Imperative Characterization of BSP Algorithms

Yoann Marquer & Frédéric Gava

Laboratory of Algorithms, Complexity and Logic (LACL) University of Paris-East









2 Algorithmic Equivalences



Introduction

What is an algorithm?

- Not a Turing machine
- Not a programming language
- Every designer writes them in different forms

Introduction

What is an algorithm?

- Not a Turing machine
- Not a programming language
- Every designer writes them in different forms

So?

- Axiomatic definition of a sequential algorithm and ASM
- Equivalence with a core programming language

Introduction

What is an algorithm?

- Not a Turing machine
- Not a programming language
- Every designer writes them in different forms

So?

- Axiomatic definition of a sequential algorithm and ASM
- Equivalence with a core programming language

And then?

- Parallel and/or distributed ASM (Gurevich and al)
- Without equivalences with a core programming language
- Because no cost model and not a bridging model



What is a bridging model? For sequential computing



What is a bridging model? For HPC computing



The BSP computer

Defined by:

- p pairs CPU/memory
- Communication network (g)
- Synchronisation unit (L
- Super-steps execution

Properties:

The BSP computer

- Defined by:
 - p pairs CPU/memory
 - Communication network (g)
 - Synchronisation unit (L
 - Super-steps execution

Properties:

The BSP computer

Defined by:

- p pairs CPU/memory
- Communication network (g)
- Synchronisation unit (L)
- Super-steps execution

Properties:

BSP Programming

The BSP computer

Defined by:

- p pairs CPU/memory
- Communication network (g)
- Synchronisation unit (L)
- Super-steps execution

Properties:

"Confluent"

"Deadlock-free"



BSP Equivalences

Bridging model: Bulk Synchronous Parallelism (BSP)

local computations

communication (⊗g) barrier (⊕L) next super-step

BSP Equivalences

 $p_0 | p_1$

The BSP computer

Defined by:

- p pairs CPU/memory
- Communication network (g)
- Synchronisation unit (L)
- Super-steps execution

Properties:

- "Confluent"
- "Deadlock-free"
- Predictable performances



local

computations



local

computations

BSP Programming

Bridging model: Bulk Synchronous Parallelism (BSP)

BSP Equivalences

The BSP computer

Defined by:

- p pairs CPU/memory
- Communication network (g)
- Synchronisation unit (L)
- Super-steps execution

Properties:

- "Confluent"
- "Deadlock-free"
- Predictable performances





7/28

BSP Programming

Bridging model: Bulk Synchronous Parallelism (BSP)

BSP Equivalences

 p_0

The BSP computer

Defined by:

- p pairs CPU/memory
- Communication network (g)
- Synchronisation unit (L)
- Super-steps execution

Properties:

- "Confluent"
- "Deadlock-free"
- Predictable performances



local

communication (⊗g) barrier (⊕L)

next super-step

Conclusion

Bridging model: Bulk Synchronous Parallelism (BSP)

The BSP computer Defined by:	
Pro and cons	
Pro Cost model Structured parallelism Easy to learn	Cons Not all parallel patterns Too regular No asynchronous warnings
 "Confluent" "Deadlock-free" Predictable performances	E E E E next super-step



Conclusion

Examples: broadcasting a value



BSP	Programming	
000000000		

Examples: broadcasting a value



 $Cost \equiv 2 \times g \times n + 2 \times L$



- Dedicated languages: NestStep, BSP++, BSP-Python, ...
- 2 BSPLib for C and Java
- BSPGPU, Ct, Hamma, JBSP, JPUB, ...
- MPI collective operations

- Dedicated languages: NestStep, BSP++, BSP-Python, ...
- BSPLib for C and Java
- BSPGPU, Ct, Hamma, JBSP, JPUB, ...
 - MPI collective operations

- Dedicated languages: NestStep, BSP++, BSP-Python, ...
- BSPLib for C and Java
- BSPGPU, Ct, Hamma, JBSP, JPUB, ...
- MPI collective operations

- Dedicated languages: NestStep, BSP++, BSP-Python, ...
- BSPLib for C and Java
- BSPGPU, Ct, Hamma, JBSP, JPUB, ...
- MPI collective operations



- Dedicated languages: NestStep, BSP++, BSP-Python, ...
- BSPLib for C and Java
- BSPGPU, Ct, Hamma, JBSP, JPUB, ...
- MPI collective operations



BSMP and DRM

- Typical **BSMP** routines:
 - bsp_send(dest,buffer,size)
 - bsp_nmsgs()
 - msg* bsp_findmsg(proc_id, index)
- Typical DRMA routines:
 - bsp_push_reg(ident, size)
 - bsp_get(srcPID, src, offset, dest, nbytes)
- bsp_sync() (barrier)

collective operations

MPI_Scatter(sendbuf,sendcount,sendtype,recvbuf,recvcount,recvtype,root,comm

BSMP and DRM

- Typical **BSMP** routines:
 - bsp_send(dest,buffer,size)
 - bsp_nmsgs()
 - msg* bsp_findmsg(proc_id, index)
- Typical DRMA routines:
 - bsp_push_reg(ident, size)
 - bsp_get(srcPID, src, offset, dest, nbytes)
- bsp_sync() (barrier)

collective operations

MPI_Scatter(sendbuf,sendcount,sendtype,recvbuf,recvcount,recvtype,root,comm)

PI_Gather(sendbuf,sendcount,sendtype,recvbuf,recvcount,recvtype,root,comm)

BSMP and DRM

- Typical **BSMP** routines:
 - bsp_send(dest,buffer,size)
 - bsp_nmsgs()
 - msg* bsp_findmsg(proc_id, index)
- Typical DRMA routines:
 - bsp_push_reg(ident, size)
 - bsp_get(srcPID, src, offset, dest, nbytes)

• bsp_sync() (barrier)

collective operations

MPI_Scatter(sendbuf,sendcount,sendtype,recvbuf,recvcount,recvtype,root,comm)

MPI_Gather(sendbuf,sendcount,sendtype,recvbuf,recvcount,recvtype,root,comm)

BSMP and DRM

- Typical **BSMP** routines:
 - bsp_send(dest,buffer,size)
 - bsp_nmsgs()
 - msg* bsp_findmsg(proc_id, index)
- Typical DRMA routines:
 - bsp_push_reg(ident, size)
 - bsp_get(srcPID, src, offset, dest, nbytes)

• bsp_sync() (barrier)

MPI collective operations

MPI_Scatter(sendbuf,sendcount,sendtype,recvbuf,recvcount,recvtype,root,comm)



MPI_Gather(sendbuf, sendcount, sendtype, recvbuf, recvcount, recvtype, root, comm)

One example: PSRS sorting

```
PSRS Sorting(array<T> tab)
begin
   seq_sort(tab);
                                    (* local sorting *)
   s1←sample(tab);
                                     (* first sampling *)
   for i←0 to nprocs-1 do
     bsp send(i,s1[i],size(s1[i]));
   done;
   bsp_sync();
   s all \leftarrow [];
                                     (* second sampling *)
   for i←0 to nprocs-1 do
     merge(s all,bsp findmsg(i,0));
   done:
   s2 \leftarrow sample(s all);
   ...
   bsp sync();
   ...
end
```



One example: PSRS sorting



end







BSP	Programming	

Conclusion

Execution using the axiomatic definition

Sequential algorithm: **BSP** algorithm: X1 X1 Z1 X2 Z2 Y2 X2 Xn Yn Zn X3 Xn+1 Yn+1 Zn+1

Axiomatic definition (1)

Sequential:



3 a transition function $\tau_{A_{seq}}: S(A_{seq}) \rightarrow S(A_{seq})$

BSP	Programming	
000	000000	

Axiomatic definition (1)



Axiomatic definition (1)

Sequential:

Axiom 1; Sequential Time

- a set of states $S(A_{seq})$
- **2** a set of initial states $I(A_{seq}) \subseteq S(A_{seq})$
- 3 a transition function $\tau_{A_{seq}}: S(A_{seq}) \rightarrow S(A_{seq})$

Axiom 2; Abstract States

- states are first-order structures and closed by isomorphism
- τ_{Aseq} commutes with the isomorphism

BSP	Programming	

Conclusion

Axiomatic definition (1)

Sequential:



2 \(\tau_{A_{seq}}\) commutes with the isomorphism

Imperative Characterization of BSP Algorithms

Axiomatic definition (1)

Sequential:

Axiom 1; Sequential Time

- a set of states $S(A_{seq})$
- **2** a set of initial states $I(A_{seq}) \subseteq S(A_{seq})$
- 3 a transition function $\tau_{A_{seq}}: S(A_{seq}) \rightarrow S(A_{seq})$

Axiom 2; Abstract States

- states are first-order structures and closed by isomorphism
- $au_{A_{seq}}$ commutes with the isomorphism

BSP:

Axiom 1; Sequential Time

- a set of states S(A_{BSP})
- a set of initial states $I(A_{BSP}) \subseteq S(A_{BSP})$
- 3 a transition function $\tau_{A_{BSP}}: S(A_{BSP}) \rightarrow S(A_{BSP})$

Axiom 2; Abstract States

- states are *p*-tuples and closed by *p*-isomorphism

Axiomatic definition (2)

Sequential:

Axiom 3; Bounded Exploration

For every algorithm *A* there exists a finite set *T* of terms such that for every state *X* and *Y*, if the elements of *T* have the same interpretations on *X* and *Y* then $\Delta(A, X) = \Delta(A, Y)$.

BSP	Programming	

Conclusion

Axiomatic definition (2)



Axiomatic definition (2)

Sequential:

Axiom 3; Bounded Exploration

For every algorithm *A* there exists a finite set *T* of terms such that for every state *X* and *Y*, if the elements of *T* have the same interpretations on *X* and *Y* then $\Delta(A, X) = \Delta(A, Y)$.

BSP:

Axiom 3; Bounded Exploration

Idem but
$$X = (X^1, \dots, X^p)$$
 and $Y = (Y^1, \dots, Y^q)$ and $p = q$
and $\vec{\Delta}(A, X) = \vec{\Delta}(A, Y)$

Axiomatic definition (2)

Sequential:

Axiom 3; Bounded Exploration

For every algorithm *A* there exists a finite set *T* of terms such that for every state *X* and *Y*, if the elements of *T* have the same interpretations on *X* and *Y* then $\Delta(A, X) = \Delta(A, Y)$.

BSP:

Axiom 3; Bounded Exploration

Idem but
$$X = (X^1, \dots, X^p)$$
 and $Y = (Y^1, \dots, Y^q)$ and $p = q$
and $\vec{\Delta}(A, X) = \vec{\Delta}(A, Y)$

Axiom 4; Barrier

For every BSP algorithm A there exists two applications $\operatorname{comp}_A : M(A) \to M(A)$ and $\operatorname{sync}_A : S(A) \to S(A)$ such that

Conclusion

Axiomatic definition (2)



ASM and BSP-ASM

Sequential:

Definition ASM

 $(\Pi, S(B), I(B))$

BSP:

Definition BSP-ASM

 $(\Pi, S(B), I(B), \text{sync}_B)$ where S(B) contains *p*-tuples of structures.

ASM and BSP-ASM

Sequential:

Definition ASM $(\Pi, S(B), I(B))$

ASM's programs П

$$\begin{aligned} \Pi =_{def} & f(t_1, \dots, t_{\alpha}) := t_0 \\ & | & \text{if } F \text{ then } \Pi_1 \\ & & \text{else } \Pi_2 \text{ endif} \\ & | & & \text{par } \Pi_1 \| \dots \| \Pi_n \text{ endpane} \end{aligned}$$

BSP:

Definition BSP-ASM

 $(\Pi, S(B), I(B), \text{sync}_B)$ where S(B) contains *p*-tuples of structures.

ASM and BSP-ASM

Sequential:

Definition ASM $(\Pi, S(B), I(B))$

ASM's programs П

$$\begin{aligned} \Pi =_{def} & f(t_1, \dots, t_{\alpha}) := t_0 \\ & | & \text{if } F \text{ then } \Pi_1 \\ & & \text{else } \Pi_2 \text{ endif} \\ & | & & \text{par } \Pi_1 \| \dots \| \Pi_n \text{ endpare} \end{aligned}$$

BSP:

Definition BSP-ASM

 $(\Pi, S(B), I(B), \text{sync}_B)$ where S(B) contains *p*-tuples of structures.

Extension of П

$$\begin{array}{ll} \Pi =_{def} & \dots \\ & \mid & f_{send}(t_1, \dots, t_{\alpha}) := t^{send} \\ & \mid & t^{rcv} := f_{rcv}(t_1, \dots, t_{\alpha}) \end{array}$$

Operational semantics

....

Sequential

$$\Delta(f(t_1, \dots, t_{\alpha}) := t_0, X) =_{def} \{(f, \overline{t_1}^X, \dots, \overline{t_{\alpha}}^X, \overline{t_0}^X)\}$$

$$\Delta(\text{if } F \text{ then } \Pi_1 \text{ else } \Pi_2 \text{ endif}, X) =_{def} \Delta(\Pi_i, X)$$

where $i = \begin{cases} 1 \text{ if } F \text{ is true on } X \\ 2 \text{ else} \end{cases}$
$$\Delta(\text{par } \Pi_1 \| \dots \| \Pi_n \text{ endpar}, X) =_{def} \Delta(\Pi_1, X) \cup \dots \cup \Delta(\Pi_n, X)$$

Operational semantics

Sequential

$$\Delta(f(t_1, \dots, t_\alpha) := t_0, X) =_{def} \{(f, \overline{t_1}^X, \dots, \overline{t_\alpha}^X, \overline{t_0}^X)\}$$

$$\Delta(\text{if } F \text{ then } \Pi_1 \text{ else } \Pi_2 \text{ endif, } X) =_{def} \Delta(\Pi_i, X)$$

where
$$i = \begin{cases} 1 \text{ if } F \text{ is true on } X \\ 2 \text{ else} \end{cases}$$

$$\Delta(\operatorname{par} \Pi_1 \| \dots \| \Pi_n \operatorname{endpar}, X) =_{def} \Delta(\Pi_1, X) \cup \dots \cup \Delta(\Pi_n, X)$$

BSP

$$\overrightarrow{\Delta}(\Pi, (X^1, \cdots, X^p)) =_{def} \begin{cases} (\Delta(\Pi, X^1), \cdots \Delta(\Pi, X^p)) & \text{if } \exists 1 \leq i \leq p \\ & \text{such that } \Delta(\Pi, X^i) \neq \emptyset \\ & \text{sync}_{\Pi}(X^1, \cdots, X^p) & \text{else} \end{cases}$$

IMP and BSP-IMP

Sequential:

$$\begin{array}{l} c =_{def} f(t_1, \ldots, t_{\alpha}) := t_0 \\ \mid \text{if } F \{ P_1 \} \text{ else } \{ P_2 \} \\ \mid \text{while } F \{ P \} \end{array}$$

$$\begin{array}{l} \mathcal{C} =_{def} & \dots \\ \mid f_{send}(t_1, \dots, t_{\alpha}) := t^{send} \\ \mid t^{r_{CV}} := f_{r_{CV}}(t_1, \dots, t_{\alpha}) \\ \mid f_{sync}() \end{array}$$

$$P =_{def} \epsilon \mid c; P$$

Small-Step Semantics



Small-Step Semantics





Small-Step Semantics rules

Local rules:

$$\begin{array}{l} f(t_1,\ldots,t_{\alpha}) := t_0; P \star X \succ^i P \star X \oplus (f,\overline{t_1}^X,\ldots,\overline{t_{\alpha}}^X,\overline{t_0}^X) \\ \text{while } F \{P_1\}; P_2 \star X \succ^i P_1 \text{ "while } F \{P_1\}; P_2 \star X \\ & \text{ if } F \text{ is true in } X \end{array}$$

...

One global rule:

$$\frac{\exists 1 \leq i \leq \mathbf{p} \quad P^{i} \star X^{i} \succ^{i} P^{\prime i} \star X^{\prime i}}{\langle P^{1} \star X^{1}, \dots, P^{\mathbf{p}} \star X^{\mathbf{p}} \rangle \succ \langle P^{\prime 1} \star X^{\prime 1}, \dots, P^{\prime \mathbf{p}} \star X^{\prime \mathbf{p}} \rangle}$$

Fair Simulation

A computation model M_1 simulates M_2 if $\forall P_2 \in M_2 \exists P_1 \in M_1$:

What too many intermediate variables

② ∃ $d \in \mathbb{N} \setminus \{0\}$ and $e \in \mathbb{N}$ (depending only on P_2) such that, for every execution \vec{Y} of P_2 , ∃ an execution \vec{X} of P_1 :

$$time(P_1, X_0) = d \times time(P_2, Y_0) + e$$



Imperative Characterization of BSP Algorithms

Conclusion

Compilation: from BSP-IMP to BSP-ASM

Main code

$\Pi_{P} \equiv \text{if } \neg b_{\text{wait}} \text{ then } \max_{P_{j} \in \mathcal{G}(P)} \text{ if } b_{P_{j}} \text{ then } \llbracket P_{j} \rrbracket_{asm} \text{ endpar endif}$

Conclusion

Compilation: from BSP-IMP to BSP-ASM

Main code

 $\Pi_P \equiv \text{if} \neg b_{\text{wait}} \text{ then } \underset{P_j \in \mathcal{G}(P)}{\text{par if } b_{P_j} \text{ then } \llbracket P_j \rrbracket_{asm} \text{ endpar endif}$

The control flow graph

$$\mathcal{G}(m{c};m{P})\equiv\mathcal{G}(m{c});m{P}\cup\mathcal{G}(m{P})\ \mathcal{G}(m{while}\;m{F}\;\{m{P}\})\equiv\mathcal{G}(m{P});m{while}\;m{F}\;\{m{P}\}$$

...

Conclusion

Compilation: from BSP-IMP to BSP-ASM

Main code

 $\Pi_{P} \equiv \text{if} \neg b_{\text{wait}} \text{ then } \max_{P_{j} \in \mathcal{G}(P)} \text{ if } b_{P_{j}} \text{ then } \llbracket P_{j} \rrbracket_{asm} \text{ endpar endif}$

The control flow graph

$$\mathcal{G}(m{c};m{P})\equiv\mathcal{G}(m{c});m{P}\cup\mathcal{G}(m{P})\ \mathcal{G}(while\;m{F}\;\{m{P}\})\equiv\mathcal{G}(m{P});while\;m{F}\;\{m{P}\}$$

. . .

Compilation of structures (example)

$$\llbracket f_{sync}(); Q \rrbracket_{asm} \equiv par$$

 $b_{sync;Q} := false$
 $\lVert b_{wait} := true$
 $\lVert b_Q := true$
endpar

Compilation: from BSP-ASM to BSP-IMP

Translation for the whole machine

$$b_{stop} := false;$$

while $\neg b_{stop}$
 $P_{step};$
while $\neg F_{\Pi} \{P_{step};\}$
 $f_{sync}();$

Compilation: from BSP-ASM to BSP-IMP

Translation for the whole machine

$$D_{stop} := false;$$

while $\neg b_{stop}$
 $P_{step};$
while $\neg F_{\Pi} \{P_{step};\}$
 $f_{sync}();$

Translation of one step

par
$$x := y \parallel y := x$$
 endpar

 \Rightarrow \Rightarrow

$$V_y := y; V_x := x; x := V_y; y := V_x;$$



Sequential:

$$\mathsf{Algo}_{seq} = \mathsf{ASM} \simeq \mathsf{Imp}_{seq}$$

BSP:

$$\mathsf{Algo}_{BSP} = BSP - \mathsf{ASM} \simeq \mathsf{Imp}_{BSP}$$



2 Algorithmic Equivalences



BSP-ASM

- Axiomatic definition of BSP algorithms and BSP-ASM
- Using fair simulations
- Algo_{BSP} ≃ BSP−ASM ≃ Imp_{BSP}
- BSPlib is algorithmic complete

BSP-ASM

Axiomatic definition of BSP algorithms and BSP – ASM

- Using fair simulations
- Algo $_{BSP} \simeq BSP ASM \simeq Imp_{BSP}$
- BSPlib is algorithmic complete

BSP-ASM

- Axiomatic definition of BSP algorithms and BSP ASM
- Using fair simulations
- Algo $_{BSP} \simeq BSP \mathsf{ASM} \simeq Imp_{BSP}$
- BSPlib is algorithmic complete

BSP-ASM

- Axiomatic definition of BSP algorithms and BSP ASM
- Using fair simulations
- Algo_{BSP} \simeq BSP-ASM \simeq Imp_{BSP}
- BSPlib is algorithmic complete

BSP-ASM

- Axiomatic definition of BSP algorithms and BSP ASM
- Using fair simulations
- Algo_{BSP} \simeq BSP-ASM \simeq Imp_{BSP}
- BSPlib is algorithmic complete

Perspectives (Ongoing/Future Work)

Application to other bridging models (Multi-BSP, etc.)

Enumerating what is BSP (BSML) and what is not (Prege

BSP-ASM

- Axiomatic definition of BSP algorithms and BSP ASM
- Using fair simulations
- Algo_{BSP} \simeq BSP-ASM \simeq Imp_{BSP}
- BSPlib is algorithmic complete

- Application to other bridging models (Multi-BSP, etc.)
- Enumerating what is BSP (BSML) and what is not (Pregel?)

BSP-ASM

- Axiomatic definition of BSP algorithms and BSP ASM
- Using fair simulations
- Algo_{BSP} \simeq BSP-ASM \simeq Imp_{BSP}
- BSPlib is algorithmic complete

- Application to other bridging models (Multi-BSP, etc.)
- Enumerating what is BSP (BSML) and what is not (Pregel?)

BSP-ASM

- Axiomatic definition of BSP algorithms and BSP ASM
- Using fair simulations
- Algo_{BSP} \simeq BSP-ASM \simeq Imp_{BSP}
- BSPlib is algorithmic complete

- Application to other bridging models (Multi-BSP, etc.)
- Enumerating what is BSP (BSML) and what is not (Pregel?)

Merci !