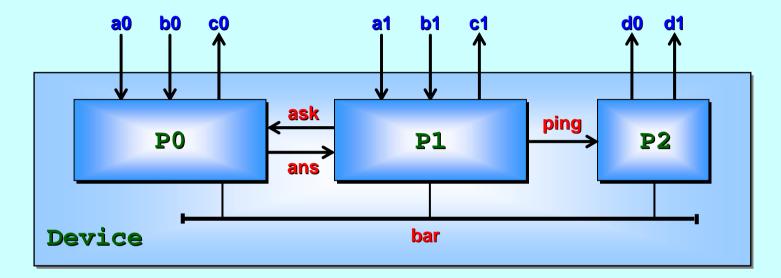


**CSP** semantics apply. **Channel communication** is unbuffered (sender waits for receiver and vice-versa). Any process **reaching a barrier** waits for **all** processes to **reach the barrier**.

They exchange information over internal *channels* (ask, ans, **ping**) and all coordinate actions with an internal *barrier* (bar).



## Behaviour: two representations

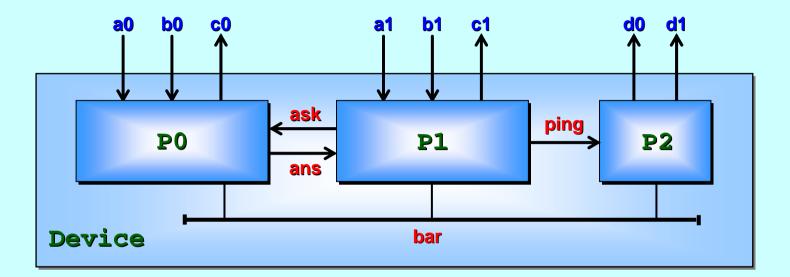


**occam-π**: for compiling to a runnable system. [memory overheads <= 32 bytes per process / synchronisation overheads of order tens of nanoseconds / eats multicore nodes for breakfast.]

**CSP**: for formal analysis.

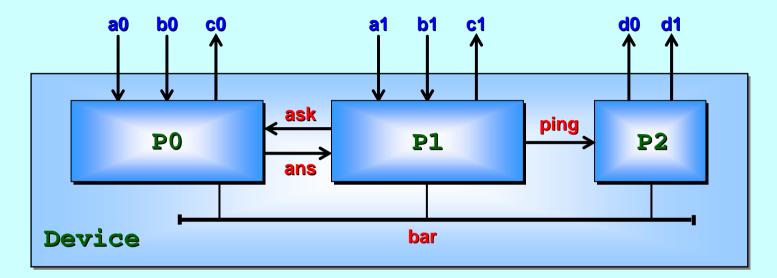
[FDR2 model checker + other (simple) formal reasoning.]

## Behaviour: two representations



occam-π syntax / semantics has an injective mapping to **occam-** $\pi$ : for compiling to a runnable exp CSP. Our students had little trouble shifting between [memory overheads <= 22. VS CSP. Our students had into a CSP automatically from them. A tool exists to generate CSP automatically from the classroom. occam-π ... not yet ready for use in the classroom.

## Behaviour: what are we looking for?

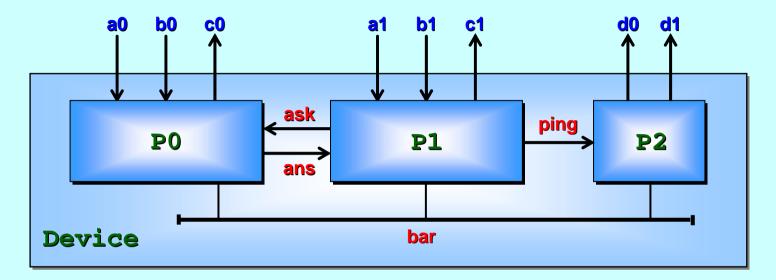


**deadlock**: *might* it ever stop?

[e.g. **PO** and **P2** want to synchronise on **bar**, but **P1** wants to **ping**.]

**livelock**: *might* it get busy ... but refuse all external signals? [e.g. **P0**, **P1** and **P2** start engaging in an infinite sequence of internal channel or barrier synchronisations (on **ask**, **ans**, **ping** and **bar**).]

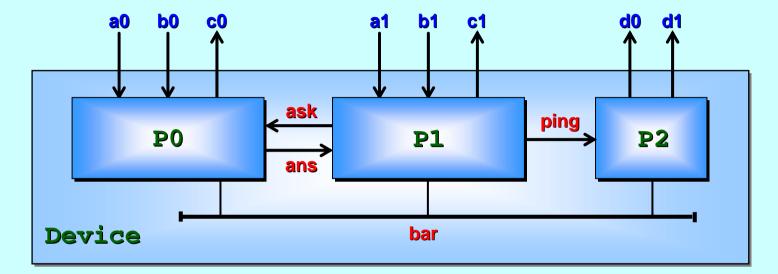
## Behaviour: what are we looking for?

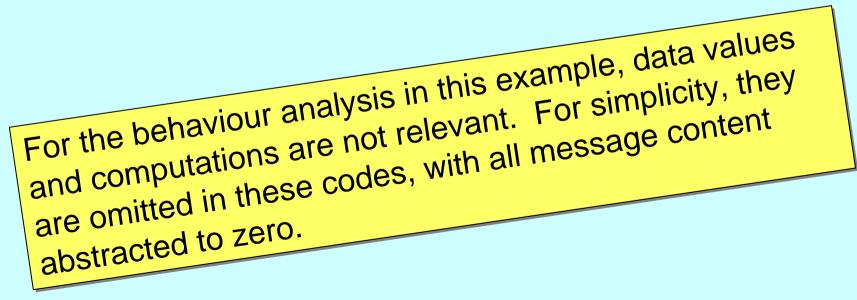


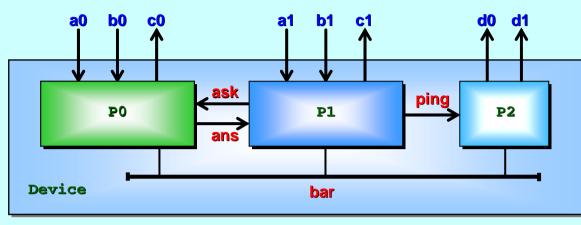
safety: might it ever engage in an incorrect sequence of
external signals?

**liveness:** *will* it engage in correct sequences of external signals, as required?

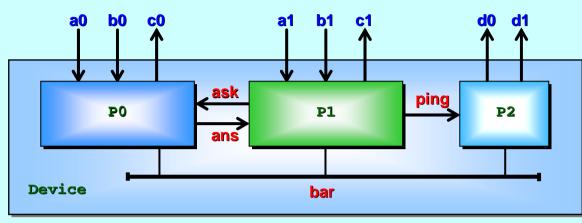
[Some specs allow alternative sequences to be performed – all are correct, but an implementation must only do one and is free to choose.]



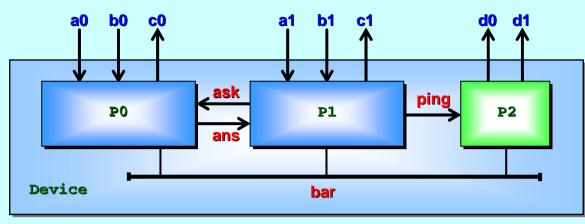




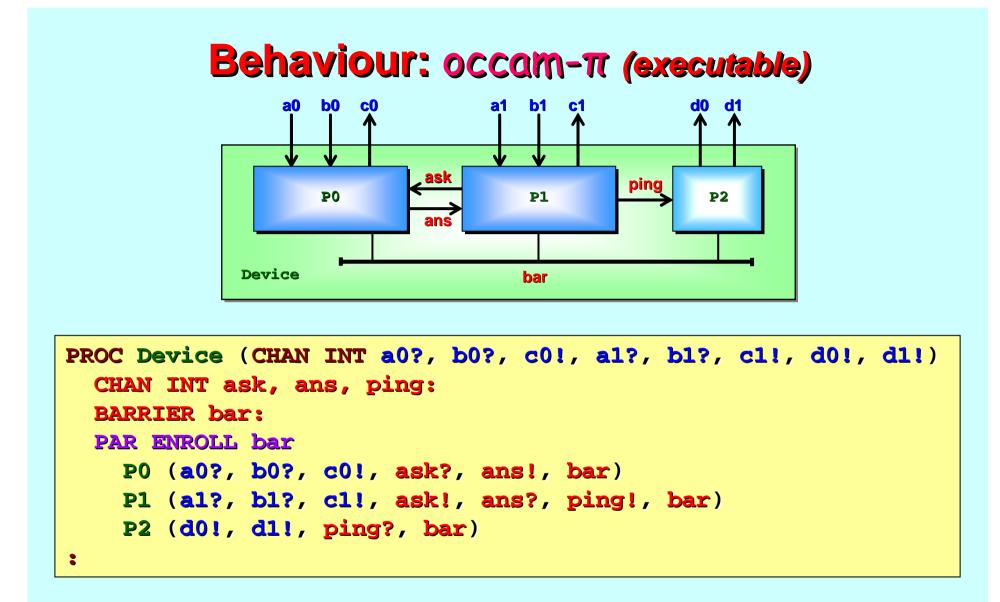
```
PROC P0 (CHAN INT a0?, b0?, c0!, ask?, ans!, BARRIER bar)
WHILE TRUE
INT x, y, z:
SEQ
ask ? x -- take question
a0 ? y
ans ! 0 -- return answer (will depend on x and y)
b0 ? z
SYNC bar -- wait for the others
c0 ! 0
;
```



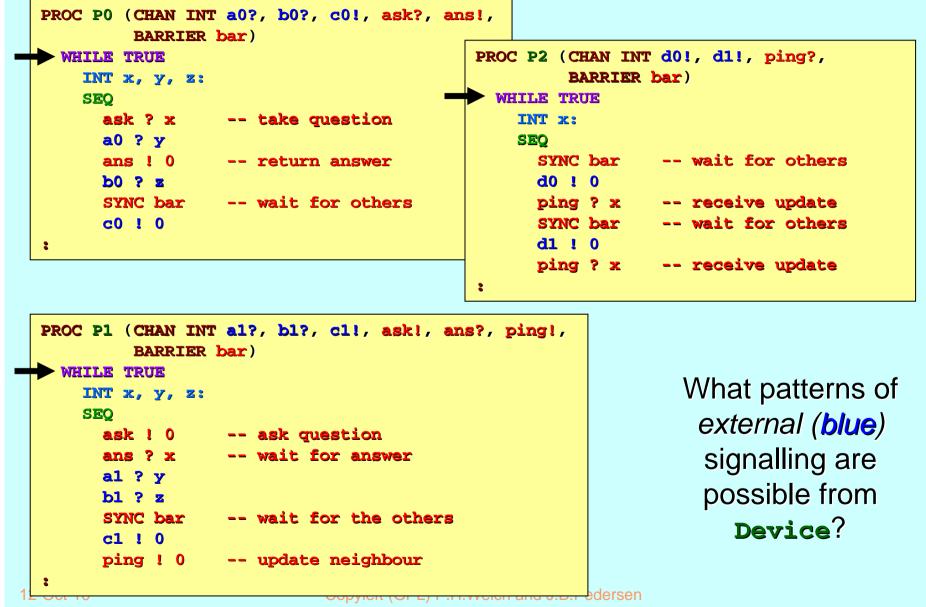
```
PROC Pl (CHAN INT al?, bl?, cl!, ask!, ans?, ping!,
BARRIER bar)
WHILE TRUE
INT Z, Y, Z:
SEQ
ask ! 0 -- ask question
ans ? x -- wait for answer
al ? y
bl ? z
SYNC bar -- wait for the others
cl ! 0
ping ! 0 -- update neighbour
```



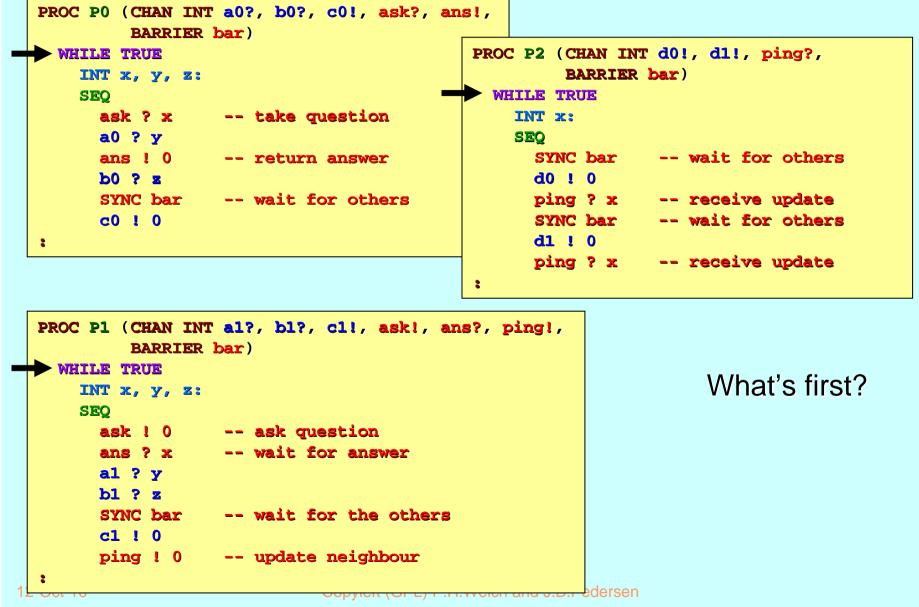
| PROC P2 (CHAN INT d0!, d1!, ping?, BARRIER bar)  |
|--|
| FROC FZ (CIRAT INT GO?, GI?, PING?, DRIVIDA DGI) |
| WHILE TRUE                                       |
| INT x:   |
| SEQ  |
| SYNC bar wait for the others                     |
| d0 ! 0   |
| ping ? x receive update                          |
| SYNC bar wait for the others                     |
| d1 ! 0   |
| ping ? x receive update                          |
|  |

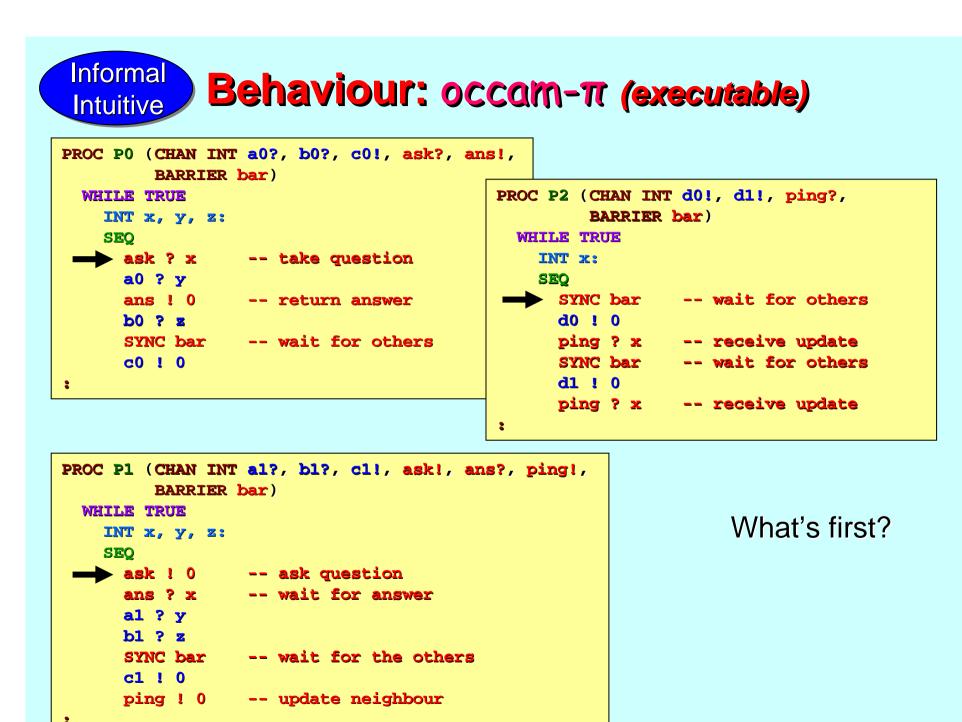






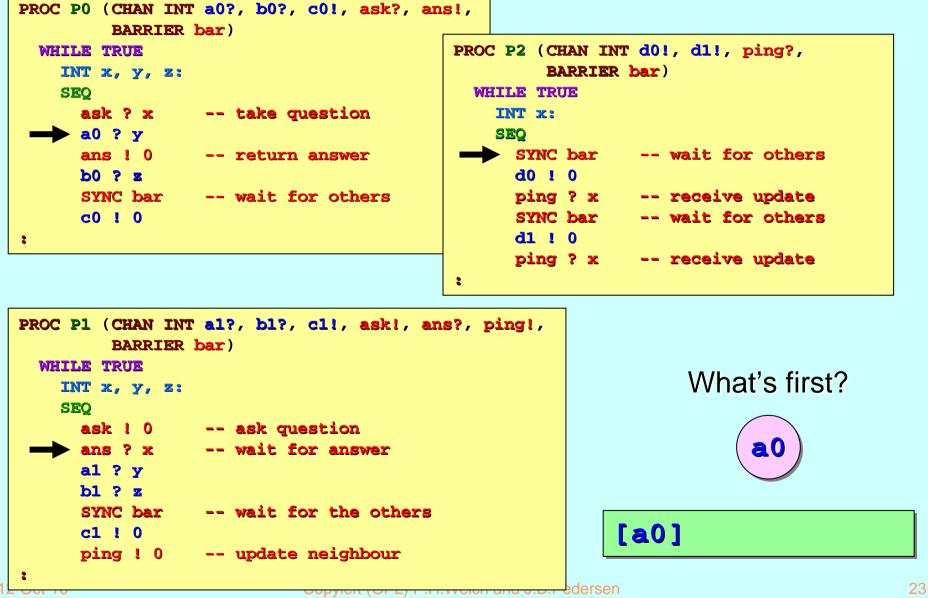




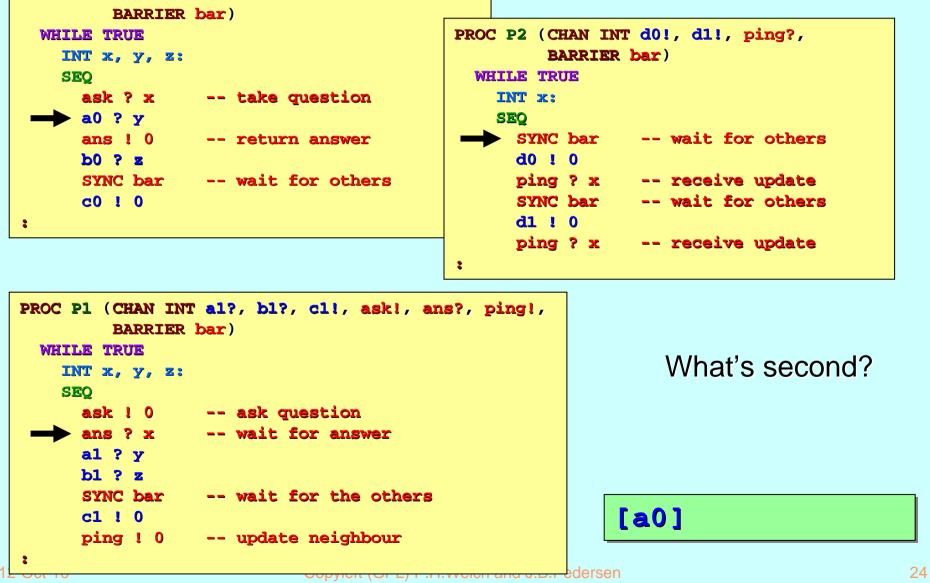


edersen

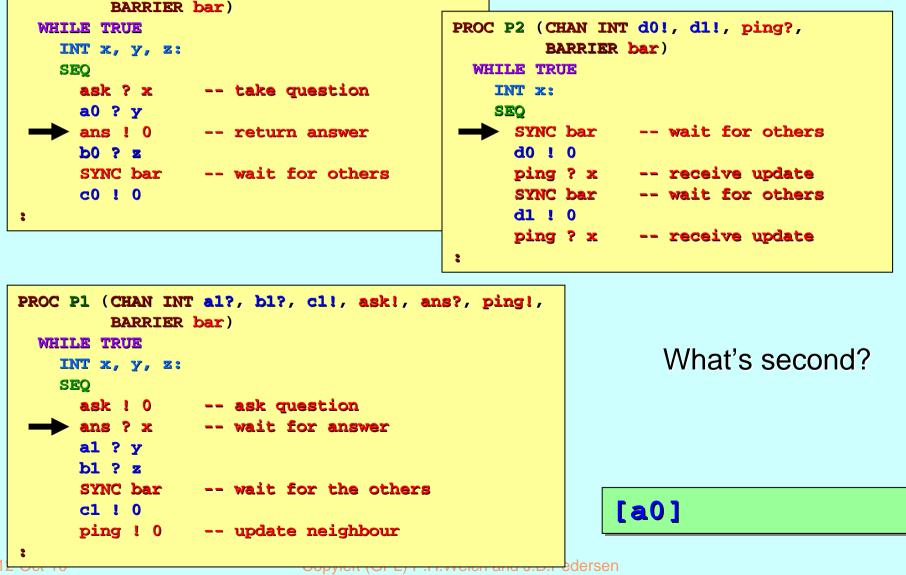




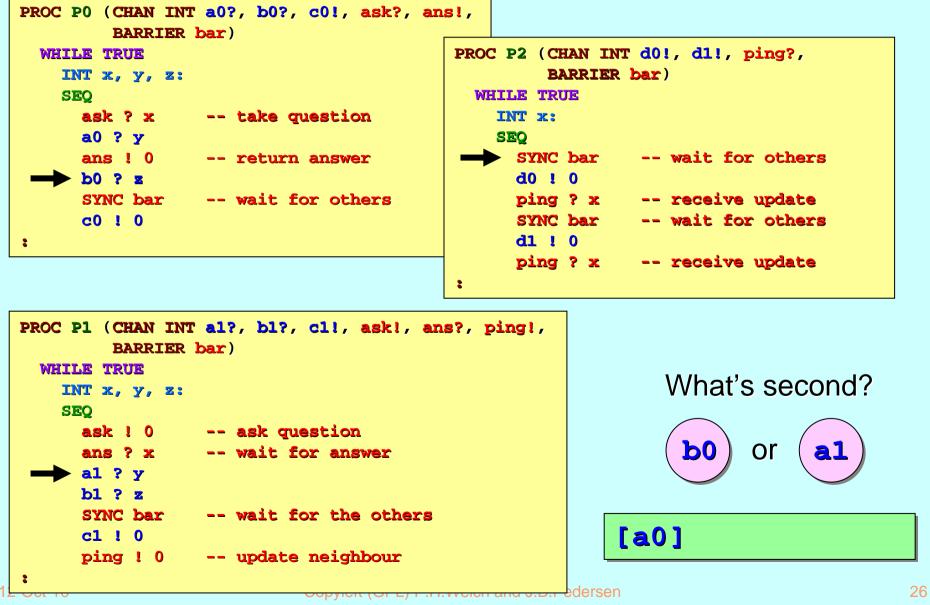




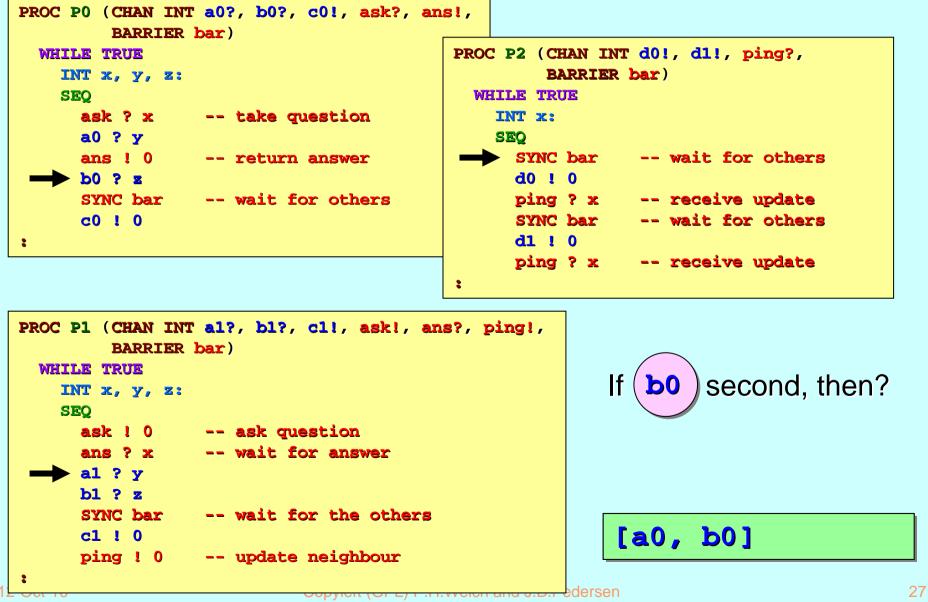




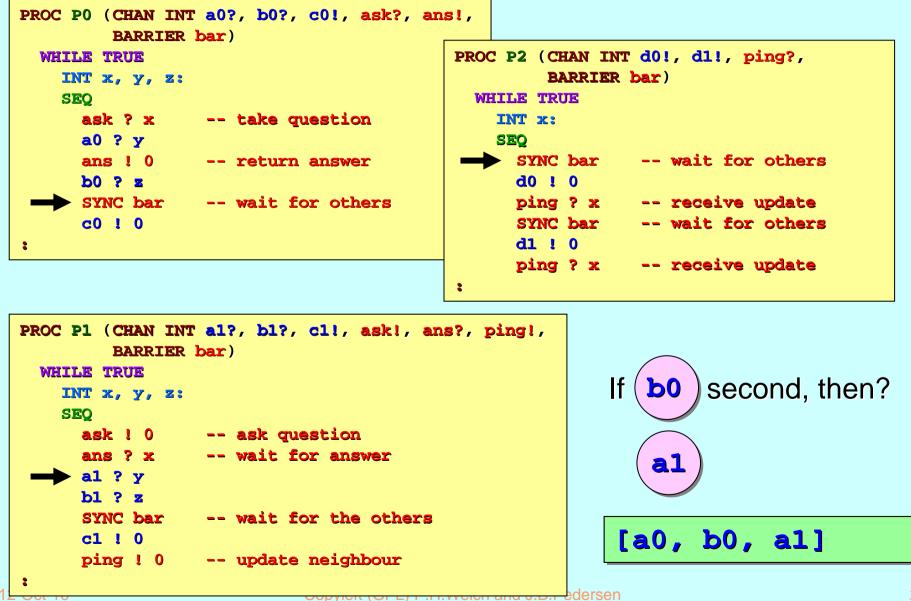




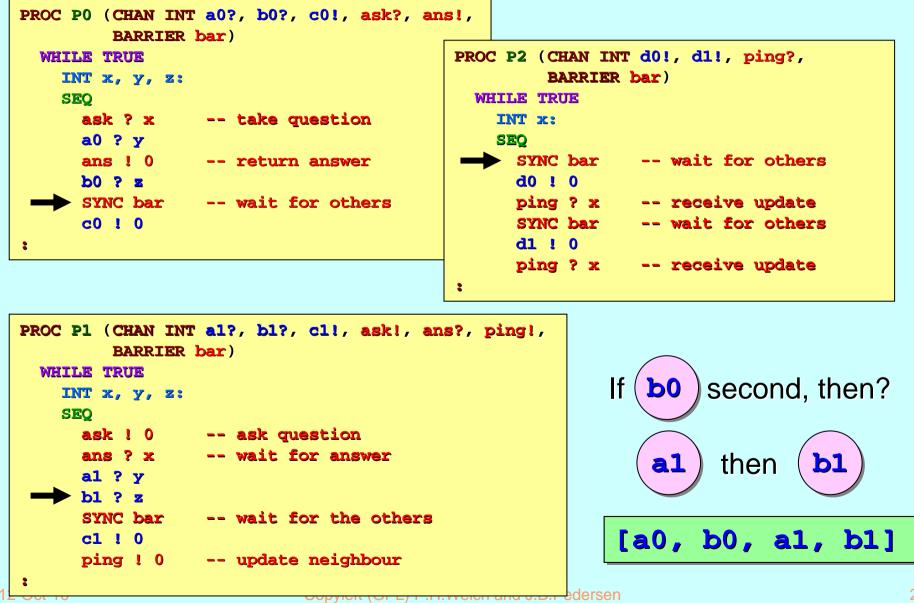




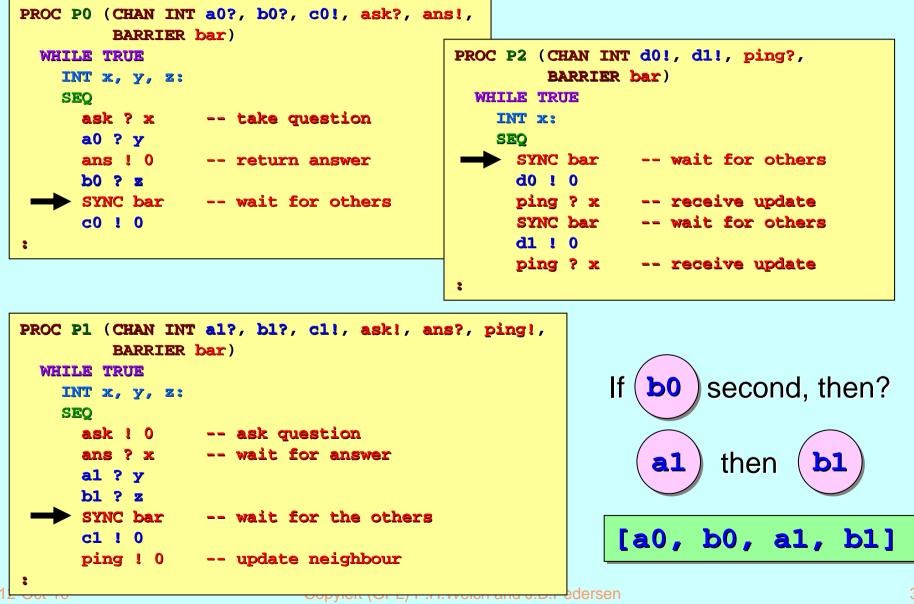




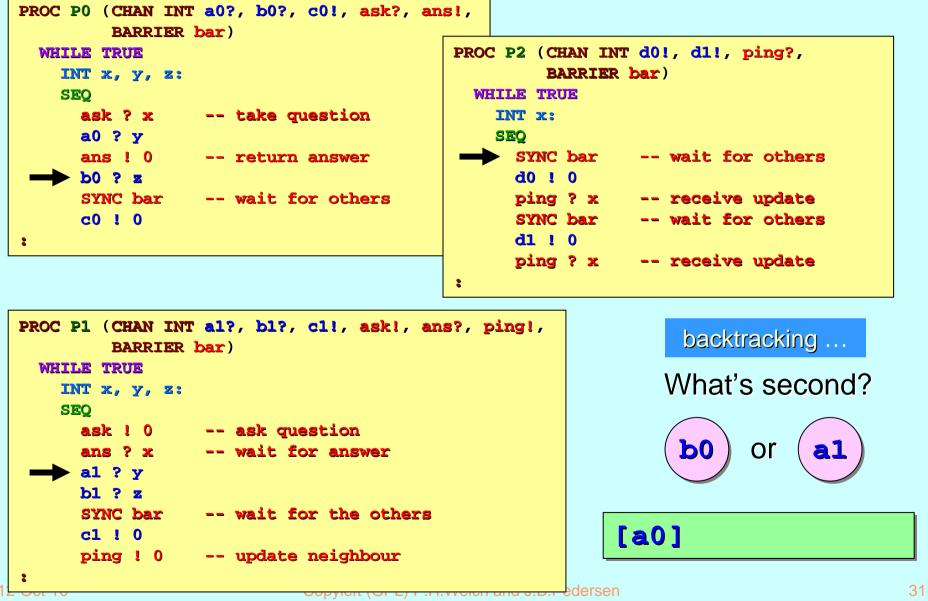




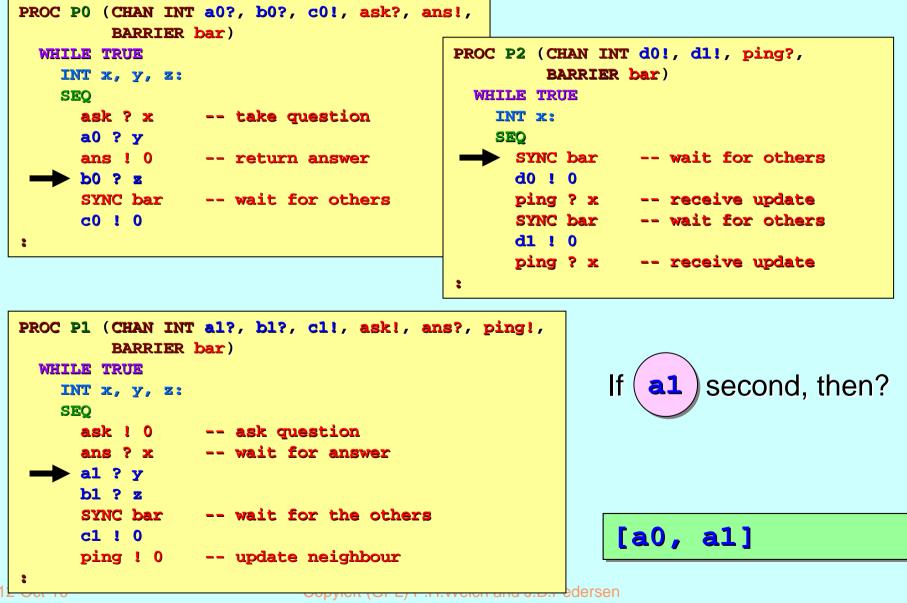




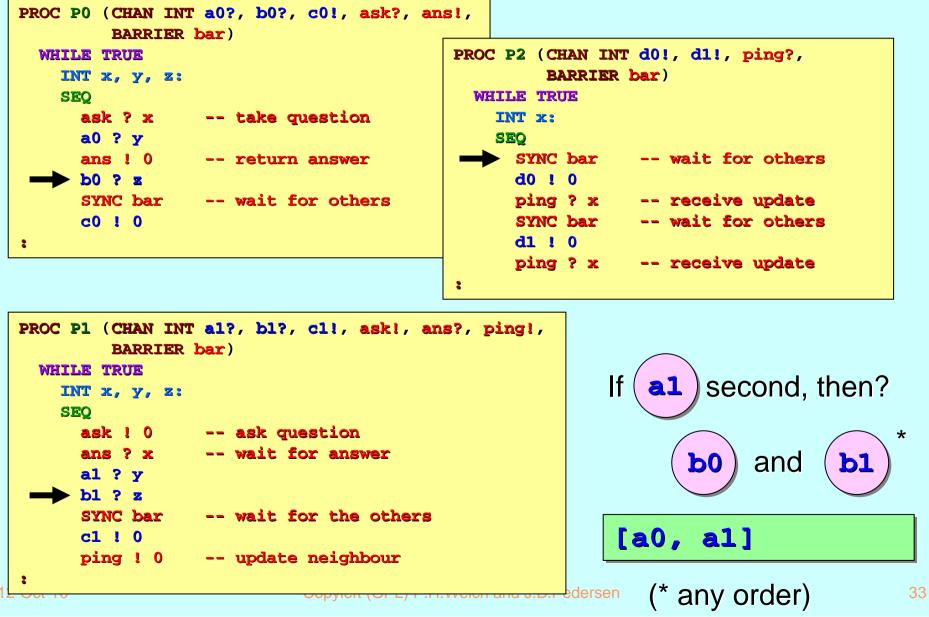




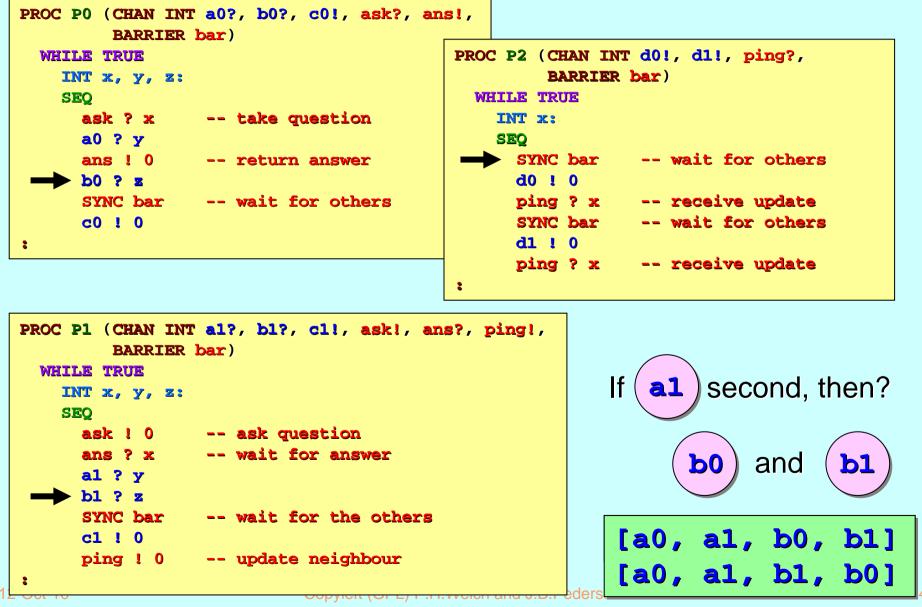




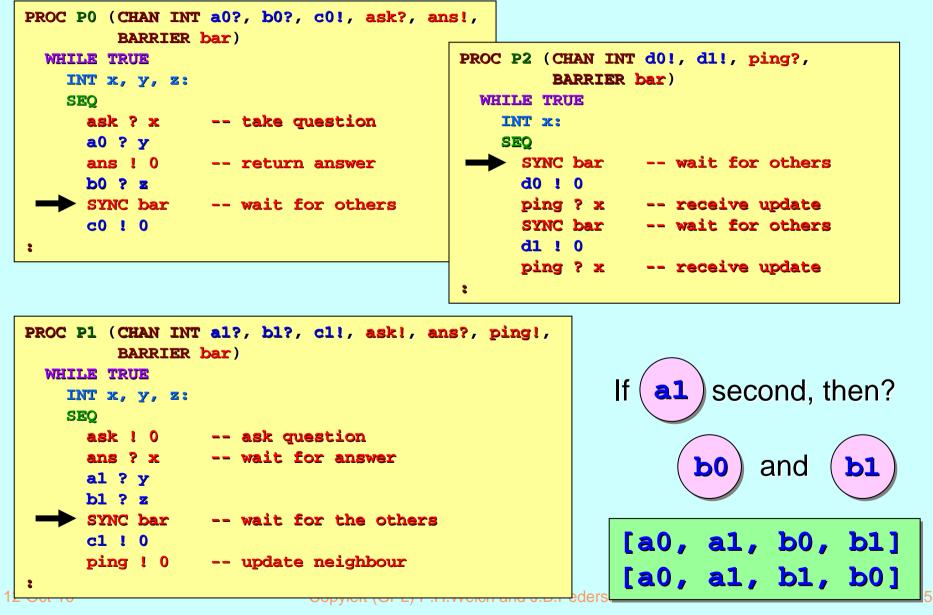




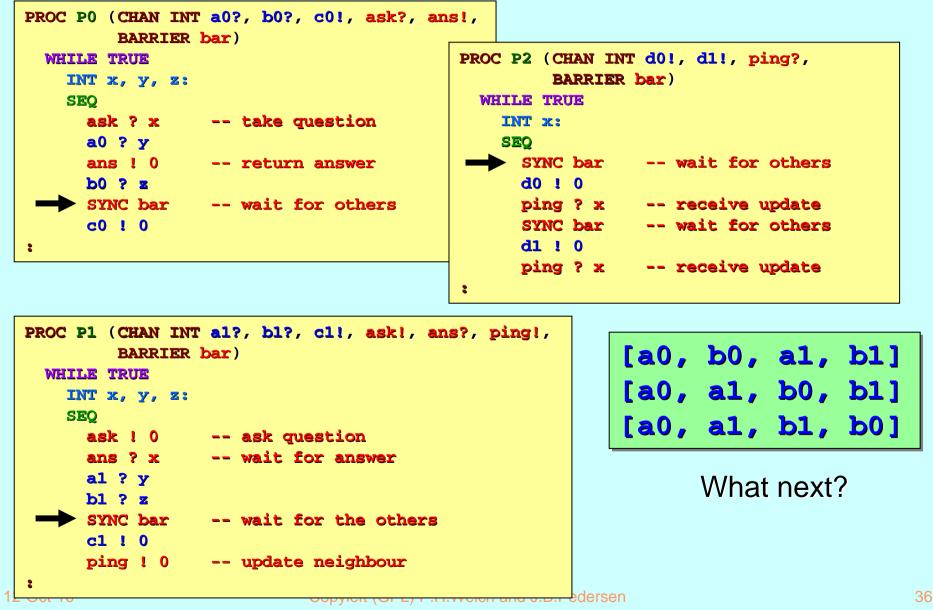




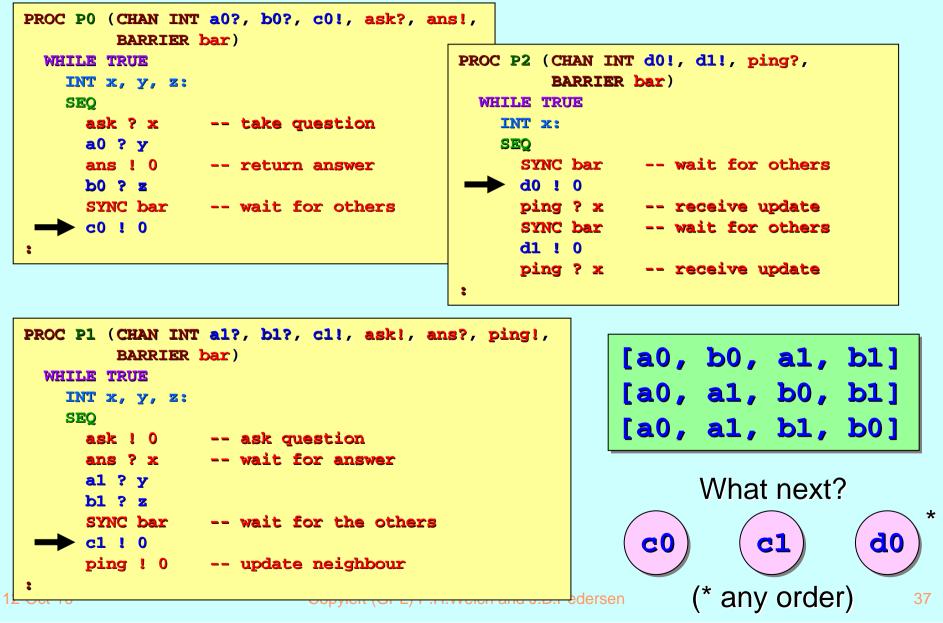




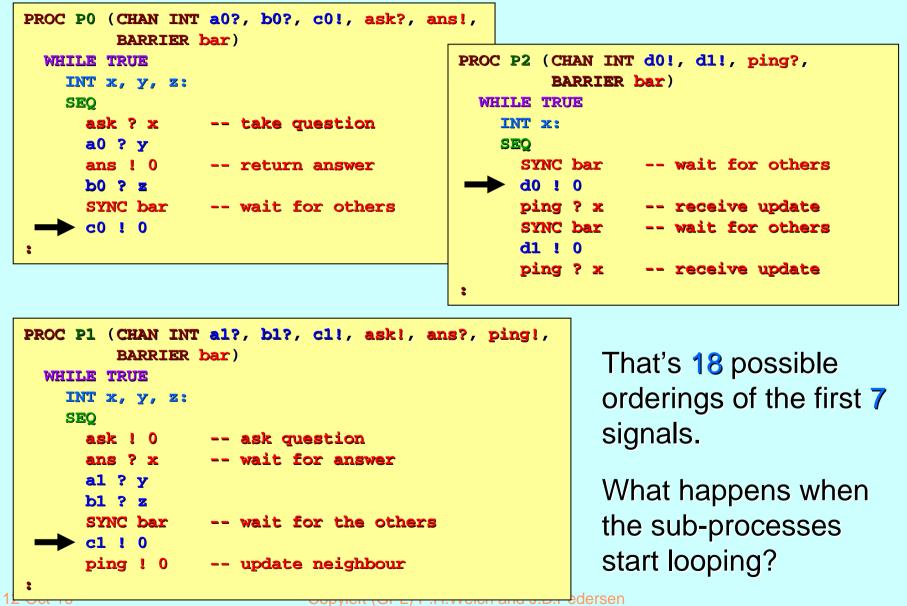




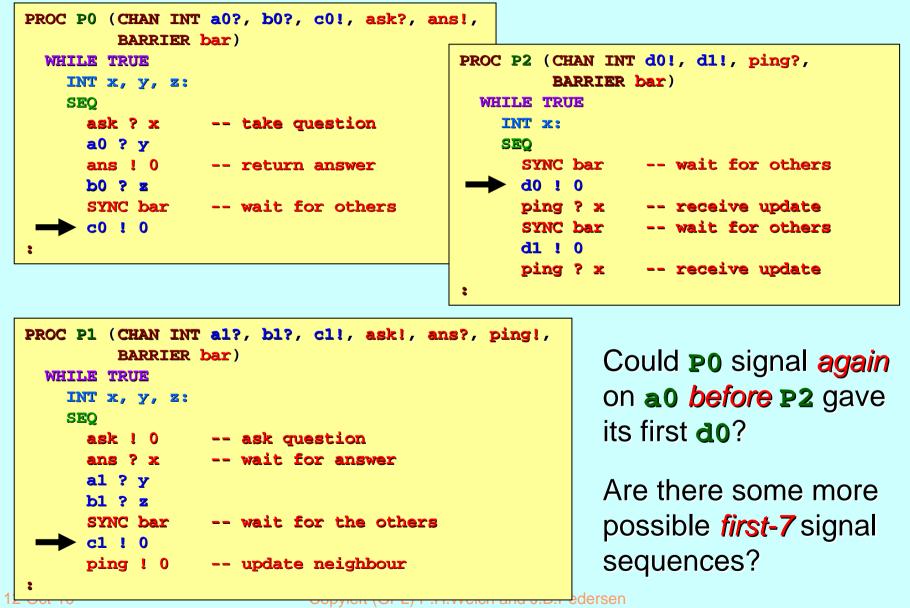


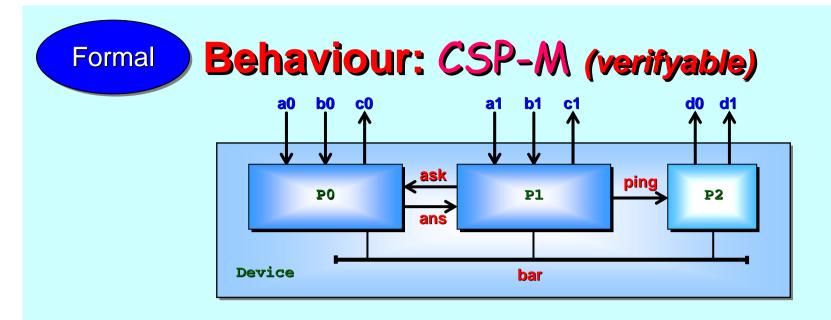






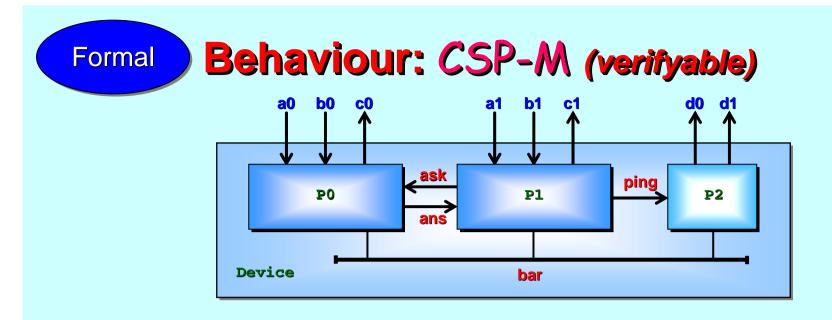






We can formally verify the previous intuition (which was only about the opening behaviour of the system) and answer the open questions (and more) about its continuous behaviour with a **CSP** representation.

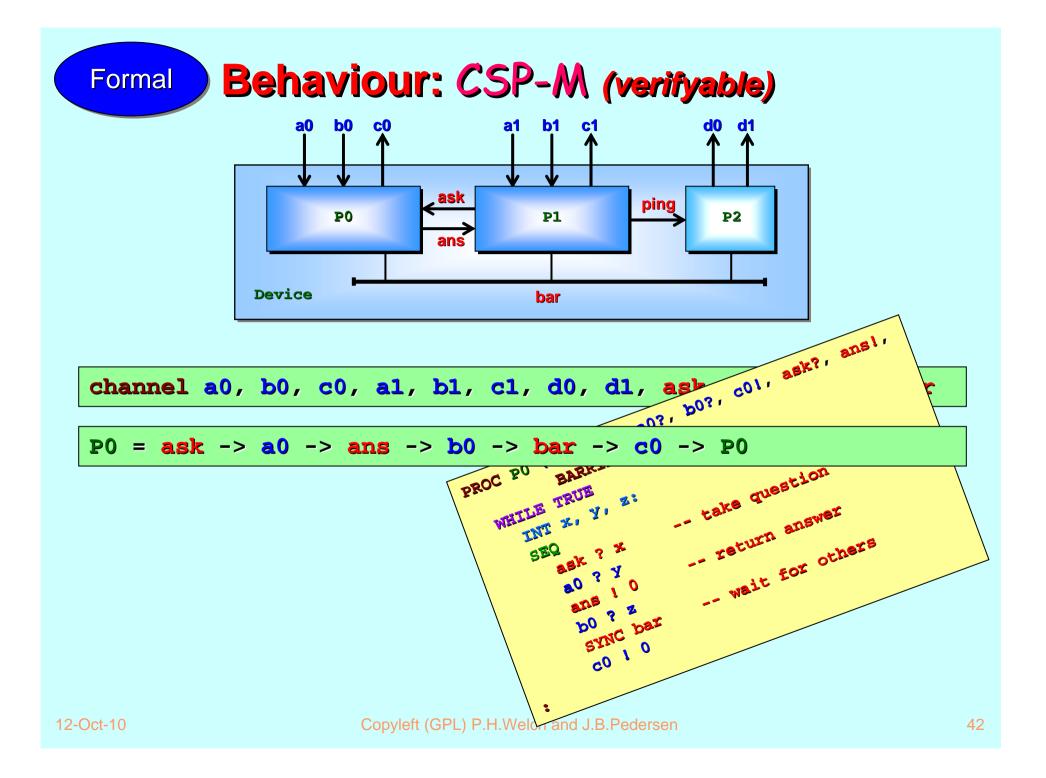
We use **CSP-M**, the machine readable form used by the **FDR2** model checker. **CSP-M** is a *declarative (functional)* language – *loops* map to *tail recursions*. Students who enjoy programming have no problem learning new syntax *(it's particularly easy when the semantics remain unchanged)* – but they need to be told why!

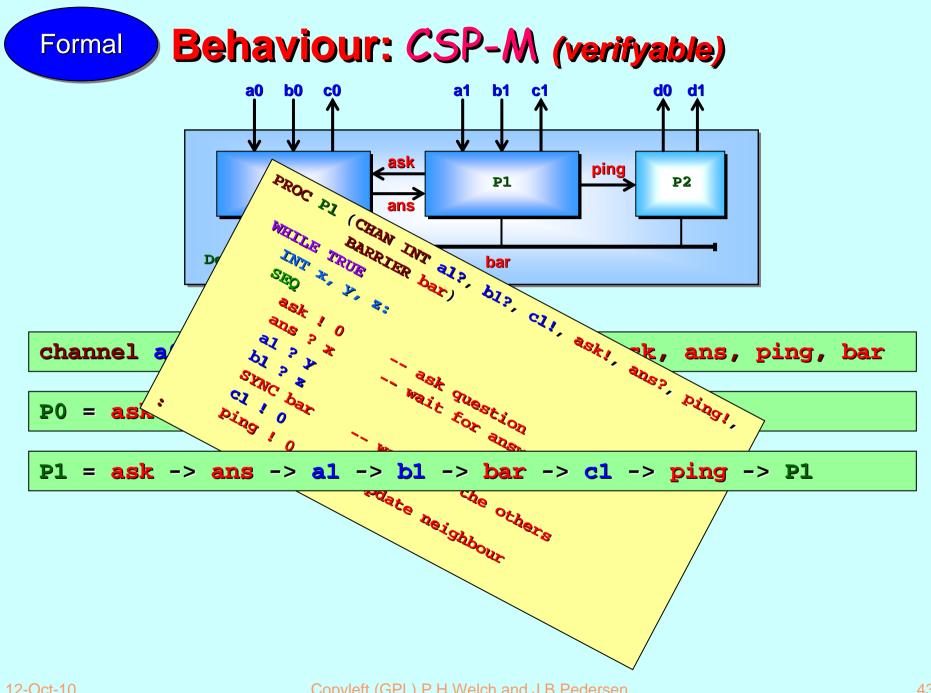


**CSP-M** lets us abstract the channel communications further by omitting the data sent (always zero in our example) and the direction of communication (irrelevant here).

**CSP** processes synchronise only on *events*, which capture the notions of point-to-point channels and multiway barriers. **CSP-M** calls them all *channels*.

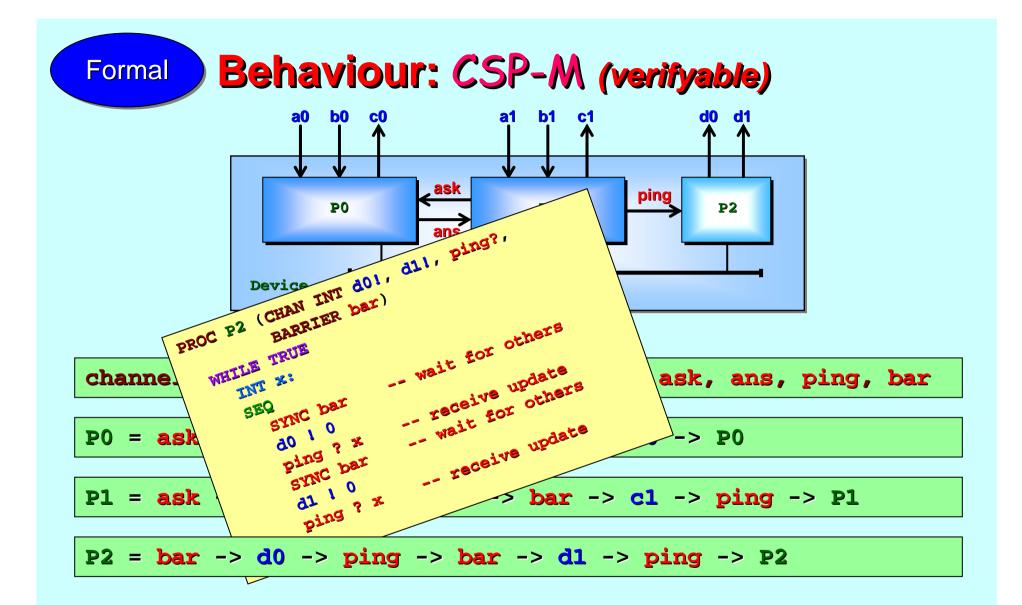
In the following **CSP-M**, we further simplify things by omitting process parameters and accessing all channels from global declaration. [We could have done this with the **occam-** $\pi$ ...]





12-Oct-10

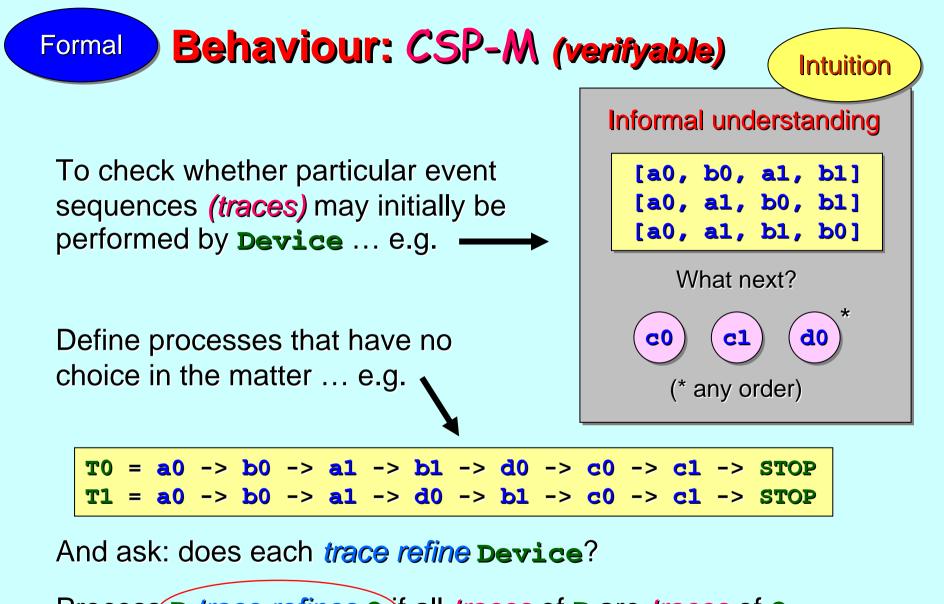
Copyleft (GPL) P.H.Welch and J.B.Pedersen

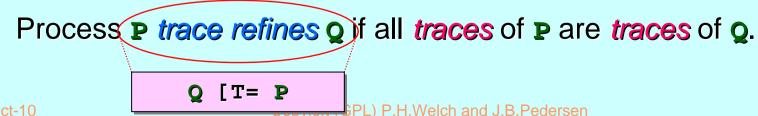


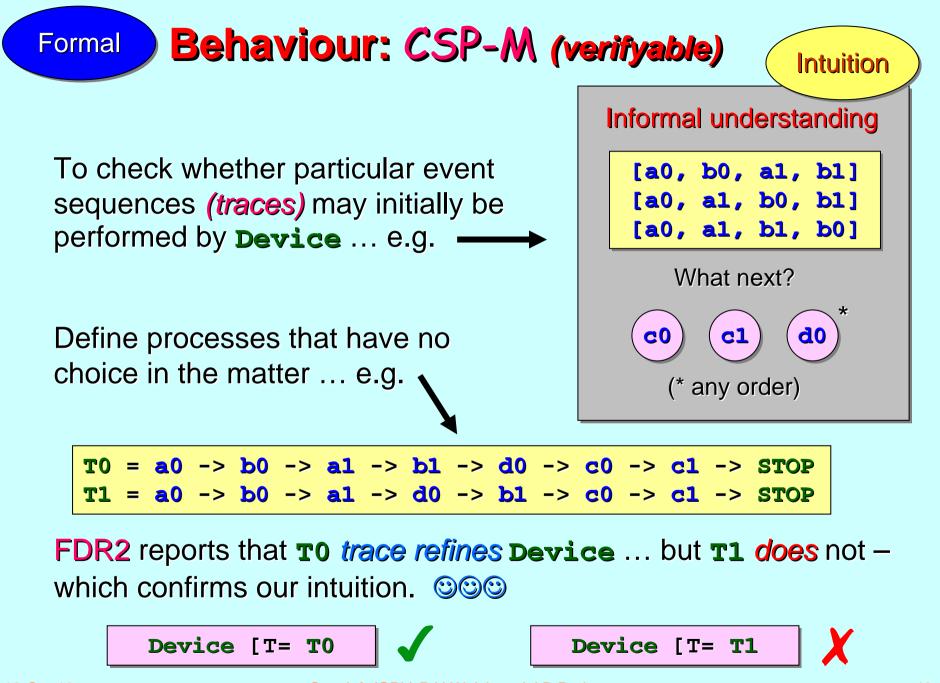
| Formal Behaviour: CSP-M (verifyable)<br>$\downarrow \downarrow $ |
|---|
| <pre>channel a0, b0, c0, a1, b1, c1, d0, d1, ask, ans, ping, bar P0 = ask -&gt; a0 -&gt; ans -&gt; b0 -&gt; bar -&gt; c0 -&gt; P0</pre>   |
| P1 = ask -> ans -> al -> bl -> bar -> cl -> ping -> P1<br>P2 = bar -> d0 -> ping -> bar -> dl -> ping -> P2   |
| <pre>P0P1 = (P0 [  {ask, ans, bar}  ] P1) \ {ask, ans} Device = (P0P1 [  {ping, bar}  ] P2) \ {ping, bar}</pre>   |

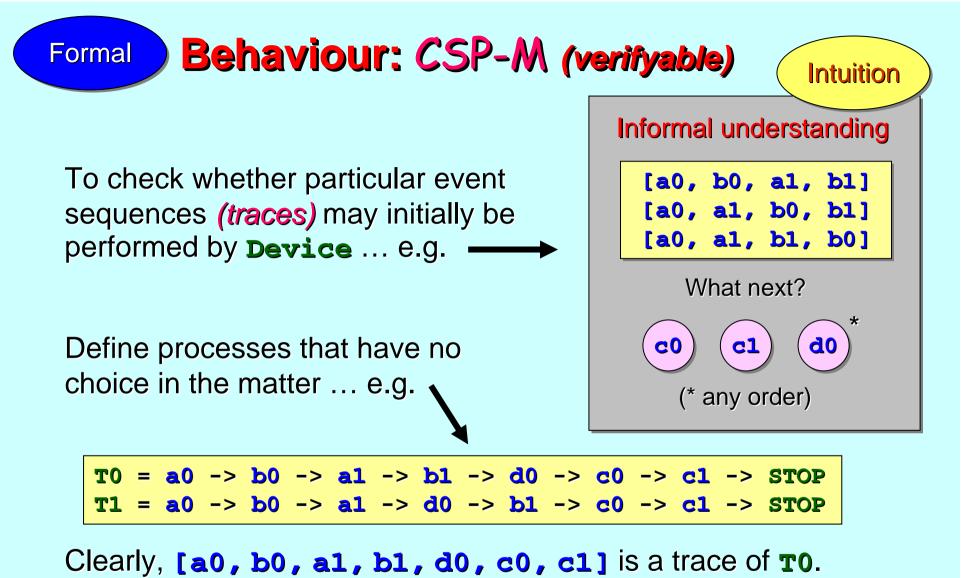


Loading the system below into FDR2, we discover straight away that **Device** is *free from deadlock and livelock* – just click the buttons!

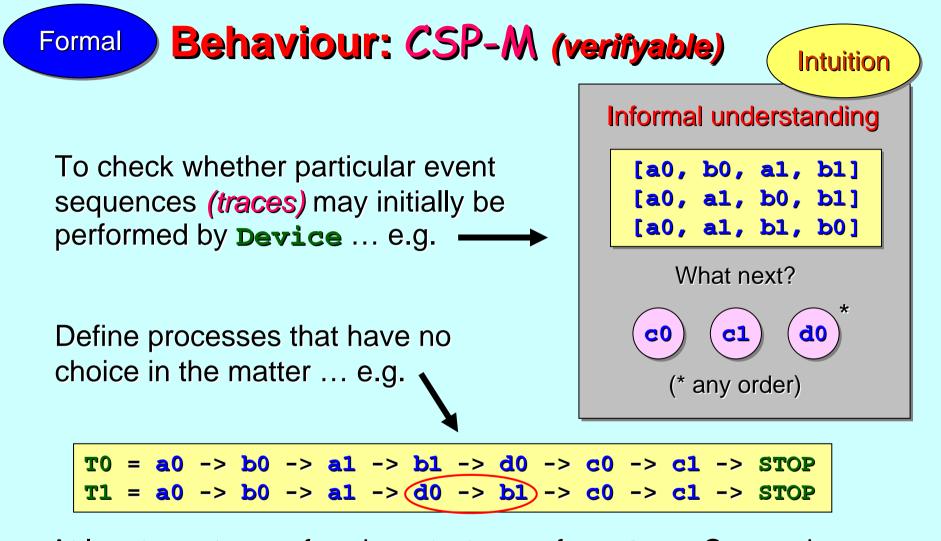






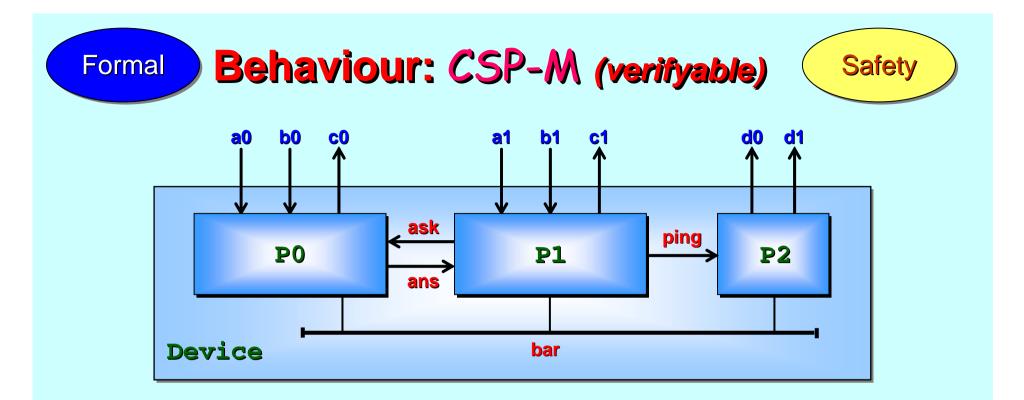


Therefore, it is also a trace of **Device**.



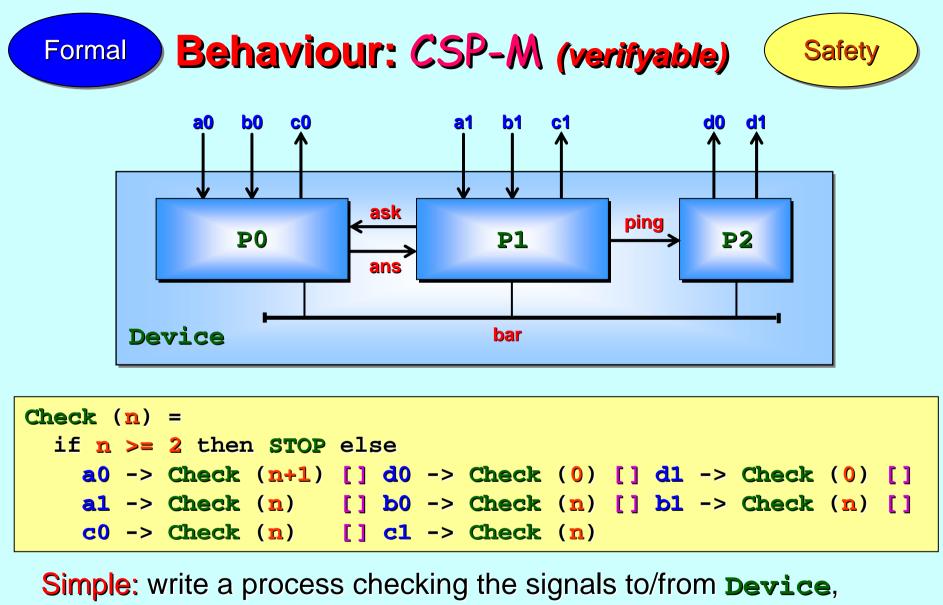
At least one trace of **T1** is **not** a trace of **Device**. Comparing **T0** and **T1**, the fault lies in the mis-ordering of **d0** and **b1**.

Device [T= T1

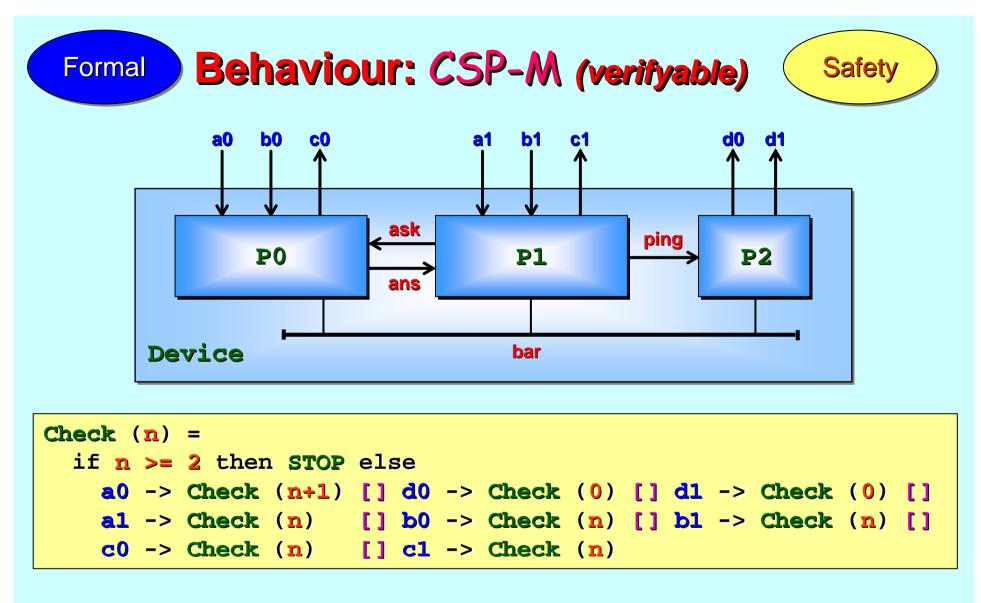


Let's ask a more difficult question about the continuous running of the system. Suppose the robot would do something *very bad* if its controller **Device** were ever to signal *twice* on **a**0 without a signal on **d**0 or **d**1 *in between*. Might this ever happen?

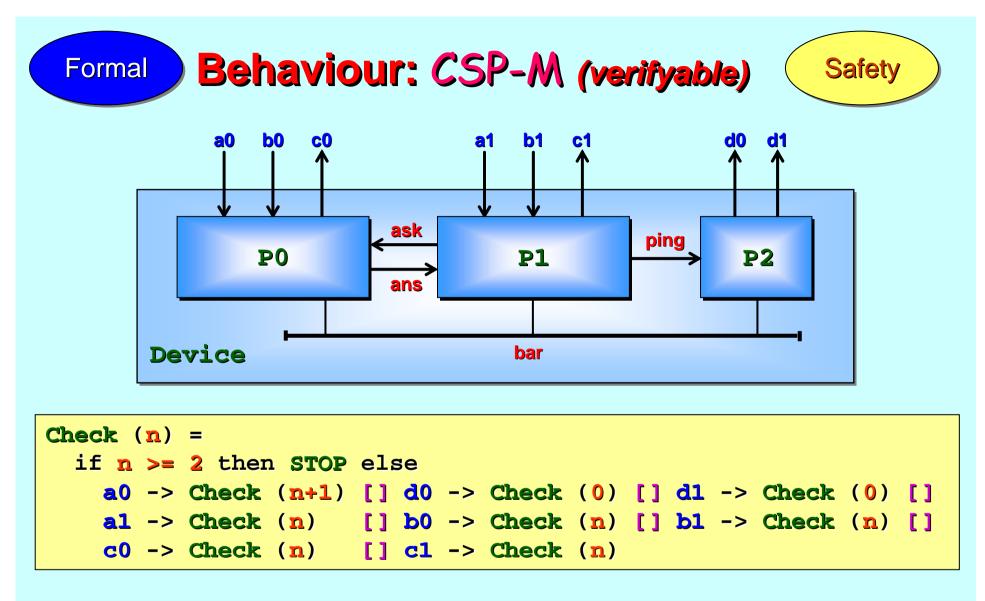
Simple: write a process checking the signals to/from **Device**, looking for the bad scenario and deadlocks if spotted. This is just programming ...



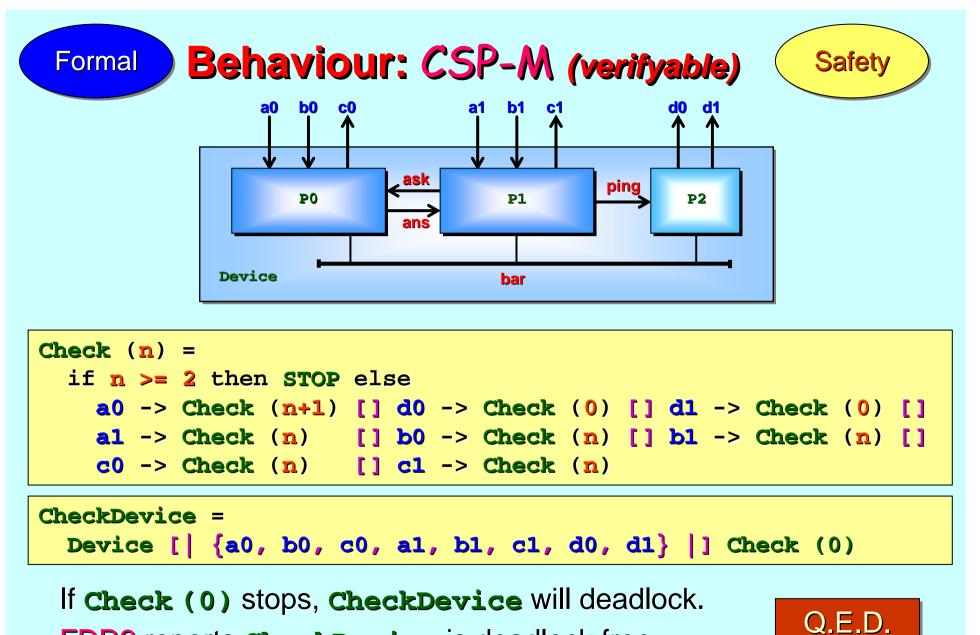
looking for the bad scenario and deadlocks if spotted. This is just programming ...



The operator "[]" means wait for one or more of the operand processes **to become able** to run ... choose one of them and run.

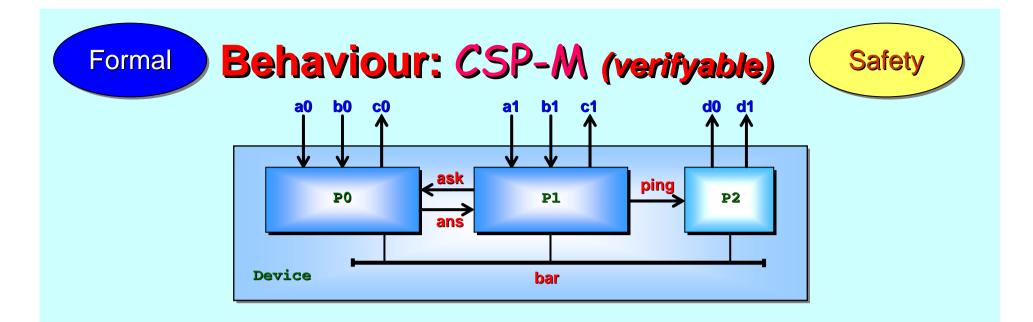


The parameter to **Check** records how many **a**0 signals have been received since the last **d**0 or **d**1, stopping if this reaches 2.



FDR2 reports **CheckDevice** is deadlock free.

Therefore, **Check (0)** never stops (& the bad thing can't happen). 12-Oct-10 Copyleft (GPL) P.H.Welch and J.B.Pedersen

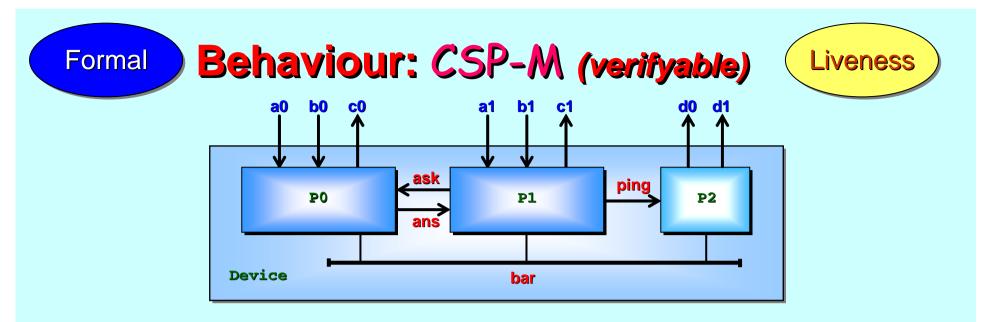


**Note:** protocol checking monitors, such as **Check**, are sometimes used live to ensure adherence at run-time (e.g. in device drivers). We are using **Check** purely for static analysis – it has no role at run-time and, therefore, no impact on performance.

If **Check (0)** stops, **CheckDevice** will deadlock. **FDR2** reports **CheckDevice** is deadlock free.



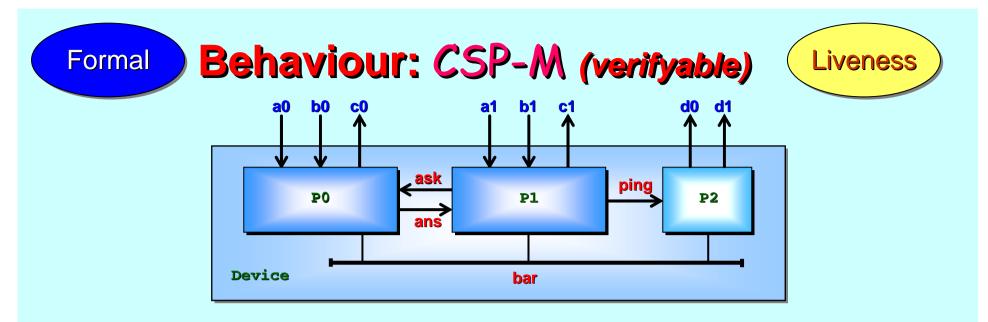
Therefore, **Check (0)** never stops **(& the bad thing can't happen)**. 12-Oct-10 Copyleft (GPL) P.H.Welch and J.B.Pedersen



So far, our checks have concerned **safety** – namely that our system will not do harm (incorrect things). This is not enough! After all, the **STOP** process does not do incorrect things – it does nothing. **STOP trace refines** every process. **Trace refinement** is not enough.

A **CSP** *failure* is a state that a system reaches (represented by its *trace* to that point) where it *may refuse to synchronise* with its environment on some given set of events.

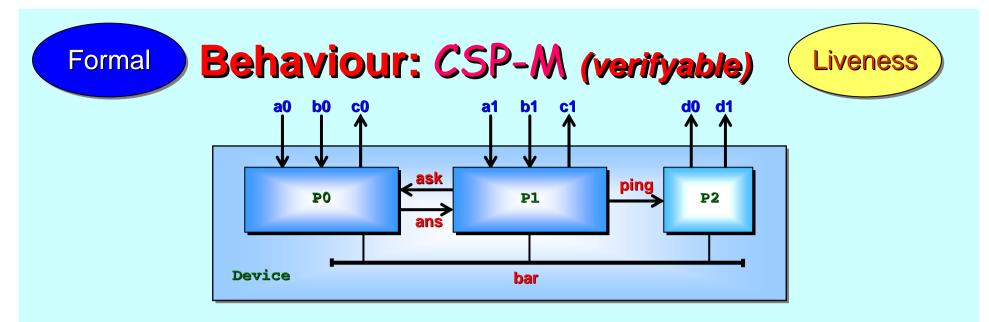
Process **p** failure refines **Q** if (all traces of **p** are traces of **Q**) and (all failures of **p** are failures of **Q**).



A **CSP** *failure* is a state that a system reaches (represented by its *trace* to that point) where it *may refuse to synchronise* with its environment on some given set of events.

Process **p** failure refines **Q** if (all its traces are traces of **Q**) and (all its failures are failures of **Q**).

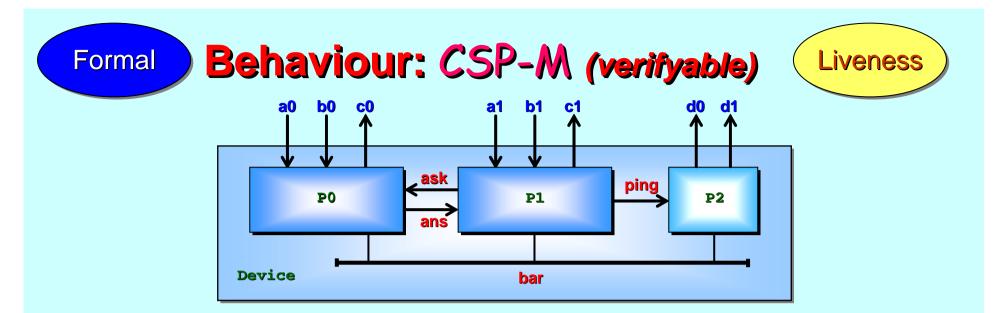
This is a powerful statement!  $\mathbf{P}$  can only do *traces* of  $\mathbf{Q}$  (so its safe). **More:** the *failures* of  $\mathbf{P}$  are allowed by  $\mathbf{Q}$ . If  $\mathbf{P}$  and  $\mathbf{Q}$  execute the same trace to a state where their environment offers a set of events that  $\mathbf{Q}$  will not refuse, then  $\mathbf{P}$  also will not refuse.



A **CSP** *failure* is a state that a system reaches (represented by its *trace* to that point) where it *may refuse to synchronise* with its environment on some given set of events.

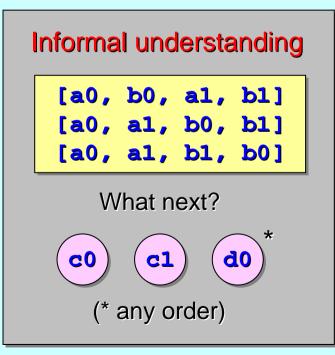
Process **p** failure refines **Q** if (all its traces are traces of **Q**) and (all its failures are failures of **Q**).

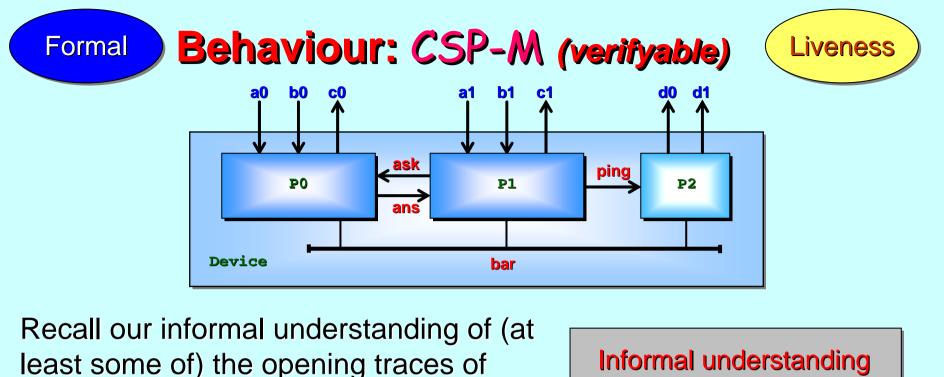
Whenever  $\mathbf{Q}$  stays alive (engaging with its environment), so does  $\mathbf{P}$  (and in the same way). If  $\mathbf{Q}$  is a specification directly written to express the required patterns of synchronisation,  $\mathbf{P}$  will fulfil them.



Recall our informal understanding of (at least some of) the opening traces of **Device (slides 20-37)**...

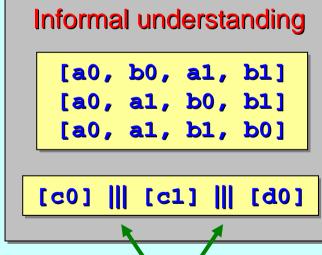
We can formalise the expression of those traces a bit better ...



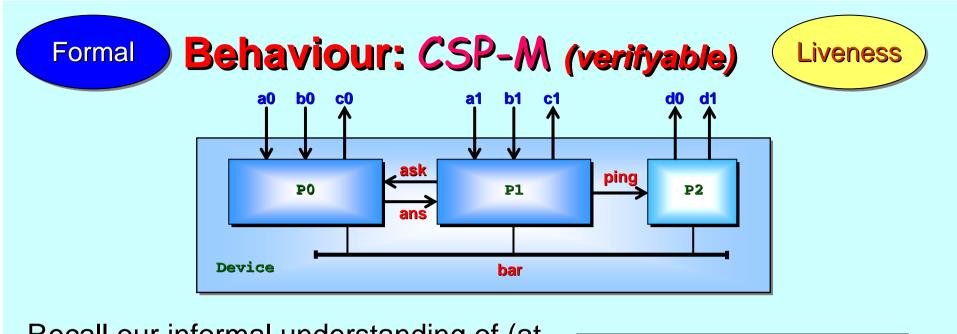


Device (slides 20-37) ...

We can formalise the expression of those traces a bit better ...

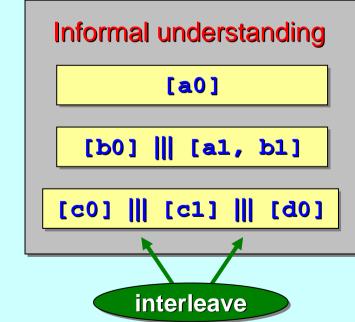


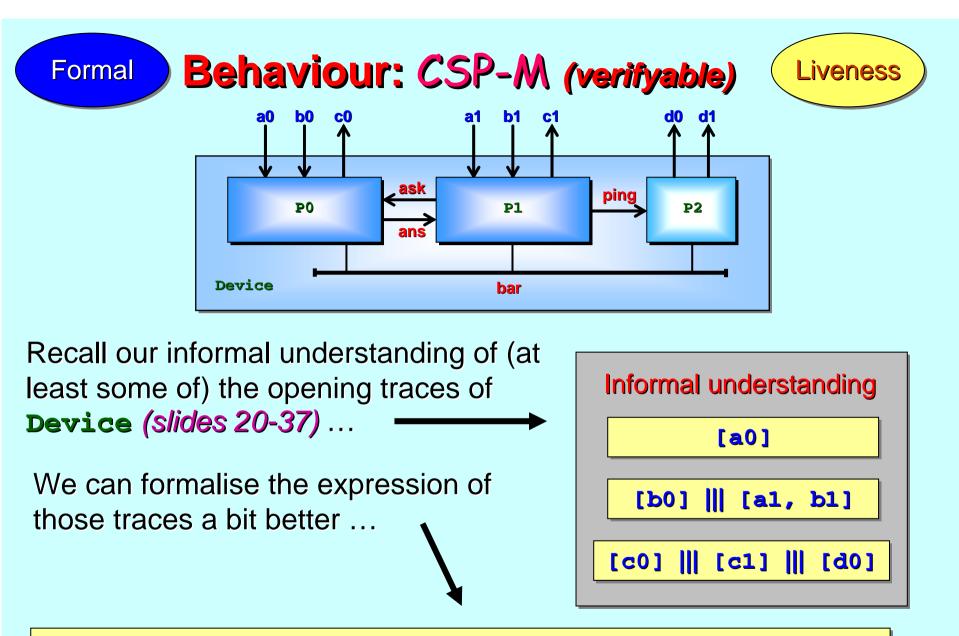
interleave



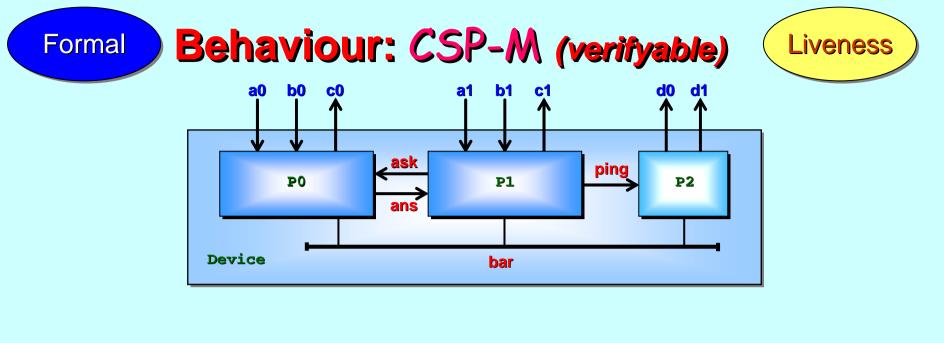
Recall our informal understanding of (at least some of) the opening traces of **Device (slides 20-37)**...

We can formalise the expression of those traces a bit better ...





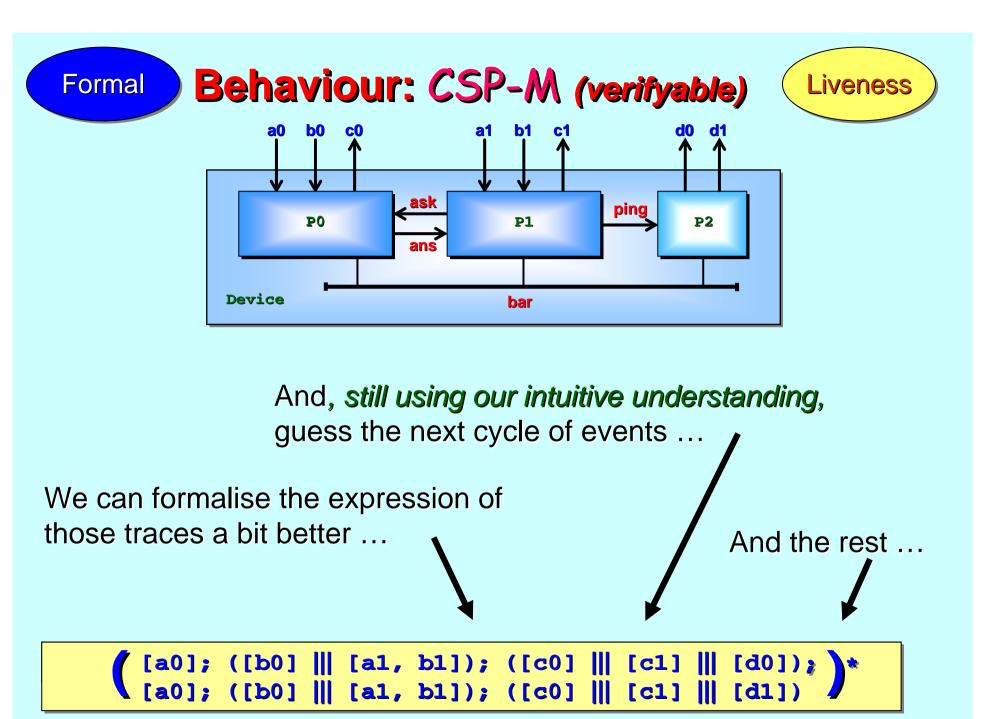
[a0]; ([b0] ||| [a1, b1]); ([c0] ||| [c1] ||| [d0])

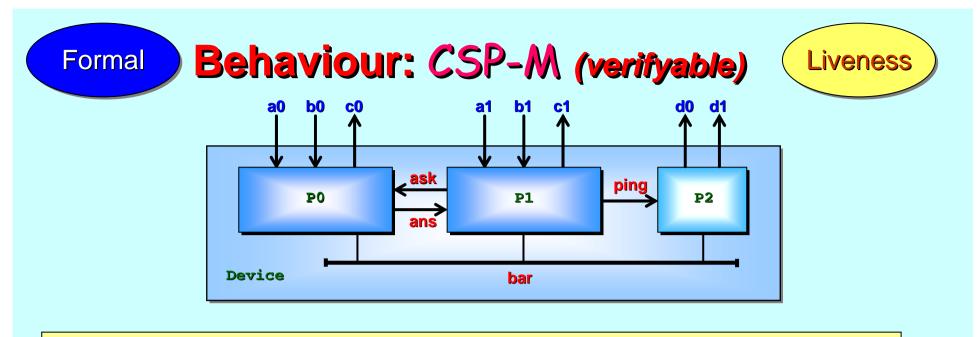


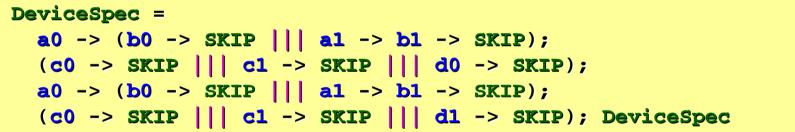
And, still using our intuitive understanding, guess the next cycle of events ...

We can formalise the expression of those traces a bit better ...

#### [a0]; ([b0] ||| [a1, b1]); ([c0] ||| [c1] ||| [d0]); [a0]; ([b0] ||| [a1, b1]); ([c0] ||| [c1] ||| [d1])

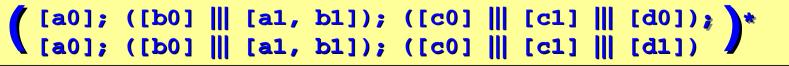


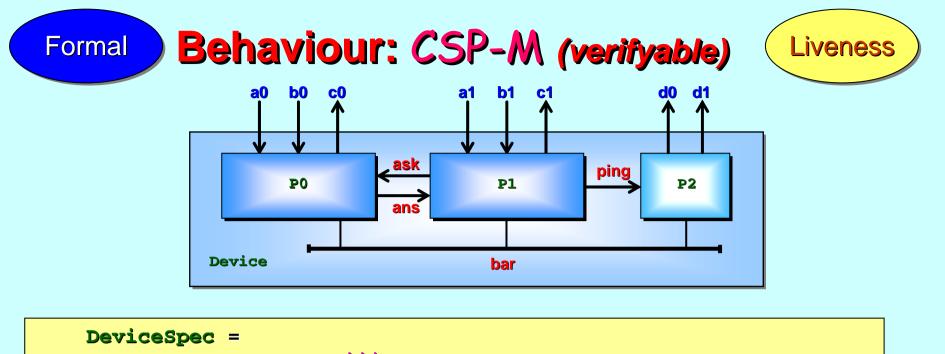


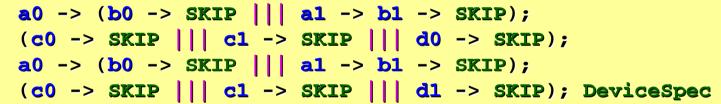


From such trace expressions, we can directly write down a **CSP** process that offers all of them to its environment ...

This generation can be automated.

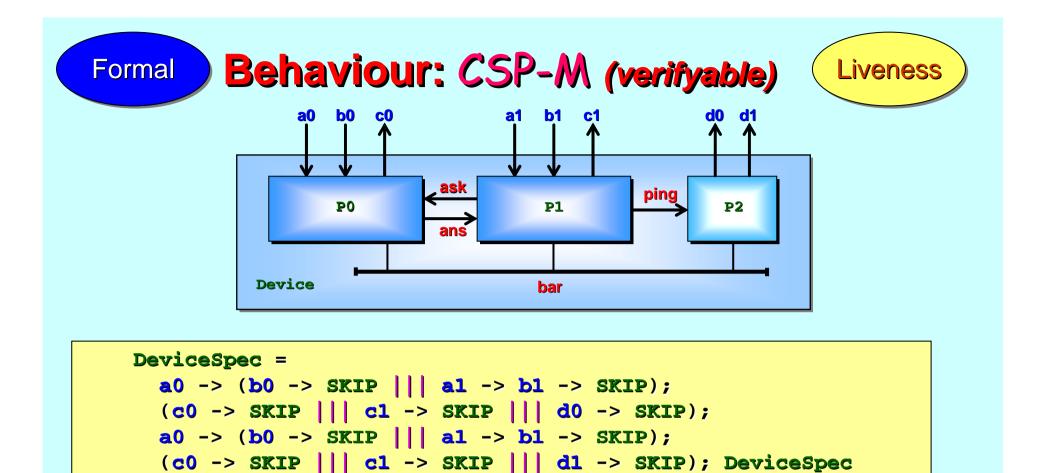




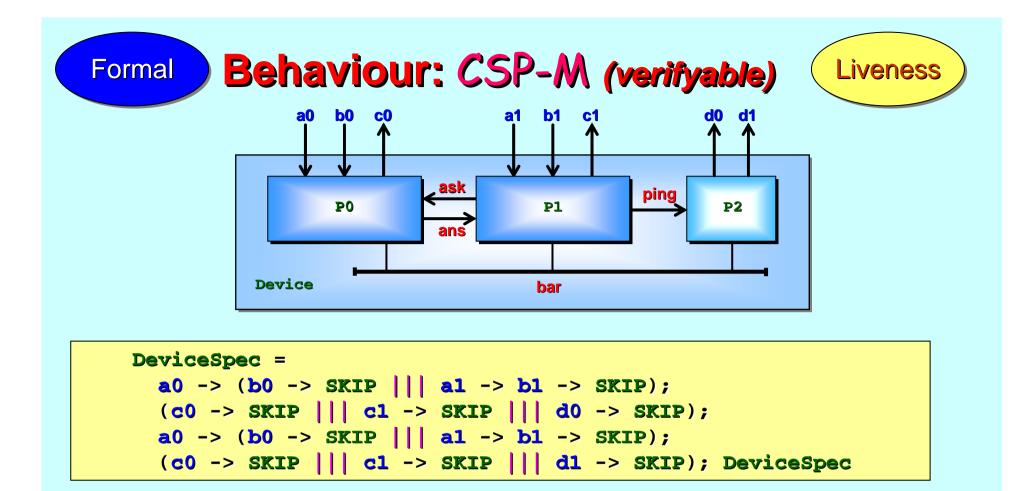


**DeviceSpec** is an explicit specification of all signal patterns we expect (or need) **Device** to be able to perform.

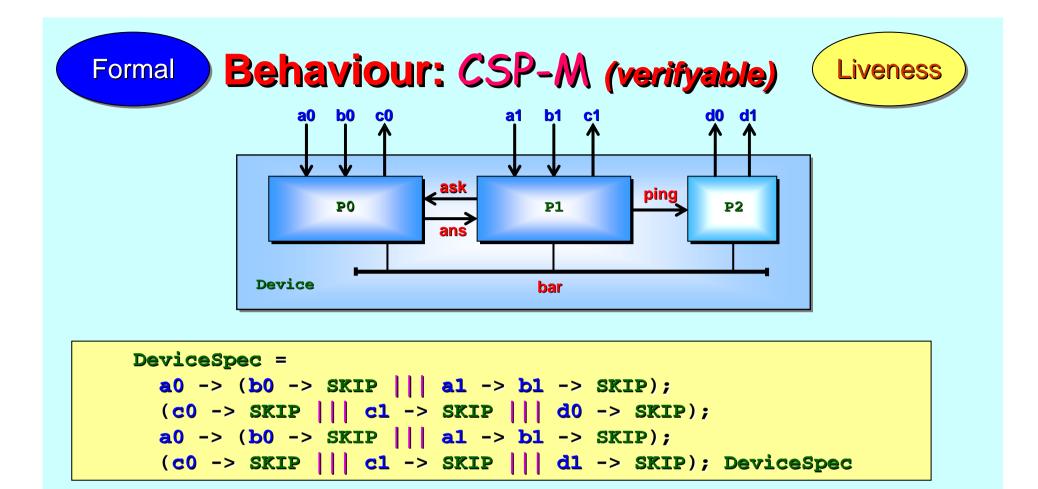
**FDR2** reports **Device** *failure refines* **CheckDevice**.  $\bigcirc$   $\bigcirc$   $\bigcirc$   $\bigcirc$   $\bigcirc$   $\bigcirc$  and *failures*.



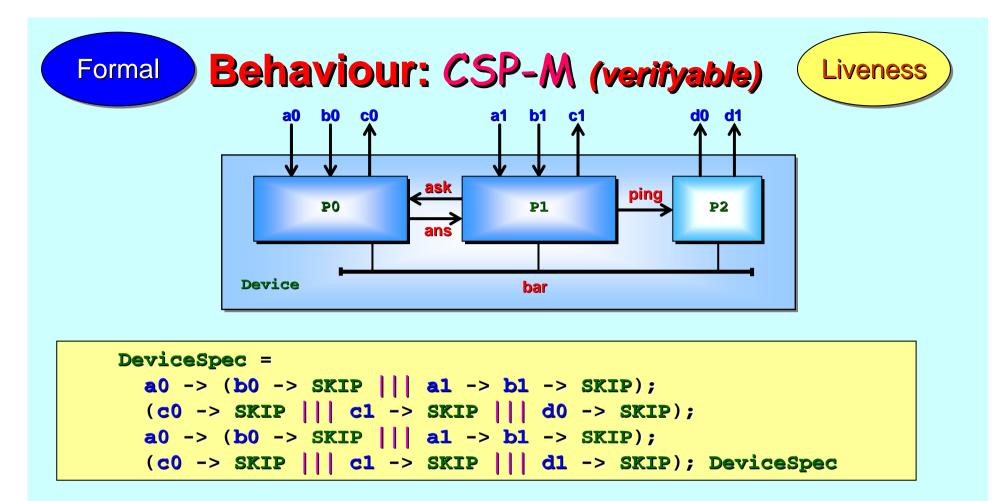
**Device** was not *implemented* as **DeviceSpec** because of the three independent functions (*weapons systems*, *vision processing* and *motion stability*) it had to perform. *Process-oriented design* led to its three communicating sub-systems.



Whilst our intuition indicated that the first two lines of **DeviceSpec** reflected the initial behaviour of **Device**, it was unclear whether the pattern repeated cleanly as its sub-components started looping.



One way to ensure this is to add another barrier (**bar**) at the end of each loop of **P0** and **P1** and half-loop of **P2**. The *failures equivalence* of **Device** and **DeviceSpec** shows that the pattern does indeed repeat cleanly and, so, this overhead is not necessary.



Rather than being deduced after implementation, **DeviceSpec** may be part of the specification for the behaviour of **Device**. We certainly need assurance of the behaviour of **Device** to use it securely with other components. All its patterns of synchronisation (for **safety** and *liveness* questions) can be trivially deduced from **DeviceSpec**.

### Class experience

The case study presented was developed from one first worked through in a single lesson of a graduate class in concurrency at UNLV in the spring of 2010.

They had previously studied a range of concurrency approaches, including *process-oriented* material from the Kent "Concurrency Design and Practice" course (presented at last year's workshop).

They were comfortable with using occam- $\pi$  in non-trivial projects (thousands of interacting processes), so the example system here would be considered fairly simple.

Nevertheless, it was appreciated that relying just on intuitive understanding is unsafe – especially if the application were safety critical.

### Class experience

During the exercise, students were given an overview (through examples) of CSP-M syntax, with semantics defined by relating back to occam- $\pi$  syntax and semantics.

The functional nature of CSP-M, compared with the imperative nature of occam- $\pi$ , was no particular problem.

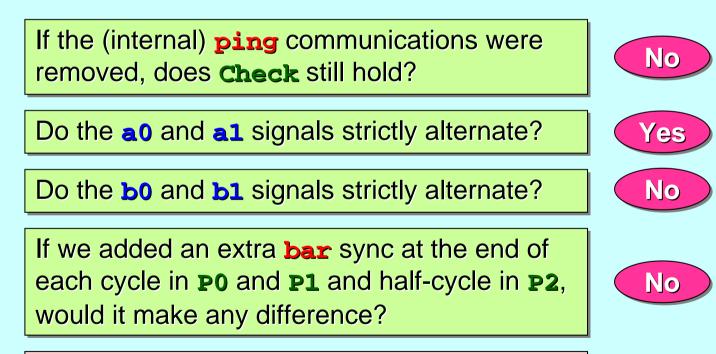
Working with **FDR2** through its **GUI** was not very sexy (by modern **GUI** standards) – but easy enough.

Checking their own (initial) test sequences for **Device** signals was very simple. Correct confirms/rejects were obtained.

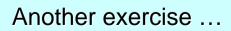
Writing safety-checking processes (like **Device**) for long term dangers was harder – but they warmed to this with practice.

### Class experience

*What-ifs* on the behaviour of the system could be explored and answered without running any code ... e.g.



If the elevator cabin is not at a floor, might the floor doors to the elevator shaft still open?



#### occam- $\pi$ / CSP-M

**occam-** $\pi$  teams well with **CSP-**M to provide efficient executables and rich formal analysis.

Of course, it would be better if only one syntactic representation were needed. We are working on extending **occam-** $\pi$  to include **verification assertions** (about **deadlock**, **livelock**, **determinism** and **refinement**). Its compiler will generate suitably abstracted CSP-M and interact with the FDR2 model checker, feeding back results in terms of the source **occam-** $\pi$  program.

Together with the ancient formal *Laws of occam Programming*<sup>\*</sup>, this moves occam- $\pi$  towards a process algebra in its own right.

http://portal.acm.org/citation.cfm?id=53255

[A.W.Roscoe and C.A.R.Hoare, 1988]

\*

#### Observation

Formal verification of the behaviour of concurrent processes has been achieved – *by students* – even though they engaged in only simple reasoning themselves.

The complexity of synchronisation and communication analysed went far beyond the *embarrassingly parallel*.

Aside: model checking found an error overlooked in developing the case study on paper (the need for ping) ... which shows the necessity for formal checks (especially when those responsible think they won't make mistakes!).

Further reading: Santa Claus: Formal Analysis of a Process Oriented Solution\*.

http:/doi.acm.org/10.1145/1734206.1734211

TOPLAS, [April, 2010]

\*

Can we teach students *(those who love to program, anyway)* concurrency so that:

they quickly develop a correct and intuitive understanding of the primitive mechanisms (e.g. processes, communication, synchronisation, networks) and higher level patterns (e.g. client-server, phased barrier, I/O-PAR) ... ?

they can use those primitives and patterns with the same fluency as they use serial computing primitives, without tripping over dark hazards ... ?

they can develop their own patterns when the standard ones don't apply ...?

they can use formal methods to verify good behaviour (e.g. freedom from deadlock and livelock, safety, liveness), without training in the underlying mathematics (process algebra, denotational semantics) ... ?

they can do this as normal everyday practice, without any sense of fear ...?

Can we teach students *(those who love to program, anyway)* concurrency so that:

they quickly develop a correct and intuitive understamechanisms (e.g. processes, communication, syand higher level patterns (e.g. client-server, ph the primitive *n, networks)* I/O-PAR) ... ?

they can use those primitives and patter serial computing primitives, without tri

they can develop their own pat

Jame fluency as they use Jark hazards ... ?

ne standard ones don't apply ... ?

they can use formal met deadlock and livelock mathematics (proc

they can do this

y good behaviour (e.g. freedom from ness), without training in the underlying denotational semantics) ... ?

al everyday practice, without any sense of fear ... ?

12-Oct-10

Copyleft (GPL) P.H.Welch and J.B.Pedersen