Fortgeschrittene Funktionale Programmierung in Haskell

Universität Bielefeld, Sommersemester 2015

Jonas Betzendahl & Stefan Dresselhaus

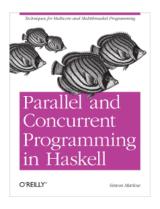
Outline I

Übersicht für Heute:

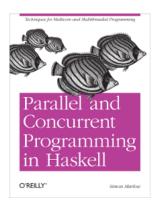
- Wiederholung
- 2 Threads, MVars, etc.
 - forkIO und MVars
 - Deadlock Detection
- Software Transactional Memory
 - Motivation
 - Beispiel: Banksoftware
- 4 Parallelism through concurrency

Wiederholung

Leseempfehlung:



Leseempfehlung:



...srsly!

Techniques for Multicore and Multibroaded Programming Parallel and Concurrent Programming in Haskell

Fortgeschrittene Funktionale Programmierung in Haskell

Überblick:

Überblick:

Was war nochmal der Unterschied zwischen Paralllelism und Nebenläufigkeit?

Parallelism:

Mehrere Hardwareelemente

Überblick:

Was war nochmal der Unterschied zwischen Paralllelism und Nebenläufigkeit?

Parallelism:

- Mehrere Hardwareelemente
- Antwort schneller kriegen

Überblick:

Was war nochmal der Unterschied zwischen Paralllelism und Nebenläufigkeit?

Parallelism:

- Mehrere Hardwareelemente
- Antwort schneller kriegen
- deterministisch (i.d.R.)

Überblick:

Was war nochmal der Unterschied zwischen Paralllelism und Nebenläufigkeit?

Parallelism:

- Mehrere Hardwareelemente
- Antwort schneller kriegen
- deterministisch (i.d.R.)
- oft deklarativ

Überblick:

Was war nochmal der Unterschied zwischen Paralllelism und Nebenläufigkeit?

Parallelism:

- Mehrere Hardwareelemente
- Antwort schneller kriegen
- deterministisch (i.d.R.)
- oft deklarativ

Concurrency:

Mehrere Threads

Überblick:

Was war nochmal der Unterschied zwischen Paralllelism und Nebenläufigkeit?

Parallelism:

- Mehrere Hardwareelemente
- Antwort schneller kriegen
- deterministisch (i.d.R.)
- oft deklarativ

Concurrency:

- Mehrere Threads
- Dinge gleichzeitig tun

Überblick:

Was war nochmal der Unterschied zwischen Paralllelism und Nebenläufigkeit?

Parallelism:

- Mehrere Hardwareelemente
- Antwort schneller kriegen
- deterministisch (i.d.R.)
- oft deklarativ

Concurrency:

- Mehrere Threads
- Dinge gleichzeitig tun
- nichtdeterministisch

Überblick:

Was war nochmal der Unterschied zwischen Paralllelism und Nebenläufigkeit?

Parallelism:

- Mehrere Hardwareelemente
- Antwort schneller kriegen
- deterministisch (i.d.R.)
- oft deklarativ

Concurrency:

- Mehrere Threads
- Dinge gleichzeitig tun
- nichtdeterministisch
- oft impertativ

Die Basics: Threads, MVars, etc.

Wir beginnen mit der Funktion, die einen neuen Thread erstellt:

```
forkIO :: IO () -> IO ThreadId
```

Wir beginnen mit der Funktion, die einen neuen Thread erstellt:

```
forkIO :: IO () -> IO ThreadId
```

Threads interagieren notwendigerweise mit der Welt, ergo ist die Berechnung, die wir übergeben vom Typ IO ().

Wir beginnen mit der Funktion, die einen neuen Thread erstellt:

```
forkIO :: IO () -> IO ThreadId
```

Threads interagieren notwendigerweise mit der Welt, ergo ist die Berechnung, die wir übergeben vom Typ IO ().

Die ThreadId kann später benutzt werden um z.B. den Thread vorzeitig zu töten oder ihm eine Exception zuzuschmeißen.

Ein kleines Beispiel:

```
import Control.Concurrent
import Control.Monad
import System.IO

main :: IO ()
main = do
   hSetBuffering stdout NoBuffering
   forkIO (replicateM_ 100000 (putChar 'A'))
   replicateM_ 100000 (putChar 'B')
```

Ein kleines Beispiel:

```
import Control.Concurrent
import Control.Monad
import System.IO

main :: IO ()
main = do
   hSetBuffering stdout NoBuffering
   forkIO (replicateM_ 100000 (putChar 'A'))
   replicateM_ 100000 (putChar 'B')

...Output?
```

Ein kleines Beispiel:

```
import Control.Concurrent
import Control.Monad
import System.IO

main :: IO ()
main = do
   hSetBuffering stdout NoBuffering
   forkIO (replicateM_ 100000 (putChar 'A'))
   replicateM_ 100000 (putChar 'B')
...Output?
```

 Aber...

Aber...wie kriegen wir jetzt Ergebnisse aus der Berechnung raus? Der Typ ist nur IO (), das liefert nichts (interessantes) zurück!

Aber...wie kriegen wir jetzt Ergebnisse aus der Berechnung raus? Der Typ ist nur IO (), das liefert nichts (interessantes) zurück!

Das gleiche Problem hatten wir schon in der Par-Monade. Lösung damals waren IVars:

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

Introducing: ...

```
Introducing: ... MVars!
```

```
data MVar a -- abstract

newEmptyMVar :: IO (MVar a)
newMVar :: a -> IO (MVar a)
takeMVar :: MVar a -> IO a
putMVar :: MVar a -> a -> IO ()
readMVar :: MVar a -> IO a
```

```
Introducing: ... MVars!
```

```
data MVar a -- abstract

newEmptyMVar :: IO (MVar a)
newMVar :: a -> IO (MVar a)
takeMVar :: MVar a -> IO a
putMVar :: MVar a -> a -> IO ()
readMVar :: MVar a -> IO a
```

Wir brauchen hier keine eigene Monade wie Par. Da Concurrency so oder so effektvoll ist, reicht IO vollkommen aus.

Unterschied zwischen IVars und MVars: erstere sind *i*mmutable, letztere sind *m*utable.

Ein Beispiel zu MVars:

```
main :: IO ()
main = do
    m <- newEmptyMVar
    forkIO $ do putMVar m 'x'; putMVar m 'y'
    r <- takeMVar m
    print r
    r <- takeMVar m
    print r</pre>
```

Ein Beispiel zu MVars:

```
main :: IO ()
main = do
    m <- newEmptyMVar
    forkIO $ do putMVar m 'x'; putMVar m 'y'
    r <- takeMVar m
    print r
    r <- takeMVar m
    print r</pre>
```

Wie wir sehen kann die gleiche MVar über Zeit mehrere Zustände annehmen und erfolgreich zur Kommunikation zwischen Threads benutzt werden.

Channel mit nur einem Slot
 Eine MVar kann als Nachrichtenkanal zwischen Threads
 benutzt werden, allerdings maximal eine Nachricht auf einmal halten.

- Channel mit nur einem Slot
 Eine MVar kann als Nachrichtenkanal zwischen Threads
 benutzt werden, allerdings maximal eine Nachricht auf einmal halten.
- Behältnis für shared mutable state
 In Concurrent Haskell brauchen oft mehrere Threads Zugriff auf einen shared state. Ein beliebtes Designpattern ist, das dieser State als normaler (immutable) Haskell-Datentyp repräsentiert und in einer MVax verpackt wird.

- Channel mit nur einem Slot
 Eine MVar kann als Nachrichtenkanal zwischen Threads
 benutzt werden, allerdings maximal eine Nachricht auf einmal halten.
- Behältnis für shared mutable state
 In Concurrent Haskell brauchen oft mehrere Threads Zugriff auf einen shared state. Ein beliebtes Designpattern ist, das dieser State als normaler (immutable) Haskell-Datentyp repräsentiert und in einer MVax verpackt wird.
- Baustein für kompliziertere Strukturen

Mehr Leckerlis:

Was passiert, wenn wir folgenden Code ausführen?

Mehr Leckerlis:

Was passiert, wenn wir folgenden Code ausführen?

Wir bekommen eine Fehlermeldung, dass das Programm hängt, statt einfach nur ein hängendes Programm.

```
$ ./mvar3
mvar3: thread blocked indefinitely in an MVar operation
```

Deadlock detection:

Threads und MVars sind Objekte auf dem Heap. Das RTS (i.e. der Garbage collector) durchläuft den Heap um alle lebendigen Objekte zu finden, angefangen bei den laufenden Threads und ihren Stacks.

Alles was so nicht erreichbar sind (z.B. ein Thread der auf eine MVar wartet, die nirgendwo sonst referenziert wird), blockiert und bekommt eine Exception geschmissen.

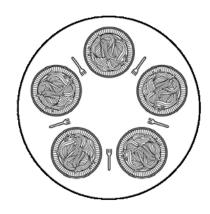


Abbildung: dining philosophers

Deadlock detection:

Dieses Vorgang funktioniert allerdings nicht immer wie man zunächst denkt. Beispiel: Was passiert mit diesem Code?

```
main :: IO ()
main = do
  lock <- newEmptyMVar
  complete <- newEmptyMVar
  forkIO $ takeMVar lock 'finally' putMVar complete ()
  takeMVar complete</pre>
```

Deadlock detection:

Dieses Vorgang funktioniert allerdings nicht immer wie man zunächst denkt. Beispiel: Was passiert mit diesem Code?

```
main :: IO ()
main = do
  lock <- newEmptyMVar
  complete <- newEmptyMVar
  forkIO $ takeMVar lock 'finally' putMVar complete ()
  takeMVar complete</pre>
```

Da nicht nur der geforkte Thread sondern auch der ursprüngliche gedeadlocked sind, wird hier die Fehlermeldung geprintet, statt die rettende Exception an das Kind zu senden.

Software Transactional Memory (STM)

Trotz aller Unterstützung durch das RTS:

Trotz aller Unterstützung durch das RTS:

Locks are absurdly hard to get right! (SPJ)

"The Future is Parallel, and the Future of Parallel is Declarative " https://www.youtube.com/watch?v=hlyQjK1qjw8

Trotz aller Unterstützung durch das RTS:

Locks are absurdly hard to get right! (SPJ)

"The Future is Parallel, and the Future of Parallel is Declarative " https://www.youtube.com/watch?v=hlyQjK1qjw8

Beliebte Fehler:

Races (vergessene Locks)

Trotz aller Unterstützung durch das RTS:

Locks are absurdly hard to get right! (SPJ)

"The Future is Parallel, and the Future of Parallel is Declarative " https://www.youtube.com/watch?v=hlyQjK1qjw8

Beliebte Fehler:

- Races (vergessene Locks)
- Deadlocks (Locks in falscher Reihenfolge genommen)

Trotz aller Unterstützung durch das RTS:

Locks are absurdly hard to get right! (SPJ)

"The Future is Parallel, and the Future of Parallel is Declarative " https://www.youtube.com/watch?v=hlyQjK1qjw8

Beliebte Fehler:

- Races (vergessene Locks)
- Deadlocks (Locks in falscher Reihenfolge genommen)
- Lost wakeups (Conditional-Variable nicht bescheid gesagt)

Trotz aller Unterstützung durch das RTS:

Locks are absurdly hard to get right! (SPJ)

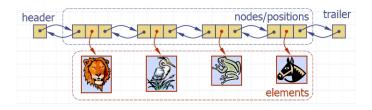
"The Future is Parallel, and the Future of Parallel is Declarative " https://www.youtube.com/watch?v=hlyQjK1qjw8

Beliebte Fehler:

- Races (vergessene Locks)
- Deadlocks (Locks in falscher Reihenfolge genommen)
- Lost wakeups (Conditional-Variable nicht bescheid gesagt)
- Error Recovery (Exceptionhandler müssen Locks freigeben und teilweise Ursprungszustand restaurieren)

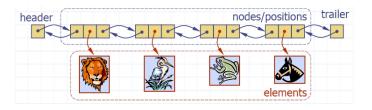
Angenommen wir möchten gerne eine Queue parallel bearbeiten:

Bildquelle: http://www.math.bas.bg/~nkirov/2015/NETB201/slides/ch04/ch04.html



Angenommen wir möchten gerne eine Queue parallel bearbeiten:

Bildquelle: http://www.math.bas.bg/~nkirov/2015/NETB201/slides/ch04/ch04.html



Problem: offensichtlich, race conditions etc.

Angenommen wir möchten gerne eine Queue parallel bearbeiten:

Bildquelle: http://www.math.bas.bg/~nkirov/2015/NETB201/slides/ch04/ch04.html



Angenommen wir möchten gerne eine Queue parallel bearbeiten:

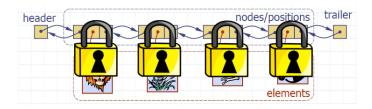
Bildquelle: http://www.math.bas.bg/~nkirov/2015/NETB201/slides/ch04/ch04.html



Problem: Nicht gerade nebenläufig...

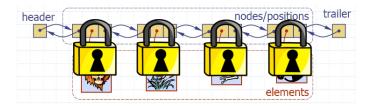
Angenommen wir möchten gerne eine Queue parallel bearbeiten:

Bildquelle: http://www.math.bas.bg/~nkirov/2015/NETB201/slides/ch04/ch04.html



Angenommen wir möchten gerne eine Queue parallel bearbeiten:

Bildquelle: http://www.math.bas.bg/~nkirov/2015/NETB201/slides/ch04/ch04.html



Problem: Fehleranfälligkeit bei kleinen Listen

Aufgabe vs. Schwierigkeit: Zweiendige Liste

Problemstellung	Schwierigkeit
sequentiell	Gymnasium oder Bachelor

Aufgabe vs. Schwierigkeit: Zweiendige Liste

Problemstellung	Schwierigkeit
sequentiell	Gymnasium oder Bachelor
Locks et al.	(gerade nicht mehr) publizierbar auf internationalen Konferenzen

Aufgabe vs. Schwierigkeit: Zweiendige Liste

Problemstellung	Schwierigkeit
sequentiell	Gymnasium oder Bachelor
Locks et al.	(gerade nicht mehr) publizierbar auf internationalen Konferenzen
atomic blocks (STM)	Bachelor

 Das bedeutet, dass Deadlocks unmöglich werden weil es keine Locks mehr gibt!

- Das bedeutet, dass Deadlocks unmöglich werden weil es keine Locks mehr gibt!
- Automatisierte error recovery. STM stellt den Ausgangszustand von selbst wieder her.

- Das bedeutet, dass Deadlocks unmöglich werden weil es keine Locks mehr gibt!
- Automatisierte error recovery. STM stellt den Ausgangszustand von selbst wieder her.
- TVars (Transaction Variables). Wie IVars und MVars, nur in der STM-Monade.

STM auf einen Blick:

```
data STM a
                                 -- abstract
instance Monad STM
                                 -- among other things
atomically :: STM a -> IO a
data TVar a
                                 -- abstract
newTVar :: a -> STM (TVar a)
readTVar :: TVar a -> STM a
writeTVar :: TVar a -> a -> STM ()
retry :: STM a
orElse :: STM a -> STM a -> STM a
throwSTM :: Exception e => e -> STM a
catchSTM :: Exception e => STM a -> (e -> STM a) -> STM a
```

atomically :: STM a -> IO a

```
atomically :: STM a -> IO a
```

• Kein Sprachkonstrukt!

```
atomically :: STM a -> IO a
```

- Kein Sprachkonstrukt!
- Führt Berechnungen in STM in der echten Welt aus

```
atomically :: STM a -> IO a
```

- Kein Sprachkonstrukt!
- Führt Berechnungen in STM in der echten Welt aus
- ... und zwar in einem Rutsch, ohne Unterbrechung!

```
atomically :: STM a -> IO a
```

- Kein Sprachkonstrukt!
- Führt Berechnungen in STM in der echten Welt aus
- ... und zwar in einem Rutsch, ohne Unterbrechung!
- Stellt bei Fehlschlag Ursprungszustand wieder her.

```
atomically :: STM a -> IO a
```

- Kein Sprachkonstrukt!
- Führt Berechnungen in STM in der echten Welt aus
- ... und zwar in einem Rutsch, ohne Unterbrechung!
- Stellt bei Fehlschlag Ursprungszustand wieder her.
- Deshalb: Kein IO in Transaktionen!

```
atomically :: STM a -> IO a
```

- Kein Sprachkonstrukt!
- Führt Berechnungen in STM in der echten Welt aus
- ... und zwar in einem Rutsch, ohne Unterbrechung!
- Stellt bei Fehlschlag Ursprungszustand wieder her.
- Deshalb: Kein IO in Transaktionen!
- Ebenfalls: Keine genesteten atomicallys (Typen!)

retry :: STM a

retry :: STM a

Kein Sprachkonstrukt!

```
retry :: STM a
```

- Kein Sprachkonstrukt!
- Rollt zurück und versucht die gleiche Transaktion erneut durchzuführen

```
retry :: STM a
```

- Kein Sprachkonstrukt!
- Rollt zurück und versucht die gleiche Transaktion erneut durchzuführen
- ...allerdings erst zu einem angebrachten Zeitpunkt.
 CPU wird nicht unnützerweise zur Heizung.

```
retry :: STM a
```

- Kein Sprachkonstrukt!
- Rollt zurück und versucht die gleiche Transaktion erneut durchzuführen
- ...allerdings erst zu einem angebrachten Zeitpunkt.
 CPU wird nicht unnützerweise zur Heizung.

orElse :: STM a -> STM a -> STM a

```
orElse :: STM a -> STM a -> STM a
```

Kein Sprachkonstrukt!

```
orElse :: STM a -> STM a -> STM a
```

- Kein Sprachkonstrukt!
- orElse a b führt b aus, wenn a retry aufruft.

```
orElse :: STM a -> STM a -> STM a
```

- Kein Sprachkonstrukt!
- orElse a b führt b aus, wenn a retry aufruft.
- Komposition von STM-Berechnungen:
 - »= ist AND; orElse ist OR

```
orElse :: STM a -> STM a -> STM a
```

- Kein Sprachkonstrukt!
- orElse a b führt b aus, wenn a retry aufruft.
- Komposition von STM-Berechnungen:

```
»= ist AND; orElse ist OR
```

Beispiel:

```
takeEitherTMVar :: TMVar a -> TMVar b -> STM (Either a b)
takeEitherTMVar ma mb =
  fmap Left (takeTMVar ma)
    'orElse'
fmap Right (takeTMVar mb)
```

Stellen wir uns vor, wir wollen eine simplifizierte Bankensoftware schreiben, die in der Lage sein soll, Konten und Überweisungen zu repräsentieren.

Stellen wir uns vor, wir wollen eine simplifizierte Bankensoftware schreiben, die in der Lage sein soll, Konten und Überweisungen zu repräsentieren.

Eine naive Implementation wäre die folgende:

```
type Account = IORef Integer

transfer :: Integer -> Account -> Account -> IO ()
transfer amount from to = do
    fromVal <- readIORef from
    toVal <- readIORef to
    writeIORef from (fromVal - amount)
    writeIORef to (toVal + amount)</pre>
```

Diese Implementation hätte jedoch in einem Nebenläufigen Setting einige Probleme. Man beachte folgende Zeile:

fromVal <- readIORef from</pre>

Diese Implementation hätte jedoch in einem Nebenläufigen Setting einige Probleme. Man beachte folgende Zeile:

fromVal <- readIORef from</pre>

Finden nun mehrere Aktionen gleichzeitig statt, so könnte es sein, dass mehrere Threads denselben Wert als fromVal lesen, bevor die jeweils andere Transaktion durchgeführt wurde.

Dies hätte zur Folge, dass später ein inkorrekter Kontostand berechnet und gesetzt würde.

Führen wir also Locks ein (durch Benutzung von MVars), um diese race condition zu verhindern.

Führen wir also Locks ein (durch Benutzung von MVars), um diese race condition zu verhindern.

```
type Account = MVar Integer
credit :: Integer -> Account -> IO ()
credit amount account = do
    current <- takeMVar account
    putMVar account (current + amount)

debit :: Integer -> Account -> IO ()
debit amount account = do
    current <- takeMVar account
    putMVar account (current - amount)</pre>
```

In dieser Implementation sähe eine Funktion für Überweisungen etwa so aus:

```
transfer :: Integer -> Account -> Account -> IO ()
transfer amount from to = do
   debit amount from
   credit amount to
```

In dieser Implementation sähe eine Funktion für Überweisungen etwa so aus:

```
transfer :: Integer -> Account -> Account -> IO ()
transfer amount from to = do
    debit amount from
    credit amount to
```

Diese verhindert, dass durch fehlerhaftes Zusammenspiel von Threads Geld geschaffen oder vernichtet wird.

In dieser Implementation sähe eine Funktion für Überweisungen etwa so aus:

```
transfer :: Integer -> Account -> Account -> IO ()
transfer amount from to = do
   debit amount from
   credit amount to
```

Diese verhindert, dass durch fehlerhaftes Zusammenspiel von Threads Geld geschaffen oder vernichtet wird.

Es existiert allerdings immer noch eine race condition: Der Thread, der die Überweisung ausführt, könnte direkt nach dem debit-Schritt von der CPU verdrängt werden und die Bankensoftware dadurch in einem inkonsistenten Zustand zurück lassen.

Wie sähe diese Software mit STM aus?

```
type Account = TVar Integer
credit :: Integer -> Account -> STM ()
credit amount account = do
    current <- readTVar account
    writeTVar account (current + amount)
debit :: Integer -> Account -> STM ()
debit amount account = do
    current <- readTVar account
    writeTVar account (current - amount)
transfer :: Integer -> Account -> Account -> STM ()
transfer amount from to = do
    debit amount from
    credit amount to
```

Vergleich zur Variante mit MVars:

```
type Account = MVar Integer
credit :: Integer -> Account -> IO ()
credit amount account = do
    current <- takeMVar account
    putMVar account (current + amount)
debit :: Integer -> Account -> IO ()
debit amount account = do
    current <- takeMVar account
    putMVar account (current - amount)
transfer :: Integer -> Account -> Account -> IO ()
transfer amount from to = do
    debit amount from
    credit amount to
```

```
transfer :: Integer -> Account -> Account -> IO ()
```

Diese Funktion führt die Überweisung vollkommen in der echten Welt durch, mit allen möglichen Fehlern, die dabei auftreten können.

```
transfer :: Integer -> Account -> Account -> IO ()
```

Diese Funktion führt die Überweisung vollkommen in der echten Welt durch, mit allen möglichen Fehlern, die dabei auftreten können.

```
transfer :: Integer -> Account -> Account -> STM ()
```

Diese Funktion stellt uns nur eine Berechnung in der STM-Monade bereit, die wir später entweder direkt oder auch als Baustein einer größeren Transaktion ausführen können.

```
transfer :: Integer -> Account -> Account -> IO ()
```

Diese Funktion führt die Überweisung vollkommen in der echten Welt durch, mit allen möglichen Fehlern, die dabei auftreten können.

```
transfer :: Integer -> Account -> Account -> STM ()
```

Diese Funktion stellt uns nur eine Berechnung in der STM-Monade bereit, die wir später entweder direkt oder auch als Baustein einer größeren Transaktion ausführen können.

Der Vorteil ist, dass wir nur einen Weg haben, diese Berechnung haben, das zu tun:

```
atomically :: STM a -> IO a
```

Ja! Hier sind einige Gründe:

Ja! Hier sind einige Gründe:

• Performance: STM ist eine Abstraktion, und wie (fast) alle Abstraktionen hat es Laufzeitkosten.

Ja! Hier sind einige Gründe:

- Performance: STM ist eine Abstraktion, und wie (fast) alle Abstraktionen hat es Laufzeitkosten.
- Fairness: Blockieren mehrere Threads auf einer MVar, werden sie garantiert FIFO wieder aufgeweckt. STM hingegen hat keine Garantie für Fairness.

Wiederholung Threads, MVars, etc. Software Transactional Memory Parallelism through concurrency

Parallelism through Concurrency