Parallel Functional Programming Lecture 3

Mary Sheeran

with thanks to Simon Marlow for use of slides
and to Koen Claessen for the guest appearance

http://www.cse.chalmers.se/edu/course/pfp

par and pseq

MUST

Pass an unevaluated computation to par

It must be somewhat expensive

Make sure the result is not needed for a bit

Make sure the result is shared by the rest of the program

par and pseq

MUST

Pass an unevaluated computatio to par

It must be somewhat expensive

Make sure the result is not need r a bit

Make sure the result is shared b st of the program

Demands an operational understanding of program execution

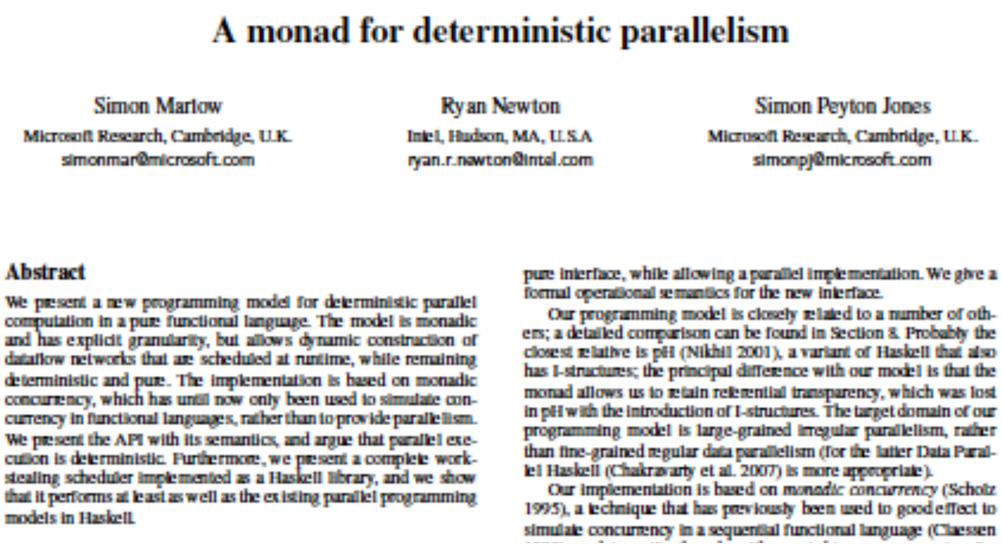
Eval monad plus Strategies

Eval monad enables expressing ordering between instances of par and pseq

Strategies separate algorithm from parallelisation

Provide useful higher level abstractions

But still demand an understanding of laziness



A monad for deterministic parallelism

Simon Marlow
Microsoft Research, Cambridge, U.K.
simonmar@microsoft.com

Ryan Newton
Intel, Hudson, MA, U.S.A
ryan.r.newton@intel.com

Simon Peyton Jones
Microsoft Research, Cambridge, U.K.
simonpj@microsoft.com

Abstract

We present a new programming model for deterministic parallel computation in a pure functional language. The model is monadic and has explicit granularity, but allows dynamic construction of dataflow networks that are scheduled at runtime, while remaining deterministic and pure. The implementation is based on monadic concurrency, which has until now only been used to simulate concurrency in functional languages, rather than to provide parallelism. We present the API with its semantics, and argue that parallel execution is deterministic. Furthermore, we present a complete work-stealing scheduler implemented as a Haskell library, and we show that it performs at least as well as the existing parallel programming models in Haskell.

pure interface, while allowing a parallel implementation. We give a formal operational semantics for the new interface.

Our programming model is closely related to a number of others; a detailed comparison can be found in Section 8. Probably the closest relative is pH (Nikhil 2001), a variant of Haskell that also has I-structures; the principal difference with our model is that the monad allows us to retain referential transparency, which was lost in pH with the introduction of I-structures. The target domain of our programming model is large-grained irregular parallelism, rather than fine-grained regular data parallelism (for the latter Data Parallel Haskell (Chakravarty et al. 2007) is more appropriate).

Our implementation is based on *monadic concurrency* (Scholz 1995), a technique that has previously been used to good effect to simulate concurrency in a sequential functional language (Claessen

Builds on Koen's paper

FUNCTIONAL PEARLS

A Poor Man's Concurrency Monad

Koen Claessen

Chalmers University of Technology

email: koen@cs.chalmers.se

Abstract

Without adding any primitives to the language, we define a concurrency monad transformer in Haskell. This allows us to add a limited form of concurrency to any existing monad. The atomic actions of the new monad are lifted actions of the underlying monad. Some extra operations, such as `fork`, to initiate new processes, are provided. We discuss the implementation, and use some examples to illustrate the usefulness of this construction.

JFP'99

Call this PMC

the Par Monad

Our goal with this work is to find a parallel programming model that is expressive enough to subsume Strategies, robust enough to reliably express parallelism, and accessible enough that non-expert programmers can achieve parallelism with little effort.

The Par Monad

```
data Par
instance Monad Par

runPar :: Par a -> a

fork :: Par () -> Par ()

data IVar
new :: Par (IVar a)
get :: IVar a -> Par a
put :: NFData a => IVar a -> a -> Par ()
```

Par is a monad for parallel computation

Parallel computations are pure (and hence deterministic)

forking is *explicit*

results are communicated through IVars

Slide by Simon Marlow

IVar

a write-once mutable reference cell

supports two operations: put and get

put assigns a value to the IVar, and may only be executed once per Ivar   Subsequent puts are an error

get waits until the IVar has been assigned a value, and then returns the value

the Par Monad

Implemented as a Haskell library

- surprisingly little code!
- includes a work stealing scheduler
- You get to roll your own schedulers!

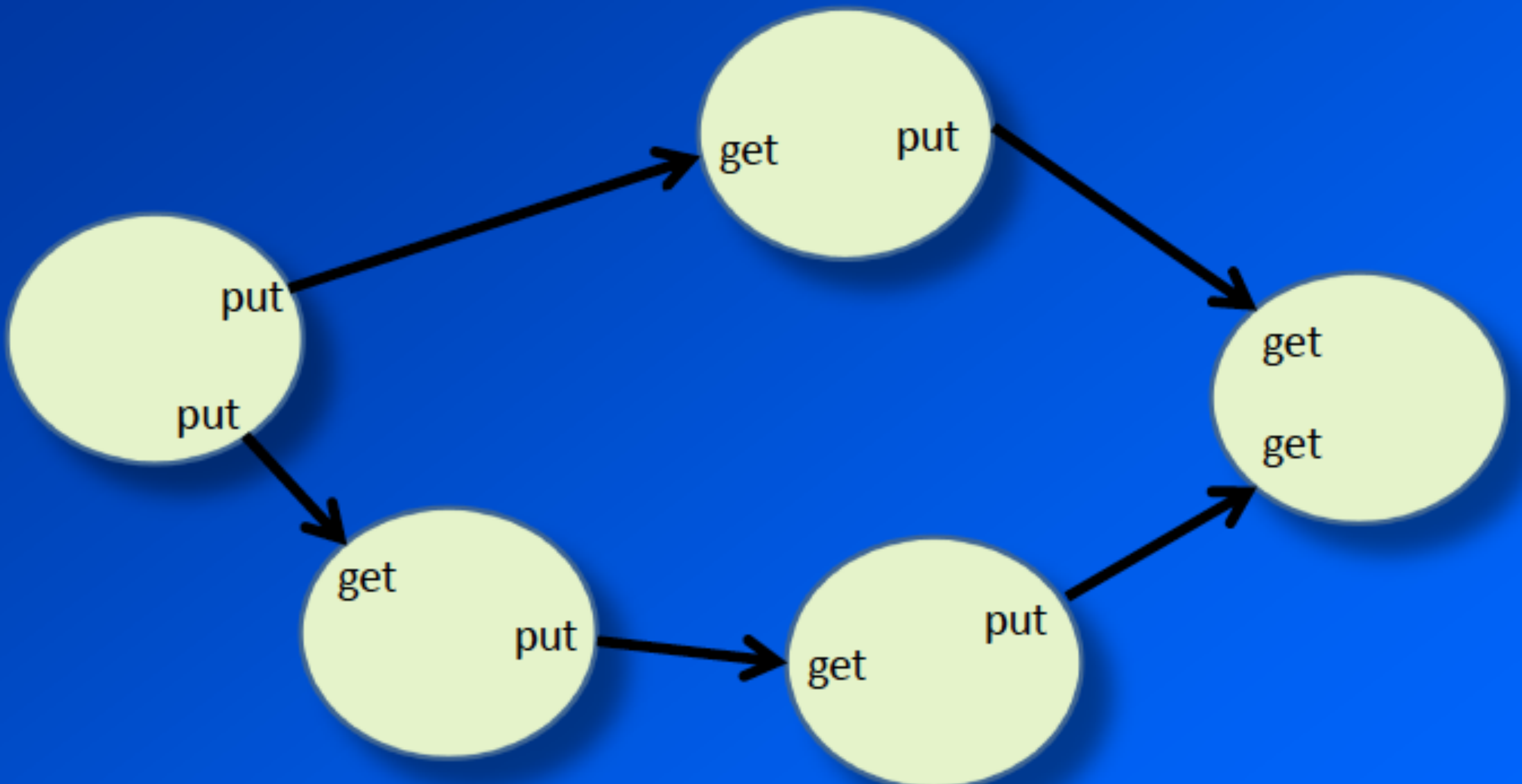
Programmer has more control than with Strategies

- => less error prone?

Good performance (comparable to Strategies)

- particularly if granularity is not too small

Par expresses dynamic dataflow

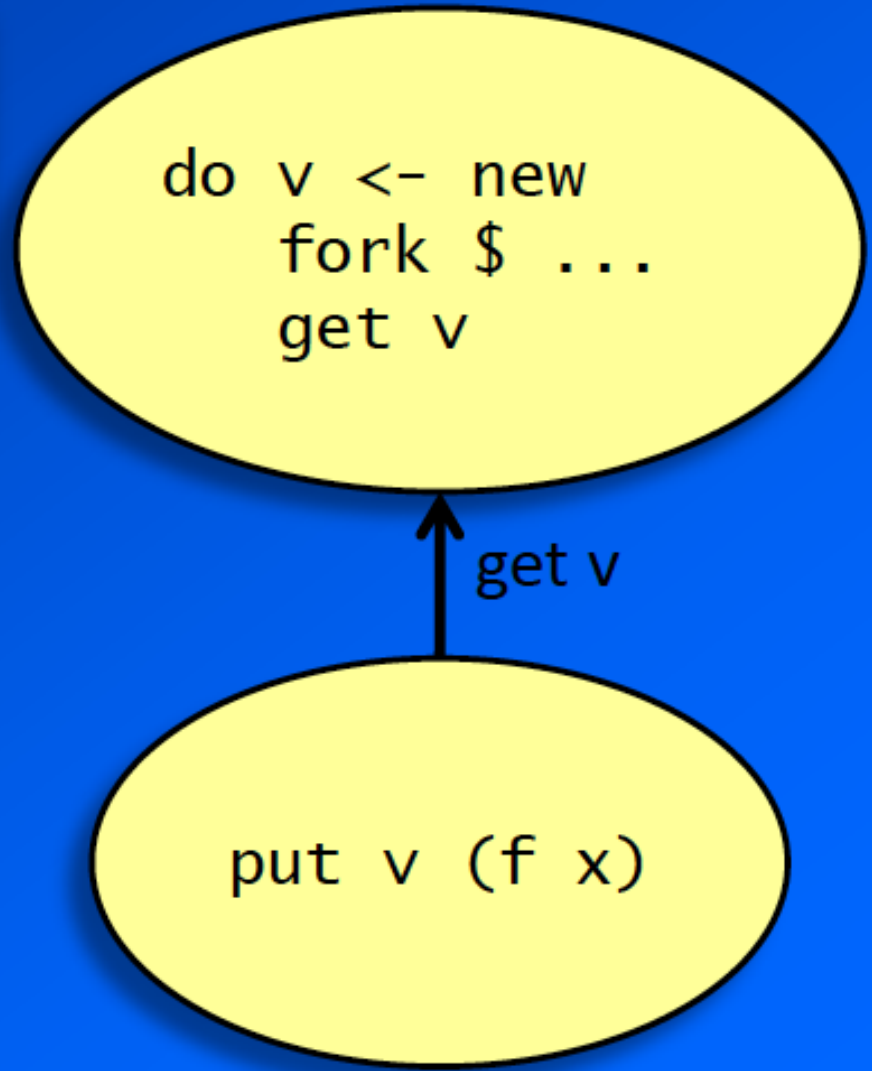


Slide by Simon Marlow

```haskell
spawn :: NFData a => Par a -> Par (IVar a)
spawn p = do
     i <- new
     fork (do x <- p; put i x)
     return i
```

How does this make a dataflow graph?

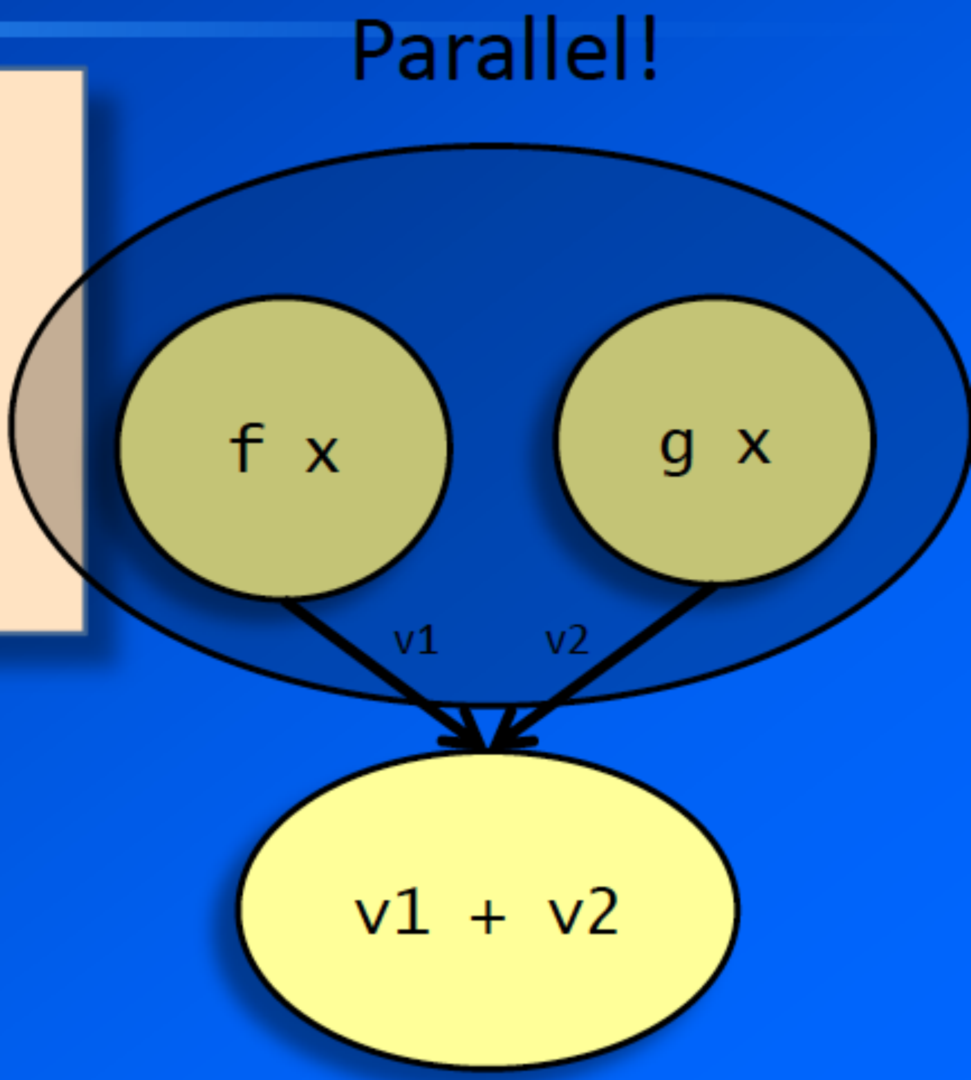
```
do v <- new
   fork $ put v (f x)
   get v
```



Slide by Simon Marlow

A bit more complex...

```
do v1 <- new
   v2 <- new
   fork $ put v1 (f x)
   fork $ put v2 (g x)
   get v1
   get v2
   return (v1 + v2)
```



Slide by Simon Marlow

```
parMapM :: NFData b => (a -> Par b) -> [a] -> Par [b]
parMapM f as = do
   ibs <- mapM (spawn . f) as
   mapM get ibs
```

Dataflow problems

- Par really shines when the problem is easily expressed as a dataflow graph, particularly an irregular or dynamic graph (e.g. shape depends on the program input)
- Identify the nodes and edges of the graph
  - each node is created by fork
  - each edge is an IVar

Slide by Simon Marlow

Example

- Consider typechecking (or inferring types for) a set of non-recursive bindings.
- Each binding is of the form `x = e` for variable x, expression e
- To typecheck a binding:
  - input: the types of the identifiers mentioned in e
  - output: the type of x
- So this is a dataflow graph
  - a node represents the typechecking of a binding
  - the types of identifiers flow down the edges

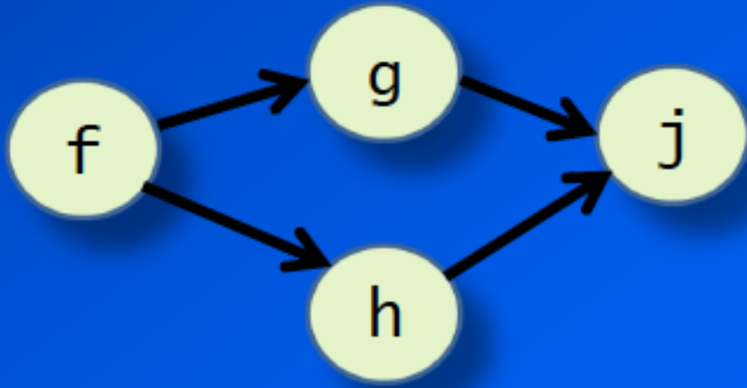
Slide by Simon Marlow

Dataflow

- Consider typechecking a set of (non-recursive) bindings:

```
f = ...
g = ... f ...
h = ... f ...
j = ... g ... h ...
```

- treat this as a dataflow graph:



Slide by Simon Marlow

Implementation

- We parallelised an existing type checker (nofib/infer).
- Algorithm works on a single term:

```
data Term = Let VarId Term Term | ...
```

- So we parallelise checking of the top-level Let bindings.

```
let x1 = e1 in
let x2 = e2 in
let x3 = e3 in
...
```

The parallel type inferencer

- Given:

```
inferTopRhs :: Env -> Term -> PolyType
makeEnv :: [(VarId,Type)] -> Env
```

- We need a type environment:

```
type TopEnv = Map VarId (IVar PolyType)
```

- The top-level inferencer has the following type:

```
inferTop :: TopEnv -> Term -> Par MonoType
```

Parallel type inference

```
inferTop :: TopEnv -> Term -> Par MonoType
inferTop topenv (Let x u v) = do
    vu <- new

    fork $ do
      let fu = Set.toList (freeVars u)
      tfu <- mapM (get . fromJust . flip Map.lookup topenv) fu
      let aa = makeEnv (zip fu tfu)
      put vu (inferTopRhs aa u)

    inferTop (Map.insert x vu topenv) v


inferTop topenv t = do
  -- the boring case: invoke the normal sequential
  -- type inference engine
```

Create nodes and edges and let the scheduler do the work

No dependency analysis required!

Maximum parallelism for little programmer effort

Dynamic parallelism

Very nice ☺

Implementation

- Starting point: A Poor Man's Concurrency Monad (Claessen JFP'99)
- PMC was used to *simulate* concurrency in a sequential Haskell implementation. We are using it as a way to implement very lightweight non-preemptive threads, with a parallel scheduler.
- Following PMC, the implementation is divided into two:
  - Par computations produce a lazy Trace
  - A scheduler consumes the Traces, and switches between multiple threads

Slide by Simon Marlow

Traces

- A “thread” produces a lazy stream of operations:

```
data Trace
  = Fork Trace Trace
  | Done
  | forall a . Get (IVar a) (a -> Trace)
  | forall a . Put (IVar a) a Trace
  | forall a . New (IVar a -> Trace)
```

Slide by Simon Marlow

The Par monad

- Par is a CPS monad:

```
newtype Par a = Par {
    runCont :: (a -> Trace) -> Trace
 }

instance Monad Par where
    return a = Par $ \c -> c a
    m >>= k  = Par $ \c -> runCont m $
                          \a -> runCont (k a) c
```

Slide by Simon Marlow

Operations

```
fork :: Par () -> Par ()
fork p = Par $ \c ->
          Fork (runCont p (\_ -> Done)) (c ())

new :: Par (IVar a)
new  = Par $ \c -> New c

get :: IVar a -> Par a
get v = Par $ \c -> Get v c

put :: NFData a => IVar a -> a -> Par ()
put v a = deepseq a (Par $ \c -> Put v a (c ()))
```

Slide by Simon Marlow

e.g.

- This code:

```
do
   x <- new
   fork (put x 3)
   r <- get x
   return (r+1)
```

- will produce a trace like this:

```
New (\x ->
  Fork (Put x 3 $ Done)
       (Get x (\r ->
          c (r + 1))))
```

Slide by Simon Marlow

The scheduler

- First, a sequential scheduler.

The currently running thread

```
sched :: SchedState -> Trace -> IO ()

type SchedState = [Trace]
```

The work pool, "runnable threads"

Why IO?
Because we're going to extend it to be a parallel scheduler in a moment.

Slide by Simon Marlow

Representation of IVar

```
newtype IVar a = IVar (IORef (IVarContents a))

data IVarContents a = Full a | Blocked [a -> Trace]
```

set of threads
blocked in **get**

Slide by Simon Marlow

Fork and Done

```
sched state Done = reschedule state
```

```
reschedule :: SchedState -> IO ()
reschedule []       = return ()
reschedule (t:ts) = sched ts t
```

```
sched state (Fork child parent) =
  sched (child:state) parent
```

Slide by Simon Marlow

New and Get

```
sched state (New f) = do
  r <- newIORef (Blocked [])
  sched state (f (IVar r))
```

```
sched state (Get (IVar v) c) = do
  e <- readIORef v
  case e of
    Full a -> sched state (c a)
    Blocked cs -> do
      writeIORef v (Blocked (c:cs))
      reschedule state
```

Slide by Simon Marlow

Put

```
sched state (Put (IVar v) a t) = do
  cs <- modifyIORef v $ \e -> case e of
           case e of
             Full _      -> error "multiple put"
             Blocked cs -> (Full a, cs)
  let state' = map ($ a) cs ++ state
  sched state' t
```

Wake up all the blocked threads, add them to the work pool

```
modifyIORef :: IORef a -> (a -> (a,b)) -> IO b
```

Slide by Simon Marlow

Finally... runPar

```
runPar :: Par a -> a
runPar x = unsafePerformIO $ do
   rref <- newIORef (Blocked [])
   sched [] $
      runCont (x >>= put_ (IVar rref))
              (const Done)
   r <- readIORef rref
   case r of
     Full a -> return a
     _      -> error "no result"
```

rref is an IVar to hold the return value

the "main thread" stores the result in rref

if the result is empty, the main thread must have deadlocked

- that's the complete sequential scheduler

Slide by Simon Marlow

A real parallel scheduler

- We will create one scheduler thread per core
- Each scheduler has a local work pool
  - when a scheduler runs out of work, it tries to steal from the other work pools
- The new state:

```
data SchedState = SchedState
    { no       :: Int,
      workpool :: IORef [Trace],
      idle     :: IORef [MVar Bool],
      scheds   :: [SchedState]
    }
```

CPU number

Local work pool

Idle schedulers (shared)

Other schedulers (for stealing)

Slide by Simon Marlow

New/Get/Put

- New is the same
- Mechanical changes to Get/Put:
  - use atomicModifyIORef to operate on IVars
  - use atomicModifyIORef to modify the work pool (now an IORef [Trace], was previously [Trace]).

Slide by Simon Marlow

reschedule

```
reschedule :: SchedState -> IO ()
reschedule state@SchedState{ workpool } = do
  e <- atomicModifyIORef workpool $ \ts ->
         case ts of
           []      -> ([], Nothing)
           (t:ts') -> (ts', Just t)
  case e of
    Just t  -> sched state t
    Nothing -> steal state
```

Here's where we go stealing

Slide by Simon Marlow

stealing

```
steal :: SchedState -> IO ()
steal state@SchedState{ scheds, no=me } = go scheds
  where
    go (x:xs)
      | no x == me    = go xs
      | otherwise     = do
         r <- atomicModifyIORef (workpool x) $ \ ts ->
                 case ts of
                    []     -> ([], Nothing)
                    (x:xs) -> (xs, Just x)
         case r of
           Just t  -> sched state t
           Nothing -> go xs
    go [] = do
      -- failed to steal anything; add ourself to the
      -- idle queue and wait to be woken up
```

Slide by Simon Marlow

runPar

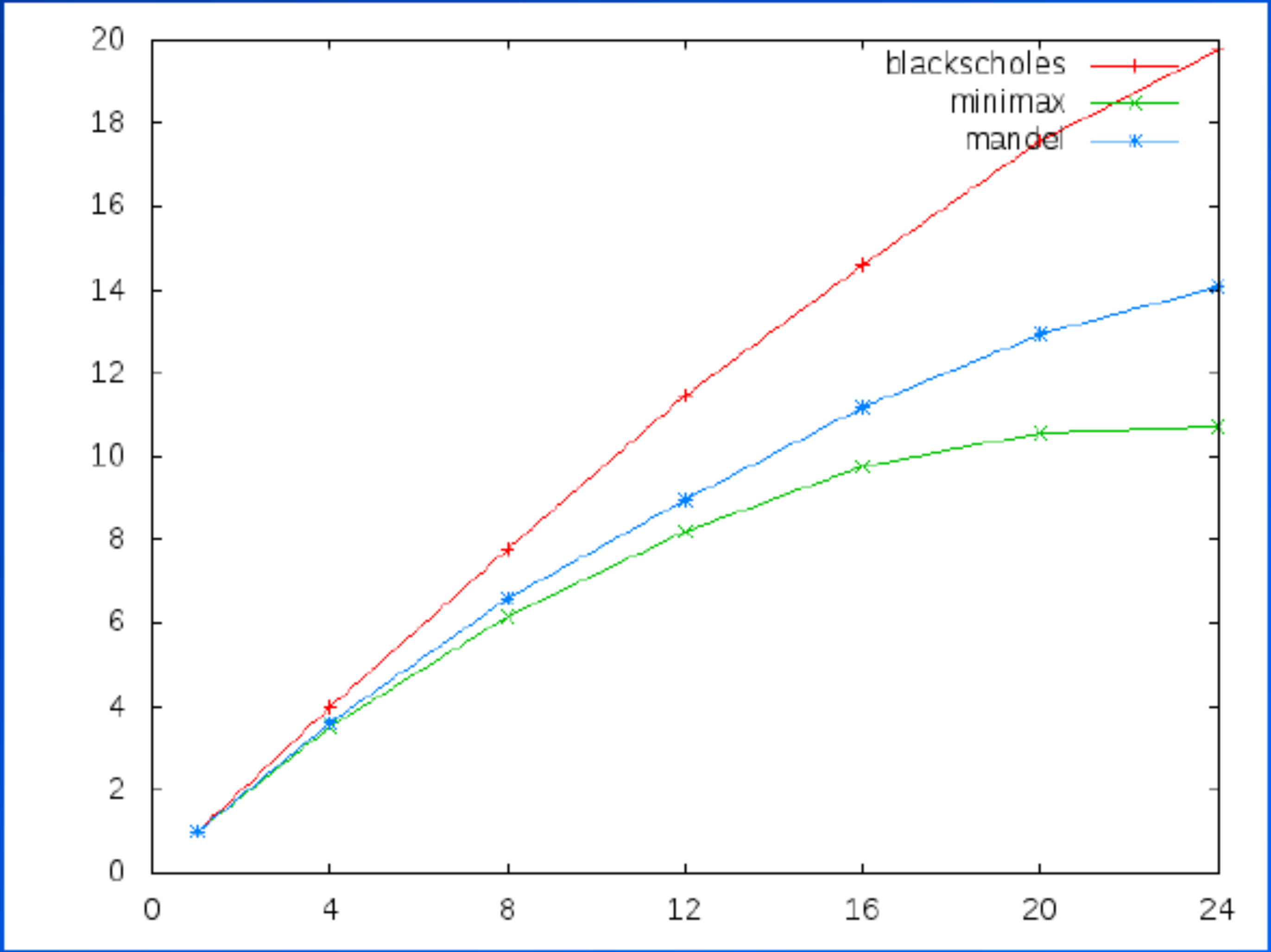
```
runPar :: Par a -> a
runPar x = unsafePerformIO $ do
   let states = ...
   main_cpu <- getCurrentCPU
   m <- newEmptyMVar
   forM_ (zip [0..] states) $ \(cpu,state) ->
     forkOnIO cpu $
       if (cpu /= main_cpu)
          then reschedule state
          else do
               rref <- newIORef Empty
               sched state $
                  runCont (x >>= put_ (IVar rref))
                          (const Done)
               readIORef rref >>= putMVar m

   r <- takeMVar m
   case r of Full a -> return a
             _ -> error "no result"
```

The "main thread" runs on the current CPU, all other CPUs run workers

An MVar communicates the result back to the caller of runPar

Slide by Simon Marlow

Results



speedup

99%

95%

50%

cores

Slide by Simon Marlow

Modularity

- Key property of Strategies is modularity

```
parMap f xs = map f xs `using` parList rwhnf
```

- Relies on lazy evaluation
  - fragile
  - not always convenient to build a lazy data structure
- Par takes a different approach to modularity:
  - the Par monad is for *coordination* only
  - the application code is written separately as pure Haskell functions
  - The “parallelism guru” writes the coordination code
  - Par performance is not critical, as long as the grain size is not too small

Slide by Simon Marlow

Par monad compared to Strategies

Separation of function and parallelisation done differently

Eval monad and Strategies are advisory

Par monad does not support speculative parallelism as Stategies do

Par monad supports stream processing pipelines well

Note: Par monad and Strategies can be combined…

Par Monad easier to use than par?

fork creates one parallel task

Dependencies between tasks represented by Ivars

No need to reason about laziness

put is hyperstrict by default

Final suggestion in Par Monad paper is that maybe par is suitable for automatic parallelisation

Next

Continue working on Lab A (due 11.59 April 18)

Friday 15.15 EC   Nikita on GHC Heap Internals, garbage collection etc.

Erlang starts next week (mon and fri)

Don't miss David Duke next Thursday on
Skeletons for Parallel Scientific Computing (very cool)