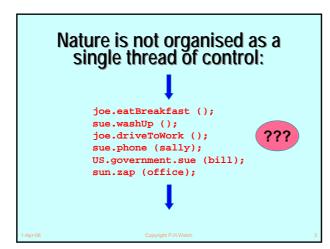


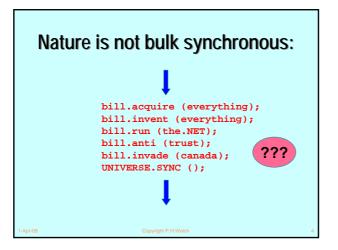
Motivation and Applications

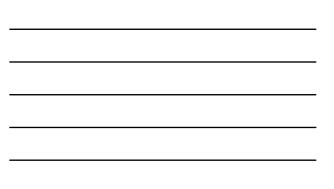
Thesis

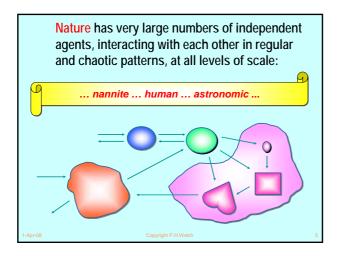
- Natural systems are robust, efficient, long-lived and continuously evolving. We should take the hint!
- Look on concurrency as a core design mechanism not as something difficult, used only to boost performance.
- Some applications
 - Hardware design and modelling.
 - Static embedded systems and parallel supercomputing.
 Field-programmable embedded systems and dynamic
 - supercomputing (e.g. SETI-at-home).
 Dynamic distributed systems, eCommerce, operating
 - systems and games.
 Biological system and *nanite* modelling.

Copyright P.H.Wel

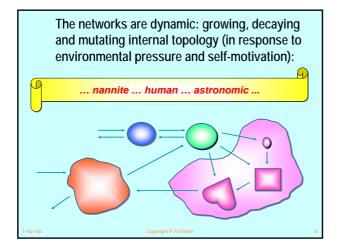


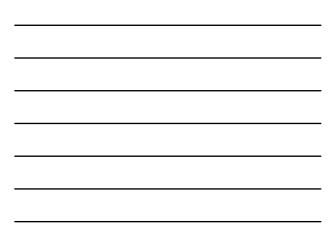


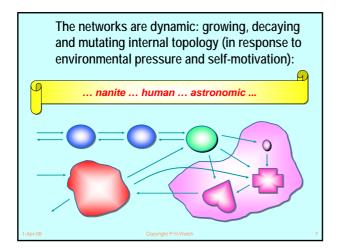














The Real World and Concurrency

Computer systems - to be of use in this world - need to model that part of the world for which it is to be used.

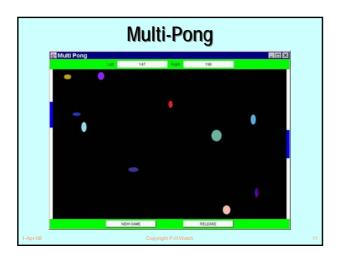
If that modeling can reflect the natural concurrency in the system ... it should be *simpler*.

Yet concurrency is thought to be an *advanced* topic, *harder* than serial computing (which therefore needs to be mastered first).

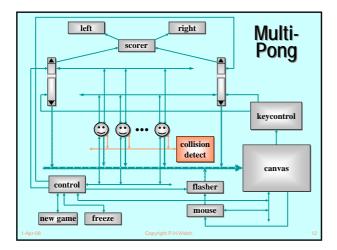


What we want from Parallelism

- A powerful tool for *simplifying* the description of systems.
- Performance that spins out from the above, but is not the primary focus.
- A model of concurrency that is *mathematically clean*, yields no engineering surprises and scales well with system complexity.









Good News!

The good news is that we can worry about each process on its own. A process interacts with its environment through its channels. It does not interact directly with other processes.

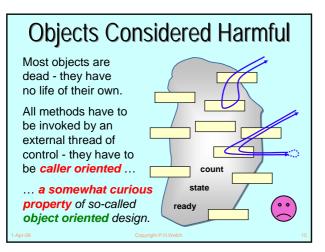
Some processes have *serial* implementations - *these are just like traditional serial programs*.

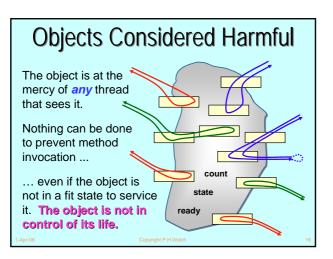
Some processes have *parallel* implementations - *networks of sub-processes (think hardware).*

Our skills for serial logic sit happily alongside our new skills for concurrency - there is no conflict. *This will scale!*

Java Monitors - CONCERNS

- Easy to learn but very difficult to apply ... safely ...
- Monitor methods are *tightly interdependent* their semantics compose in *complex ways* ... the whole skill lies in setting up and staying in control of these complex interactions ...
- Threads have no structure ... there are no threads within threads ...
- Big problems when it comes to scaling up complexity ...



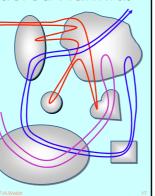




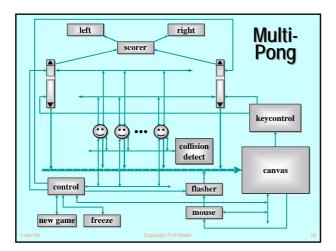


Each single thread of control snakes around objects in the system, bringing them to life *transiently* as their methods are executed.

Threads cut across object boundaries leaving spaghetti-like trails, paying no regard to the underlying structure.









Java Monitors - CONCERNS

- Almost all multi-threaded codes making direct use of the Java monitor primitives that we have seen (*including our own*) contained race or deadlock hazards.
- Sun's Swing classes are not thread-safe ... why not?
- One of our codes contained a race hazard that did not trip for two years. This had been in daily use, its sources published on the web and its algorithms presented without demur to several Java literate audiences.

Java Monitors - CONCERNS

<java.sun.com/products/jfc/tsc/articles/threads/threads1.html>

- "If you can get away with it, avoid using threads. Threads can be difficult to use, and they make programs harder to debug."
- "Component developers do not have to have an in-depth understanding of threads programming: toolkits in which all components must fully support multithreaded access, can be difficult to extend, particularly for developers who are not expert at threads programming."

Java Monitors - CONCERNS

<java.sun.com/products/jfc/tsc/articles/threads/threads1.html>

"It is our basic belief that extreme caution is warranted when designing and building multi-threaded applications ... use of threads can be very deceptive ... in almost all cases they make debugging, testing, and maintenance vastly more difficult and sometimes impossible. Neither the training, experience, or actual practices of most programmers, nor the tools we have to help us, are designed to cope with the non-determinism ... this is particularly true in Java ... we urge you to think twice about using threads in cases where they are not absolutely necessary ..."

	Java Monitors - CONCERNS	
	Sala monitorio Solito El mo	
	No guarantee that any synchronized method will ever be executed (e.g. <i>stacking</i> JVMs)	
	Even if we had above promise (e.g. <i>queueing</i> JVMs), standard design patterns for wait() / notify() fail to guarantee liveness (<i>"Wot, no chickens?"</i>)	
Se	96:	
	http://www.hensa.ac.uk/parallel/groups/wotug/java/discussion/3.html	
	http://www.nist.gov/itl/div896/emaildir/rt-j/msg00385.html	
	http://www.nist.gov/itl/div896/emaildir/rt-j/msg00363.html	
1-Apr-0	08 Copyright P.H.Welch	22

Java Monitors - CONCERNS

- Threads yield non-determinacy (and, therefore, scheduling sensitivity) straight away ...
- No help provided to guard against race hazards ...
- Overheads too high (> 30 times ???)
- Tyranny of Magic Names (e.g for listener callbacks)
- Learning curve is long ...
- Scalability (both in logic and performance) ???
- Theoretical foundations ???
 - (deadlock / livelock / starvation analysis ???)
 - (rules / tools ???)

Java Monitors - CONCERNS

- So, Java monitors are not something with which we want to think certainly not on a daily basis.
- But concurrency should be a powerful tool for **simplifying** the description of systems ...
- So it needs to be something I want to use and am comfortable with - on a daily basis!



A mathematical theory for specifying and verifying complex patterns of behaviour arising from interactions between concurrent objects.

CSP has a formal, and *compositional*, semantics that is in line with our informal intuition about the way things work.



Why CSP?

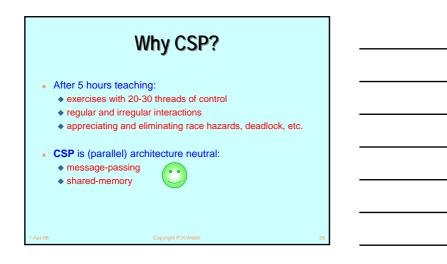
- Encapsulates fundamental principles of communication.
- Semantically defined in terms of structured mathematical model.
- Sufficiently expressive to enable reasoning about deadlock and livelock.
- Abstraction and refinement central to underlying theory.
- Robust and commercially supported software engineering tools exist for formal verification.

Why CSP?

• CSP libraries available for Java (JCSP, CTJ).

- Ultra-lightweight kernels* have been developed yielding sub-microsecond overheads for context switching, process startup/shutdown, synchronized channel communication and high-level shared-memory locks.
- Easy to learn and easy to apply ...

* not yet available for JVMs (or Core JVMs!)

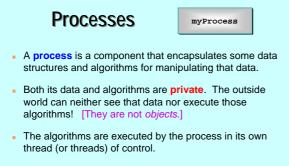


So, what is CSP?

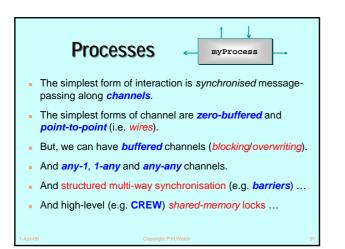
CSP deals with *processes*, *networks* of processes and various forms of *synchronisation / communication* between processes.

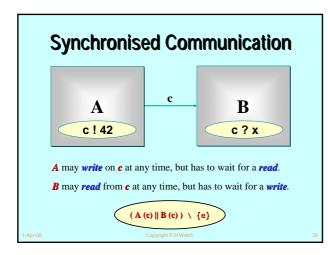
A network of processes is also a process - so **CSP** naturally accommodates layered network structures (*networks of networks*).

We do not need to be mathematically sophisticated to work with **CSP**. That sophistication is pre-engineered into the model. We benefit from this simply by using it.

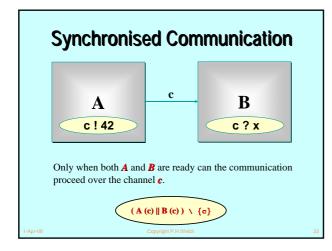


So, how does one process interact with another?

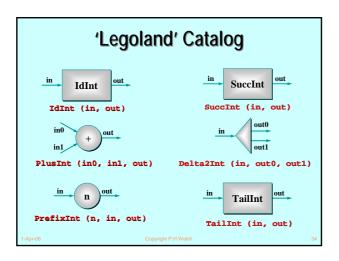










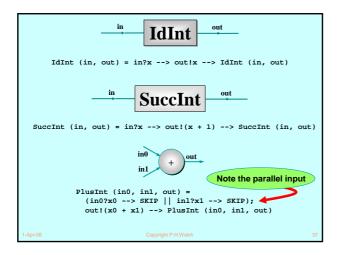




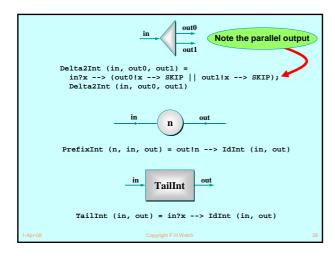
'Legoland' Catalog

- This is a catalog of fine-grained processes think of them as pieces of hardware (e.g. chips). They process data (ints) flowing through them.
- They are presented not because we suggest working at such fine levels of granularity ...
- They are presented in order to build up fluency in working with parallel logic.

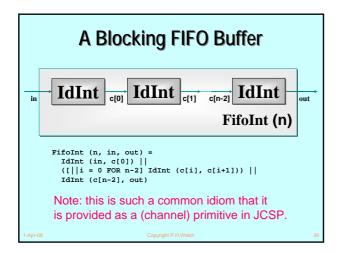
'Legoland' Catalog Parallel logic should become just as easy to manage as serial logic. This is not the traditionally held view ... But that tradition is wrong. CSP/occam people have always known this. Let's look at some CSP pseudo-code for these processes ...



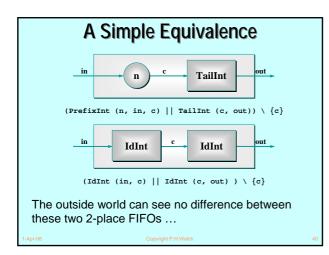




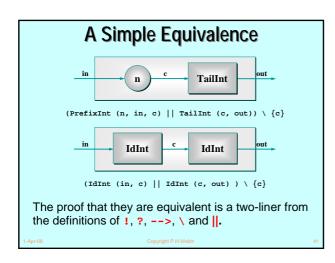






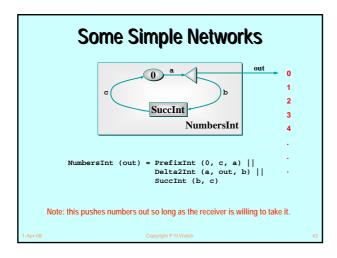




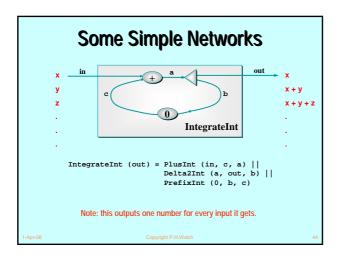




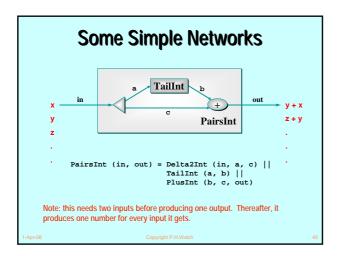




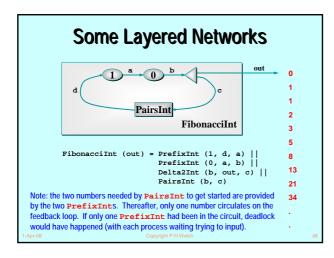




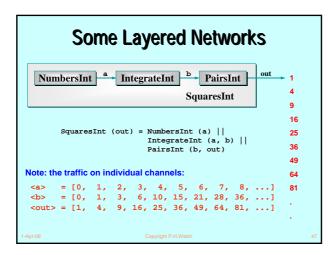




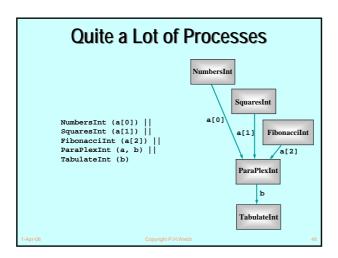




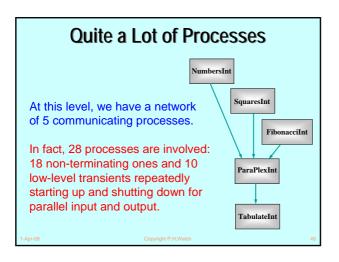




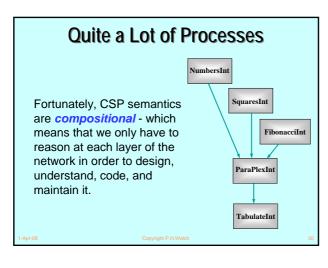


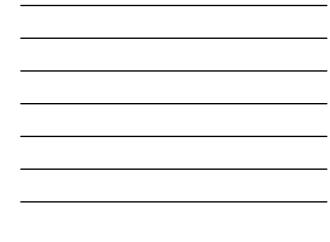


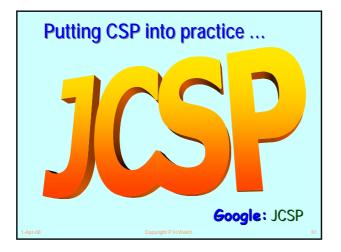


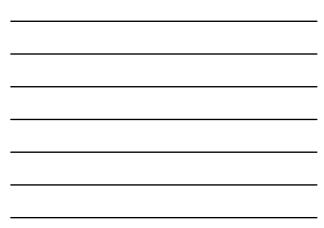


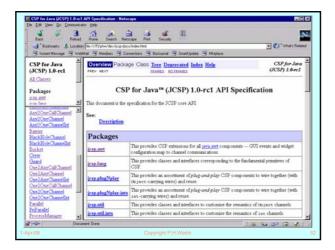




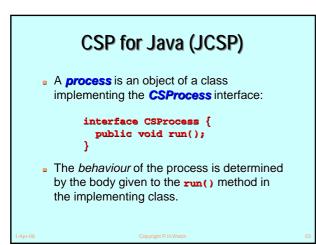


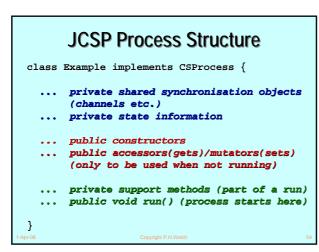


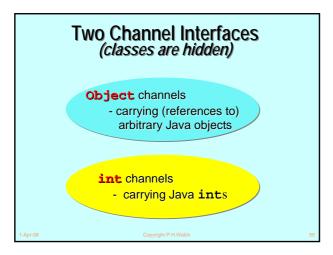












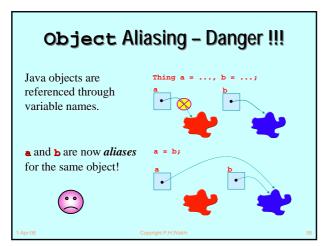


Channel-End Interfaces

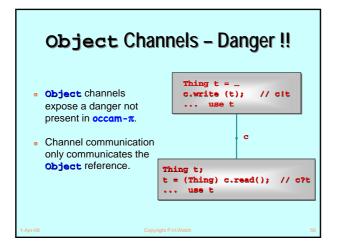
- Channel-end interfaces are what the processes see. Processes only need to care what kind of data they carry (ints or Objects) and whether the channels are for output, input or ALTing (i.e. choice) input.
- It is the network builder's concern to choose the variety of channel (e.g. synchronous, buffered, shared) to use when connecting processes together.

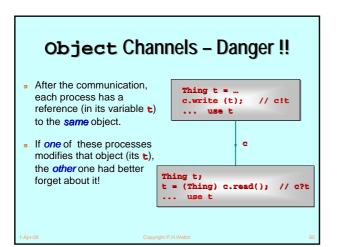
int Channels

- The **int** channels are convenient and secure.
- As with occam-π, it's difficult to introduce race hazards.
- For completeness, JCSP should provide channels for carrying all of the Java primitive data-types. These would be trivial to add. So far, there has been no pressing need.



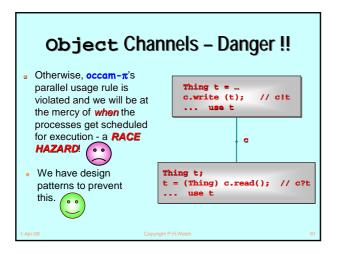




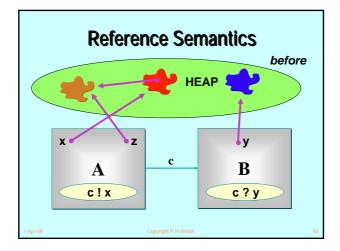




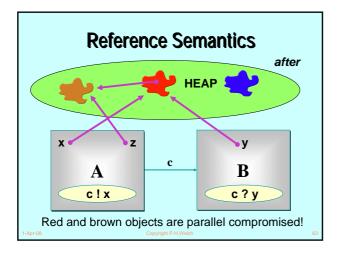




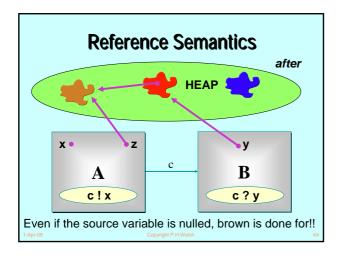


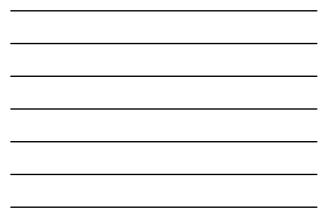


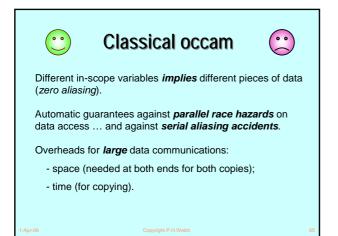


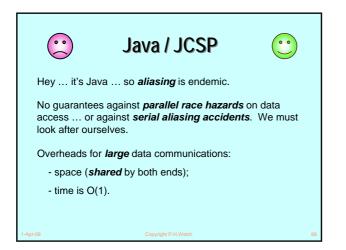




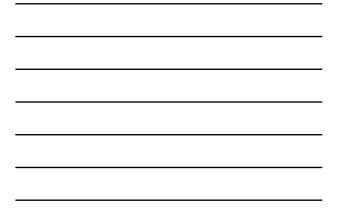






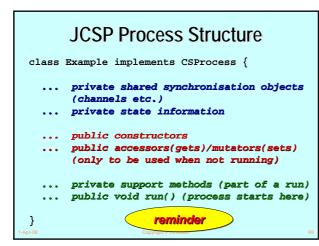


Object and Int Channels (<i>interfaces</i>)		
<pre>interface ChannelOutput { public void write (Object o); }</pre>	<pre>interface ChannelOutputInt { public void write (int i); }</pre>	
<pre>interface ChannelInput { public Object read (); }</pre>	<pre>interface ChannelInputInt { public int read (); }</pre>	
abstract class AltingChannelInput extends Guard implements ChannelInput { }	abstract class AltingChannelInputInt extends Guard implements ChannelInputInt { }	
1-Apr-08 Copyright P.	H.Welch 67	



Channel-End Interfaces

- Channel-ends are what the processes see they only care what kind of data they carry (ints or Objects) and whether the channels are for output, input or *ALTing* (i.e. choice) input.
- It will be the network builder's concern to decide the kinds of *channels* to be used and construct them for connecting processes.
- Let's review some of the *Legoland* processes this time in **JCSP**.





	in SuccInt	
c	lass SuccInt implements CSProcess {	
	private final ChannelInputInt in;	
	private final ChannelOutputInt out;	
	<pre>public SuccInt (ChannelInputInt in,</pre>	
	}	
	,	
}		
1-Apr-08	Copyright P.H.Welch	7

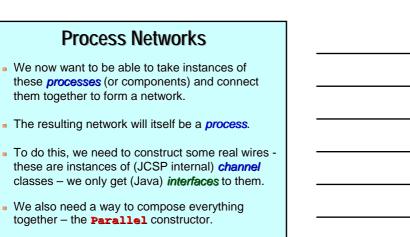


	in0 + out in1 +
	class PlusInt implements CSProcess {
	private final ChannelInputInt in0; private final ChannelInputInt in1; private final ChannelOutputInt out;
	<pre>public PlusInt (ChannelInputInt in0,</pre>
	} public void run () }
1-Apr-08	Copyright P.H.Welch

in0 + out in1 +
class PlusInt implements CSProcess {
private final channels (in0, in1, out)
public PlusInt (ChannelInputInt in0,)
<pre>public void run () { while (true) { int n0 = in0.read (); int n1 = in1.read (); out.write (n0 + n1); } } Note: the inputs really need to be done in parallel - later!</pre>
1-Apr-08 Copyright P.H.Welch 72

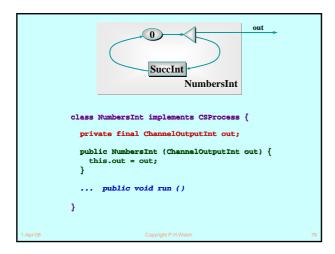


in out	
<pre>class PrefixInt implements CSProcess {</pre>	
private final int n;	
private final ChannelInputInt in;	
private final ChannelOutputInt out;	
public PrefixInt (int n, ChannelInputInt in,	
ChannelOutputInt out) {	
this.n = n;	
this.in = in;	
this.out = out;	
}	
<pre>public void run () {</pre>	
out.write (n);	
<pre>new IdInt (in, out).run ();</pre>	
}	
}	
1-Apr-08 Copyright P.H.Welch	73

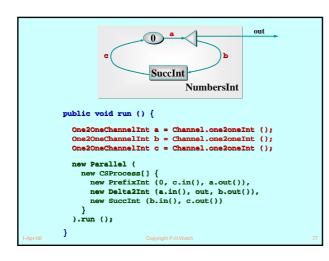


Parallel

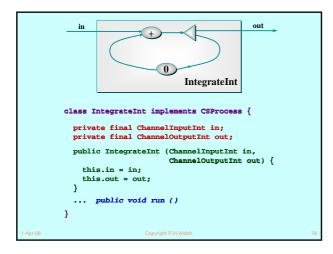
- Parallel is a CSProcess whose constructor takes an array of CSProcesses.
- Its run() method is the parallel composition of its given CSProcesses.
- The semantics is the same as for the occam-π PAR (or CSP ||).
- The *run()* terminates when and only when all of its component processes have terminated.



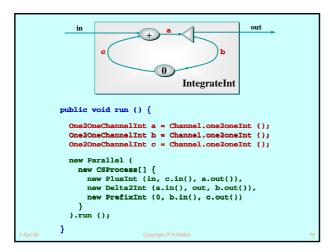




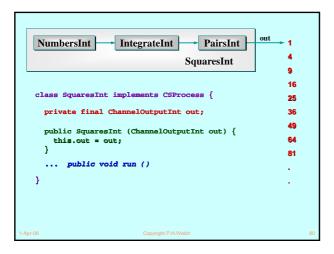




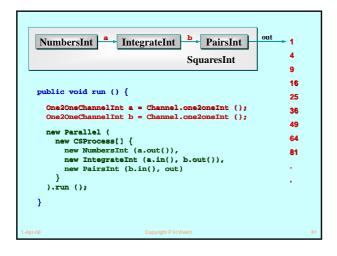




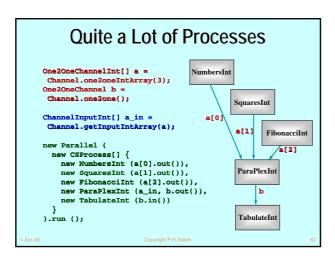




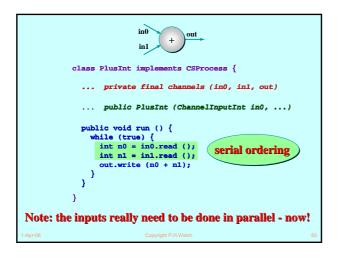




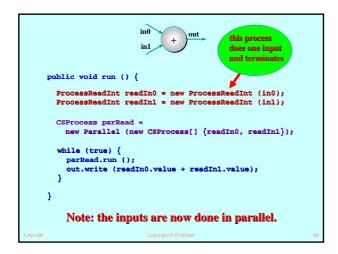




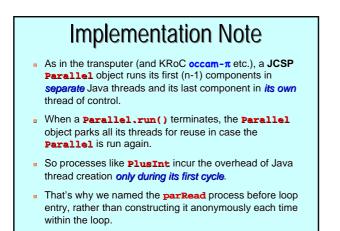


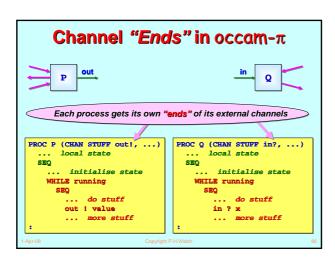




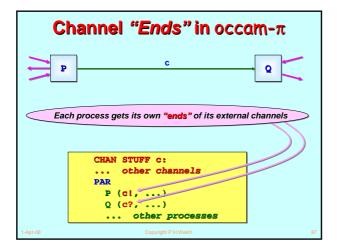




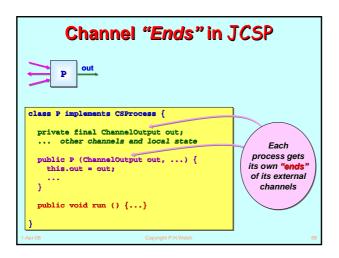




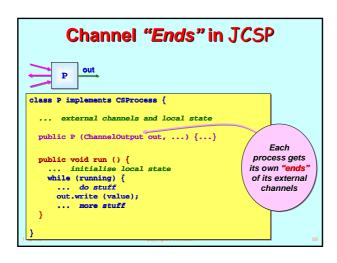




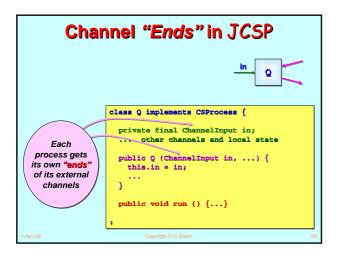




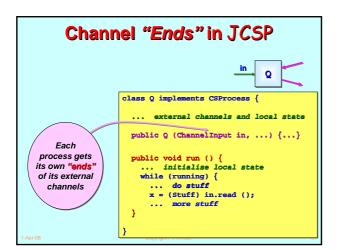




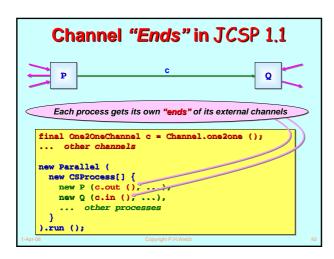




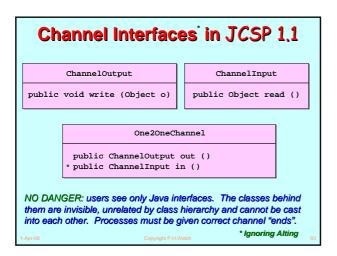




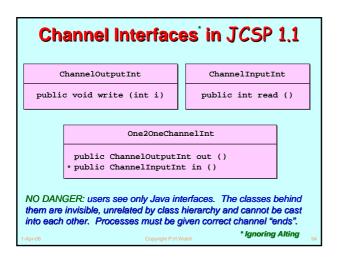




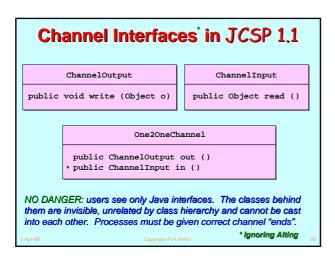




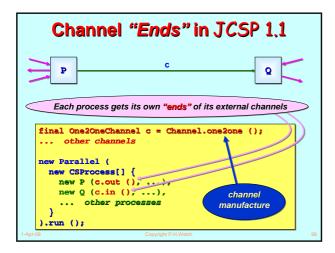




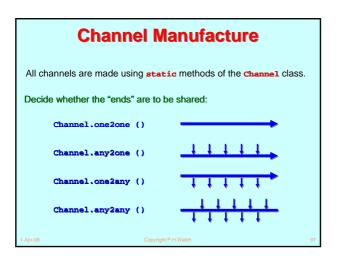




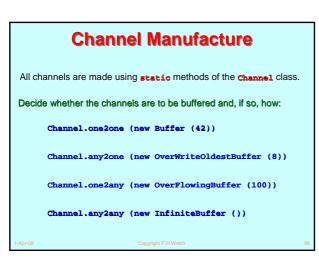


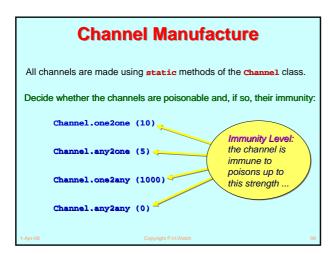




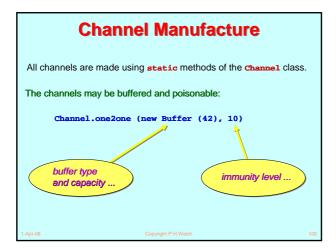




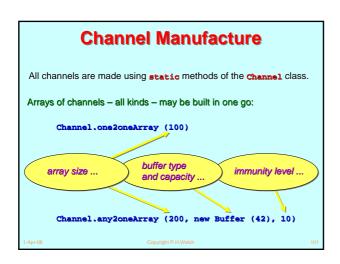














Channel Summary
The JCSP process view and use of its external channels:
Sees: ChannelInput, AltingChannelInput, ChannelOutput, ChannelInputInt, &C.
Increased safety - cannot violate "endianness"
A process does not (usually*) care about the kind of channel - whether it is shared, buffered, poisonable,
* If a process needs to share an external channel-end between many sub-processes, it must be given one that is shareable – i.e. an Any end. JCSP 1.1 does cater for this.

Channel Summary

The JCSP network view of channels:

The correct channel "ends" must be extracted from channels and plugged into the processes using them ...

Increased safety - cannot violate "endianness" ...

A wide range of channel kinds (fully synchronised, buffered, poisonable, typed) are built from the **Channel** class...

JCSP processes work only with *interfaces* both for channels (whatever their kind) and for channel-ends. We think this will prove safer than providing *classes*.

Copyright P.H.We

Deterministic Processes (CSP)

So far, our parallel systems have been *deterministic*:

- the values in the output streams depend only on the values in the input streams;
- the semantics is scheduling independent;
- no race hazards are possible.

CSP parallelism, on its own, *does not introduce non-determinism*.

This gives a firm foundation for exploring real-world models which cannot always behave so simply.

Non-Deterministic Processes (CSP)

In the real world, it is sometimes the case that things happen as a result of:

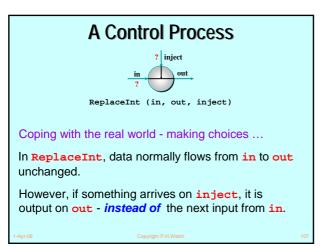
- what happened in the past;
- when (or, at least, in what order) things happened.

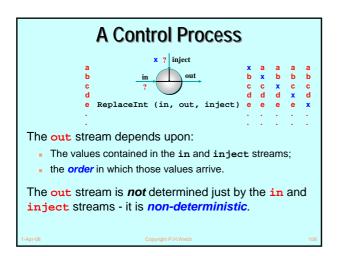
In this world, things are scheduling dependent.

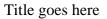
CSP (and JCSP) addresses these issues explicitly.

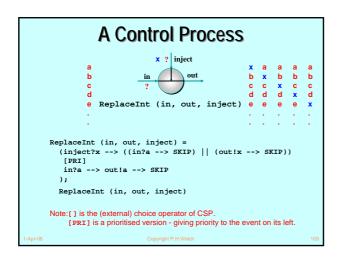
Non-determinism does not arise by default.

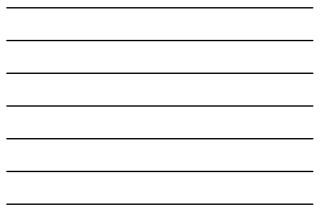


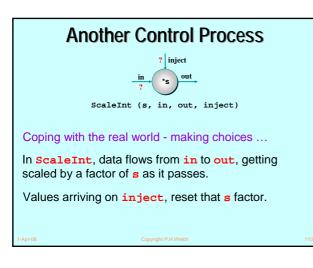


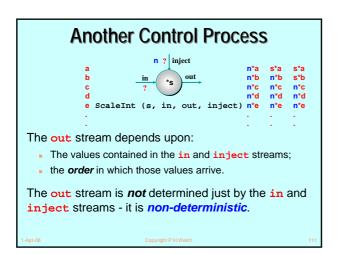




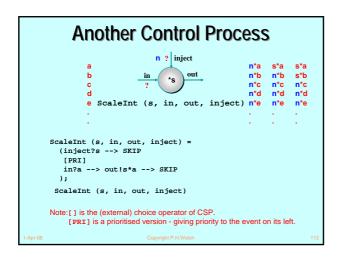




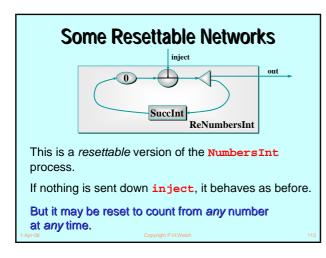


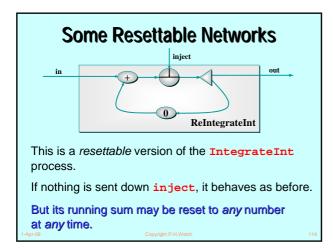




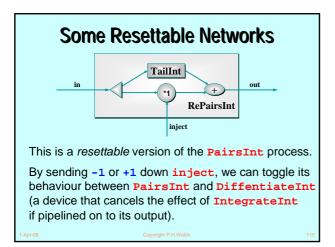




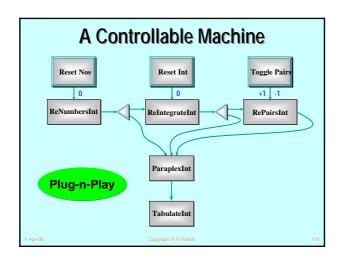




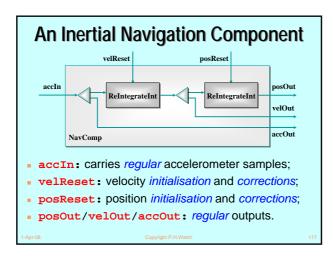














Deterministic Processes (JCSP)

So far, our JCSP systems have been *determistic*:

- the values in the output streams depend only on the values in the input streams;
- the semantics is scheduling independent;
- no race hazards are possible.

CSP parallelism, on its own, *does not introduce non-determinism*.

This gives a firm foundation for exploring real-world models which cannot always behave so simply.

Non-Deterministic Processes (JCSP)

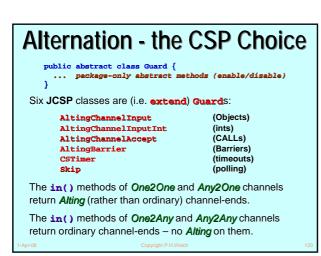
In the real world, it is sometimes the case that things happen as a result of:

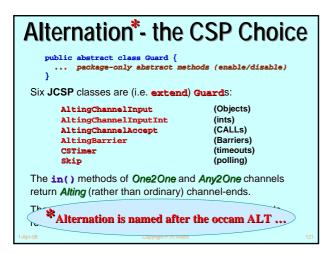
- what happened in the past;
- when (or, at least, in what order) things happened.

In this world, things are scheduling dependent.

CSP (JCSP) addresses these issues explicitly.

Non-determinism does not arise by default.





Ready/Unready Guards

- A channel guard is ready iff data is pending - i.e. a process at the other end has output to (or called) the channel and this has not yet been input (or accepted).
- A timer guard is ready iff its timeout has expired.
- A skip guard is always ready.

Alternation

For **ALTing**, a **JCSP** process must have a **Guard[]** array - this can be any mix of channel inputs, call channel accepts, timeouts or skips:

```
final Guard[] guards = {...};
```

It must construct an **Alternative** object for each such guard array:

```
final Alternative alt =
    new Alternative (guards);
```

The **ALT** is carried out by invoking one of the three varieties of select methods on the alternative.

alt.select()

This blocks passively until one or more of the guards are ready. Then, it makes an **ARBITRARY** choice of one of these ready guards and returns the index of that chosen one. If that guard is a **channel**, the **ALTing** process must then **read** from (or **accept**) it.

alt.priSelect()

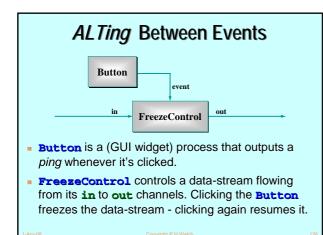
Same as above - except that if there is more than one ready guard, it chooses the one with the *lowest index*.

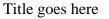
alt.fairSelect()

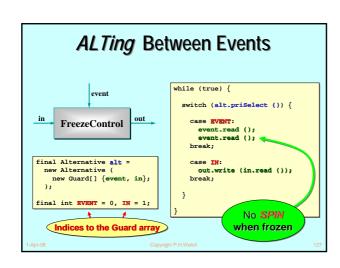
Same as above - except that if there are more than one ready guards, it makes a **FAIR** choice.

This means that, in successive invocations of **alt.fairSelect ()**, no ready guard will be chosen twice if another ready guard is available. At worst, no ready guard will miss out on **n** successive selections (where **n** is the number of guards).

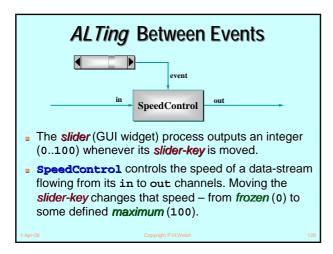
Fair alternation is possible because an *Alternative* object is tied to *one* set of guards.

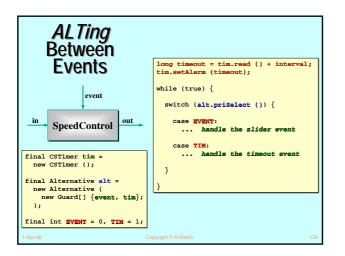




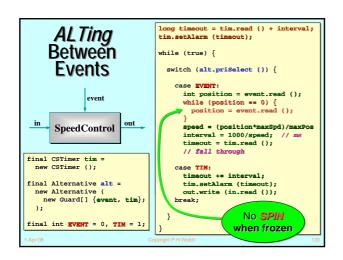




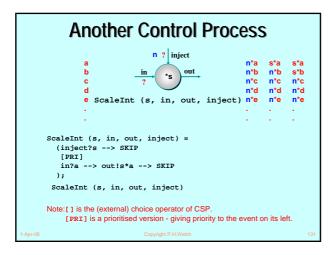




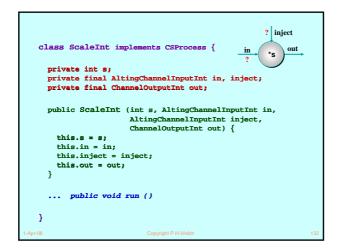




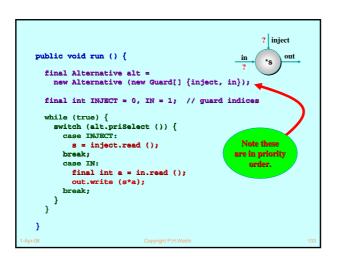




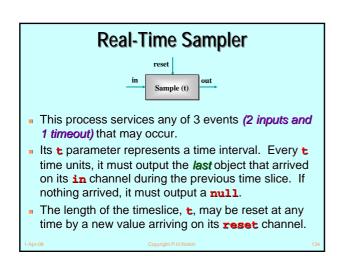


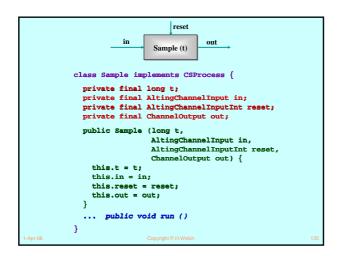


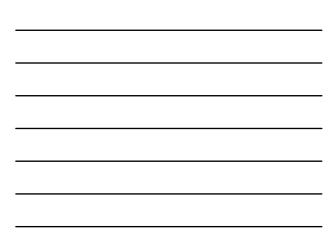


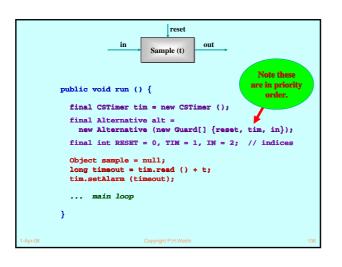




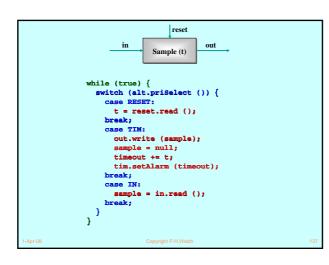




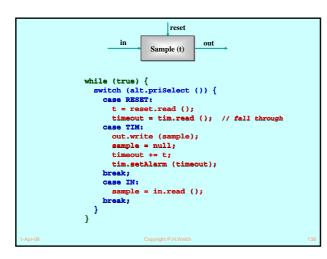




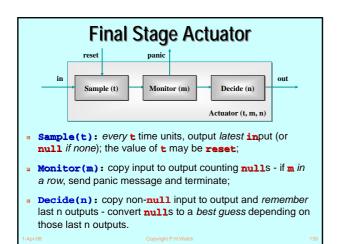




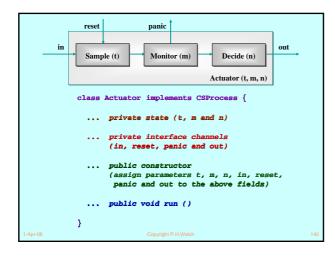




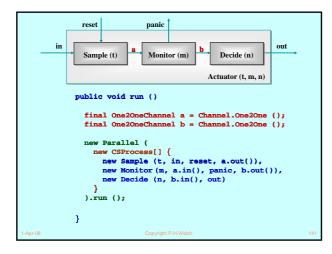














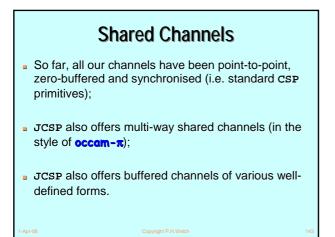
Pre-conditioned Alternation

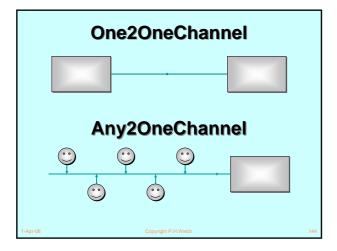
We may set an array of **boolean** *pre-conditions* on any of the **select** operations of an Alternative:

switch (alt.fairSelect (depends)) {...}

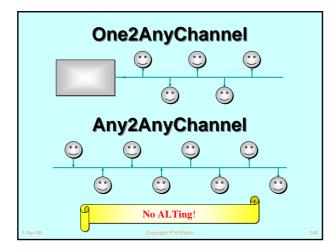
The **depends** array must have the same length as the **Guard** array to which the **alt** is bound.

The **depends** array, set at run-time, **enables/disables** the guards at corresponding indices. If **depends[i]** is **false**, that guard will be ignored - even if *ready*. This gives considerable flexibility to how we program the willingness of a process to service events.

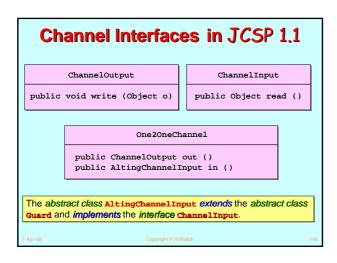




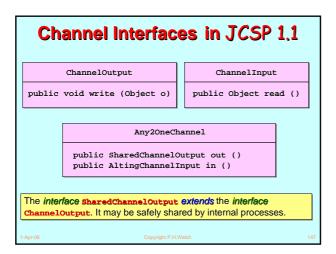




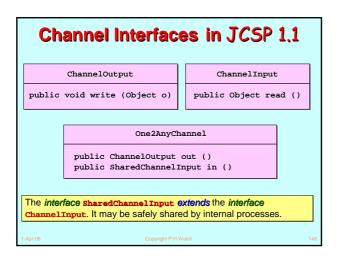




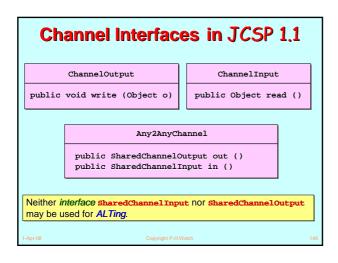




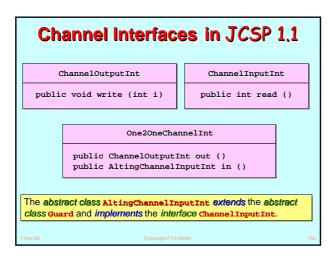




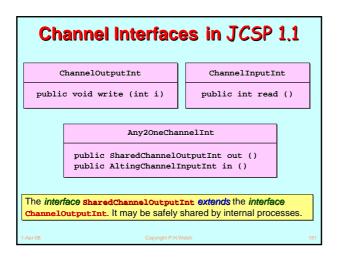




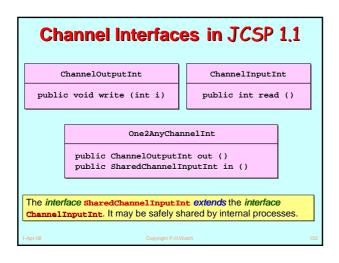




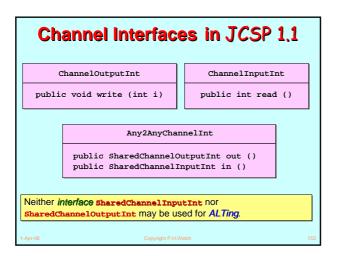




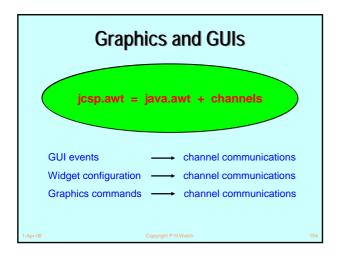




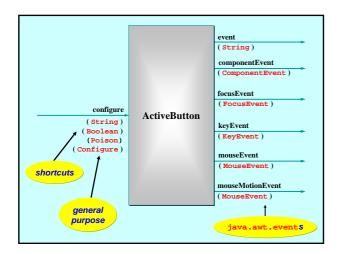




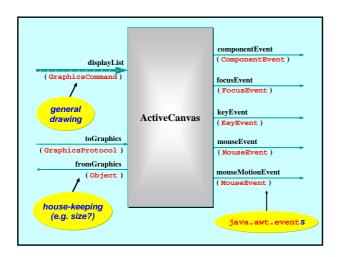




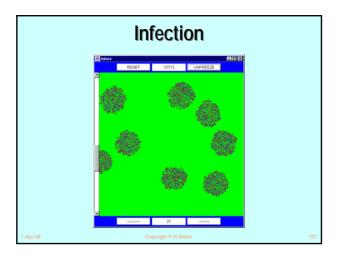




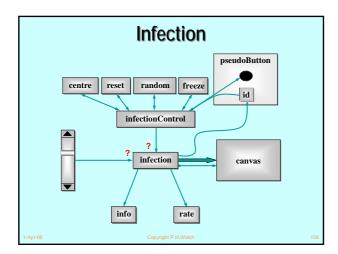




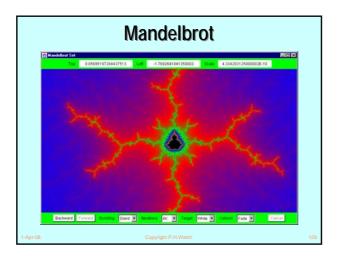




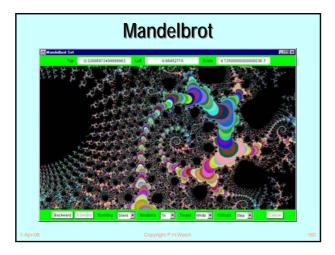




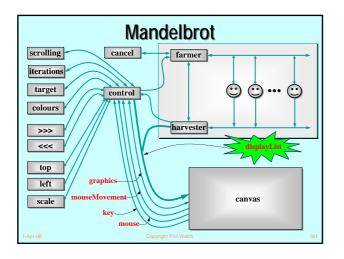




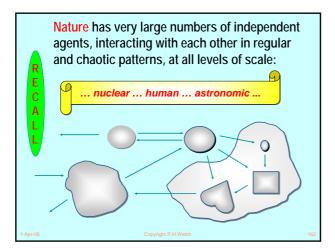














R

Ε

С

Α

Good News!

The good news is that we can worry about each process on its own. *A process interacts with its environment through its channels.* It does not interact directly with other processes.

Some processes have serial implementations - these are just like traditional serial programs.

Some processes have *parallel* implementations - *networks of sub-processes*.

Our skills for serial logic sit happily alongside our new skills for concurrency there is no conflict. *This will scale!*

Other Work

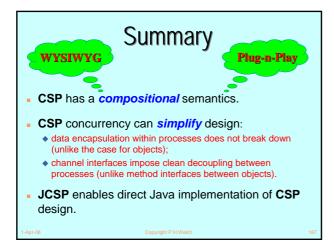
- A CSP model for the Java monitor mechanisms (synchronized, wait, notify, notifyAll) has been built.
- This enables any Java threaded system to be analysed in CSP terms - e.g. for formal verification of freedom from deadlock/livelock.
- Confidence gained through the formal proof of correctness of the JCSP channel implementation:
 - a JCSP channel is a non-trivial monitor the CSP model for monitors transforms this into an even more complex system of CSP processes and channels;
 - using FDR, that system has been proven to be a refinement of a single CSP channel and vice versa - Q.E.D.

Other Work

- Higher level synchronisation primitives (e.g. JCSP CALL channels, barriers, buckets, ...) that capture good patterns of working with low level CSP events.
- Proof rules and design tool support for the above.
- **CSP** *kernels* and their binding into JVMs to support **JCSP**.
- Communicating Threads for Java (CTJ):
 - this is another Java class library based on CSP principles;
 developed at the University of Twente (Netherlands) with
 - special emphasis on real-time applications it's excellent;
 - CTJ and JCSP share a common heritage and reinforce each other's on-going development - we do talk to each other!

Distributed JCSP.net Network channels + plus simple brokerage service for letting JCSP systems find and connect to each other transparently (from anywhere on the *Internet*). Virtual channel infrastructure to support this. All application channels auto-multiplexed over *single* (auto-generated) TCP/IP link between any two JVMs. Channel Name Server (CNS) provided. Participating ICSP systems just paed to know where this is More

- JCSP systems just need to know where this is. More sophisticated brokers are easily bootstrapped on top of the CNS (using JCSP).
- Killer Application Challenge:
 second generation Napster (no central control or database) ...



Summary

- CSP kernel overheads are sub-100-nanosecond (KRoC/CCSP). Currently, JCSP depends on the underlying Java threads/monitor implementation.
- Rich mathematical foundation:
 - 20 years mature recent extensions include simple priority semantics;
 - higher level design rules (e.g. *client-server, resource allocation priority, IO-par)* with formally proven guarantees (e.g. freedom from deadlock, livelock, process starvation);
 - commercially supported tools (e.g. FDR).
- We don't need to be mathematically sophisticated to take advantage of CSP. It's built-in. Just use it!

Summary Process Oriented Design (processes, syncs, alts, parallel, layered networks).

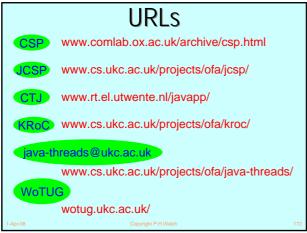
- WYSIWYG:
 - each process considered individually (own data, own control threads, external synchronisation);
 - leaf processes in network hierarchy are ordinary serial programs - all our past skills and intuition still apply;
 - concurrency skills sit happily alongside the old serial ones.
- Race hazards, deadlock, livelock, starvation problems: we have a rich set of design patterns, theory, intuition and tools to apply.

Conclusions

- We are *not* saying that Java's threading mechanisms need changing.
- Java is sufficiently flexible to allow many concurrency paradigms to be captured.
- JCSP is just a *library* Java needs no language change to support CSP.
- CSP rates serious consideration as a basis for any real-time specialisation of Java:
 - quality (robustness, ease of use, scalability, management of complexity, formalism);
 - *lightness* (overheads do not invalidate the above benefits they encourage them).

Acknowledgements

- Paul Austin the original developer of JCSP (p_d_austin@hotmail.com).
- Andy Bakkers and Gerald Hilderink the CTJ library (bks@el.utwente.nl, G.H.Hilderink@el.utwente.nl).
- Jeremy Martin for the formal proof of correctness of the JCSP channel (Jeremy.Martin@comlab.ox.ac.uk)
- Nan Schaller (ncs@cs.rit.edu), Chris Nevison (chris@cs.colgate.edu) and Dyke Stiles (dyke.stiles@ece.usu.edu) - for pioneering the teaching.
- The WoTUG community its workshops, conferences and people.



Stop Press

JCSP.net

