Under consideration for publication in J. Functional Programming

Arrows for Parallel Computations

MARTIN BRAUN
University Bayreuth, 95440 Bayreuth, Germany
PHIL TRINDER
Glasgow University, Glasgow, G12 8QQ, Scotland
and OLEG LOBACHEV
University Bayreuth, 95440 Bayreuth, Germany

Abstract

Arrows are a general interface for computation and therefore form an alternative to Monads for API design. We express parallelism using this concept in a novel way: We define an arrows-based language for parallelism and implement it using multiple parallel Haskells. In this manner we are able to bridge across various parallel Haskells.

Additionally, our way of writing parallel programs has the benefit of being portable across flavours of parallel Haskells. Furthermore, as each parallel computation is an arrow, which means that they can be composed and transformed as such. We introduce some syntactic sugar to provide parallelism-aware arrow combinators.

To show that our arrow-based language is on par with the existing parallel languages, we also define several parallel skeletons with our framework. Benchmarks show that our framework does not induce too much overhead performance-wise. **OL: Summarize conclusions**

MB: Jedes Kapitel soll einmal ins Abstract. Conclusions sollen mit ins Abstract

Contents

Contents

1 Introduction

OL: todo, reuse 5.5, and more

MB: Haskell is Spielwiese für Parallelität; verschiedene Ansätze (Par, Multicore, Eden); Orthogonale Ansätze; Verwenden höchstens eine Monade - manchmal auch nur intern; Wir wollen Parallelität mit Arrows abbilden, was noch niemand gemacht hat; Statt einer eigenen Implementierung definieren wir ein "shallow embedded DSL" (ACHTUNG, ist das der richtige Name? effektiv API); Umsetzung mit verschiedenen parallelen Haskells; We tame the zoo of parallel Haskells und vergewissern uns dass es nicht viel Overhead bringt

blablabla arrows, parallel, haskell.

Contribution **OL: HIT HERE REALLY STRONG**

MB: different, how? We wrap parallel Haskells inside of our *ArrowParallel* interface, but why do we aim to abstract parallelism this way and what does this approach do better than the other parallel Haskells?

- **Arrow DSL benefits**: With the *ArrowParallel* typeclass we do not lose any benefits of using arrows as *parEvalN* is a yet another Arrow combinator. The resulting Arrow can be used in the same way a potential serial version could be used. This is a big advantage of this approach, especially compared to the Monad solutions as we do not introduce any new types. We can just ‘plug’ in parallel parts into sequential Arrow-based programs without having to change anything.
- **Abstraction**: With the *ArrowParallel* typeclass, we abstract all parallel implementation logic away from the business logic. This leaves us in the beautiful situation of being able to write our code against the interface of the typeclass without being bound to any parallel Haskell. So as an example, during development, we can run the program in a simple GHC-compiled variant and afterwards deploy it on a cluster by converting it into an Eden version, by just replacing the actual *ArrowParallel* instance.

Structure The remaining text is structures as follows. Section 2 briefly introduces known parallel Haskell flavours (Sec. 2.1) and gives an overview of Arrows to the reader (Sec. 2.2). Section 3 discusses related work. Section 4 defines Parallel Arrows and presents a basic

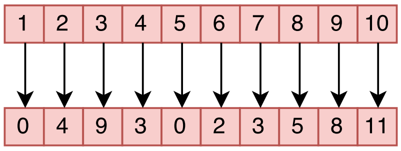


Figure 1: Schematic illustration of *parEvalN*.

interface. Section 5 defines Futures for Parallel Arrows, this concept enables better communication. Section 6 presents some basic algorithmic skeletons in our newly defined dialect: parallel *map* with and without load balancing. More advanced skeletons are showcased in Section 7 (*pipe*, *ring*, *torus*). Section 8 shows the benchmark results. Section 9 discusses future work and concludes.

2 Background

2.1 Short introduction to parallel Haskells

There are already several ways to write parallel programs in Haskell. As we base our parallel arrows on existing parallel Haskells, we will now give a short introduction to the ones we use as backends in this paper.

In its purest form, parallel computation (on functions) can be looked at as the execution of some functions $a \to b$ in parallel or $parEvalN :: [a \to b] \to [a] \to [b]$, as also Figure 1 symbolically shows. As a demonstration, we implement here the non-Arrows *parEvalN* in multiple parallel Haskells.

2.1.1 Multicore Haskell

Multicore Haskell (Marlow *et al.*, 2009; Trinder *et al.*, 1998) is a way to do parallel processing found in standard GHC.[1] It ships with parallel evaluation strategies for several types which can be applied with $using :: a \to Strategy\ a \to a$. Let:

$$\begin{aligned}
&parEvalN :: (NFData\ b) \Rightarrow [a \to b] \to [a] \to [b]\\
&parEvalN\ fs\ as = \mathbf{let}\ bs = zipWith\ (\$)\ fs\ as\\
&\quad \mathbf{in}\ (bs\ `using`\ parList\ rdeepseq)\ `pseq`\ bs
\end{aligned}$$

In the above definition of *parEvalN* we just apply the list of functions $[a \to b]$ to the list of inputs $[a]$ by zipping them with the application operator \$. We then evaluate this lazy list $[b]$ according to a $Strategy\ [b]$ with the $using :: a \to Strategy\ a \to a$ operator. We construct this strategy with $parList :: Strategy\ a \to Strategy\ [a]$ and $rdeepseq :: NFData\ a \Rightarrow Strategy\ a$ where the latter is a strategy which evalutes to normal form. To ensure that programs that use *parEvalN* have the correct evaluation order, we annotate the computation with $pseq :: a \to b \to b$ which forces the compiler to not reorder multiple *parEvalN* computations.

[1] Multicore Haskell on Hackage is available under `https://hackage.haskell.org/package/parallel-3.2.1.0`, compiler support is integrated in the stock GHC.

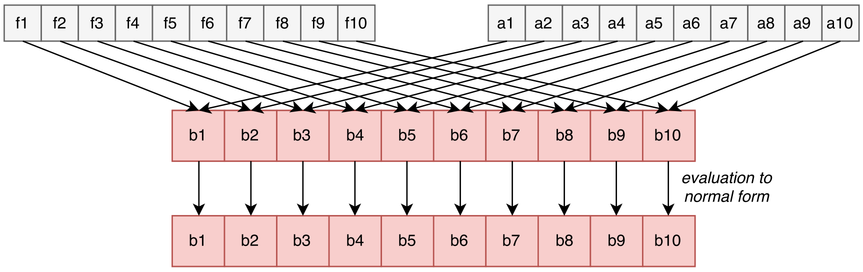


Figure 2: Dataflow of the Multicore Haskell *parEvalN* version.

This is particularly necessary in circular communication topologies like in the *torus* or *ring* (Chap. 7), where a wrong execution order would result in deadlock scenarios when executed without *pseq*. Fig. 2 shows a visual representation of this code.

2.1.2 Par Monad

The *Par* Monad[2] introduced by Marlow *et al.* (2011), is a Monad designed for composition of parallel programs. Let:

$$
\begin{aligned}
&\mathit{parEvalN} :: (\mathit{NFData}\ b) \Rightarrow [a \to b] \to [a] \to [b] \\
&\mathit{parEvalN}\ \mathit{fs}\ \mathit{as} = \mathit{runPar}\ \$ \\
&\quad (\mathit{sequenceA}\ \$\ \mathit{map}\ (\mathit{spawnP})\ \$\ \mathit{zipWith}\ (\$)\ \mathit{fs}\ \mathit{as}) \gg\!= \mathit{mapM}\ \mathit{get}
\end{aligned}
$$

The *Par* Monad version of our parallel evaluation function *parEvalN* is defined by zipping the list of $[a \to b]$ with the list of inputs $[a]$ with the application operator \$ just like with Multicore Haskell. Then, we map over this not yet evaluated lazy list of results $[b]$ with $\mathit{spawnP} :: \mathit{NFData}\ a \Rightarrow a \to \mathit{Par}\ (\mathit{IVar}\ a)$ to transform them to a list of not yet evaluated forked away computations $[\mathit{Par}\ (\mathit{IVar}\ b)]$, which we convert to $\mathit{Par}\ [\mathit{IVar}\ b]$ with *sequenceA*. We wait for the computations to finish by mapping over the *IVar b* values inside the *Par* Monad with *get*. This results in $\mathit{Par}\ [b]$. We execute this process with *runPar* to finally get $[b]$. Fig. 3 shows a graphical representation.

2.1.3 Eden

Eden (Loogen *et al.*, 2005; Loogen, 2012) is a parallel Haskell for distributed memory and comes with a MPI and a PVM backends.[3] It is targeted towards clusters, but also functions well in a shared-memory setting with a further simple backend. However, in contrast to many other parallel Haskells, in Eden each process has its own heap. This seems to be a waste of memory, but with distributed programming paradigm and individual GC per process, Eden yields good performance results also on multicores (Berthold *et al.*, 2009a; Aswad *et al.*, 2009).

[2] It can be found in the `monad-par` package on hackage under `https://hackage.haskell.org/package/monad-par-0.3.4.8/`.

[3] See also `http://www.mathematik.uni-marburg.de/~eden/` and `https://hackage.haskell.org/package/edenmodules-1.2.0.0/`.

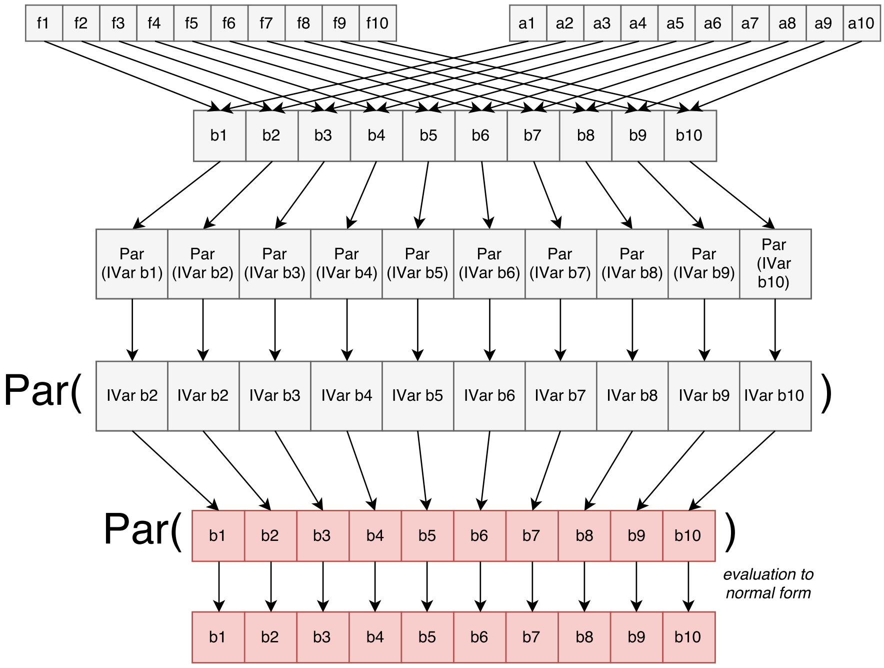


Figure 3: Dataflow of the *Par* Monad *parEvalN* version.

While Eden also comes with a Monad *PA* for parallel evaluation, it also ships with a completely functional interface that includes a $spawnF :: (Trans\ a, Trans\ b) \Rightarrow [a \to b] \to [a] \to [b]$ function that allows us to define *parEvalN* directly:

$$parEvalN :: (Trans\ a, Trans\ b) \Rightarrow [a \to b] \to [a] \to [b]$$
$$parEvalN = spawnF$$

Eden TraceViewer. To comprehend the efficiency and the lack thereof in a parallel program, an inspection of its execution is extremely helpful. While some large-scale solutions exist (Geimer *et al.*, 2010), the parallel Haskell community mainly utilises the tools Threadscope (Wheeler & Thain, 2009) and Eden TraceViewer[4] (Berthold & Loogen, 2007). In the next sections we will present some *trace visualizations*, the post-mortem process diagrams of Eden processes and their activity.

The trace visualizations are color-coded. In such a visualization (Fig. 13), the x axis shows the time, the y axis enumerates the machines and processes. The visualization shows a running process in green, a blocked process is red. If the process is 'runnable', i.e. it may run, but does not, it is yellow. The typical reason for thus is GC. An inactive machine, where no processes are started yet, or all are already terminated, shows as a blue bar. A comminication from one process to another is represented with a black arrow. A stream

[4] See `http://hackage.haskell.org/package/edentv` on Hackage for the last available version of Eden TraceViewer.

```
class Arrow arr where
  arr :: (a → b) → arr a b
  (>>>) :: arr a b → arr b c → arr a c
  first :: arr a b → arr (a,c) (b,c)
instance Arrow (→) where
  arr f = f
  f >>> g = g ∘ f
  first f = λ(a,c) → (f a,c)
data Kleisli m a b = Kleisli { run :: a → m b }
instance Monad m ⇒ Arrow (Kleisli m) where
  arr f = Kleisli (return ∘ f)
  f >>> g = Kleisli (λa → f a >>= g)
  first f = Kleisli (λ(a,c) → f a >>= λb → return (b,c))
```

Figure 4: The definition of *Arrow* type class and its two most typical instances.

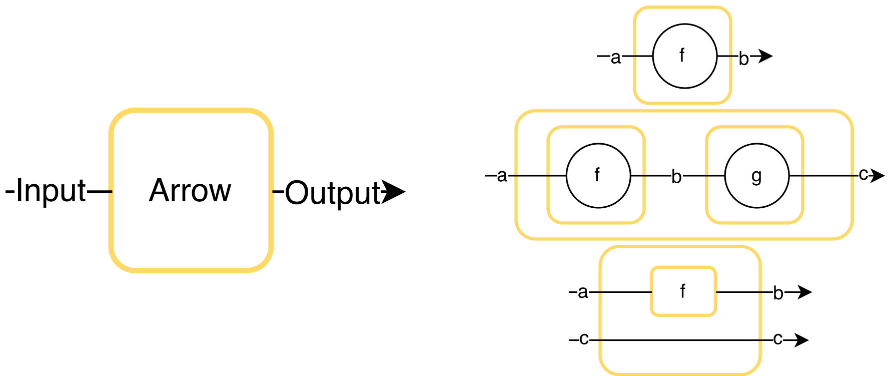


Figure 5: Schematic depiction of Arrow (left) and its basic combinators *arr*, >>> and *first* (right).

of communications, e. g. a transmitted list is shows as a dark shading between sender and receiver processes.

2.2 Arrows

Arrows were introduced by Hughes (2000) as a general interface for computation. An Arrow *arr a b* represents a computation that converts an input *a* to an output *b*. This is defined in the *Arrow* type class shown in Fig. 4. Its *arr* operation is used to lift an ordinary function to the specified arrow type, similarly to the monadic *return*. The >>> operator is analogous to the monadic composition >>= and combines two arrows *arr a b* and *arr b c* by "wiring" the outputs of the first to the inputs to the second to get a new arrow *arr a c*. Lastly, the *first* operator takes the input arrow from *b* to *c* and converts it into an arrow on pairs with the second argument untouched. It allows us to to save input across arrows. Figure 5 shows a graphical representation of the basic Arrow combinators. The most prominent instances of this interface are regular functions ($\rightarrow$) and the Kleisli type (Fig. 4).

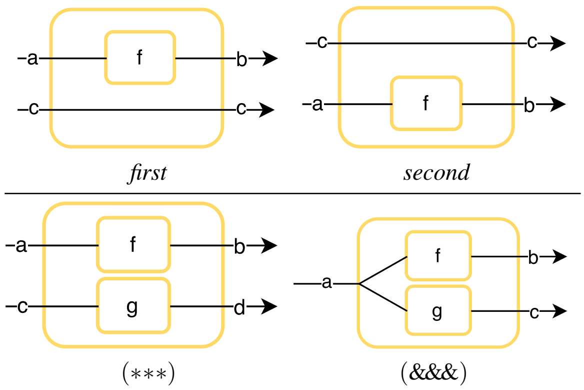


Figure 6: Visual depiction of syntactic sugar for Arrows.

With this typeclass in place, Hughes also defined some syntactic sugar (Fig. 6): The combinators *second*, $***$ and &&&. The combinator *second* is the mirrored version of *first* (Appendix A). The $***$ function combines *first* and *second* to handle two inputs in one arrow, is defined as

$$(***) :: Arrow\ arr \Rightarrow arr\ a\ b \rightarrow arr\ c\ d \rightarrow arr\ (a,c)\ (b,d)$$
$$f *** g = first\ f >>> second\ g$$

while the &&& combinator, that constructs an arrow which outputs two different values like $***$, but takes only one input, is:

$$(\&\&\&) :: Arrow\ arr \Rightarrow arr\ a\ b \rightarrow arr\ a\ c \rightarrow a\ a\ (b,c)$$
$$f\ \&\&\&\ g = arr\ (\lambda a \rightarrow (a,a)) >>> (f *** g)$$

A first short example given by Hughes on how to use arrows is addition with arrows:

$$add :: Arrow\ arr \Rightarrow arr\ a\ Int \rightarrow arr\ a\ Int \rightarrow arr\ a\ Int$$
$$add\ f\ g = (f\ \&\&\&\ g) >>> arr\ (\lambda(u,v) \rightarrow u+v)$$

The more restrictive interface of Arrows allows for more elaborate composition and transformation combinators—a Monad can be *anything*, an Arrow is a process of doing something, a *computation*. One of the major problems in parallel computing is, however, composition of parallel processes.

3 Related Work

Parallel Haskells. Of course, the three parallel Haskell flavours we have presented above: the GpH (Trinder *et al.*, 1996, 1998) parallel Haskell dialect and its multicore version (Marlow *et al.*, 2009), the *Par* Monad (Marlow *et al.*, 2011; Foltzer *et al.*, 2012), and Eden (Loogen *et al.*, 2005; Loogen, 2012) are related to this work. We use these languages as backends: our DSL can switch from one to other at user's command.

HdpH (Maier *et al.*, 2014; Stewart *et al.*, 2016) is an extension of *Par* Monad to heterogeneous clusters. LVish (Kuper *et al.*, 2014) is a communication-centred extension of *Par* Monad. Further parallel Haskell approaches include pH (Nikhil & Arvind, 2001), research work done on distributed variants of GpH (Trinder *et al.*, 1996; Aljabri *et al.*, 2014, 2015), and low-level Eden implementation (Berthold, 2008; Berthold *et al.*, 2016). Skeleton composition (Dieterle *et al.*, 2016), communication (Dieterle *et al.*, 2010a), and generation of process networks (Horstmeyer & Loogen, 2013) are recent in-focus research topics in Eden. This also includes the definitions of new skeletons (Hammond *et al.*, 2003; Berthold & Loogen, 2006; Berthold *et al.*, 2009b,c; Dieterle *et al.*, 2010b; de la Encina *et al.*, 2011; Dieterle *et al.*, 2013; Janjic *et al.*, 2013).

More different approaches include data parallelism (Chakravarty *et al.*, 2007; Keller *et al.*, 2010), GPU-based approaches (Mainland & Morrisett, 2010; Svensson, 2011), software transactional memory (Harris *et al.*, 2005; Perfumo *et al.*, 2008). The Haskell–GPU bridge Accelerate (Chakravarty *et al.*, 2011; Clifton-Everest *et al.*, 2014; McDonell *et al.*, 2015) deserves a special mention. Accelerate is completely orthogonal to our approach. Marlow authored a recent book in 2013 on parallel Haskells.

Algorithmic skeletons. Algorithmic skeletons were introduced by Cole (1989). Early publications on this topic include (Darlington *et al.*, 1993; Botorog & Kuchen, 1996; Danelutto *et al.*, 1997; Gorlatch, 1998; Lengauer *et al.*, 1997). Rabhi & Gorlatch (2003) consolidated early reports on high-level programming approaches. The effort is ongoing, including topological skeletons (Berthold & Loogen, 2006), special-purpose skeletons for computer algebra (Berthold *et al.*, 2009c; Lobachev, 2011, 2012; Janjic *et al.*, 2013), iteration skeletons (Dieterle *et al.*, 2013). The idea of Linton *et al.* (2010) is to use a parallel Haskell to orchestrate further software systems to run in parallel. Dieterle *et al.* (2016) compare the composition of skeletons to stable process networks.

Arrows. Arrows were introduced by Hughes (2000), basically they are a generalised function arrow $\rightarrow$. Hughes (2005a) presents a tutorial on Arrows. Some theoretical details on Arrows (Jacobs *et al.*, 2009; Lindley *et al.*, 2011; Atkey, 2011) are viable. Paterson (2001) introduced a new notation for Arrows. Arrows have applications in information flow research (Li & Zdancewic, 2006, 2010; Russo *et al.*, 2008), invertible programming (Alimarine *et al.*, 2005), and quantum computer simulation (Vizzotto *et al.*, 2006). But probably most prominent application of Arrows is Arrow-based functional reactive programming, AFRP (Nilsson *et al.*, 2002; Hudak *et al.*, 2003; Czaplicki & Chong, 2013). Liu *et al.* (2009) formally define a more special kind of Arrows that capsule the computation more than regular arrows do and thus enable optimizations. Their approach would allow parallel composition, as their special Arrows would not interfere with each other in concurrent execution. In contrast, we capture a whole parallel computation as a single entity: our main instantiation function *parEvalN* makes a single (parallel) Arrow out of list of Arrows.Huang *et al.* (2007) utilise Arrows for parallelism, but strikingly different from our approach. They basically use Arrows to orchestrate several tasks in robotics. We, however, propose a general interface for parallel programming, while remaining completely in Haskell.

Other languages. Although this work is centred on Haskell implementation of arrows, it is applicable to any functional programming language where parallel evaluation and arrows can be defined. Experiments with our approach in Frege language[5] (which is basically Haskell on the JVM) were quite successful. However, it is beyond the scope of this work.

Achten *et al.* (2004, 2007) use an arrow implementation in Clean for better handling of typical GUI tasks. Dagand *et al.* (2009) used arrows in OCaml in the implementation of a distributed system.

4 Parallel Arrows

Arrows are a general interface to computation. Here we introduce special Arrows as general interface to *parallel computations*. First, we present the interface and explain the reasonings behind it. Then, we discuss some implementations using exisiting parallel Haskells. Finally, we explain why using Arrows for expressing parallelism is beneficial.

4.1 The ArrowParallel typeclass

A parallel computation (on functions) in its purest form can be seen as execution of some functions $a \to b$ in parallel, as our *parEvalN* prototype shows (Sec. 2.1). Translating this into arrow terms gives us a new operator *parEvalN* that lifts a list of arrows $[arr\ a\ b]$ to a parallel arrow $arr\ [a]\ [b]$. This combinator is similar to our utility function *listApp* from Appendix A, but does parallel instead of serial evaluation.

$$parEvalN :: (Arrow\ arr) \Rightarrow [arr\ a\ b] \to arr\ [a]\ [b]$$

With this definition of *parEvalN*, parallel execution is yet another arrow combinator. But as the implementation may differ depending on the actual type of the arrow *arr* and we want this to be an interface for different backends, we introduce a new typeclass *ArrowParallel arr a b*:

```
class Arrow arr ⇒ ArrowParallel arr a b where
  parEvalN :: [arr a b] → arr [a] [b]
```

Sometimes parallel Haskells require or allow for additional configuration parameters, e. g. an information about the execution environment or the level of evaluation (weak head normal form vs. normal form). For this reason we also introduce an additional *conf* parameter to the function. We also do not want *conf* to be a fixed type, as the configuration parameters can differ for different instances of *ArrowParallel*.

```
class Arrow arr ⇒ ArrowParallel arr a b conf where
  parEvalN :: conf → [arr a b] → arr [a] [b]
```

We do not require the *conf* parameter in every implementation. If it is not needed, we usually just default the *conf* type parameter to () and even blank it out in the parameter list of the implemented *parEvalN*.

[5] GitHub project page at `https://github.com/Frege/frege`

```
data Conf a = Conf (Strategy a)
instance (NFData b, ArrowApply arr, ArrowChoice arr) => ArrowParallel arr a b () where
  parEvalN _ fs =
    listApp fs >>>
    arr (withStrategy (parList rdeepseq)) &&& arr id >>>
    arr (uncurry pseq)
```

Figure 7: Fully evaluating *ArrowParallel* instance for the Multicore Haskell backend.

```
instance (NFData b, ArrowApply arr, ArrowChoice arr) =>
  ArrowParallel arr a b (Conf b) where
  parEvalN (Conf strat) fs =
    listApp fs >>>
    arr (withStrategy (parList strat)) &&& arr id >>>
    arr (uncurry pseq)
```

Figure 8: Configurable *ArrowParallel* instance for the Multicore Haskell backend.

4.2 *ArrowParallel* instances

4.2.1 Multicore Haskell

The Multicore Haskell implementation of *ArrowParallel* is implemented in a straightforward manner by using *listApp* (Appendix A) combined with the $withStrategy :: Strategy\ a \to a \to a$ and $pseq :: a \to b \to b$ combinators from Multicore Haskell, where *withStrategy* is the same as $using :: a \to Strategy\ a \to a$ but with flipped parameters. For most cases a fully evaluating version like in Fig. 7 would probably suffice, but as the Multicore Haskell interface allows the user to specify the level of evaluation to be done via the *Strategy* interface, our DSL should allow for this. We therefore introduce the *Conf a* data-type that simply wraps a *Strategy a*.

4.2.2 Par Monad

OL: introduce a newcommand for par-monad, "arrows", "parrows" and replace all mentions to them to ensure uniform typesetting done!, we write Arrows. also "Monad"? done! The *Par* Monad implementation (Fig. 9) makes use of Haskells laziness and *Par* Monad's $spawnP :: NFData\ a \Rightarrow a \to Par\ (IVar\ a)$ function. The latter forks away the computation of a value and returns an *IVar* containing the result in the *Par* Monad.

We therefore apply each function to its corresponding input value with and then fork the computation away with *arr spawnP* inside a *zipWithArr* (Fig. A 3) call. This yields a list $[Par\ (IVar\ b)]$, which we then convert into $Par\ [IVar\ b]$ with *arr sequenceA*. In order to wait for the computation to finish, we map over the *IVar*s inside the *Par* Monad with *arr* $(\gg\!= mapM\ get)$. The result of this operation is a $Par\ [b]$ from which we can finally remove the Monad again by running *arr runPar* to get our output of $[b]$.

$$
\begin{array}{l}
\textbf{instance}\ (\mathit{NFData}\ b, \mathit{ArrowApply}\ \mathit{arr}, \mathit{ArrowChoice}\ \mathit{arr}) \Rightarrow \\
\quad \mathit{ArrowParallel}\ \mathit{arr}\ a\ b\ \mathit{conf}\ \textbf{where} \\
\qquad \mathit{parEvalN}\ \_\ \mathit{fs} = \\
\qquad\quad (\mathit{arr}\ \$\ \lambda \mathit{as} \rightarrow (\mathit{fs}, \mathit{as})) \ggg \\
\qquad\quad \mathit{zipWithArr}\ (\mathit{app} \ggg \mathit{arr}\ \mathit{spawnP}) \ggg \\
\qquad\quad \mathit{arr}\ \mathit{sequenceA} \ggg \\
\qquad\quad \mathit{arr}\ (\gg\!= \mathit{mapM}\ \mathit{get}) \ggg \\
\qquad\quad \mathit{arr}\ \mathit{runPar}
\end{array}
$$

Figure 9: ArrowParallel instance for the Par Monad backend.

4.2.3 Eden

For both the Multicore Haskell and *Par* Monad implementations we could use general instances of *ArrowParallel* that just require the *ArrowApply* and *ArrowChoice* typeclasses. With Eden this is not the case as we can only spawn a list of functions and we cannot extract simple functions out of arrows. While we could still manage to have only one class in the module by introducing a typeclass:

$$
\begin{array}{l}
\textbf{class}\ (\mathit{Arrow}\ \mathit{arr}) \Rightarrow \mathit{ArrowUnwrap}\ \mathit{arr}\ \textbf{where} \\
\quad \mathit{arr}\ a\ b \rightarrow (a \rightarrow b)
\end{array}
$$

However, we avoid doing so for aesthetic resons. For now, we just implement *ArrowParallel* for normal functions:

$$
\begin{array}{l}
\textbf{instance}\ (\mathit{Trans}\ a, \mathit{Trans}\ b) \Rightarrow \mathit{ArrowParallel}\ (\rightarrow)\ a\ b\ \mathit{conf}\ \textbf{where} \\
\mathit{parEvalN}\ \_\ \mathit{fs}\ \mathit{as} = \mathit{spawnF}\ \mathit{fs}\ \mathit{as}
\end{array}
$$

and the Kleisli type:

$$
\begin{array}{l}
\textbf{instance}\ (\mathit{Monad}\ m, \mathit{Trans}\ a, \mathit{Trans}\ b, \mathit{Trans}\ (m\ b)) \Rightarrow \\
\quad \mathit{ArrowParallel}\ (\mathit{Kleisli}\ m)\ a\ b\ \mathit{conf}\ \textbf{where} \\
\mathit{parEvalN}\ \mathit{conf}\ \mathit{fs} = \\
\quad (\mathit{arr}\ \$\ \mathit{parEvalN}\ \mathit{conf}\ (\mathit{map}\ (\lambda (\mathit{Kleisli}\ f) \rightarrow f)\ \mathit{fs})) \ggg \\
\quad (\mathit{Kleisli}\ \$\ \mathit{sequence})
\end{array}
$$

4.3 Extending the Interface

With the *ArrowParallel* typeclass in place and implemented, we can now implement some further basic parallel interface functions. These are algorithmic skeletons that, however, mostly serve as a foundation to further, more specific algorithmic skeletons.

4.3.1 Lazy parEvalN

The function *parEvalN* is 100% strict, which means that it fully evaluates all passed arrows. Sometimes this might not be feasible, as it will not work on infinite lists of functions like e.g. $\mathit{map}\ (\mathit{arr} \circ (+))\ [1\,..]$ or just because we need the arrows evaluated in chunks. *parEvalNLazy* (Figs. 10, 11) fixes this. It works by first chunking the input from $[a]$ to $[[a]]$

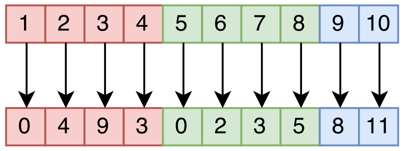


Figure 10: Schematic depiction of *parEvalNLazy*.

```
parEvalNLazy :: (ArrowParallel arr a b conf, ArrowChoice arr, ArrowApply arr) ⇒
  conf → ChunkSize → [arr a b] → (arr [a] [b])
parEvalNLazy conf chunkSize fs =
  arr (chunksOf chunkSize) >>>
  listApp fchunks >>>
  arr concat
  where fchunks = map (parEvalN conf) $ chunksOf chunkSize fs
```

Figure 11: Definition of *parEvalNLazy*.

with the given *ChunkSize* in *arr* (*chunksOf chunkSize*). These chunks are then fed into a list [*arr* [*a*] [*b*]] of parallel arrows created by feeding chunks of the passed *ChunkSize* into the regular parEvalN by using *listApp*. The resulting [[*b*]] is lastly converted into [*b*] with *arr concat*.

4.3.2 Heterogenous tasks

We have only talked about the paralellization arrows of the same type until now. But sometimes we want to paralellize heterogenous types as well. However, we can implement such a *parEval2* combinator (Figs. 12, B 11) which combines two arrows *arr a b* and *arr c d* into a new parallel arrow *arr* (a,c) (b,d) quite easily with the help of the *ArrowChoice* typeclass. The idea is to use the +++ combinator which combines two arrows *arr a b* and *arr c d* and transforms them into *arr* (*Either a c*) (*Either b d*) to get a common arrow type that we can then feed into *parEvalN*.

5 Futures

Consider a mock-up parallel arrow combinator:

```
someCombinator :: (Arrow arr) ⇒ [arr a b] → [arr b c] → arr [a] [c]
someCombinator fs1 fs2 = parEvalN () fs1 >>> rightRotate >>> parEvalN () fs2
```

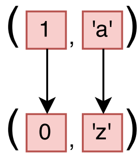


Figure 12: Schematic depiction of *parEval2*.

In a distributed environment, a resulting arrow of this combinator first evaluates all $[arr\ a\ b]$ in parallel, sends the results back to the master node, rotates the input once and then evaluates the $[arr\ b\ c]$ in parallel to then gather the input once again on the master node. Such situations arise, e. g. in scientific computations when the data distributed across the nodes needs to be transposed. A concrete example is 2D FFT computation (Gorlatch & Bischof, 1998; Berthold *et al.*, 2009c).

While the above example could be rewritten into only one *parEvalN* call by directly wiring the arrows properly together, it illustrates an important problem. When using a *ArrowParallel* backend that resides on multiple computers, all communication between the nodes is done via the master node, as shown in the Eden trace in Figure 13. This can become a serious bottleneck for larger amount of data and number of processes (as e. g. Berthold *et al.*, 2009c, showcases).

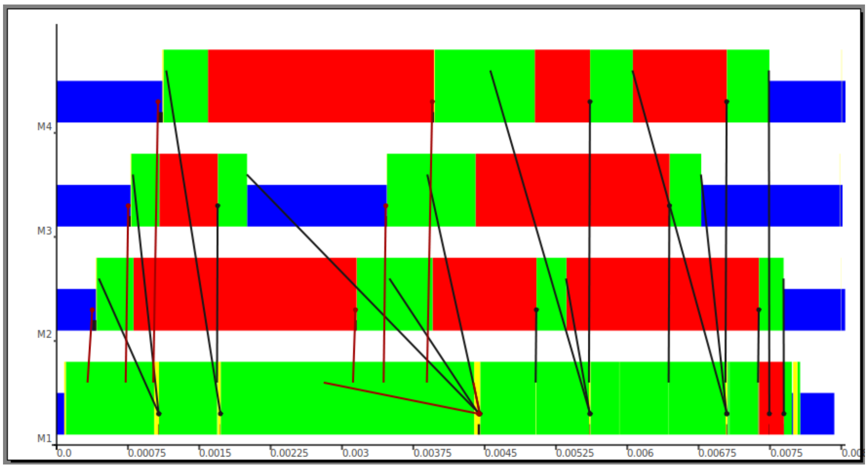


Figure 13: Communication between 4 Eden processes without Futures. All communication goes through the master node. Each bar represents one process. Black lines between processes represent communication. Colors: blue ≙ idle, green ≙ running, red ≙ blocked, yellow ≙ suspended.

OL: more practical and heavy-weight example! fft (I have the code)?
MB: Depends... Are the communications easy to read in such an example?
MB: Keep the description for the different colours, or link to the EdenTV description in 2.1.3
OL: ok as is OL: use the fft example (when it works)?

We should allow the nodes to communicate directly with each other. Eden already ships with "remote data" that enable this (Alt & Gorlatch, 2003; Dieterle *et al.*, 2010a). But as we want code with our DSL to be implementation agnostic, we have to wrap this context. We do this with the *Future* typeclass (Fig. 14). Since *RD* is only a type synonym for a communication type that Eden uses internally, we have to use some wrapper classes to fit that definition, though, as Fig. B 5 shows. Technical details are in Appendix, in Section B.

For our *Par* Monad and Multicore Haskell backends, we can simply use *MVar*s (Jones *et al.*, 1996) (Fig. 15), as in a shared memory setting we do not require Eden's sophisticated communication channels. **MB: explain MVars**

$$
\begin{array}{l}
\textbf{class } \mathit{Future\ fut\ a} \mid a \rightarrow \mathit{fut} \textbf{ where} \\
\quad \mathit{put} :: (\mathit{Arrow\ arr}) \Rightarrow \mathit{arr\ a}\ (\mathit{fut\ a}) \\
\quad \mathit{get} :: (\mathit{Arrow\ arr}) \Rightarrow \mathit{arr}\ (\mathit{fut\ a})\ a
\end{array}
$$

Figure 14: Definition of the *Future* typeclass.

$$
\begin{array}{l}
\{\text{-\# NOINLINE putUnsafe \#-}\} \\
\mathit{putUnsafe} :: a \rightarrow \mathit{MVar\ a} \\
\mathit{putUnsafe\ a} = \mathit{unsafePerformIO}\ \$\ \textbf{do} \\
\quad \mathit{mVar} \leftarrow \mathit{newEmptyMVar} \\
\quad \mathit{putMVar\ mVar\ a} \\
\quad \mathit{return\ mVar} \\
\textbf{instance } (\mathit{NFData\ a}) \Rightarrow \mathit{Future\ MVar\ a} \textbf{ where} \\
\quad \mathit{put} = \mathit{arr\ putUnsafe} \\
\quad \mathit{get} = \mathit{arr\ takeMVar} \ggg \mathit{arr\ unsafePerformIO}
\end{array}
$$

Figure 15: A *MVar* instance of the *Future* typeclass for the *Par* Monad and Multicore Haskell backends.

In our communication example we can use this *Future* concept for direct communications between the nodes as shown in Fig. 16. In a distributed environment, this gives us a

$$
\begin{array}{l}
\mathit{someCombinator} :: (\mathit{Arrow\ arr}) \Rightarrow [\mathit{arr\ a\ b}] \rightarrow [\mathit{arr\ b\ c}] \rightarrow \mathit{arr}\ [a]\ [c] \\
\mathit{someCombinator\ fs1\ fs2} = \\
\quad \mathit{parEvalN}\ ()\ (\mathit{map}\ ({\ggg}\mathit{put})\ \mathit{fs1}) \ggg \\
\quad \mathit{rightRotate} \ggg \\
\quad \mathit{parEvalN}\ ()\ (\mathit{map}\ (\mathit{get}{\ggg})\ \mathit{fs2})
\end{array}
$$

Figure 16: The mock-up combinator in parallel.

communication scheme with messages going through the master node only if it is needed—similar to what is shown in the trace visualization in Fig. 17.**OL: Fig. is not really clear. Do Figs with a lot of load? — fft?**

6 Map-based Skeletons

Now we have developed Parallel Arrows far enough to define some useful algorithmic skeletons that abstract typical parallel computations.

Parallel *map* and laziness. The *parMap* skeleton (Figs. B 1, B 2) is probably the most common skeleton for parallel programs. We can implement it with *ArrowParallel* by repeating an arrow *arr a b* and then passing it into *parEvalN* to obtain an arrow *arr* $[a]$ $[b]$. Just like *parEvalN*, *parMap* is 100% strict. As *parMap* is 100% strict it has the same restrictions as *parEvalN* compared to *parEvalNLazy*. So it makes sense to also have a

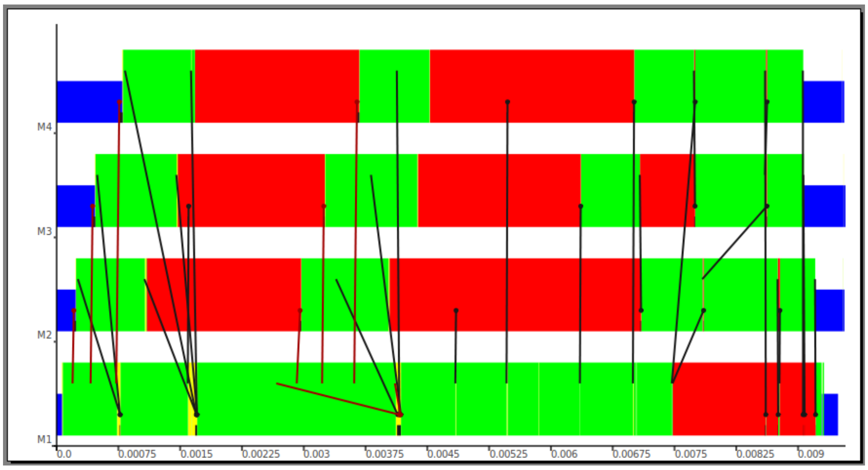


Figure 17: Communication between 4 Eden processes with Futures. Other than in Fig. 13, processes communicate directly (black lines between the bars) instead of always going through the master node (bottom bar).

parMapStream (Figs. B 3, B 4) which behaves like *parMap*, but uses *parEvalNLazy* instead of *parEvalN*. The code is quite straightforward, we show it in Appendix.

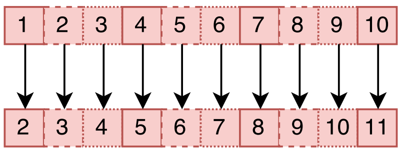


Figure 18: Schematic depiction of a *farm*, a statically load-balanced *parMap*.

```
farm :: (ArrowParallel arr a b conf,
   ArrowParallel arr [a] [b] conf, ArrowChoice arr) ⇒
   conf → NumCores → arr a b → arr [a] [b]
farm conf numCores f =
   unshuffle numCores >>>
   parEvalN conf (repeat (mapArr f)) >>>
   shuffle
```

Figure 19: The definition of *farm*.

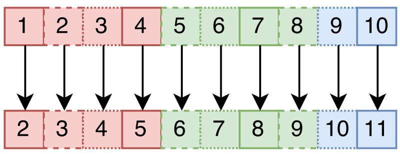


Figure 20: Schematic depiction of *farmChunk*.

Statically load-balancing parallel *map*. Our *parMap* spawns every single computation in a new thread (at least for the instances of *ArrowParallel* we gave in this paper). This can be quite wasteful and a statically load-balancing *farm* (Figs. 18, 19) that equally distributes the workload over *numCores* workers seems useful. The definitions of the helper functions *unshuffle*, *takeEach*, *shuffle* (Fig. B 6) originate from an Eden skeleton[6].

Since a *farm* is basically just *parMap* with a different work distribution, it is, again, 100% strict. So we can define *farmChunk* (Figs. 20, B 9) which uses *parEvalNLazy* instead of *parEvalN*. It is basically the same definition as for *farm*, with *parEvalN* replaced with *parEvalNLazy*.

7 Topological Skeletons

Even though many algorithms can be expressed by parallel maps, some problems require more sophisticated skeletons. The Eden library leverages this problem and already comes with more predefined skeletons[7], among them a *pipe*, a *ring*, and a *torus* implementations (Loogen, 2012). These seem like reasonable candidates to be ported to our Arrow-based parallel Haskell. We aim to showcase that we can express more sophisticated skeletons with Parallel Arrows as well.

7.1 Parallel pipe

The parallel *pipe* skeleton is semantically equivalent to folding over a list $[arr\ a\ a]$ of arrows with $\ggg$, but does this in parallel, meaning that the arrows do not have to reside on the same thread/machine. We implement this skeleton using the *ArrowLoop* typeclass which gives us the $loop :: arr\ (a,b)\ (c,b) \rightarrow arr\ a\ c$ combinator which allows us to express recursive fix-point computations in which output values are fed back as input. For example

$$loop\ (arr\ (\lambda(a,b) \rightarrow (b, a:b)))$$

which is the same as

$$loop\ (arr\ snd\ \&\&\&\ arr\ (uncurry\ (:)))$$

defines an arrow that takes its input a and converts it into an infinite stream $[a]$ of it. Using this to our advantage gives us a first draft of a pipe implementation (Fig. 21) by plugging in the parallel evaluation call *parEvalN conf fs* inside the second argument of &&& and then only picking the first element of the resulting list with *arr last*.

However, using this definition directly will make the master node a potential bottleneck in distributed environments as described in Section 5. Therefore, we introduce a more sophisticated version that internally uses Futures and obtain the final definition of *pipe* in Fig. 22.

[6] Available on Hackage under `https://hackage.haskell.org/package/edenskel-2.1.0.0/docs/src/Control-Parallel-Eden-Map.html`.

[7] Available on Hackage: `https://hackage.haskell.org/package/edenskel-2.1.0.0/docs/Control-Parallel-Eden-Topology.html`.

$$
\begin{aligned}
&\mathit{pipeSimple} :: (\mathit{ArrowLoop\ arr}, \mathit{ArrowParallel\ arr\ a\ a\ conf}) \Rightarrow \\
&\quad \mathit{conf} \to [\mathit{arr\ a\ a}] \to \mathit{arr\ a\ a} \\
&\mathit{pipeSimple\ conf\ fs} = \\
&\quad \mathit{loop}\ (\mathit{arr\ snd} \,\&\&\&\, \\
&\qquad (\mathit{arr}\ (\mathit{uncurry}\ (:) \ggg \mathit{lazy}) \ggg \mathit{parEvalN\ conf\ fs})) \ggg \\
&\quad \mathit{arr\ last}
\end{aligned}
$$

Figure 21: A first implementation of the *pipe* skeleton expressed with Parallel Arrows. Note that the use of *lazy* (Fig. B 7) is essential as without it programs using this definition would never halt. We need to enforce that the evaluation of the input $[a]$ terminates before passing it into *parEvalN*.

$$
\begin{aligned}
&\mathit{pipe} :: (\mathit{ArrowLoop\ arr}, \mathit{ArrowParallel\ arr}\ (\mathit{fut\ a})\ (\mathit{fut\ a})\ \mathit{conf}, \mathit{Future\ fut\ a}) \Rightarrow \\
&\quad \mathit{conf} \to [\mathit{arr\ a\ a}] \to \mathit{arr\ a\ a} \\
&\mathit{pipe\ conf\ fs} = \mathit{unliftFut}\ (\mathit{pipeSimple\ conf}\ (\mathit{map\ liftFut\ fs}))
\end{aligned}
$$

Figure 22: Final definition of the *pipe* skeleton with Futures.

$$
\begin{aligned}
&\mathit{pipe2} :: (\mathit{ArrowLoop\ arr}, \mathit{ArrowChoice\ arr}, \mathit{Future\ fut}\ (([a],[b]),[c]), \\
&\quad \mathit{ArrowParallel\ arr}\ (\mathit{fut}\ (([a],[b]),[c]))\ (\mathit{fut}\ (([a],[b]),[c]))\ \mathit{conf}) \Rightarrow \\
&\quad \mathit{conf} \to \mathit{arr\ a\ b} \to \mathit{arr\ b\ c} \to \mathit{arr\ a\ c} \\
&\mathit{pipe2\ conf\ f\ g} = \\
&\quad (\mathit{arr\ return} \,\&\&\&\, \mathit{arr}\ (\mathit{const}\ [\,]))\,\&\&\&\, \mathit{arr}\ (\mathit{const}\ [\,]) \ggg \\
&\quad \mathit{pipe\ conf}\ (\mathit{replicate}\ 2\ (\mathit{unify\ f\ g})) \ggg \\
&\quad \mathit{arr\ snd} \ggg \mathit{arr\ head}\ \textbf{where} \\
&\qquad \mathit{unify} :: (\mathit{ArrowChoice\ arr}) \Rightarrow \mathit{arr\ a\ b} \to \mathit{arr\ b\ c} \to \mathit{arr}\ (([a],[b]),[c])\ (([a],[b]),[c]) \\
&\qquad \mathit{unify\ f\ g} = \\
&\qquad\quad (\mathit{mapArr\ f} *\!*\!* \mathit{mapArr\ g}) *\!*\!* \mathit{arr}\ (\backslash\_ \to [\,]) \ggg \\
&\qquad\quad \mathit{arr}\ (\lambda((a,b),c) \to ((c,a),b)) \\
&(|\ggg|) :: (\mathit{ArrowLoop\ arr}, \mathit{ArrowChoice\ arr}, \mathit{Future\ fut}\ (([a],[b]),[c]), \\
&\quad \mathit{ArrowParallel\ arr}\ (\mathit{fut}\ (([a],[b]),[c]))\ (\mathit{fut}\ (([a],[b]),[c]))\ ()) \Rightarrow \\
&\quad \mathit{arr\ a\ b} \to \mathit{arr\ b\ c} \to \mathit{arr\ a\ c} \\
&(|\ggg|) = \mathit{pipe2}\ ()
\end{aligned}
$$

Figure 23: Definition of *pipe2* and a parallel $\ggg$.

Sometimes, this *pipe* definition can be a bit inconvenient, especially if we want to pipe arrows of mixed types together, i.e. *arr a b* and *arr b c*. By wrapping these two arrows inside a common type we obtain *pipe2* (Fig. 23).

Note that extensive use of *pipe2* over *pipe* with a hand-written combination data type will probably result in worse performance because of more communication overhead from the many calls to parEvalN. Nonetheless, we can define a version of parallel piping operator $|\ggg|$, which is semantically equivalent to $\ggg$ similarly to other parallel syntactic sugar from Appendix C.

Another version of $\ggg$ is:

$$f\ \mathit{parparcomp}\ g = (f \circ \mathit{put}) \ggg (\mathit{get} \circ g)$$

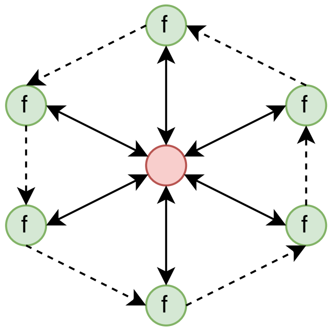


Figure 24: Schematic depiction of the ring skeleton.

It does not launch both arrows f and g in parallel, but allows for more smooth data commuinication between them. Basically, it is a *Future*-lifted *sequential* $\ggg$, a way to compose parallel Arrows efficiently.

7.2 Ring skeleton

Eden comes with a ring skeleton[8] (Fig. 24) implementation that allows the computation of a function $[i] \to [o]$ with a ring of nodes that communicate in a ring topology with each other. Its input is a node function $i \to r \to (o,r)$ in which r serves as the intermediary output that gets send to the neighbour of each node. This data is sent over direct communication channels, the so called 'remote data'. We depict it in Appendix, Fig. B 10.

We can rewrite this functionality easily with the use of *loop* as the definition of the node function, $arr\ (i,r)\ (o,r)$, after being transformed into an arrow, already fits quite neatly into the *loop*'s $arr\ (a,b)\ (c,b) \to arr\ a\ c$. In each iteration we start by rotating the intermediary input from the nodes $[fut\ r]$ with $second\ (rightRotate \ggg lazy)$ (Fig. B 7). Similarly to the *pipe* from Section 7.1 (Fig. 21), we have to feed the intermediary input into our *lazy* (Fig. B 7) arrow here, or the evaluation would fail to terminate. The reasoning is explained by Loogen (2012) as a demand problem.

Next, we zip the resulting $([i],[fut\ r])$ to $[(i,fut\ r)]$ with $arr\ (uncurry\ zip)$ so we can feed that into a our input arrow $arr\ (i,r)\ (o,r)$, which we transform into $arr\ (i,fut\ r)\ (o,fut\ r)$ before lifting it to $arr\ [(i,fut\ r)]\ [(o,fut\ r)]$ to get a list $[(o,fut\ r)]$. Finally we unzip this list into $([o],[fut\ r])$. Plugging this arrow $arr\ ([i],[fut\ r])\ ([o],fut\ r)$ into the definition of *loop* from earlier gives us $arr\ [i]\ [o]$, our ring arrow (Fig. 25). This combinator can, for example, be used to calculate the shortest paths in a graph using Warshall's algorithm.

Note that the Par Monad version of the *ArrowParallel* version presented in this paper (Fig. 9) does not work with the looping behaviour in the *ring* skeletons. Whenever *runPar* is invoked, strictness is enforced, which means that the *lazy* operation in the definition of *ring* will not work.

[8] Available on Hackage: `https://hackage.haskell.org/package/edenskel-2.1.0.0/docs/Control-Parallel-Eden-Topology.html`

$$
\begin{aligned}
&\mathit{ring} :: (\mathit{ArrowLoop}\ \mathit{arr}, \mathit{Future}\ \mathit{fut}\ r, \mathit{ArrowParallel}\ \mathit{arr}\ (i, \mathit{fut}\ r)\ (o, \mathit{fut}\ r)\ \mathit{conf}) \Rightarrow\\
&\quad \mathit{conf} \rightarrow \mathit{arr}\ (i,r)\ (o,r) \rightarrow \mathit{arr}\ [i]\ [o]\\
&\mathit{ring}\ \mathit{conf}\ f =\\
&\quad \mathit{loop}\ (\mathit{second}\ (\mathit{rightRotate} \ggg \mathit{lazy}) \ggg \mathit{arr}\ (\mathit{uncurry}\ \mathit{zip}) \ggg\\
&\quad \mathit{parMap}\ \mathit{conf}\ (\mathit{second}\ \mathit{get} \ggg f \ggg \mathit{second}\ \mathit{put}) \ggg \mathit{arr}\ \mathit{unzip})
\end{aligned}
$$

Figure 25: Final definition of the *ring* skeleton.

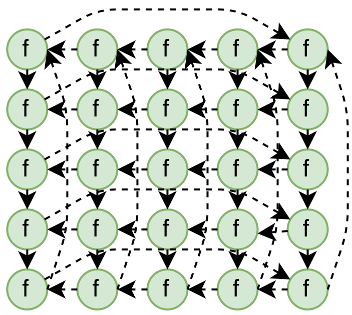


Figure 26: Schematic depiction of the *torus* skeleton.

7.3 Torus skeleton

If we take the concept of a *ring* from Section 7.2 one dimension further, we obtain a *torus* skeleton (Fig. 26, 27). Every node sends ands receives data from horizontal and vertical neighbours in each communication round. With our Parallel Arrows we re-implement the *torus* combinator[9] from Eden—yet again with the help of the *ArrowLoop* typeclass.

Similar to the *ring*, we once again start by rotating the input (Fig. B 7), but this time not only in one direction, but in two. This means that the intermediary input from the neighbour nodes has to be stored in a tuple $([[\mathit{fut}\ a]], [[\mathit{fut}\ b]])$ in the second argument (loop only allows for two arguments) of our looped arrow $\mathit{arr}\ ([[c]], ([[\mathit{fut}\ a]], [[\mathit{fut}\ b]]))\ ([[d]], ([[\mathit{fut}\ a]], [[\mathit{fut}\ b]]))$ and our rotation arrow becomes

$$\mathit{second}\ ((\mathit{mapArr}\ \mathit{rightRotate} \ggg \mathit{lazy}) *** (\mathit{arr}\ \mathit{rightRotate} \ggg \mathit{lazy}))$$

instead of the singular rotation in the ring as we rotate $[[\mathit{fut}\ a]]$ horizontally and $[[\mathit{fut}\ b]]$ vertically. Then, we once again zip the inputs for the input arrow with

$$\mathit{arr}\ (\mathit{uncurry3}\ \mathit{zipWith3}\ \mathit{lazyzip3})$$

from $([[c]], ([[\mathit{fut}\ a]], [[\mathit{fut}\ b]]))$ to $[[(c, \mathit{fut}\ a, \mathit{fut}\ b)]]$, which we then feed into our parallel execution.

This action is, however, more complicated than in the ring case as we have one more dimension of inputs to be transformed. We first have to *shuffle* all the inputs to then pass it into $\mathit{parMap}\ \mathit{conf}\ (\mathit{ptorus}\ f)$ which yields $[(d, \mathit{fut}\ a, \mathit{fut}\ b)]$. We can then unpack this shuffled

[9] Available on Hackage: `https://hackage.haskell.org/package/edenskel-2.1.0.0/docs/Control-Parallel-Eden-Topology.html`.

$$
\begin{aligned}
&\mathit{torus} :: (\mathit{ArrowLoop\ arr}, \mathit{ArrowChoice\ arr}, \mathit{ArrowApply\ arr}, \mathit{Future\ fut\ a}, \mathit{Future\ fut\ b},\\
&\quad \mathit{ArrowParallel\ arr}\ (c, \mathit{fut\ a}, \mathit{fut\ b})\ (d, \mathit{fut\ a}, \mathit{fut\ b})\ \mathit{conf}) \Rightarrow\\
&\quad \mathit{conf} \to \mathit{arr}\ (c, a, b)\ (d, a, b) \to \mathit{arr}\ [[c]]\ [[d]]\\
&\mathit{torus\ conf\ f} =\\
&\quad \mathit{loop}\ (\mathit{second}\ ((\mathit{mapArr\ rightRotate} \ggg \mathit{lazy}) *\!*\!* (\mathit{arr\ rightRotate} \ggg \mathit{lazy})) \ggg\\
&\quad \mathit{arr}\ (\mathit{uncurry3}\ (\mathit{zipWith3\ lazyzip3})) \ggg\\
&\quad (\mathit{arr\ length} \ggg \mathit{arr\ unshuffle})\ \&\&\&\ (\mathit{shuffle} \ggg \mathit{parMap\ conf}\ (\mathit{ptorus\ f})) \ggg \mathit{app} \ggg\\
&\quad \mathit{arr}\ (\mathit{map\ unzip3}) \ggg \mathit{arr\ unzip3} \ggg \mathit{threetotwo})\\
&\mathit{ptorus} :: (\mathit{Arrow\ arr}, \mathit{Future\ fut\ a}, \mathit{Future\ fut\ b}) \Rightarrow\\
&\quad \mathit{arr}\ (c, a, b)\ (d, a, b) \to \mathit{arr}\ (c, \mathit{fut\ a}, \mathit{fut\ b})\ (d, \mathit{fut\ a}, \mathit{fut\ b})\\
&\mathit{ptorus\ f} = \mathit{arr}\ (\lambda{\sim}(c, a, b) \to (c, \mathit{get\ a}, \mathit{get\ b})) \ggg f \ggg \mathit{arr}\ (\lambda{\sim}(d, a, b) \to (d, \mathit{put\ a}, \mathit{put\ b}))
\end{aligned}
$$

Figure 27: Definition of the *torus* skeleton. The definitions of *lazyzip3*, *uncurry3* and *threetotwo* have been omitted and can be found in Fig. B 8

.

$$
\begin{aligned}
&\mathit{nodefunction} :: \mathit{Int} \to ((\mathit{Matrix}, \mathit{Matrix}), [\mathit{Matrix}], [\mathit{Matrix}]) \to ([\mathit{Matrix}], [\mathit{Matrix}], [\mathit{Matrix}])\\
&\mathit{nodefunction\ n}\ ((\mathit{bA}, \mathit{bB}), \mathit{rows}, \mathit{cols}) = ([\mathit{bSum}], \mathit{bA} : \mathit{nextAs}, \mathit{bB} : \mathit{nextBs})\\
&\quad \mathbf{where}\ \mathit{bSum} = \mathit{foldl}'\ \mathit{matAdd}\ (\mathit{matMult\ bA\ bB})\ (\mathit{zipWith\ matMult\ nextAs\ nextBs})\\
&\qquad \mathit{nextAs} = \mathit{take}\ (n-1)\ \mathit{rows}\\
&\qquad \mathit{nextBs} = \mathit{take}\ (n-1)\ \mathit{cols}
\end{aligned}
$$

Figure 28: Adapted *nodefunction* for matrix multiplication with the *torus* from Fig. 27.

list back to its original ordering by feeding it into the specific unshuffle arrow we created one step earlier with $\mathit{arr\ length} \ggg \mathit{arr\ unshuffle}$ with the use of $\mathit{app} :: \mathit{arr}\ (\mathit{arr\ a\ b}, a)\ c$ from the *ArrowApply* typeclass. Finally, we unpack this matrix $[[[(d, \mathit{fut\ a}, \mathit{fut\ b})]]$ with $\mathit{arr}\ (\mathit{map\ unzip3}) \ggg \mathit{arr\ unzip3} \ggg \mathit{threetotwo}$ to get $([[d]], ([[\mathit{fut\ a}]], [[\mathit{fut\ b}]]))$.

As an example of using this skeleton (Loogen, 2012) showed the matrix multiplication using the Gentleman algorithm (1978). Their instantiation of the skeleton *nodefunction* can be adapted as shown in Fig. 28. If we compare the trace from a call using our arrow definition of the torus (Fig. 29) with the Eden version (Fig. 30) we can see that the behaviour of the arrow version and execution times are comparable. We discuss further examples on larger clusters and in a more detail in the next section.

Again, the implementation of the Par Monad version prevents this definition of the *torus* skeleton to work due to the reasons explained in the section about the *ring* skeleton in section 7.2.

OL: more nodes!!!

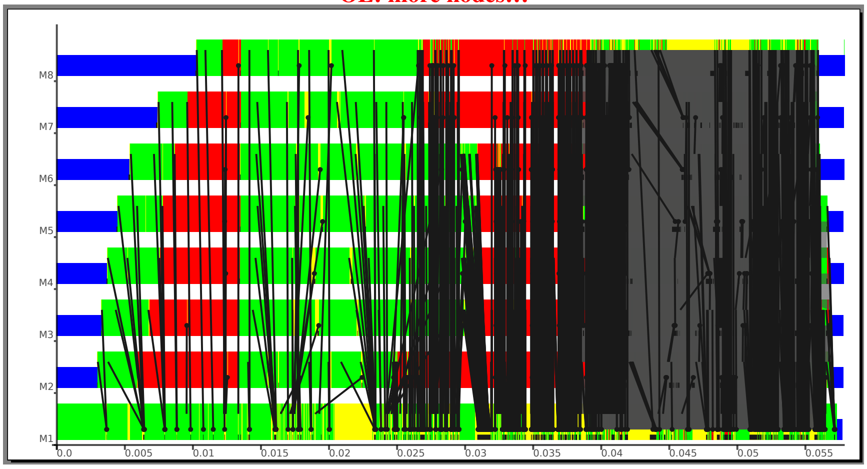


Figure 29: Matrix Multiplication with *torus* (PArrows).

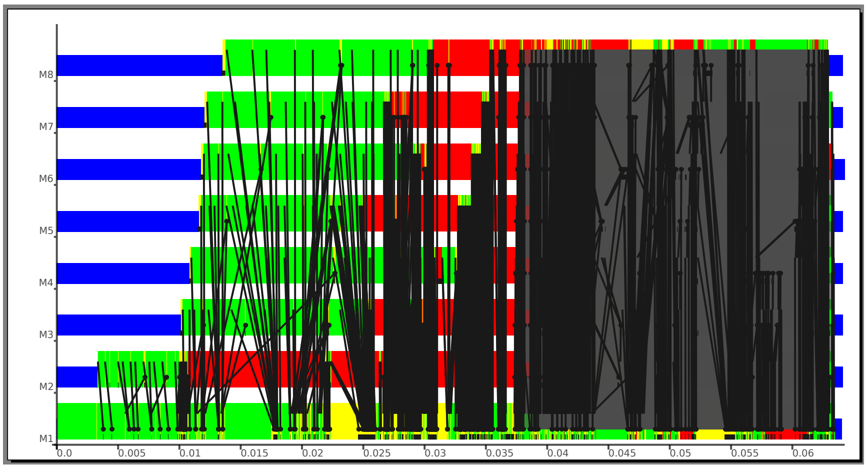
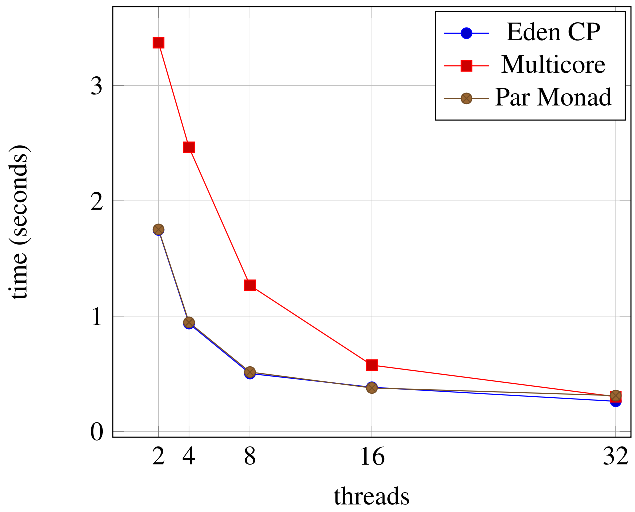


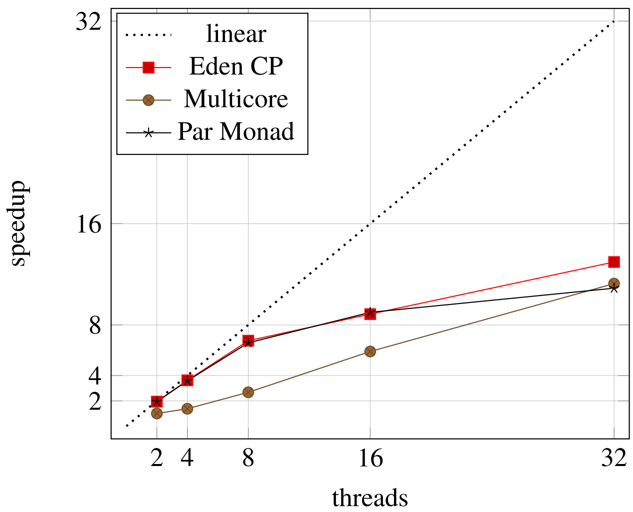
Figure 30: Matrix Multiplication with *torus* (Eden).

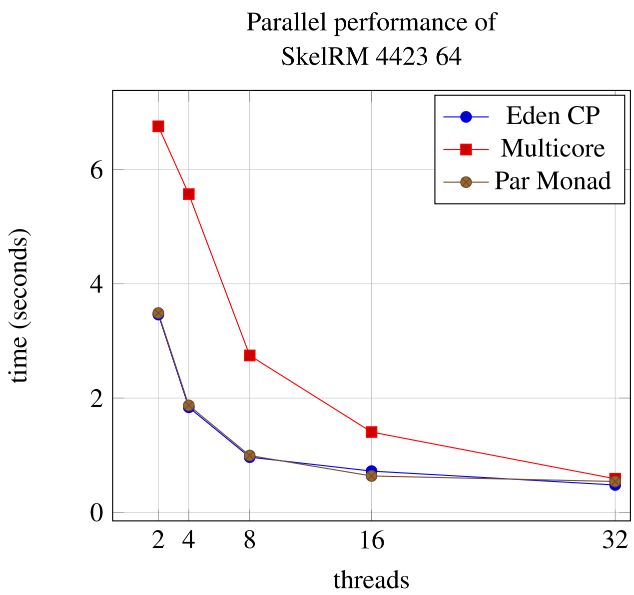
8 Benchmarks

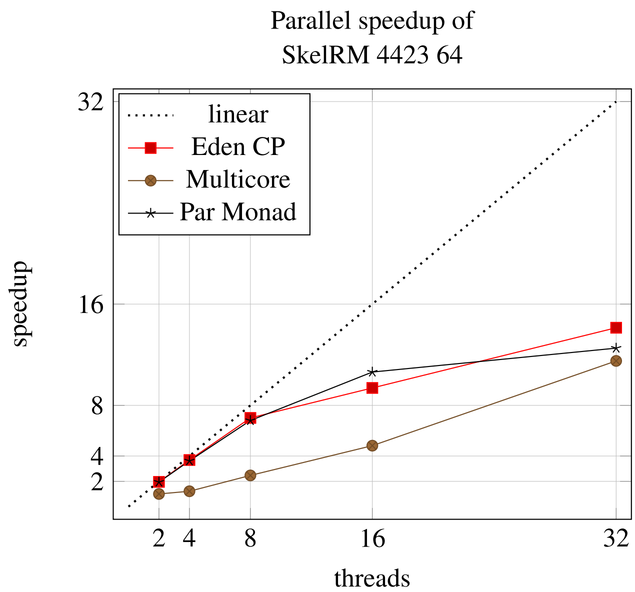
Parallel performance of
SkelRM 4423 32



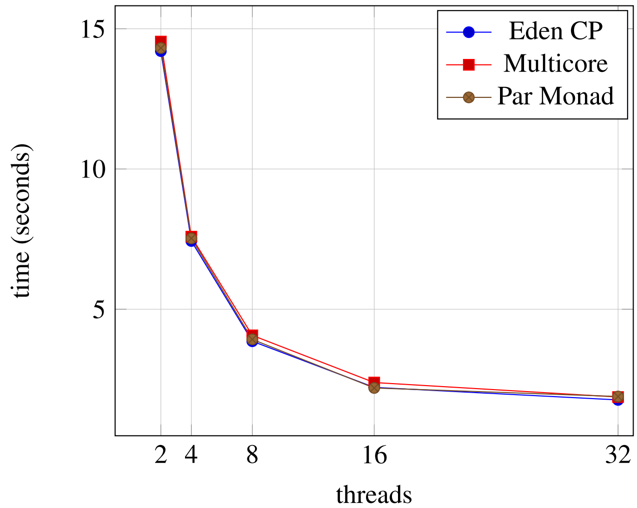
Parallel speedup of
SkelRM 4423 32



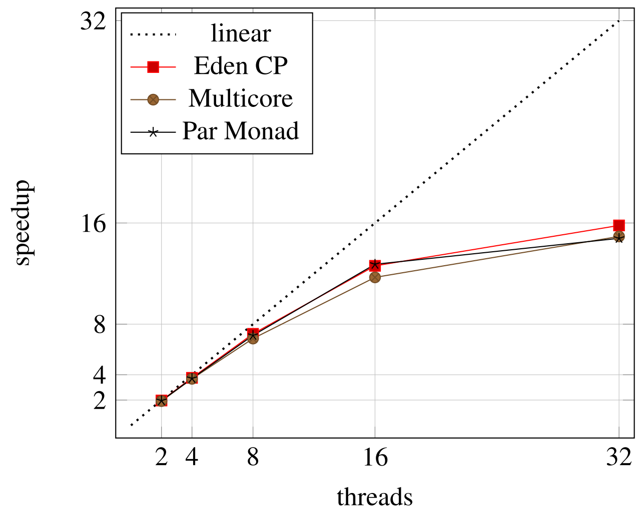

Parallel performance of
SkelRM 4423 64
time (seconds)
threads
Eden CP
Multicore
Par Monad
0
2
4
6
2
4
8
16
32


Parallel speedup of
SkelRM 4423 64
speedup
threads
linear
Eden CP
Multicore
Par Monad
2
4
8
16
32
2
4
8
16
32

Parallel performance of
SkelRM 9941 32

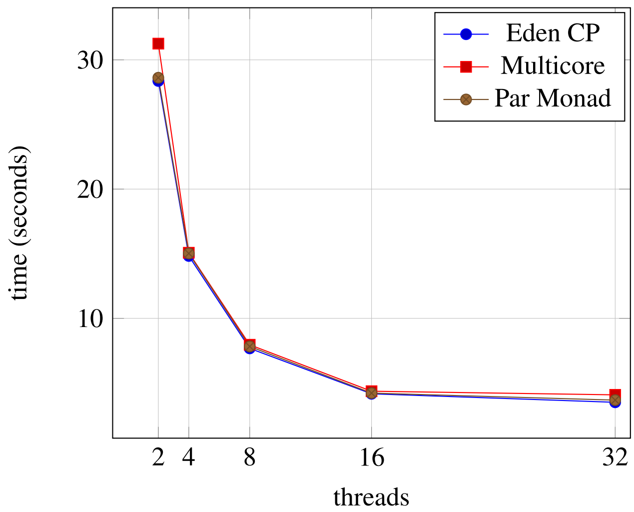


Parallel speedup of
SkelRM 9941 32

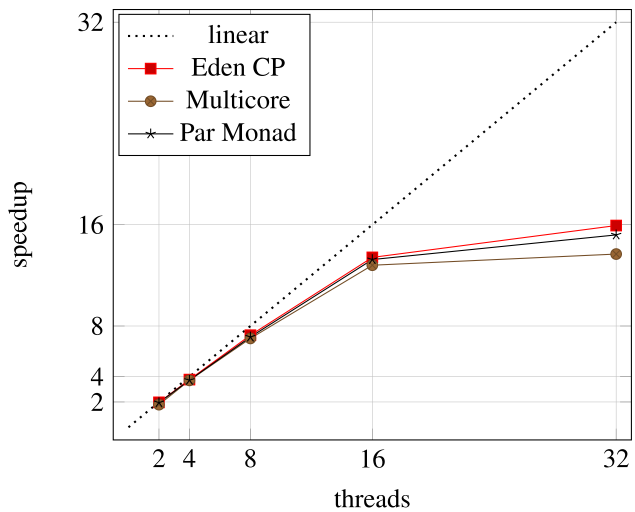


32

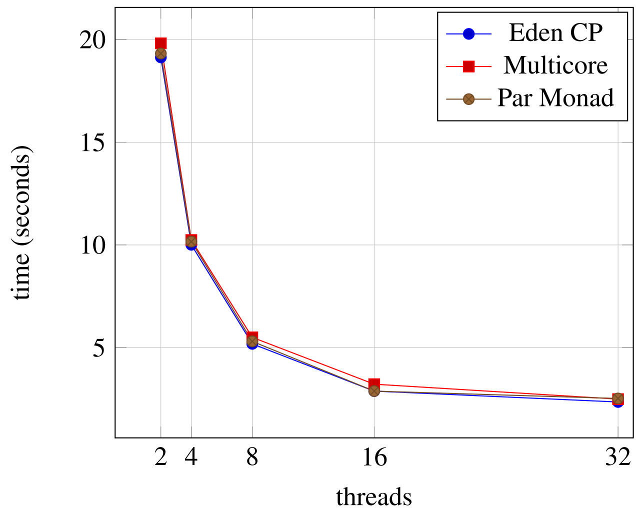
Parallel performance of
SkelRM 9941 64



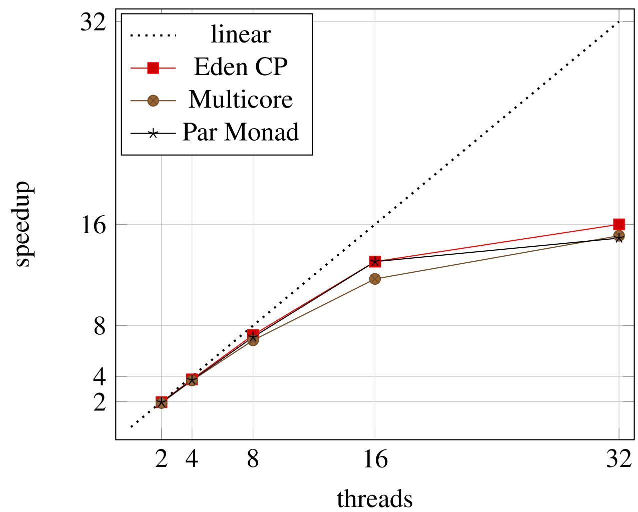
Parallel speedup of
SkelRM 9941 64



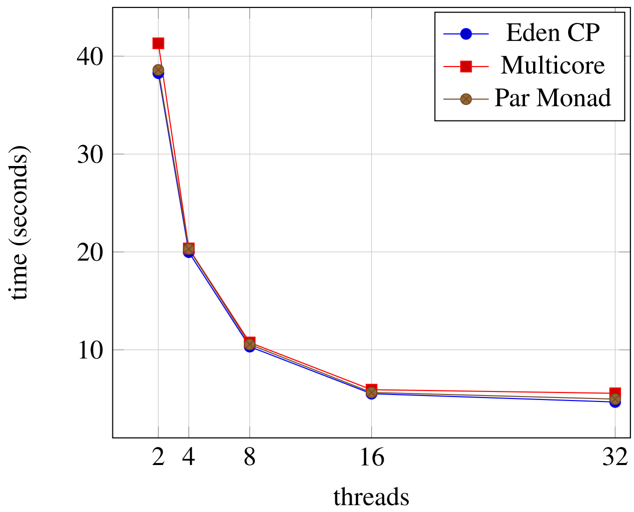
Parallel performance of
SkelRM 11213 32



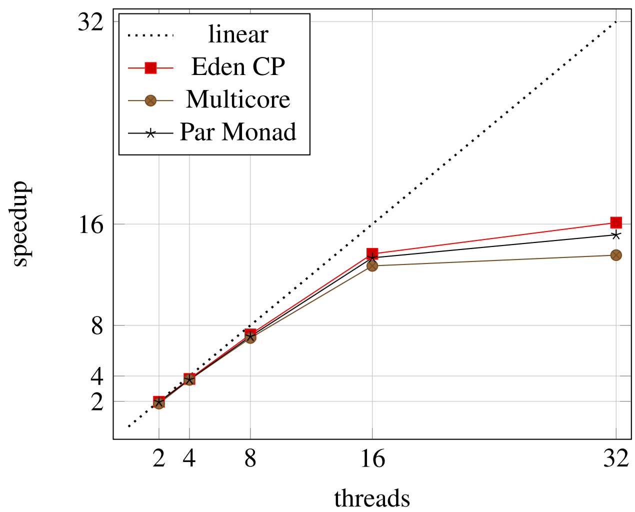
Parallel speedup of
SkelRM 11213 32

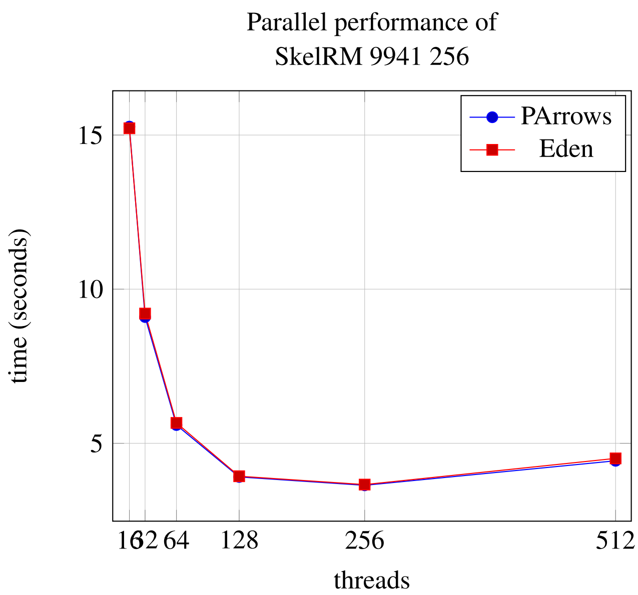


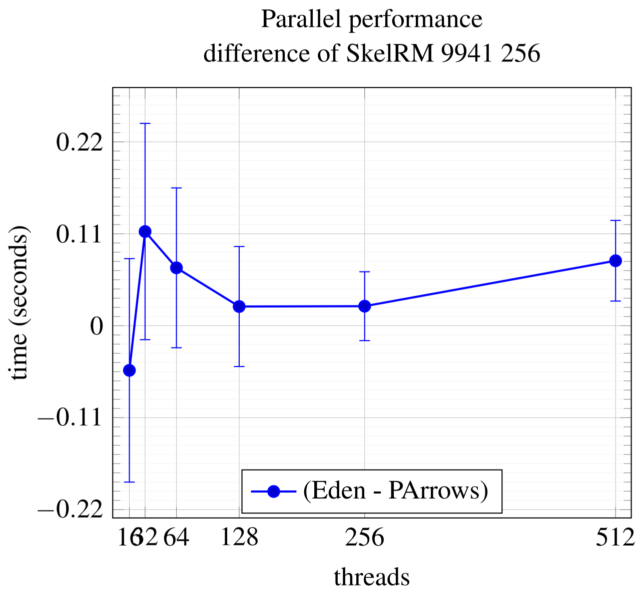
Parallel performance of
SkelRM 11213 64

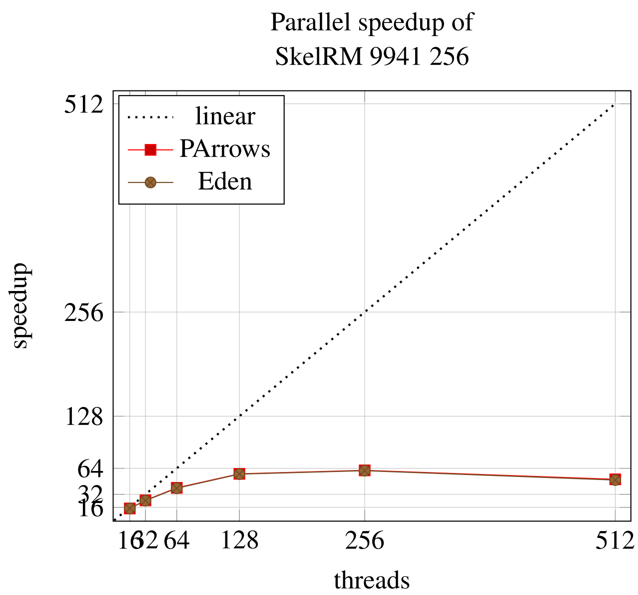


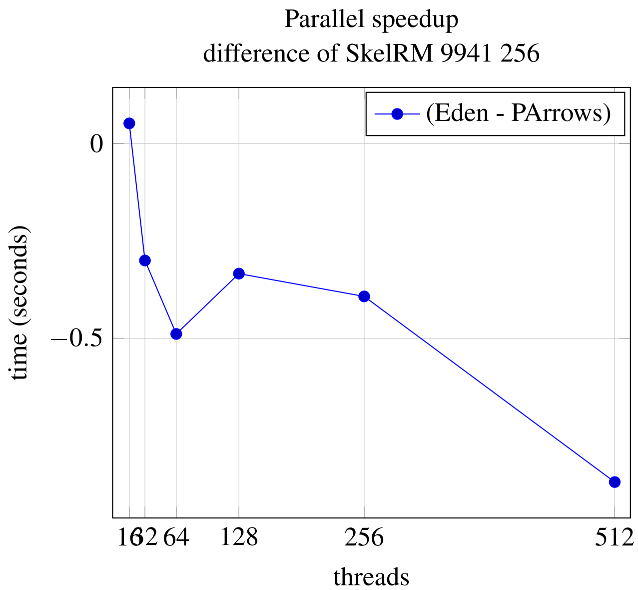
Parallel speedup of
SkelRM 11213 64

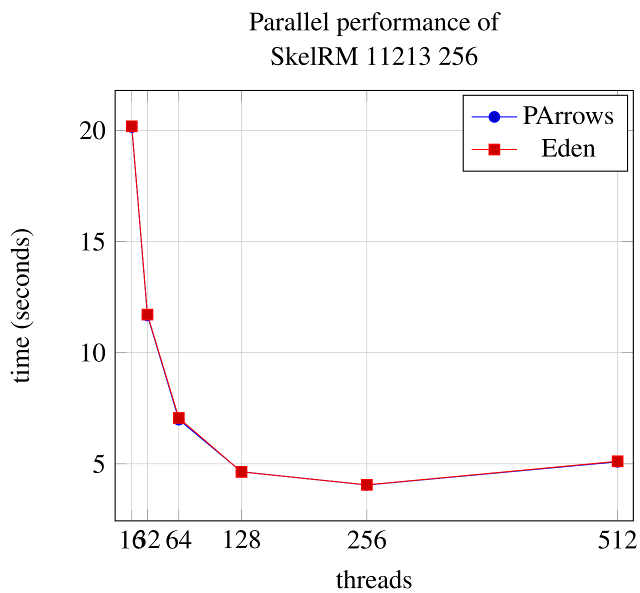


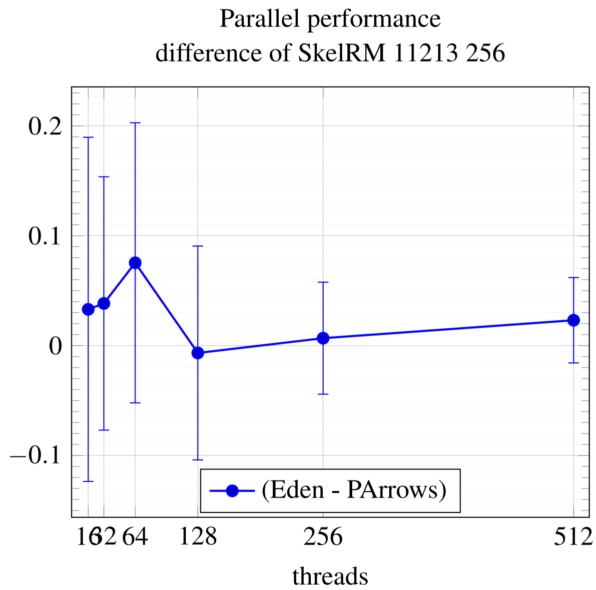

Parallel performance of
SkelRM 9941 256
PArrows
Eden
time (seconds)
15
10
5
16
32
64
128
256
512
threads

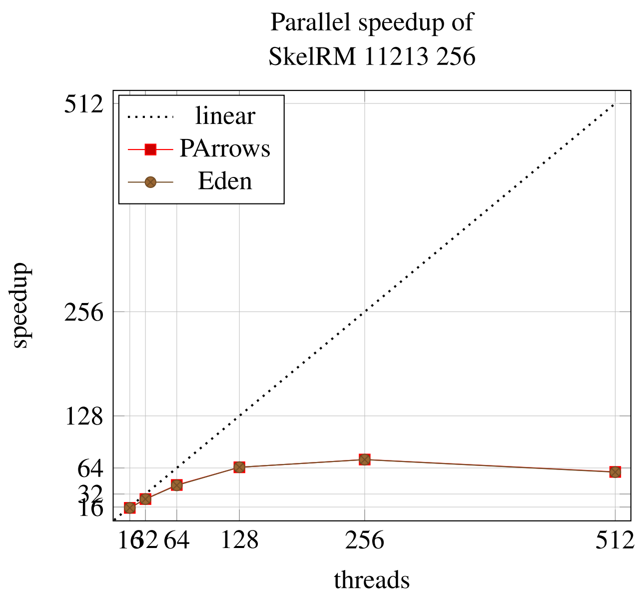

Parallel performance
difference of SkelRM 9941 256
(Eden - PArrows)
time (seconds)
0.22
0.11
0
−0.11
−0.22
16
32
64
128
256
512
threads

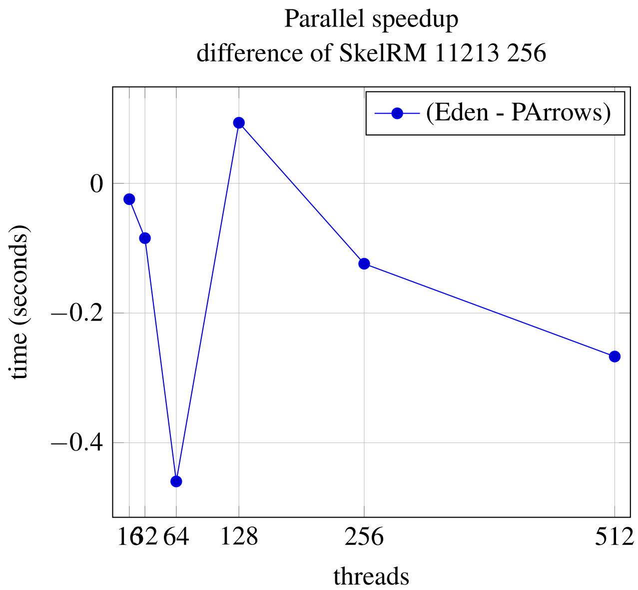

Parallel speedup of
SkelRM 9941 256
linear
PArrows
Eden
speedup
512
256
128
64
32
16
16
32
64
128
256
512
threads

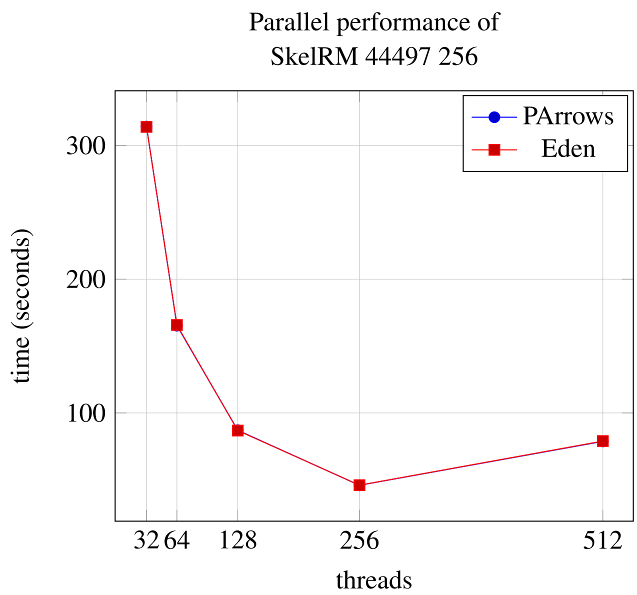

Parallel speedup
difference of SkelRM 9941 256
(Eden - PArrows)
time (seconds)
0
−0.5
16
32
64
128
256
512
threads

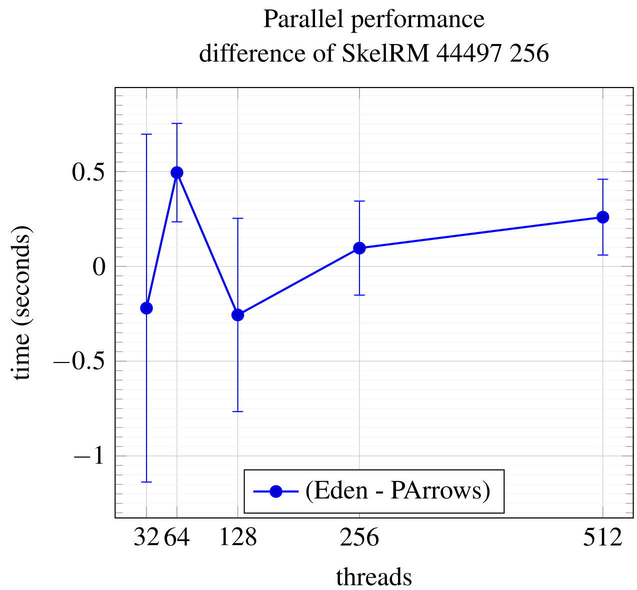

Parallel performance of
SkelRM 11213 256
PArrows
Eden
time (seconds)
20
15
10
5
16
32
64
128
256
512
threads


Parallel performance
difference of SkelRM 11213 256
(Eden - PArrows)
time (seconds)
0.2
0.1
0
−0.1
16
32
64
128
256
512
threads

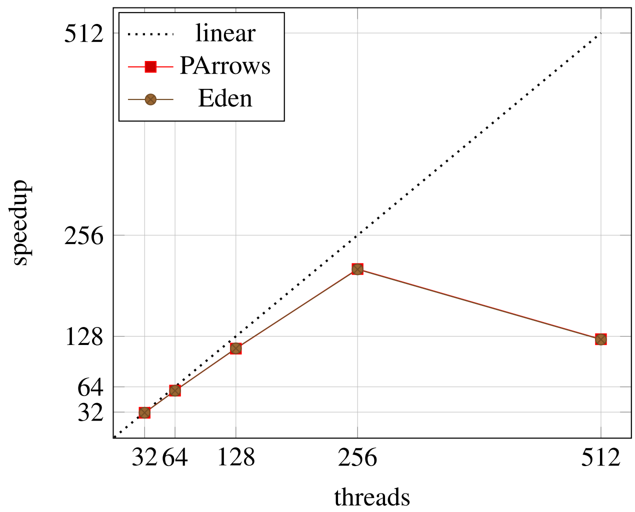

Parallel speedup of
SkelRM 11213 256
linear
PArrows
Eden
speedup
512
256
128
64
32
16
16
32
64
128
256
512
threads


Parallel speedup
difference of SkelRM 11213 256
(Eden - PArrows)
time (seconds)
0
−0.2
−0.4
16
32
64
128
256
512
threads

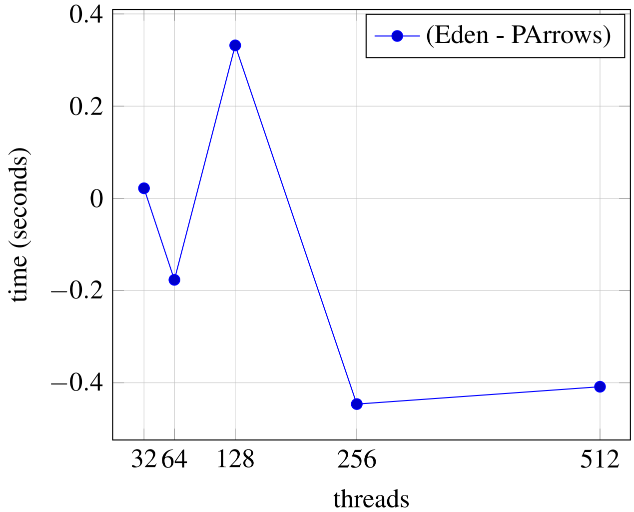

Parallel performance of
SkelRM 44497 256
PArrows
Eden
time (seconds)
300
200
100
32
64
128
256
512
threads


Parallel performance
difference of SkelRM 44497 256
time (seconds)
0.5
0
−0.5
−1
(Eden - PArrows)
32
64
128
256
512
threads

Parallel speedup of
SkelRM 44497 256



Parallel speedup
difference of SkelRM 44497 256



9 Conclusion

Arrows are a generic concept that allows for powerful composition combinators. To our knowledge we are the first ones to represent parallel computation with arrows.**OL: that strange arrows-based robot interaction paper from 1993 or so! clearly discuss in related work done!**

Arrows turn out to be a useful tool for composing in parallel programs. We do not have to introduce new monadic types that wrap the computation. Instead use arrows just like regular sequential pure functions. This work features multiple parallel backends: the already available parallel Haskell flavours. Parallel Arrows feature an implementation of the *ArrowParallel* interface for Multicore Haskell, *Par* Monad, and Eden. With our approach parallel programs can be ported across these flavours with little to no effort. Performance-wise, Parallel Arrows are on par with existing parallel Haskells, as they do not introduce any notable overhead.

The strictness problems of the *Par* Monad backend materializes with *ArrowLoop*. Additional work is required to ensure correct behaviour of the *Par* Monad in this context.

MB: mention ArrowLoop in Torus and Ring chapters OL: Parrows + accelerate = love? Metion port to Frege.

9.1 Future Work

Our PArrows DSL can be expanded to futher parallel Haskells. More specifically we target HdpH (Maier *et al.*, 2014), a modern distributed Haskell that would benefit from our Arrows notation. More Future-aware versions of Arrow combinators can be defined and existing can be further improved. We would look into more transparency of the DSL, it should basically infuse as little overhead as possible.

We are looking into more experiences with seamless porting of parallel PArrow-based programs across the backends. Of course, we are working on expanding both our skeleton library and the number of skeleton-based parallel programs that use our DSL to be portable across flavours of parallel Haskells. It would also be interesting to see a hybrid of PArrows and Accelerate. Ports of our approach to other languages like Frege or Java directly are in an early development stage.

References

Achten, Peter, van Eekelen, Marko, de Mol, Maarten, & Plasmeijer, Rinus. (2007). An arrow based semantics for interactive applications. *Draft Proceedings of the Symposium on Trends in Functional Programming*. TFP '07.

Achten, PM, van Eekelen, Marko CJD, Plasmeijer, MJ, & Weelden, A van. (2004). *Arrows for generic graphical editor components*. Tech. rept. Nijmegen Institute for Computing and Information Sciences, Faculty of Science, University of Nijmegen, The Netherlands. Technical Report NIII-R0416.

Alimarine, Artem, Smetsers, Sjaak, van Weelden, Arjen, van Eekelen, Marko, & Plasmeijer, Rinus. (2005). There and back again: Arrows for invertible programming. *Pages 86–97 of: Proceedings of the 2005 ACM SIGPLAN Workshop on Haskell*. Haskell '05. ACM.

Aljabri, Malak, Loidl, Hans-Wolfgang, & Trinder, Phil W. (2014). The design and implementation of GUMSMP: A multilevel parallel haskell implementation. *Pages 37:37–37:48 of: Proceedings of the 25th Symposium on Implementation and Application of Functional Languages*. IFL '13. ACM.

Aljabri, Malak, Loidl, Hans-Wolfgang, & Trinder, Phil. (2015). *Balancing shared and distributed heaps on NUMA architectures*. Springer. Pages 1–17.

Alt, Martin, & Gorlatch, Sergei. (2003). Future-Based RMI: Optimizing compositions of remote method calls on the Grid. *Pages 682–693 of:* Kosch, Harald, Böszörményi, László, & Hellwagner, Hermann (eds), *Euro-Par 2003*. LNCS 2790. Springer-Verlag.

Aswad, Mustafa, Trinder, Phil, Al Zain, Abdallah, Michaelson, Greg, & Berthold, Jost. (2009). Low pain vs no pain multi-core Haskells. *Pages 49–64 of: Trends in Functional Programming*.

Atkey, Robert. (2011). What is a categorical model of arrows? *Electronic notes in theoretical computer science*, **229**(5), 19–37.

Berthold, Jost. (2008). *Explicit and implicit parallel functional programming — concepts and implementation*. Ph.D. thesis, Philipps-Universität Marburg.

Berthold, Jost, & Loogen, Rita. (2006). Skeletons for recursively unfolding process topologies. Joubert, Gerhard R., Nagel, Wolfgang E., Peters, Frans J., Plata, Oscar G., Tirado, P., & Zapata, Emilio L. (eds), *Parallel Computing: Current & Future Issues of High-End Computing, ParCo 2005, Malaga, Spain*. NIC Series 33. Central Institute for Applied Mathematics, Jülich, Germany.

Berthold, Jost, & Loogen, Rita. (2007). Visualizing Parallel Functional Program Executions: Case Studies with the Eden Trace Viewer. *ParCo '07. Parallel Computing: Architectures, Algorithms and Applications*. IOS Press.

Berthold, Jost, Dieterle, Mischa, Lobachev, Oleg, & Loogen, Rita. (2009a). Distributed Memory Programming on Many-Cores – A Case Study Using Eden Divide-&-Conquer Skeletons. *Pages 47–55 of:* Großpitsch, K.-E., Henkersdorf, A., Uhrig, S., Ungerer, T., & Hähner, J. (eds), *Workshop on Many-Cores at ARCS '09 – 22nd International Conference on Architecture of Computing Systems 2009*. VDE-Verlag.

Berthold, Jost, Dieterle, Mischa, & Loogen, Rita. (2009b). Implementing parallel Google map-reduce in Eden. *Pages 990–1002 of:* Sips, Henk, Epema, Dick, & Lin, Hai-Xiang (eds), *Euro-Par 2009 Parallel Processing*. LNCS 5704. Springer Berlin Heidelberg.

Berthold, Jost, Dieterle, Mischa, Lobachev, Oleg, & Loogen, Rita. (2009c). *Parallel FFT with Eden skeletons*. PaCT '09. Springer. Pages 73–83.

Berthold, Jost, Loidl, Hans-Wolfgang, & Hammond, Kevin. (2016). PAEAN: Portable and scalable runtime support for parallel Haskell dialects. *Journal of functional programming*, **26**.

Botorog, G. H., & Kuchen, H. (1996). *Euro-Par'96 Parallel Processing*. LNCS 1123. Springer-Verlag. Chap. Efficient parallel programming with algorithmic skeletons, pages 718–731.

Chakravarty, Manuel M. T., Leshchinskiy, Roman, Peyton Jones, Simon L., Keller, Gabriele, & Marlow, Simon. (2007). Data Parallel Haskell: a status report. *Pages 10–18 of: DAMP '07*. ACM Press.

Chakravarty, Manuel M.T., Keller, Gabriele, Lee, Sean, McDonell, Trevor L., & Grover, Vinod. (2011). Accelerating Haskell array codes with multicore GPUs. *Pages 3–14 of: Proceedings of the 6th Workshop on Declarative Aspects of Multicore Programming*. DAMP '11. ACM.

Clifton-Everest, Robert, McDonell, Trevor L, Chakravarty, Manuel M T, & Keller, Gabriele. (2014). Embedding Foreign Code. *PADL '14: The 16th International Symposium on Practical Aspects of Declarative Languages*. LNCS. Springer-Verlag.

Cole, M. I. (1989). Algorithmic skeletons: Structured management of parallel computation. *Research Monographs in Parallel and Distributed Computing*. Pitman.

Czaplicki, Evan, & Chong, Stephen. (2013). Asynchronous functional reactive programming for guis. *Sigplan not.*, **48**(6), 411–422.

Dagand, Pierre-Évariste, Kostić, Dejan, & Kuncak, Viktor. (2009). Opis: Reliable distributed systems in OCaml. *Pages 65–78 of: Proceedings of the 4th International Workshop on Types in Language Design and Implementation*. TLDI '09. ACM.

Danelutto, M., Pasqualetti, F., & Pelagatti, S. (1997). Skeletons for Data Parallelism in P^3L. *Pages 619–628 of:* Lengauer, C., Griebl, M., & Gorlatch, S. (eds), *Euro-Par'97*. LNCS 1300. Springer-Verlag.

Darlington, J., Field, AJ, Harrison, PG, Kelly, PHJ, Sharp, DWN, Wu, Q., & While, RL. (1993). Parallel programming using skeleton functions. *Pages 146–160 of: Parallel architectures and languages Europe*. Springer-Verlag.

de la Encina, Alberto, Hidalgo-Herrero, Mercedes, Rabanal, Pablo, & Rubio, Fernando. (2011). *A parallel skeleton for genetic algorithms*. IWANN '11. Springer. Pages 388–395.

Dieterle, M., Horstmeyer, T., & Loogen, R. (2010a). Skeleton composition using remote data. *Pages 73–87 of:* Carro, M., & Peña, R. (eds), *12th International Symposium on Practical Aspects of Declarative Languages*. PADL '10, vol. 5937. Springer-Verlag.

Dieterle, M., Horstmeyer, T., Loogen, R., & Berthold, J. (2016). Skeleton composition versus stable process systems in Eden. *Journal of functional programming*, **26**.

Dieterle, Mischa, Berthold, Jost, & Loogen, Rita. (2010b). *A skeleton for distributed work pools in Eden*. FLOPS '10. Springer. Pages 337–353.

Dieterle, Mischa, Horstmeyer, Thomas, Berthold, Jost, & Loogen, Rita. (2013). *Iterating skeletons*. IFL '12. Springer. Pages 18–36.

Foltzer, Adam, Kulkarni, Abhishek, Swords, Rebecca, Sasidharan, Sajith, Jiang, Eric, & Newton, Ryan. (2012). A meta-scheduler for the Par-monad: Composable scheduling for the heterogeneous cloud. *Sigplan not.*, **47**(9), 235–246.

Geimer, M., Wolf, F., Wylie, B. J. N., Ábrahám, E., Becker, D., & Mohr, B. (2010). The Scalasca performance toolset architecture. *Concurrency and computation: Practice and experience*, **22**(6).

Gentleman, W. Morven. (1978). Some complexity results for matrix computations on parallel processors. *Journal of the acm*, **25**(1), 112–115.

Gorlatch, Sergei. (1998). Programming with divide-and-conquer skeletons: A case study of FFT. *Journal of supercomputing*, **12**(1-2), 85–97.

Gorlatch, Sergei, & Bischof, Holger. (1998). A generic MPI implementation for a data-parallel skeleton: Formal derivation and application to FFT. *Parallel processing letters*, **8**(4).

Hammond, Kevin, Berthold, Jost, & Loogen, Rita. (2003). Automatic skeletons in Template Haskell. *Parallel processing letters*, **13**(03), 413–424.

Harris, Tim, Marlow, Simon, Peyton Jones, Simon, & Herlihy, Maurice. (2005). Composable memory transactions. *Pages 48–60 of: Proceedings of the 10th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming*. PPoPP '05. ACM.

Horstmeyer, Thomas, & Loogen, Rita. (2013). Graph-based communication in Eden. *Higher-order and symbolic computation*, **26**(1), 3–28.

Huang, Liwen, Hudak, Paul, & Peterson, John. (2007). *HPorter: Using arrows to compose parallel processes*. Berlin, Heidelberg: Springer Berlin Heidelberg. Pages 275–289.

Hudak, Paul, Courtney, Antony, Nilsson, Henrik, & Peterson, John. (2003). *Arrows, robots, and functional reactive programming*. Springer. Pages 159–187.

Hughes, John. (2000). Generalising monads to arrows. *Science of computer programming*, **37**(1–3), 67–111.

Hughes, John. (2005a). *Programming with arrows*. AFP '04. Springer. Pages 73–129.

Hughes, John. (2005b). *Programming with arrows*. Berlin, Heidelberg: Springer Berlin Heidelberg. Pages 73–129.

Jacobs, Bart, Heunen, Chris, & Hasuo, Ichiro. (2009). Categorical semantics for arrows. *Journal of functional programming*, **19**(3-4), 403–438.

Janjic, Vladimir, Brown, Christopher Mark, Neunhoeffer, Max, Hammond, Kevin, Linton, Stephen Alexander, & Loidl, Hans-Wolfgang. (2013). Space exploration using parallel orbits: a study in parallel symbolic computing. *Parallel computing*.

Jones, Simon Peyton, Gordon, Andrew, & Finne, Sigbjorn. (1996). Concurrent haskell. *Pages 295–308 of: POPL*, vol. 96.

Keller, Gabriele, Chakravarty, Manuel M.T., Leshchinskiy, Roman, Peyton Jones, Simon, & Lippmeier, Ben. (2010). Regular, shape-polymorphic, parallel arrays in haskell. *Sigplan not.*, **45**(9), 261–272.

Kuper, Lindsey, Todd, Aaron, Tobin-Hochstadt, Sam, & Newton, Ryan R. (2014). Taming the parallel effect zoo: Extensible deterministic parallelism with LVish. *Sigplan not.*, **49**(6), 2–14.

Lengauer, Christian, Gorlatch, Sergei, & Herrmann, Christoph. (1997). The static parallelization of loops and recursions. *The journal of supercomputing*, **11**(4), 333–353.

Li, Peng, & Zdancewic, S. (2006). Encoding information flow in Haskell. *Pages 12–16 of: 19th IEEE Computer Security Foundations Workshop*. CSFW '06.

Li, Peng, & Zdancewic, Steve. (2010). Arrows for secure information flow. *Theoretical computer science*, **411**(19), 1974–1994.

Lindley, Sam, Wadler, Philip, & Yallop, Jeremy. (2011). Idioms are oblivious, arrows are meticulous, monads are promiscuous. *Electronic notes in theoretical computer science*, **229**(5), 97–117.

Linton, S., Hammond, K., Konovalov, A., Al Zain, A. D., Trinder, P., Horn, P., & Roozemond, D. (2010). Easy composition of symbolic computation software: a new lingua franca for symbolic computation. *Pages 339–346 of: Proceedings of the 2010 International Symposium on Symbolic and Algebraic Computation*. ISSAC '10. ACM Press.

Liu, Hai, Cheng, Eric, & Hudak, Paul. (2009). Causal commutative arrows and their optimization. *Sigplan not.*, **44**(9), 35–46.

Lobachev, Oleg. (2011). *Implementation and evaluation of algorithmic skeletons: Parallelisation of computer algebra algorithms*. Ph.D. thesis, Philipps-Universität Marburg.

Lobachev, Oleg. (2012). *Parallel computation skeletons with premature termination property*. FLOPS 2012. Springer. Pages 197–212.

Loogen, Rita. (2012). *Eden – parallel functional programming with Haskell*. Springer. Pages 142–206.

Loogen, Rita, Ortega-Mallén, Yolanda, & Peña-Marí, Ricardo. (2005). Parallel Functional Programming in Eden. *Journal of Functional Programming*, **15**(3), 431–475. Special Issue on Functional Approaches to High-Performance Parallel Programming.

Maier, Patrick, Stewart, Robert, & Trinder, Phil. (2014). The HdpH DSLs for scalable reliable computation. *Sigplan not.*, **49**(12), 65–76.

Mainland, Geoffrey, & Morrisett, Greg. (2010). Nikola: Embedding compiled GPU functions in Haskell. *Sigplan not.*, **45**(11), 67–78.

Marlow, S., Peyton Jones, S., & Singh, S. (2009). Runtime support for multicore Haskell. *Acm sigplan notices*, **44**(9), 65–78.

Marlow, Simon. (2013). *Parallel and concurrent programming in Haskell: Techniques for multicore and multithreaded programming*. "O'Reilly Media, Inc.".

Marlow, Simon, Newton, Ryan, & Peyton Jones, Simon. (2011). A monad for deterministic parallelism. *Sigplan not.*, **46**(12), 71–82.

McDonell, Trevor L., Chakravarty, Manuel M. T., Grover, Vinod, & Newton, Ryan R. (2015). Type-safe runtime code generation: Accelerate to LLVM. *Sigplan not.*, **50**(12), 201–212.

Nikhil, R., & Arvind, L. A. (2001). *Implicit parallel programming in pH*. Morgan Kaufmann.

Nilsson, Henrik, Courtney, Antony, & Peterson, John. (2002). Functional reactive programming, continued. *Pages 51–64 of: Proceedings of the 2002 ACM SIGPLAN Workshop on Haskell*. Haskell '02. New York, NY, USA: ACM.

Paterson, Ross. (2001). A new notation for arrows. *Sigplan not.*, **36**(10), 229–240.

Perfumo, Cristian, Sönmez, Nehir, Stipic, Srdjan, Unsal, Osman, Cristal, Adrián, Harris, Tim, & Valero, Mateo. (2008). The limits of software transactional memory (STM): Dissecting Haskell STM applications on a many-core environment. *Pages 67–78 of: Proceedings of the 5th Conference on Computing Frontiers*. CF '08. ACM.

Rabhi, F. A., & Gorlatch, S. (eds). (2003). *Patterns and Skeletons for Parallel and Distributed Computing*. Springer-Verlag.

Russo, Alejandro, Claessen, Koen, & Hughes, John. (2008). A library for light-weight information-flow security in Haskell. *Pages 13–24 of: Proceedings of the 1st ACM SIGPLAN Symposium on Haskell*. Haskell '08. ACM.

Stewart, Robert, Maier, Patrick, & Trinder, Phil. (2016). Transparent fault tolerance for scalable functional computation. *Journal of functional programming*, **26**.

Svensson, Joel. (2011). *Obsidian: Gpu kernel programming in haskell*. Ph.D. thesis, Chalmers University of Technology.

Trinder, Phil W., Hammond, Kevin, Mattson Jr., James S., Partridge, Andrew S., & Peyton Jones, Simon L. (1996). GUM: a Portable Parallel Implementation of Haskell. *PLDI'96*. ACM Press.

Trinder, P.W., Hammond, K., Loidl, H-W., & Peyton Jones, S. (1998). Algorithm + Strategy = Parallelism. *Journal of functional programming*, **8**(1), 23–60.

Vizzotto, Juliana, Altenkirch, Thorsten, & Sabry, Amr. (2006). Structuring quantum effects: superoperators as arrows. *Mathematical structures in computer science*, **16**(3), 453–468.

Wheeler, K. B., & Thain, D. (2009). Visualizing massively multithreaded applications with ThreadScope. *Concurrency and computation: Practice and experience*, **22**(1), 45–67.

A Utility Arrows

MB: Text außenrum bauen, siehe Omitted Function Definitions The *second* combinator:

$$
\begin{aligned}
&\mathit{second} :: \mathit{Arrow}\ \mathit{arr} \Rightarrow \mathit{arr}\ a\ b \rightarrow \mathit{arr}\ (c,a)\ (c,b)\\
&\mathit{second}\ f = \mathit{arr}\ \mathit{swap} \ggg \mathit{first}\ f \ggg \mathit{arr}\ \mathit{swap}\\
&\quad \mathbf{where}\ \mathit{swap}\ (x,y) = (y,x)
\end{aligned}
$$

is a mirrored version of *first*, used, e.g. in definition of $***$.

Utility combinators for Parallel Arrows follow. We define *map*, *foldl* and *zipWith* on Arrows. The *mapArr* combinator (Fig. A 1) lifts any arrow *arr a b* to an arrow *arr* $[a]$ $[b]$ (Hughes, 2005b). Similarly, we can also define *foldlArr* (Fig. A 2) that lifts any arrow *arr* (b,a) b with a neutral element b to *arr* $[a]$ b.

$$
\begin{aligned}
&\mathit{mapArr} :: \mathit{ArrowChoice}\ \mathit{arr} \Rightarrow \mathit{arr}\ a\ b \rightarrow \mathit{arr}\ [a]\ [b]\\
&\mathit{mapArr}\ f =\\
&\quad \mathit{arr}\ \mathit{listcase} \ggg\\
&\quad \mathit{arr}\ (\mathit{const}\ [\,])\ |||\ (f *** \mathit{mapArr}\ f \ggg \mathit{arr}\ (\mathit{uncurry}\ (:)))\\
&\mathit{listcase}\ [\,] = \mathit{Left}\ ()\\
&\mathit{listcase}\ (x:xs) = \mathit{Right}\ (x,xs)
\end{aligned}
$$

Figure A 1: The definition of *map* over Arrows and the *listcase* helper function.

$$
\begin{aligned}
&\mathit{foldlArr} :: (\mathit{ArrowChoice}\ \mathit{arr}, \mathit{ArrowApply}\ \mathit{arr}) \Rightarrow \mathit{arr}\ (b,a)\ b \rightarrow b \rightarrow \mathit{arr}\ [a]\ b\\
&\mathit{foldlArr}\ f\ b =\\
&\quad \mathit{arr}\ \mathit{listcase} \ggg\\
&\quad \mathit{arr}\ (\mathit{const}\ b)\ |||\\
&\qquad (\mathit{first}\ (\mathit{arr}\ (\lambda a \rightarrow (b,a)) \ggg f \ggg \mathit{arr}\ (\mathit{foldlArr}\ f)) \ggg \mathit{app})
\end{aligned}
$$

Figure A 2: The definition of *foldl* over Arrows.

Finally, with the help of *mapArr* (Fig. A 1), we can define *zipWithArr* (Fig. A 3) that lifts any arrow *arr* (a,b) c to an arrow *arr* $([a],[b])$ $[c]$.

$$
\begin{aligned}
&\mathit{zipWithArr} :: \mathit{ArrowChoice}\ \mathit{arr} \Rightarrow \mathit{arr}\ (a,b)\ c \rightarrow \mathit{arr}\ ([a],[b])\ [c]\\
&\mathit{zipWithArr}\ f = (\mathit{arr}\ \$\ \lambda(as,bs) \rightarrow \mathit{zipWith}\ (,)\ \mathit{as}\ \mathit{bs}) \ggg \mathit{mapArr}\ f
\end{aligned}
$$

Figure A 3: *zipWith* over arrows.

These combinators make use of the *ArrowChoice* type class which provides the $|||$ combinator. It takes two arrows *arr a c* and *arr b c* and combines them into a new arrow *arr* $(\mathit{Either}\ a\ b)$ c which pipes all *Left a*'s to the first arrow and all *Right b*'s to the second arrow:

$$(|||) :: \mathit{ArrowChoice}\ \mathit{arr}\ a\ c \rightarrow \mathit{arr}\ b\ c \rightarrow \mathit{arr}\ (\mathit{Either}\ a\ b)\ c$$

With the zipWithArr combinator we can also write a combinator *listApp*, that lifts a list of arrows $[\mathit{arr}\ a\ b]$ to an arrow *arr* $[a]$ $[b]$.

$$\begin{aligned}
&\mathit{listApp} :: (\mathit{ArrowChoice}\ \mathit{arr}, \mathit{ArrowApply}\ \mathit{arr}) \Rightarrow [\mathit{arr}\ a\ b] \rightarrow \mathit{arr}\ [a]\ [b] \\
&\mathit{listApp}\ \mathit{fs} = (\mathit{arr}\ \$\ \lambda \mathit{as} \rightarrow (\mathit{fs}, \mathit{as})) \ggg \mathit{zipWithArr}\ \mathit{app}
\end{aligned}$$

Note that this additionally makes use of the *ArrowApply* typeclass that allows us to evaluate arrows with $\mathit{app} :: \mathit{arr}\ (\mathit{arr}\ a\ b, a)\ c$.

B Omitted Function Definitions

We have omitted some function definitions in the main text for brevity, and redeem this here. We warp Eden's build-in Futures in PArrows as in Figure B 5, where *rd* is the accessor function for the *RD* wrapped inside *RemoteData*. Furthermore, in order for these *Future* types to fit with the *ArrowParallel* instances we gave earlier, we have to give the necessary *NFData* and *Trans* instances, the latter are only needed in Eden. The *Trans* instance does not have any functions declared as the default implementation suffices here. Furthermore, because *MVar* already ships with a *NFData* instance, we only have to supply a simple delegating *NFData* instance for our *RemoteData* type, where *rd* simply unwraps *RD*. The *Trans* instance does not have any functions declared as the default implementation suffices:

$$\begin{aligned}
&\textbf{instance}\ \mathit{NFData}\ (\mathit{RemoteData}\ a)\ \textbf{where} \\
&\quad \mathit{rnf} = \mathit{rnf} \circ \mathit{rd} \\
&\textbf{instance}\ \mathit{Trans}\ (\mathit{RemoteData}\ a)
\end{aligned}$$

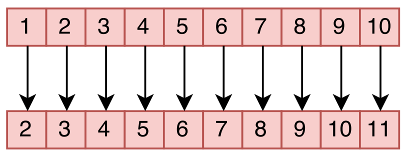


Figure B 1: Schematic depiction of *parMap*.

$$\begin{aligned}
&\mathit{parMap} :: (\mathit{ArrowParallel}\ \mathit{arr}\ a\ b\ \mathit{conf}) \Rightarrow \mathit{conf} \rightarrow (\mathit{arr}\ a\ b) \rightarrow (\mathit{arr}\ [a]\ [b]) \\
&\mathit{parMap}\ \mathit{conf}\ f = \mathit{parEvalN}\ \mathit{conf}\ (\mathit{repeat}\ f)
\end{aligned}$$

Figure B 2: Definition of parMap.

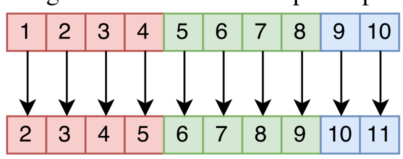


Figure B 3: Schematic depiction of *parMapStream*.

$$\begin{aligned}
&\mathit{parMapStream} :: (\mathit{ArrowParallel}\ \mathit{arr}\ a\ b\ \mathit{conf}, \mathit{ArrowChoice}\ \mathit{arr}, \mathit{ArrowApply}\ \mathit{arr}) \Rightarrow \\
&\quad \mathit{conf} \rightarrow \mathit{ChunkSize} \rightarrow \mathit{arr}\ a\ b \rightarrow \mathit{arr}\ [a]\ [b] \\
&\mathit{parMapStream}\ \mathit{conf}\ \mathit{chunkSize}\ f = \mathit{parEvalNLazy}\ \mathit{conf}\ \mathit{chunkSize}\ (\mathit{repeat}\ f)
\end{aligned}$$

Figure B 4: Definition of *parMapStream*.

Figures B 1–B 4 show the definitions and a visualizations of two parallel *map* variants, defined using *parEvalN* and its lazy counterpart.

Arrow versions of Eden's *shuffle*, *unshuffle* and the definition of *takeEach* are in Figure B 6. Similarly, Figure B 7 contains the definition of arrow versions of Eden's *lazy* and *rightRotate* utility functions. Fig. B 8 contains Eden's definition of *lazyzip3* together with the utility functions *uncurry3* and *threetotwo*. The full definition of *farmChunk* is in Figure B 9. Eden definition of *ring* skeleton is in Figure B 10. It follows Loogen (2012).

```
data RemoteData a = RD { rd :: RD a }
instance (Trans a) ⇒ Future RemoteData a where
  put = arr (λa → RD { rd = release a })
  get = arr rd >>> arr fetch
```

Figure B 5: *RD*-based *RemoteData* version of *Future* for the Eden backend.

```
shuffle :: (Arrow arr) ⇒ arr [[a]] [a]
shuffle = arr (concat ∘ transpose)
unshuffle :: (Arrow arr) ⇒ Int → arr [a] [[a]]
unshuffle n = arr (λxs → [takeEach n (drop i xs) | i ← [0 .. n − 1]])
takeEach :: Int → [a] → [a]
takeEach n [] = []
takeEach n (x : xs) = x : takeEach n (drop (n − 1) xs)
```

Figure B 6: Definitions of *shuffle*, *unshuffle*, *takeEach*.

```
lazy :: (Arrow arr) ⇒ arr [a] [a]
lazy = arr (λ∼(x : xs) → x : lazy xs)
rightRotate :: (Arrow arr) ⇒ arr [a] [a]
rightRotate = arr $ λlist → case list of
  [] → []
  xs → last xs : init xs
```

Figure B 7: Definitions of *lazy* and *rightRotate*.

The *parEval2* skeleton is defined in Figure B 11. We start by transforming the (a,c) input into a two-element list $[Either\ a\ c]$ by first tagging the two inputs with *Left* and *Right* and wrapping the right element in a singleton list with *return* so that we can combine them with *arr* (*uncurry* (:)). Next, we feed this list into a parallel arrow running on two instances of $f +\!\!+\!\!+ g$ as described above. After the calculation is finished, we convert the resulting $[Either\ b\ d]$ into $([b],[d])$ with *arr partitionEithers*. The two lists in this tuple contain only one element each by construction, so we can finally just convert the tuple to (b,d) in the last step.

$$
\begin{aligned}
&\mathit{lazyzip3} :: [a] \to [b] \to [c] \to [(a,b,c)]\\
&\mathit{lazyzip3}\ \mathit{as}\ \mathit{bs}\ \mathit{cs} = \mathit{zip3}\ \mathit{as}\ (\mathit{lazy}\ \mathit{bs})\ (\mathit{lazy}\ \mathit{cs})\\
&\mathit{uncurry3} :: (a \to b \to c \to d) \to (a,(b,c)) \to d\\
&\mathit{uncurry3}\ f\ (a,(b,c)) = f\ a\ b\ c\\
&\mathit{threetotwo} :: (\mathit{Arrow}\ \mathit{arr}) \Rightarrow \mathit{arr}\ (a,b,c)\ (a,(b,c))\\
&\mathit{threetotwo} = \mathit{arr}\ \$\ \lambda{\sim}(a,b,c) \to (a,(b,c))
\end{aligned}
$$

Figure B 8: Definitions of *lazyzip3*, *uncurry3* and *threetotwo*.

$$
\begin{aligned}
&\mathit{farmChunk} :: (\mathit{ArrowParallel}\ \mathit{arr}\ a\ b\ \mathit{conf}, \mathit{ArrowParallel}\ \mathit{arr}\ [a]\ [b]\ \mathit{conf},\\
&\quad \mathit{ArrowChoice}\ \mathit{arr}, \mathit{ArrowApply}\ \mathit{arr}) \Rightarrow\\
&\quad \mathit{conf} \to \mathit{ChunkSize} \to \mathit{NumCores} \to \mathit{arr}\ a\ b \to \mathit{arr}\ [a]\ [b]\\
&\mathit{farmChunk}\ \mathit{conf}\ \mathit{chunkSize}\ \mathit{numCores}\ f =\\
&\quad \mathit{unshuffle}\ \mathit{numCores} \ggg\\
&\quad \mathit{parEvalNLazy}\ \mathit{conf}\ \mathit{chunkSize}\ (\mathit{repeat}\ (\mathit{mapArr}\ f)) \ggg\\
&\quad \mathit{shuffle}
\end{aligned}
$$

Figure B 9: Definition of *farmChunk*.

C Syntactic Sugar

For basic arrows, we have the $*\!*\!*$ combinator (Fig. 6) which allows us to combine two arrows *arr a b* and *arr c d* into an arrow $\mathit{arr}\ (a,c)\ (b,d)$ which does both computations at once. This can easily be translated into a parallel version $*\!*\!*$ with the use of *parEval2*, but for this we require a backend which has an implementation that does not require any configuration (hence the $()$ as the *conf* parameter):

$$
\begin{aligned}
&(|{*}{*}{*}|) :: (\mathit{ArrowChoice}\ \mathit{arr}, \mathit{ArrowParallel}\ \mathit{arr}\ (\mathit{Either}\ a\ c)\ (\mathit{Either}\ b\ d)\ ())) \Rightarrow\\
&\quad \mathit{arr}\ a\ b \to \mathit{arr}\ c\ d \to \mathit{arr}\ (a,c)\ (b,d)\\
&(|{*}{*}{*}|) = \mathit{parEval2}\ ()
\end{aligned}
$$

We define the parallel &&& in a similar manner to its sequential pendant &&& (Fig. 6):

$$
\begin{aligned}
&(|\&\&\&|) :: (\mathit{ArrowChoice}\ \mathit{arr}, \mathit{ArrowParallel}\ \mathit{arr}\ (\mathit{Either}\ a\ a)\ (\mathit{Either}\ b\ c)\ ()) \Rightarrow\\
&\quad \mathit{arr}\ a\ b \to \mathit{arr}\ a\ c \to \mathit{arr}\ a\ (b,c)\\
&(|\&\&\&|)\ f\ g = (\mathit{arr}\ \$\ \lambda a \to (a,a)) \ggg f\ |{*}{*}{*}|\ g
\end{aligned}
$$

$ringSimple :: (Trans\ i, Trans\ o, Trans\ r) \Rightarrow (i \to r \to (o, r)) \to [i] \to [o]$
$ringSimple\ f\ is = os$
  where $(os, ringOuts) = unzip\ (parMap\ (toRD\ \$\ uncurry\ f)\ (zip\ is\ \$\ lazy\ ringIns))$
    $ringIns = rightRotate\ ringOuts$
$toRD :: (Trans\ i, Trans\ o, Trans\ r) \Rightarrow ((i, r) \to (o, r)) \to ((i, RD\ r) \to (o, RD\ r))$
$toRD\ f\ (i, ringIn) = (o, release\ ringOut)$
  where $(o, ringOut) = f\ (i, fetch\ ringIn)$
$rightRotate :: [a] \to [a]$
$rightRotate\ [\,] = [\,]$
$rightRotate\ xs = last\ xs : init\ xs$
$lazy :: [a] \to [a]$
$lazy {\sim} (x : xs) = x : lazy\ xs$

Figure B 10: Eden's definition of the *ring* skeleton.

$parEval2 :: (ArrowChoice\ arr,$
  $ArrowParallel\ arr\ (Either\ a\ c)\ (Either\ b\ d)\ conf) \Rightarrow$
  $conf \to arr\ a\ b \to arr\ c\ d \to arr\ (a, c)\ (b, d)$
$parEval2\ conf\ f\ g =$
  $arr\ Left *** (arr\ Right >>> arr\ return) >>>$
  $arr\ (uncurry\ (:)) >>>$
  $parEvalN\ conf\ (replicate\ 2\ (f +++ g)) >>>$
  $arr\ partitionEithers >>>$
  $arr\ head *** arr\ head$

Figure B 11: Definition of parEval2.