Verified Validation for Polyhedral Scheduling

Xuyang Li, Hongjin Liang, Xinyu Feng Nanjing University

Background

- Human needs highly optimized compilation techniques for modern software.
- <u>Nested loops</u> are optimization targets due to its heavy numerical computation.
- Common loop optimizations:
 - Apply combination of loop fusion, distribution, interchange, skewing, reverse, tiling, ..., to improve memory locality and/or parallelizability.
 - May further apply vectorization (SIMD) and parallelization (openmp primitives).
 - Map it automatically to domain-specific hardware like GPU.

Background

• To do such loop transformations, we need to analyze dependences between instructions of different iterations.

```
for i in [1, N]:
   for j in [1, N]:
S: A[i][j] = A[i+1][j] + A[i-1][j-1]
```

Can we parallelize it?

Example

for i in [1, N]:
 for j in [1, N]:
 S: A[i][j] = A[i+1][j] + A[i-1][j-1]





















Example



loop skewing + loop interchange



loop skewing + loop interchange



loop skewing + loop interchange







Our focus! Scheduling can be done manually or by automatic algorithms, like Pluto¹. We want to <u>ensure its correctness</u>.

Compilation correctness

- For compiler *Comp*, programs \mathcal{P}_s and \mathcal{P}_t that *Comp*(\mathcal{P}_s) = Some \mathcal{P}_t .
- If \mathcal{P}_t <u>refines</u> \mathcal{P}_s (written as $\mathcal{P}_t \sqsubseteq \mathcal{P}_s$), we say this compilation is correct.
 - It says, from the same beginning state, whenever \mathcal{P}_t terminates at some state, then \mathcal{P}_s is able to stop at the same final state.

Compilation correctness

- For compiler *Comp*, programs \mathcal{P}_s and \mathcal{P}_t that *Comp*(\mathcal{P}_s) = Some \mathcal{P}_t .
- If \mathcal{P}_t <u>refines</u> \mathcal{P}_s (written as $\mathcal{P}_t \sqsubseteq \mathcal{P}_s$), we say this compilation is correct.
 - It says, from the same beginning state, whenever \mathcal{P}_t terminates at some state, then \mathcal{P}_s is able to stop at the same final state.
- Two ways to guarantee correct compilation:
 - Compiler proof: reasoning on *Comp*'s concrete definition to prove

$$\forall \mathcal{P}_s, \mathcal{P}_t. \ \textit{Comp}(\mathcal{P}_s) = \texttt{Some} \ \mathcal{P}_t \implies \mathcal{P}_t \sqsubseteq \mathcal{P}_s.$$

• Verified validation: define a separate validator Validate and prove

$$\forall \mathcal{P}_s, \mathcal{P}_t. \ \textit{Validate}(\mathcal{P}_s, \mathcal{P}_t) = \texttt{true} \implies \mathcal{P}_t \sqsubseteq \mathcal{P}_s.$$

• And run *Validate* after *each run* of *Comp*.

Compilation correctness

Why not directly verify the scheduling algorithm?

- On the one hand, it contains complex heuristic with heavy mathematics. Hard/Impractical to verify.
- On the other hand, it has simple validation algorithm due to its simple correctness criterion: not breaking dependence. (Also called *Bernstein's Condition²*.)
- Two ways to guarantee correct compilation:
 - Compiler proof: reasoning on *Comp*'s concrete definition to prove

 $\forall \mathcal{P}_s, \mathcal{P}_t. \ \textit{Comp}(\mathcal{P}_s) = \texttt{Some} \ \mathcal{P}_t \implies \mathcal{P}_t \sqsubseteq \mathcal{P}_s.$

• Verified validation: define a separate validator Validate and prove

 $\forall \mathcal{P}_s, \mathcal{P}_t. \ \textit{Validate}(\mathcal{P}_s, \mathcal{P}_t) = \texttt{true} \implies \mathcal{P}_t \sqsubseteq \mathcal{P}_s.$

And run Validate after <u>each run</u> of Comp.



Validating Polyhedral Scheduling



Our work: Verified Validation for Polyhedral Scheduling

- We implement and verify a general validation function *Validate* for polyhedral scheduling in Coq proof assistant from scratch. It is parameterized by instruction language.
- We instantiate *Validate* with a variant of CompCert's instruction language, showing its practicality.
- We adapt *Validate* so that it works with the Pluto compiler. It successfully validates all its available test cases with reasonable overhead.

General Verified Validation function for Polyhedral Scheduling

- We define a validation function Validate that checks <u>Bernstein's</u> <u>condition</u> within polyhedral model, and mechanized its correctness with the Coq proof assistant.
- It is parametrized by instruction language to be reusable.
- Our top lemma

Definition (correctness of the validator)

$$\begin{array}{lll} \textit{Correct(Validate)} &\triangleq & \forall \mathcal{P}_s, \mathcal{P}_t. \textit{Validate}(\mathcal{P}_s, \mathcal{P}_t) = \texttt{true} \\ &\implies \mathcal{P}_t \sqsubseteq \mathcal{P}_s. \\ \mathcal{P}_t \sqsubseteq \mathcal{P}_s &\triangleq & \forall \sigma, \sigma'. \\ &\models \mathcal{P}_t, \sigma \Rightarrow \sigma' \implies \models \mathcal{P}_s, \sigma \Rightarrow \sigma' \end{array}$$

Case study on CompCert

• What is CompCert³?



- a high-assurance compiler for almost all of the C language (ISO C 2011), generating efficient code for the ARM, PowerPC, RISC-V and x86 processors.
- It is not enough optimized⁴ than production compiler like clang and gcc. So aggressive optimizers like polyhedral-based ones do help!
- We successfully instantiate Validate with CompCert's semantics model, showing the possibility towards a fully verified polyhedral extension to CompCert.

³ https://compcert.org/

⁴ https://doi.org/10.1145/3622799

Case study on Pluto compiler

- Loop optimizers like polyhedral-based ones are error prone⁵ indeed!
 So formal methods do help.
- We evaluate on Pluto, one of the famous polyhedral compiler.
 - Pluto: ACM SIGPLAN PLDI Most Influential Paper award in 2018



Figure from https://www.csa.iisc.ac.in/~udayb/publications/uday-thesis.pdf, Page 98

⁵ https://doi.org/10.1145/3591295

Case study on Pluto compiler

- Loop optimizers like polyhedral-based ones are error prone⁵ indeed!
 So formal methods do help.
- We evaluate on Pluto, one of the famous polyhedral compiler.
 - Pluto: ACM SIGPLAN PLDI Most Influential Paper award in 2018



Case study on Pluto compiler

- Loop optimizers like polyhedral-based ones are error prone⁵ indeed!
 So formal methods do help.
- We evaluate on Pluto, one of the famous polyhedral compiler.
 - Pluto: ACM SIGPLAN PLDI Most Influential Paper award in 2018
- Work and work well with Pluto compiler: successfully validate its all 62 available test cases.
 - Not just a research prototype: work as an executable on practical polyhedral compiler.
 - Soundness v.s. completeness:
 - Formal verification proves the algorithm <u>soundness</u>: guarantee <u>refinement</u> whenever it outputs <u>true</u>.
 - The evaluation shows validation's algorithm <u>completeness</u>: if refinement establishes on inputs, the algorithm tries its best to output <u>true</u> rather than <u>unknown</u>!
 - Also, the algorithm has reasonable overhead.

Ca	se	S

•	Loop opti
	So forma

- We evalue
 - Pluto: /
- Work and available
 - Not jus compile
 - Soundr
 - For out
 The
 - the
 - Also, th

Test	Time of Pluto (ms)	Time of Validation (ms,ms)	Result	
covcol	3.5	434.6, 320.7	EQ	
dsyr2k	2.6	106.0, 83.4	EQ	
fdtd-2d	46.4	1615.5, 1296.3	EQ	
gemver	7.0	247.9, 240.4	EQ	
lu	6.1	410.6, 331.2	EQ	
mvt	2.2	70.2, 56.3	EQ	
ssymm	40.7	726.0, 551.2	EQ	
tce	568.6	4442.0, 4422.5	EQ	
adi	77.5	2531.7, 2377.8	EQ	iedl
corcol	5.5	442.5, 362.1	EQ	
dct	21.8	879.4, 739.4	EQ	
dsyrk	1.8	96.8, 78.9	EQ	
floyd	12.1	502.6, 421.7	EQ	
jacobi-1d-imper	3.8	184.0, 167.8	EQ	
matmul-init	2.9	257.8, 192.4	EQ	
pca	202.5	2923.6, 2679.5	EQ	
strmm	1.9	141.4, 110.8	EQ	all 62
tmm	1.6	109.7, 89.6	EQ	
advect3d	1023.1	579.1, 498.1	EQ	
corcol3	13.6	851.3, 733.4	EQ	dral
doitgen	10.4	1069.2, 837.4	EQ	
fdtd-1d	6.0	268.7, 229.9	EQ	
jacobi-2d-imper	17.7	619.5, 543.5	EQ	
matmul	3.2	157.1, 125.5	EQ	enever it
seidel	24.5	818.1, 725.5	EQ	
strsm	6.4	209.3, 161.2	EQ	ichoc on innutc
trisolv	5.1	338.9, 248.8	EQ	isnes on inputs,
1dloop-invar	0.3	6.7, 6.0	EQ	
costfunc	0.8	47.4, 35.0	EQ	
fusion1	0.9	15.3, 13.9	EQ	
	•			

Table 1: Evaluation results on Pluto's test suits

Coq development

- Around 18000 lines of Coq, 1000 lines of OCaml. Open source at https://github.com/verif-scop/PolCert.
- Our work bases on Verified Polyhedron Library⁶ and is syntactically compatible to PolyGen⁷ (Verified Polyhedral Code Generation) in POPL'21.

Future work

- Complete verified polyhedral compilation.
 - Verified extractor.
 - Engineering in CompCert's driver & frontend.
 - Apply optimistic approach⁸ to deal with polyhedral model's heavy assumptions, like integer overflow⁹.
- Support validation for other polyhedral transformations, like index set split (as a pre-phrase), tiling (as a post-phrase), layout transformation (as an orthogonal phrase).
- Support vectorization, parallelization, GPU compilation ...

Q&A

Our work: Verified Validation for Polyhedral Scheduling

- We implement and verify a general validation function *Validate* for polyhedral scheduling in Coq proof assistant from scratch. It is parameterized by instruction language.
- We instantiate *Validate* with a variant of CompCert's instruction language, showing its practicality.
- We adapt *Validate* so that it works with the Pluto compiler. It successfully validates all its available test cases with reasonable overhead.

Original code π_{cov} for covariance matrix calculation, 1.84s

Optimized code π'_{cov} for covariance matrix calculation, 0.43s, with loop distribution and loop interchange

```
for (i = 1; i \le N; i++)
 1
      for (j1 = 1; j1 \le M; j1++) {
 2
      for (j_2 = j_1; j_2 \le M; j_2++) {
 3
            symmat[j1][j2] += data[i][j1] * data[i][j2];
 4
 5
 6
 7
   for (j1 = 1; j1 \le M; j1++)
 8
      for (j_2 = j_1; j_2 \le M; j_2++)
        symmat[j2][j1] = symmat[j1][j2];
10
11
12
    }
```

See at https://github.com/verif-scop/speed-up.

Polyhedral compilation does high-level structural tranformations and only impose a few properties of the underlying instruction language (called I). The validation function given in this work is parameterized by I.

I allows user define the syntax, types, state, semantics of the language, how it initializes, and its the non-alias proposition. It demands user to provide a verified *Checker* function to validate the consistency between the read and write access function and an instruction's semantics, and prove that any two instances that satisfy Bernstein's conditions are permutable.

We assume Pluto's instruction language satisfy this abstraction. Module Type I ≜ $\begin{array}{l} \mathsf{T},\mathsf{I},\mathsf{S}:\textit{Type} \\ \models :\mathsf{I} \to \textit{List}(\mathbb{Z}) & \to \textit{Memory Cells} \\ & \to \textit{Memory Cells} \to \mathsf{S} \to \mathsf{S} \to \textit{Prop} \end{array}$ $\mathsf{Compat}: \mathit{List}(\mathit{Identifier}) \to \mathsf{S} \to \mathit{Prop}$ Consistent : $List(Identifier \times T) \rightarrow List(\mathbb{Z}) \rightarrow S \rightarrow Prop$ NonAlias : $S \rightarrow Prop$ NonAliasPsrv : $\forall I, \sigma, \sigma'$. NonAlias $(\sigma) \land p \models I, \sigma \xrightarrow{-'} \sigma' \implies$ NonAlias (σ') . Checker : I \rightarrow Access Functions \rightarrow Access Functions \rightarrow Bool Correct(Checker) : $\forall I, W, \mathcal{R}. Checker(I, W, \mathcal{R}) = true$ $\implies (\forall \sigma, \sigma', \mathbf{p}, \Delta_r, \Delta_w, \mathbf{p} \models \mathbf{I}, \sigma \xrightarrow{\Delta_r, \Delta_w} \sigma' \implies \Delta_r \subseteq \mathcal{R}(\mathbf{p}) \land \Delta_w \subseteq \mathcal{W}(\mathbf{p})).$ BCPermut : $\forall \mathbf{I_1}, \mathbf{I_2}, \boldsymbol{p_1}, \boldsymbol{p_2}, \sigma, \sigma', \sigma'', \Delta_r, \Delta_w, \Delta_r', \Delta_w'.$ $(\mathbf{p_1} \models \mathbf{I_1}, \sigma \xrightarrow{\Delta_r, \Delta_w} \sigma' \land \mathbf{p_2} \models \mathbf{I_2}, \sigma' \xrightarrow{\Delta'_r, \Delta'_w} \sigma''$ $\wedge \Delta_r \cap \Delta'_w = \emptyset \land \Delta_w \cap \Delta'_r = \emptyset \land \Delta_w \cap \Delta'_w = \emptyset)$ $\Rightarrow \exists \sigma^* . \ p_2 \models \mathbf{I}_2, \sigma \xrightarrow{\Delta'_r, \Delta'_w} \sigma^* \land p_1 \models \mathbf{I}_1, \sigma^* \xrightarrow{\Delta_r, \Delta_w} \sigma''.$

Figure 1: Definition of Instruction Language Module I

Case study: Towards verified polyhedral compilation for CompCert

We instantiate the validation function with CompCert C's type, state and subset of its instruction language, and implement *Checker* with symbolic execution. All verified. Only differences are, affine expression is evaluated in \mathbb{Z} (no overflow), and multi-dimensional array access is sugarized.

(Base Type)	$ au_{\perp}$::=	int32s
(Type)	au	E	Base Type \times List(\mathbb{Z})
(Value)	V	::=	$I32(n) \dots$
(Iterator)	i	e	\mathbb{N}
(Unop)	op ₁	::=	
(Binop)	op ₂	::=	+ *
(May Affine Expression)	ε	::=	$z \mid i \mid op_1 \varepsilon \mid \varepsilon_1 op_2 \varepsilon_2$
(Access Expression)	ϵ	E	Identifier ×Base Type
			×List(May Affine Expression)
(Expression)	е	::=	$v \mid i \mid \epsilon \mid op_1 \mid e \mid e_1 \mid op_2 \mid e_2$
(Base Instruction)	I	::=	skip $ \epsilon := e$

• Extraction -> Scheduling -> Codegen

loop reverse + loop fusion



- An alternative integer-programming-based representation for nested loop. A polyhedral program is a multiset of polyhedral instructions, which consists of at least three parts: the **instruction**, the **domain**, and the schedule.
- "Extraction"

Lexicographically rdered! ~> for i in [0, 3]: S: B[i] = A[i] + 1 {(S, 0<=i<3, [i]~>[0,i]), $(T, 0 \le i \le 3, [i] \ge [1, i]) \}$ for i in [0, 3]: T: C[i] = B[2-i] * 3

Execution order: S[0]; S[1]; S[2]; T[0]; T[1]; T[2]



Our work's focus!

- "Scheduling": do reordering transformation on polyhedral model, guided by dependence analysis.
- Well-designed automatic algorithms, like Pluto², serve for this purpose.
- Manual scheduling is also possible.
- Recall the correctness criterion: Bernstein's condition.

{(S, 0<=i<3, [i]~>[0,i]), ~> {(S, 0<=i<3, [i]~>...), (T, 0<=i<3, [i]~>[1,i])} (T, 0<=i<3, [i]~>...)}

Old Execution order: S[0]; S[1]; S[2]; T[0]; T[1]; T[2]



Our work's focus!

- "Scheduling": do reordering transformation on polyhedral model, guided by dependence analysis.
- Well-designed automatic algorithms, like Pluto², serve for this purpose.
- Manual scheduling is also possible.

Average distance of dependence's satisfaction seems too long. Cache locality.

Old Execution order: S[0]; S[1]; S[2]; T[0]; T[1]; T[2]

S[0] S[1] S[2] model T[0] T[1] T[2] B[0] B[1] B[2]

Our work's focus!

- "Scheduling": do reordering transformation on polyhedral model, guided by dependence analysis.
- Well-designed automatic algorithms, like Pluto², serve for this purpose.
- Manual scheduling is also possible.

Average distance of dependence's satisfaction seems to long.

Try to minimize the distance without breaking dependences!

Old Execution order: S[0];S[1];S[2];T[0];T[1];T[2] **New execution order**: S[0];T[2];S[1];T[1];S[2];T[0]

² https://pluto-compiler.sourceforge.net/



Our work's focus!

- "Scheduling": do reordering transformation on polyhedral model, guided by dependence analysis.
- Well-designed automatic algorithms, like Pluto², serve for this purpose.
- Manual scheduling is also possible.
- Recall the correctness criterion: Bernstein's condition.

{(S, 0<=i<3, [i]~>[0,i]), ~> {(S, 0<=i<3, [i]~>[i,0]), (T, 0<=i<3, [i]~>[1,i])} (T, 0<=i<3, [i]~>[2-i,1])}

Old Execution order: S[0];S[1];S[2];T[0];T[1];T[2] **New execution order**: S[0];T[2];S[1];T[1];S[2];T[0]

• "Codegen": recover the imperative control structure from the optimized polyhedral model.

~>

{ (S, 0<=i<3, [i]~>[i,0]), (T, 0<=i<3, [i]~>[2-i,1]) } for i in [0, 3]:
S: B[i] = A[i] + 1
T: C[2-i] = B[i] * 3