Work Analysis with Resource-Aware Session Types

Ankush Das Jan Hoffmann Frank Pfenning

LICS, July 10, 2018



Goal for this Talk Resource Analysis for Concurrent Programs



Execution Time



Memory Usage





Complexity of Parallel Algorithms

Çiçek et. al. (ESOP '15)



Design of Optimal Scheduling Policies

Acar et. al. (JFP '16)



Complexity of Parallel Algorithms

Çiçek et. al. (ESOP '15)



Prevention of Side-Channel Attacks Ngo et. al. (S&P '17)



Design of Optimal Scheduling Policies

Acar et. al. (JFP '16)



Complexity of Parallel Algorithms

Çiçek et. al. (ESOP '15)



Prevention of Side-Channel Attacks Ngo et. al. (S&P '17)



Design of Optimal Scheduling Policies

Acar et. al. (JFP '16)



Static and Dynamic Profiling Haemmerlé et. al. (FLOPS '17)



Work Sequential Complexity Execution time on one processor





Work Sequential Complexity Execution time on one processor Span Parallel Complexity

Execution time on arbitrarily many processors





Work Sequential Complexity Execution time on one processor Span Parallel Complexity

Execution time on arbitrarily many processors





Work Sequential Complexity Execution time on one processor Span Parallel Complexity

Execution time on arbitrarily many processors



Read/Write Overhead



Read/Write Overhead





Shared Memory Read/Write Overhead







Shared Memory Read/Write Overhead









Shared Memory Read/Write Overhead









No Shared Memory









No Shared Memory





Types strictly enforce communication protocols





No Shared Memory





Types strictly enforce communication protocols



5



No Shared Memory





Types strictly enforce communication protocols



- Implement message-passing concurrent programs
- Communication via bi-directional typed channels
- Curry-Howard isomorphism with intuitionistic linear logic

- Implement message-passing concurrent programs
- Communication via bi-directional typed channels
- Curry-Howard isomorphism with intuitionistic linear logic



- Implement message-passing concurrent programs
- Communication via bi-directional typed channels
- Curry-Howard isomorphism with intuitionistic linear logic

$$\mathbf{bits} = \oplus \{\mathbf{b0} : \mathbf{bits}, \mathbf{b1} : \mathbf{bits}\}$$



- Implement message-passing concurrent programs
- Communication via bi-directional typed channels
- Curry-Howard isomorphism with intuitionistic linear logic



Design a type system to analyze work of session-typed programs

Design a type system to analyze work of session-typed programs

- based on amortized analysis
- can be parameterized to count different resources
- proved sound w.r.t. cost semantics
- conservative extension to standard session type system
- applied to all standard programs

Design a type system to analyze work of session-typed programs

based on amortized analysis



- can be parameterized to count different resources
- proved sound w.r.t. cost semantics
- conservative extension to standard session type system
- applied to all standard programs

Design a type system to analyze work of session-typed programs

- based on amortized analysis
- can be parameterized to count different resources
- proved sound w.r.t. cost semantics
- conservative extension to standard session type system
- applied to all standard programs



Design a type system to analyze work of session-typed programs

 based on amortized analysis messages exchanged • can be parameterized to count different resources processes • proved sound w.r.t. cost semantics spawned conservative extension to standard session type system instructions applied to all standard programs executed

Design a type system to analyze work of session-typed programs

- based on amortized analysis
- can be parameterized to count different resources
- proved sound w.r.t. cost semantics
- conservative extension to standard session type system
- applied to all standard programs







Example: Queues



$$\begin{array}{l} \mathsf{queue}_\mathbf{A} = \&\{\mathsf{ins}: \mathbf{A} \multimap \mathsf{queue}_\mathbf{A}, \\ & \mathsf{del}: \oplus\{\mathsf{none}: \mathbf{1}, \\ & \mathsf{some}: \mathbf{A} \otimes \mathsf{queue}_\mathbf{A}\} \} \end{array}$$


Example: Queues a b c d

























$$\begin{array}{l} (x:A) \ (t: \mathsf{queue}_A) \vdash elem :: (s: \mathsf{queue}_A) \\ s \leftarrow elem \leftarrow x \ t = \\ & \mathsf{case} \ s \ (\mathsf{ins} \Rightarrow y \leftarrow \mathsf{recv} \ s \ ; \\ & t.\mathsf{ins} \ ; \\ & s \mathsf{end} \ t \ y \ ; \\ & s \leftarrow elem \leftarrow x \ t \\ | \ \mathsf{del} \Rightarrow s.\mathsf{some} \ ; \\ & s \leftarrow t) \end{array}$$















Cost Semantics

$\mathbf{proc}(\mathbf{c},\mathbf{w},\mathbf{P})$

Process P offering along channel c and has performed work w

- Standard semantics extended with local work counters w for each process
- Total work of system is sum of local counters w
- w is incremented every time process P performs some 'work' (this talk: whenever message is sent)





Type System

Based on Amortized Analysis!

- Store potential in each process
- Potential can be transferred via messages
- Potential is consumed to perform 'work'

Type System

Based on Amortized Analysis!

- Store potential in each process
- Potential can be transferred via messages
- Potential is consumed to perform 'work'

Process P offers along channel c, acts as a client for channels in Ω and storing potential q

Concurrent Queues







Concurrent Queues Count the messages! ins(e) ins(e) ins(e) ins(e) b d a C e ins(e)

Concurrent Queues



Concurrent Queues





Count the messages!



w_i = Work done to process insertion = 2n (where n is the size of queue) = 'ins' and 'e' travel to end of queue

w_d = Work done to process deletion = 2 (sends back 'some' and 'a')

$$\begin{split} \mathsf{queue}_{\mathbf{A}}[\mathbf{n}] &= \&\{\mathsf{ins}^{\mathbf{2n}}: \mathbf{A} \multimap \mathsf{queue}_{\mathbf{A}}[\mathbf{n}+\mathbf{1}], \\ & \mathsf{del}^{\mathbf{2}}: \oplus\{\mathsf{none}: \mathbf{1}, \\ & \mathsf{some}: \mathbf{A} \otimes \mathsf{queue}_{\mathbf{A}}[\mathbf{n}-\mathbf{1}]\} \} \end{split}$$

$$\begin{split} \mathsf{queue}_{\mathbf{A}}[\mathbf{n}] &= \&\{\mathsf{ins}^{\mathbf{2n}}: \mathbf{A} \multimap \mathsf{queue}_{\mathbf{A}}[\mathbf{n}+\mathbf{1}], \\ & \mathsf{del}^{\mathbf{2}}: \oplus\{\mathsf{none}: \mathbf{1}, \\ & \mathsf{some}: \mathbf{A} \otimes \mathsf{queue}_{\mathbf{A}}[\mathbf{n}-\mathbf{1}]\} \rbrace \end{split}$$

Resource-Aware Session Types

types augmented with potential information sender pays potential with message receiver gets potential with message

$$\begin{split} \mathsf{queue}_{\mathbf{A}}[\mathbf{n}] &= \&\{\mathsf{ins}^{\mathbf{2n}}: \mathbf{A} \multimap \mathsf{queue}_{\mathbf{A}}[\mathbf{n}+\mathbf{1}], \\ & \mathsf{del}^{\mathbf{2}}: \oplus\{\mathsf{none}: \mathbf{1}, \\ & \mathsf{some}: \mathbf{A} \otimes \mathsf{queue}_{\mathbf{A}}[\mathbf{n}-\mathbf{1}]\} \} \end{split}$$

```
\begin{array}{l} (x:A) \ (t: \mathsf{queue}_A[n]) \stackrel{\wp}{\mapsto} elem :: (s: \mathsf{queue}_A[n+1]) \\ s \leftarrow elem \leftarrow x \ t = \\ & \mathsf{case} \ s \ (\mathsf{ins} \Rightarrow y \leftarrow \mathsf{recv} \ s \ ; \\ & t.\mathsf{ins} \ ; \\ & s \mathsf{end} \ t \ y \ ; \\ & s \leftarrow elem \leftarrow x \ t \\ | \ \mathsf{del} \Rightarrow s.\mathsf{some} \ ; \\ & s \leftarrow t ) \end{array}
```

 $\begin{aligned} \mathsf{queue}_{\mathbf{A}}[\mathbf{n}] &= \&\{\mathsf{ins}^{\mathbf{2n}} : \mathbf{A} \multimap \mathsf{queue}_{\mathbf{A}}[\mathbf{n}+\mathbf{1}], \\ & \mathsf{del}^{\mathbf{2}} : \oplus\{\mathsf{none} : \mathbf{1}, \\ & \mathsf{some} : \mathbf{A} \otimes \mathsf{queue}_{\mathbf{A}}[\mathbf{n}-\mathbf{1}]\} \end{aligned}$

$$(x : A) (t : queue_A[n]) \stackrel{\rho}{=} elem :: (s : queue_A[n+1])$$

$$s \leftarrow elem \leftarrow x t =$$

$$case s (ins \Rightarrow \overline{y \leftarrow recv s};$$

$$t.ins;$$

$$send t y;$$

$$s \leftarrow elem \leftarrow x t$$

$$| del \Rightarrow s.some;$$

$$send s x;$$

$$s \leftarrow t)$$

 $\begin{aligned} \mathsf{queue}_{\mathbf{A}}[\mathbf{n}] &= \&\{\mathsf{ins}^{\mathbf{2n}} : \mathbf{A} \multimap \mathsf{queue}_{\mathbf{A}}[\mathbf{n}+\mathbf{1}], \\ & \mathsf{del}^{\mathbf{2}} : \oplus\{\mathsf{none} : \mathbf{1}, \\ & \mathsf{some} : \mathbf{A} \otimes \mathsf{queue}_{\mathbf{A}}[\mathbf{n}-\mathbf{1}]\} \end{aligned}$


Type for Queues

 $\begin{aligned} \mathsf{queue}_{\mathbf{A}}[\mathbf{n}] &= \&\{\mathsf{ins}^{\mathbf{2n}} : \mathbf{A} \multimap \mathsf{queue}_{\mathbf{A}}[\mathbf{n}+\mathbf{1}], \\ \mathsf{del}^{\mathbf{2}} : \oplus\{\mathsf{none} : \mathbf{1}, \\ \mathsf{some} : \mathbf{A} \otimes \mathsf{queue}_{\mathbf{A}}[\mathbf{n}-\mathbf{1}]\} \end{aligned}$



Type for Queues

 $\begin{aligned} \mathsf{queue}_{\mathbf{A}}[\mathbf{n}] &= \&\{\mathsf{ins}^{\mathbf{2n}} : \mathbf{A} \multimap \mathsf{queue}_{\mathbf{A}}[\mathbf{n}+\mathbf{1}], \\ \mathsf{del}^{\mathbf{2}} : \oplus\{\mathsf{none} : \mathbf{1}, \\ \mathsf{some} : \mathbf{A} \otimes \mathsf{queue}_{\mathbf{A}}[\mathbf{n}-\mathbf{1}]\} \end{aligned}$



Type for Queues

 $\begin{aligned} \mathsf{queue}_{\mathbf{A}}[\mathbf{n}] &= \&\{\mathsf{ins}^{\mathbf{2n}}: \mathbf{A} \multimap \mathsf{queue}_{\mathbf{A}}[\mathbf{n}+\mathbf{1}], \\ &\quad \mathsf{del}^{\mathbf{2}}: \oplus\{\mathsf{none}: \mathbf{1}, \\ &\quad \mathsf{some}: \mathbf{A} \otimes \mathsf{queue}_{\mathbf{A}}[\mathbf{n}-\mathbf{1}]\} \end{aligned}$





Process P offers along channel c, acts as a client for channels in Ω and storing potential q





Features of Type System

- Flexible: supports counting of different resources (e.g. messages exchanged, processes spawned, etc.) by being parametric in cost model
- Compositional: types describe individual processes, not just whole programs
- Precise: potential upper bounds work accurately
- Conservative: strict extension of type system
- General: works on most standard examples
- Automatic: future work



Binary Counters

$$\begin{aligned} \mathsf{ctr}[\mathbf{n}] &= \&\{\mathsf{inc}^{\mathbf{1}}:\mathsf{ctr}[\mathbf{n}+\mathbf{1}],\\ &\quad \mathsf{val}^{\mathbf{2}\lceil \log(\mathbf{n}+\mathbf{1})\rceil+\mathbf{2}}:\mathsf{bits}\} \end{aligned}$$

- Increment:
 - requires one unit of potential
 - uses amortized analysis!
- Value:
 - requires logarithmic potential
 - precise work bound

Stacks vs Queues

$$\begin{split} \mathsf{stack}_{\mathbf{A}}[\mathbf{n}] &= \&\{\mathsf{ins}^{\mathbf{0}}: \mathbf{A} \multimap \mathsf{stack}_{\mathbf{A}}[\mathbf{n}+\mathbf{1}], \\ & \mathsf{del}^{\mathbf{2}}: \oplus\{\mathsf{none}: \mathbf{1}, \\ & \mathsf{some}: \mathbf{A} \otimes \mathsf{stack}_{\mathbf{A}}[\mathbf{n}-\mathbf{1}]\} \rbrace \end{split}$$

$$\begin{split} \mathsf{queue}_{\mathbf{A}}[\mathbf{n}] &= \&\{\mathsf{ins}^{\mathbf{2n}}: \mathbf{A} \multimap \mathsf{queue}_{\mathbf{A}}[\mathbf{n}+\mathbf{1}], \\ &\quad \mathsf{del}^{\mathbf{2}}: \oplus\{\mathsf{none}: \mathbf{1}, \\ &\quad \mathsf{some}: \mathbf{A} \otimes \mathsf{queue}_{\mathbf{A}}[\mathbf{n}-\mathbf{1}]\} \} \end{split}$$

Stacks vs Queues

- $$\begin{split} \mathsf{stack}_{\mathbf{A}}[\mathbf{n}] &= \&\{\mathsf{ins}^{\mathbf{0}}: \mathbf{A} \multimap \mathsf{stack}_{\mathbf{A}}[\mathbf{n}+\mathbf{1}], \\ & \mathsf{del}^{\mathbf{2}}: \oplus\{\mathsf{none}: \mathbf{1}, \\ & \mathsf{some}: \mathbf{A} \otimes \mathsf{stack}_{\mathbf{A}}[\mathbf{n}-\mathbf{1}]\} \rbrace \end{split}$$
- $\begin{array}{l} \mathsf{queue}_{\mathbf{A}}[\mathbf{n}] = \&\{\mathsf{ins}^{\mathbf{2n}}: \mathbf{A} \multimap \mathsf{queue}_{\mathbf{A}}[\mathbf{n}+\mathbf{1}],\\ & \mathsf{del}^{\mathbf{2}}: \oplus\{\mathsf{none}: \mathbf{1},\\ & \mathsf{some}: \mathbf{A} \otimes \mathsf{queue}_{\mathbf{A}}[\mathbf{n}-\mathbf{1}]\} \} \end{array}$

Contributions

Type System for Work Analysis based on amortized analysis types augmented with potential information work is upper bounded by potential Flexible, Compositional, Precise, Conservative, General, Automatic

Contributions

Type System for Work Analysis

based on amortized analysis

types augmented with potential information

work is upper bounded by potential

Flexible, Compositional, Precise, Conservative,

General, Automatic

Soundness Theorem

Cost Semantics

Contributions

Type System for Work Analysis

based on amortized analysis

types augmented with potential information

work is upper bounded by potential

Flexible, Compositional, Precise, Conservative, General, Automatic

Soundness Theorem

Cost Semantics

Examples

stacks, queues, binary counters efficiency comparison list examples: append, map, fold