Inlining Control-Flow Jumps in Library Usage Graphs of Legacy Code

Presentation Abstract

Ruben Opdebeeck *ropdebee@vub.be*

Johan Fabry *johan@raincode.com*

Coen De Roover *cderoove@vub.be*

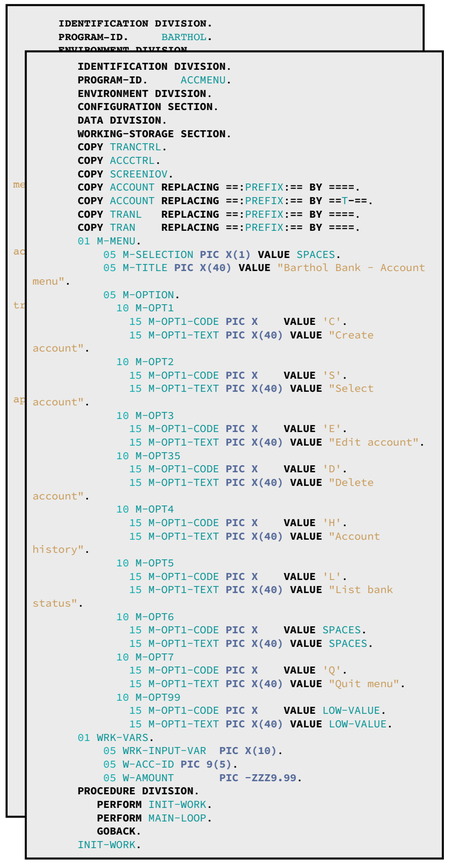
VRIJE UNIVERSITEIT BRUSSEL

Software Languages Lab

raincode LABS compiler experts

INTiMALS

COBOL Library Usage Pattern Mining



```
       IDENTIFICATION DIVISION.
       PROGRAM-ID.     BARTHOL.
       ENVIRONMENT DIVISION.
```

```
       IDENTIFICATION DIVISION.
       PROGRAM-ID.     ACCMENU.
       ENVIRONMENT DIVISION.
       CONFIGURATION SECTION.
       DATA DIVISION.
       WORKING-STORAGE SECTION.
       COPY TRANCTRL.
       COPY ACCCTRL.
       COPY SCREENIOV.
       COPY ACCOUNT REPLACING ==:PREFIX:== BY ====.
       COPY ACCOUNT REPLACING ==:PREFIX:== BY ==T-==.
       COPY TRANL   REPLACING ==:PREFIX:== BY ====.
       COPY TRAN    REPLACING ==:PREFIX:== BY ====.
       01 M-MENU.
           05 M-SELECTION PIC X(1) VALUE SPACES.
           05 M-TITLE PIC X(40) VALUE "Barthol Bank - Account
menu".
           05 M-OPTION.
             10 M-OPT1
               15 M-OPT1-CODE PIC X    VALUE 'C'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Create
account".
             10 M-OPT2
               15 M-OPT1-CODE PIC X    VALUE 'S'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Select
account".
             10 M-OPT3
               15 M-OPT1-CODE PIC X    VALUE 'E'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Edit account".
             10 M-OPT35
               15 M-OPT1-CODE PIC X    VALUE 'D'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Delete
account".
             10 M-OPT4
               15 M-OPT1-CODE PIC X    VALUE 'H'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Account
history".
             10 M-OPT5
               15 M-OPT1-CODE PIC X    VALUE 'L'.
               15 M-OPT1-TEXT PIC X(40) VALUE "List bank
status".
             10 M-OPT6
               15 M-OPT1-CODE PIC X    VALUE SPACES.
               15 M-OPT1-TEXT PIC X(40) VALUE SPACES.
             10 M-OPT7
               15 M-OPT1-CODE PIC X    VALUE 'Q'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Quit menu".
             10 M-OPT99
               15 M-OPT1-CODE PIC X    VALUE LOW-VALUE.
               15 M-OPT1-TEXT PIC X(40) VALUE LOW-VALUE.
       01 WRK-VARS.
           05 WRK-INPUT-VAR  PIC X(10).
           05 W-ACC-ID PIC 9(5).
           05 W-AMOUNT       PIC -ZZZ9.99.
       PROCEDURE DIVISION.
          PERFORM INIT-WORK.
          PERFORM MAIN-LOOP.
          GOBACK.
       INIT-WORK.
```

	# Programs	# KLOC
Case 1	305	662.2
Case 2	3926	22889

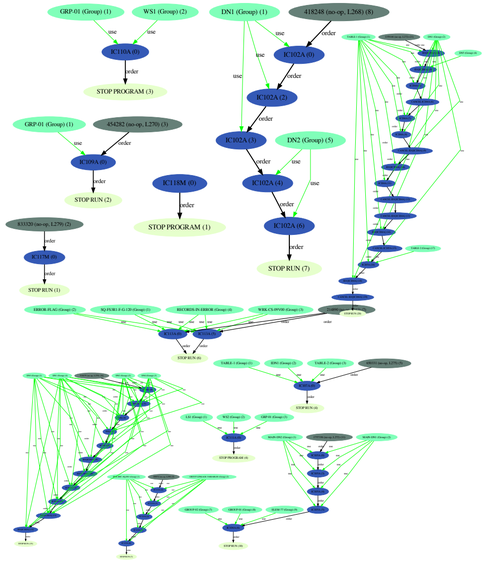
Goal: Understand library usages
to estimate modernisation effort

COBOL Library Usage Pattern Mining

```
       IDENTIFICATION DIVISION.
       PROGRAM-ID.     ACCMENU.
       ENVIRONMENT DIVISION.
       CONFIGURATION SECTION.
       DATA DIVISION.
       WORKING-STORAGE SECTION.
       COPY TRANCTRL.
       COPY ACCCTRL.
       COPY SCREENIOV.
       COPY ACCOUNT REPLACING ==:PREFIX:== BY ====.
       COPY ACCOUNT REPLACING ==:PREFIX:== BY ==T-==.
       COPY TRANL   REPLACING ==:PREFIX:== BY ====.
       COPY TRAN    REPLACING ==:PREFIX:== BY ====.
       01 M-MENU.
           05 M-SELECTION PIC X(1) VALUE SPACES.
           05 M-TITLE PIC X(40) VALUE "Barthol Bank - Account
menu".
           05 M-OPTION.
             10 M-OPT1
               15 M-OPT1-CODE PIC X    VALUE 'C'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Create
account".
             10 M-OPT2
               15 M-OPT1-CODE PIC X    VALUE 'S'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Select
account".
             10 M-OPT3
               15 M-OPT1-CODE PIC X    VALUE 'E'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Edit account".
             10 M-OPT35
               15 M-OPT1-CODE PIC X    VALUE 'D'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Delete
account".
             10 M-OPT4
               15 M-OPT1-CODE PIC X    VALUE 'H'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Account
history".
             10 M-OPT5
               15 M-OPT1-CODE PIC X    VALUE 'L'.
               15 M-OPT1-TEXT PIC X(40) VALUE "List bank
status".
             10 M-OPT6
               15 M-OPT1-CODE PIC X    VALUE SPACES.
               15 M-OPT1-TEXT PIC X(40) VALUE SPACES.
             10 M-OPT7
               15 M-OPT1-CODE PIC X    VALUE 'Q'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Quit menu".
             10 M-OPT99
               15 M-OPT1-CODE PIC X    VALUE LOW-VALUE.
               15 M-OPT1-TEXT PIC X(40) VALUE LOW-VALUE.
       01 WRK-VARS.
           05 WRK-INPUT-VAR  PIC X(10).
           05 W-ACC-ID PIC 9(5).
           05 W-AMOUNT       PIC -ZZZ9.99.
       PROCEDURE DIVISION.
          PERFORM INIT-WORK.
          PERFORM MAIN-LOOP.
          GOBACK.
       INIT-WORK.
```

	# Programs	# KLOC
Case 1	305	662.2
Case 2	3926	22889

Library Usages



Case	Time (s)	# Groums	Avg. size	Max. size
Case 1	43.8	273	244	4634
Case 2 (full)	1019	3925	4769	1217422
Case 2 (limited)	N/A	3573	125	9644

Goal: Understand library usages
to estimate modernisation effort

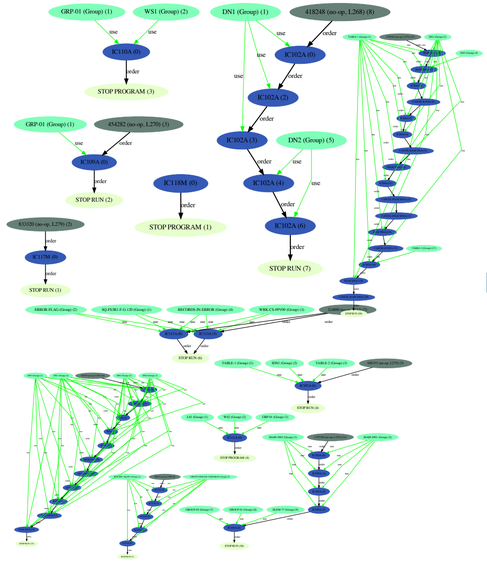
COBOL Library Usage Pattern Mining

```
IDENTIFICATION DIVISION.
PROGRAM-ID.     BARTHOL.
ENVIRONMENT DIVISION.
```

```
       IDENTIFICATION DIVISION.
       PROGRAM-ID.     ACCMENU.
       ENVIRONMENT DIVISION.
       CONFIGURATION SECTION.
       DATA DIVISION.
       WORKING-STORAGE SECTION.
       COPY TRANCTRL.
       COPY ACCCTRL.
       COPY SCREENIOV.
       COPY ACCOUNT REPLACING ==:PREFIX:== BY ====.
       COPY ACCOUNT REPLACING ==:PREFIX:== BY ==T-==.
       COPY TRANL   REPLACING ==:PREFIX:== BY ====.
       COPY TRAN    REPLACING ==:PREFIX:== BY ====.
       01 M-MENU.
           05 M-SELECTION PIC X(1) VALUE SPACES.
           05 M-TITLE PIC X(40) VALUE "Barthol Bank - Account
menu".
           05 M-OPTION.
             10 M-OPT1
               15 M-OPT1-CODE PIC X    VALUE 'C'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Create
account".
             10 M-OPT2
               15 M-OPT1-CODE PIC X    VALUE 'S'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Select
account".
             10 M-OPT3
               15 M-OPT1-CODE PIC X    VALUE 'E'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Edit account".
             10 M-OPT35
               15 M-OPT1-CODE PIC X    VALUE 'D'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Delete
account".
             10 M-OPT4
               15 M-OPT1-CODE PIC X    VALUE 'H'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Account
history".
             10 M-OPT5
               15 M-OPT1-CODE PIC X    VALUE 'L'.
               15 M-OPT1-TEXT PIC X(40) VALUE "List bank
status".
             10 M-OPT6
               15 M-OPT1-CODE PIC X    VALUE SPACES.
               15 M-OPT1-TEXT PIC X(40) VALUE SPACES.
             10 M-OPT7
               15 M-OPT1-CODE PIC X    VALUE 'Q'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Quit menu".
             10 M-OPT99
               15 M-OPT1-CODE PIC X    VALUE LOW-VALUE.
               15 M-OPT1-TEXT PIC X(40) VALUE LOW-VALUE.
       01 WRK-VARS.
           05 WRK-INPUT-VAR  PIC X(10).
           05 W-ACC-ID PIC 9(5).
           05 W-AMOUNT       PIC -ZZZ9.99.
       PROCEDURE DIVISION.
          PERFORM INIT-WORK.
          PERFORM MAIN-LOOP.
          GOBACK.
       INIT-WORK.
```

	# Programs	# KLOC
Case 1	305	662.2
Case 2	3926	22889

Library Usages



Case	Time (s)	# Groums	Avg. size	Max. size
Case 1	43.8	273	244	4634
Case 2 (full)	1019	3925	4769	1217422
Case 2 (limited)	N/A	3573	125	9644

Pattern Mining

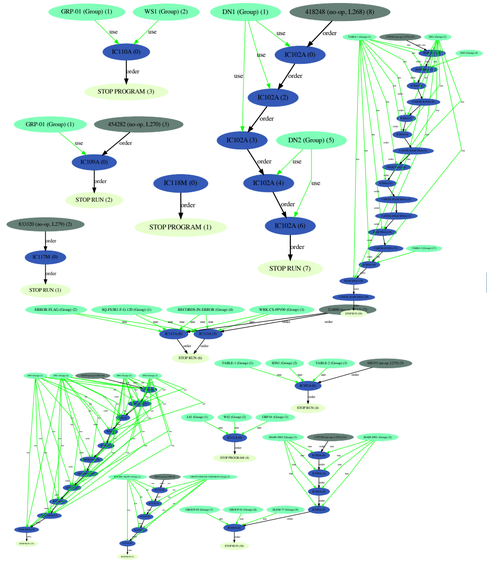
Goal: Understand library usages
to estimate modernisation effort

COBOL Library Usage Pattern Mining

```
      IDENTIFICATION DIVISION.
      PROGRAM-ID.     BARTHOL.
      ENVIRONMENT DIVISION.
      CONFIGURATION SECTION.
      DATA DIVISION.
      WORKING-STORAGE SECTION.
      COPY TRANCTRL.
      COPY ACCCTRL.
      COPY SCREENIOV.
      COPY ACCOUNT REPLACING ==:PREFIX:== BY ====.
      COPY ACCOUNT REPLACING ==:PREFIX:== BY ==T-==.
      COPY TRANL   REPLACING ==:PREFIX:== BY ====.
      COPY TRAN    REPLACING ==:PREFIX:== BY ====.
      01 M-MENU.
          05 M-SELECTION PIC X(1) VALUE SPACES.
          05 M-TITLE PIC X(40) VALUE "Barthol Bank - Account
menu".
          05 M-OPTION.
            10 M-OPT1
              15 M-OPT1-CODE PIC X    VALUE 'C'.
              15 M-OPT1-TEXT PIC X(40) VALUE "Create
account".
            10 M-OPT2
              15 M-OPT1-CODE PIC X    VALUE 'S'.
              15 M-OPT1-TEXT PIC X(40) VALUE "Select
account".
            10 M-OPT3
              15 M-OPT1-CODE PIC X    VALUE 'E'.
              15 M-OPT1-TEXT PIC X(40) VALUE "Edit account".
            10 M-OPT35
              15 M-OPT1-CODE PIC X    VALUE 'D'.
              15 M-OPT1-TEXT PIC X(40) VALUE "Delete
account".
            10 M-OPT4
              15 M-OPT1-CODE PIC X    VALUE 'H'.
              15 M-OPT1-TEXT PIC X(40) VALUE "Account
history".
            10 M-OPT5
              15 M-OPT1-CODE PIC X    VALUE 'L'.
              15 M-OPT1-TEXT PIC X(40) VALUE "List bank
status".
            10 M-OPT6
              15 M-OPT1-CODE PIC X    VALUE SPACES.
              15 M-OPT1-TEXT PIC X(40) VALUE SPACES.
            10 M-OPT7
              15 M-OPT1-CODE PIC X    VALUE 'Q'.
              15 M-OPT1-TEXT PIC X(40) VALUE "Quit menu".
            10 M-OPT99
              15 M-OPT1-CODE PIC X    VALUE LOW-VALUE.
              15 M-OPT1-TEXT PIC X(40) VALUE LOW-VALUE.
      01 WRK-VARS.
          05 WRK-INPUT-VAR  PIC X(10).
          05 W-ACC-ID PIC 9(5).
          05 W-AMOUNT       PIC -ZZZ9.99.
      PROCEDURE DIVISION.
         PERFORM INIT-WORK.
         PERFORM MAIN-LOOP.
         GOBACK.
      INIT-WORK.
```

	# Programs	# KLOC
Case 1	305	662.2
Case 2	3926	22889

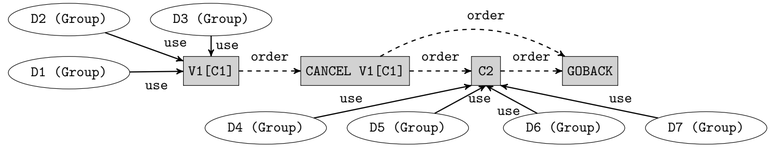
Library Usages



Case	Time (s)	# Groums	Avg. size	Max. size
Case 1	43.8	273	244	4634
Case 2 (full)	1019	3925	4769	1217422
Case 2 (limited)	N/A	3573	125	9644

Pattern Mining

Library Usage Patterns



Corpus	GrouMiner*		BigGroum		
	Time (min)	# P	Time (min)	# Part.	# P
Corpus 1	0.95	3	10.4	3	4
Corpus 2	121.8	7	618.2	1	0

Goal: Understand library usages
to estimate modernisation effort

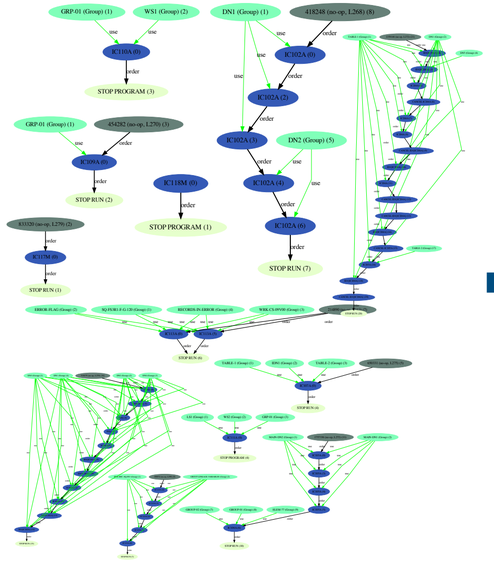
COBOL Library Usage Pattern Mining

```
IDENTIFICATION DIVISION.
PROGRAM-ID.     BARTHOL.
ENVIRONMENT DIVISION.
```

```
       IDENTIFICATION DIVISION.
       PROGRAM-ID.     ACCMENU.
       ENVIRONMENT DIVISION.
       CONFIGURATION SECTION.
       DATA DIVISION.
       WORKING-STORAGE SECTION.
       COPY TRANCTRL.
       COPY ACCCTRL.
       COPY SCREENIOV.
       COPY ACCOUNT REPLACING ==:PREFIX:== BY ====.
       COPY ACCOUNT REPLACING ==:PREFIX:== BY ==T-==.
       COPY TRANL   REPLACING ==:PREFIX:== BY ====.
       COPY TRAN    REPLACING ==:PREFIX:== BY ====.
       01 M-MENU.
           05 M-SELECTION PIC X(1) VALUE SPACES.
           05 M-TITLE PIC X(40) VALUE "Barthol Bank - Account
menu".
           05 M-OPTION.
             10 M-OPT1
               15 M-OPT1-CODE PIC X    VALUE 'C'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Create
account".
             10 M-OPT2
               15 M-OPT1-CODE PIC X    VALUE 'S'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Select
account".
             10 M-OPT3
               15 M-OPT1-CODE PIC X    VALUE 'E'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Edit account".
             10 M-OPT35
               15 M-OPT1-CODE PIC X    VALUE 'D'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Delete
account".
             10 M-OPT4
               15 M-OPT1-CODE PIC X    VALUE 'H'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Account
history".
             10 M-OPT5
               15 M-OPT1-CODE PIC X    VALUE 'L'.
               15 M-OPT1-TEXT PIC X(40) VALUE "List bank
status".
             10 M-OPT6
               15 M-OPT1-CODE PIC X    VALUE SPACES.
               15 M-OPT1-TEXT PIC X(40) VALUE SPACES.
             10 M-OPT7
               15 M-OPT1-CODE PIC X    VALUE 'Q'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Quit menu".
             10 M-OPT99
               15 M-OPT1-CODE PIC X    VALUE LOW-VALUE.
               15 M-OPT1-TEXT PIC X(40) VALUE LOW-VALUE.
       01 WRK-VARS.
           05 WRK-INPUT-VAR  PIC X(10).
           05 W-ACC-ID PIC 9(5).
           05 W-AMOUNT       PIC -ZZZ9.99.
       PROCEDURE DIVISION.
          PERFORM INIT-WORK.
          PERFORM MAIN-LOOP.
          GOBACK.
       INIT-WORK.
```

	# Programs	# KLOC
Case 1	305	662.2
Case 2	3926	22889

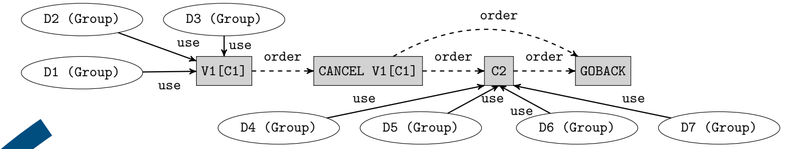
Library Usages



Case	Time (s)	# Groums	Avg. size	Max. size
Case 1	43.8	273	244	4634
Case 2 (full)	1019	3925	4769	1217422
Case 2 (limited)	N/A	3573	125	9644

Pattern Mining

Library Usage Patterns



Corpus	GrouMiner*		BigGroum		
	Time (min)	# P	Time (min)	# Part.	# P
Corpus 1	0.95	3	10.4	3	4
Corpus 2	121.8	7	618.2	1	0

Goal: Understand library usages
to estimate modernisation effort

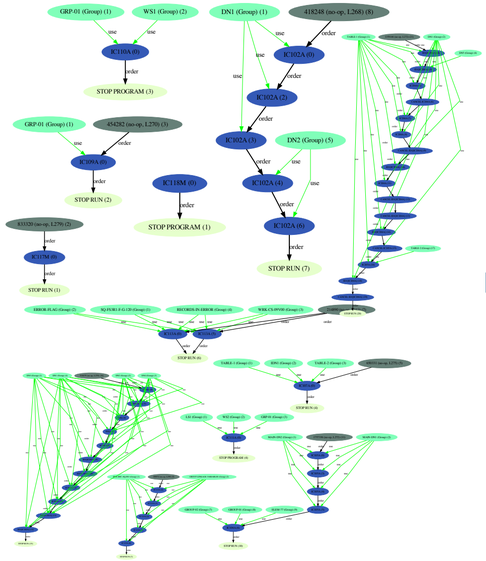
COBOL Library Usage Pattern Mining

```
       IDENTIFICATION DIVISION.
       PROGRAM-ID.     BARTHOL.
       ENVIRONMENT DIVISION.
```

```
       IDENTIFICATION DIVISION.
       PROGRAM-ID.     ACCMENU.
       ENVIRONMENT DIVISION.
       CONFIGURATION SECTION.
       DATA DIVISION.
       WORKING-STORAGE SECTION.
       COPY TRANCTRL.
       COPY ACCCTRL.
       COPY SCREENIOV.
       COPY ACCOUNT REPLACING ==:PREFIX:== BY ====.
       COPY ACCOUNT REPLACING ==:PREFIX:== BY ==T-==.
       COPY TRANL   REPLACING ==:PREFIX:== BY ====.
       COPY TRAN    REPLACING ==:PREFIX:== BY ====.
       01 M-MENU.
           05 M-SELECTION PIC X(1) VALUE SPACES.
           05 M-TITLE PIC X(40) VALUE "Barthol Bank - Account
menu".
           05 M-OPTION.
             10 M-OPT1
               15 M-OPT1-CODE PIC X     VALUE 'C'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Create
account".
             10 M-OPT2
               15 M-OPT1-CODE PIC X     VALUE 'S'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Select
account".
             10 M-OPT3
               15 M-OPT1-CODE PIC X     VALUE 'E'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Edit account".
             10 M-OPT35
               15 M-OPT1-CODE PIC X     VALUE 'D'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Delete
account".
             10 M-OPT4
               15 M-OPT1-CODE PIC X     VALUE 'H'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Account
history".
             10 M-OPT5
               15 M-OPT1-CODE PIC X     VALUE 'L'.
               15 M-OPT1-TEXT PIC X(40) VALUE "List bank
status".
             10 M-OPT6
               15 M-OPT1-CODE PIC X     VALUE SPACES.
               15 M-OPT1-TEXT PIC X(40) VALUE SPACES.
             10 M-OPT7
               15 M-OPT1-CODE PIC X     VALUE 'Q'.
               15 M-OPT1-TEXT PIC X(40) VALUE "Quit menu".
             10 M-OPT99
               15 M-OPT1-CODE PIC X     VALUE LOW-VALUE.
               15 M-OPT1-TEXT PIC X(40) VALUE LOW-VALUE.
       01 WRK-VARS.
           05 WRK-INPUT-VAR  PIC X(10).
           05 W-ACC-ID PIC 9(5).
           05 W-AMOUNT       PIC -ZZZ9.99.
       PROCEDURE DIVISION.
          PERFORM INIT-WORK.
          PERFORM MAIN-LOOP.
          GOBACK.
       INIT-WORK.
```

	# Programs	# KLOC
Case 1	305	662.2
Case 2	3926	22889

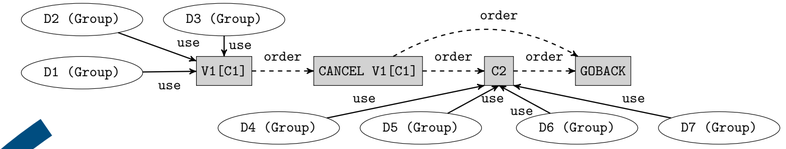
Library Usages



Case	Time (s)	# Groums	Avg. size	Max. size
Case 1	43.8	273	244	4634
Case 2 (full)	1019	3925	4769	1217422
Case 2 (limited)	N/A	3573	125	9644

Pattern Mining

Library Usage Patterns



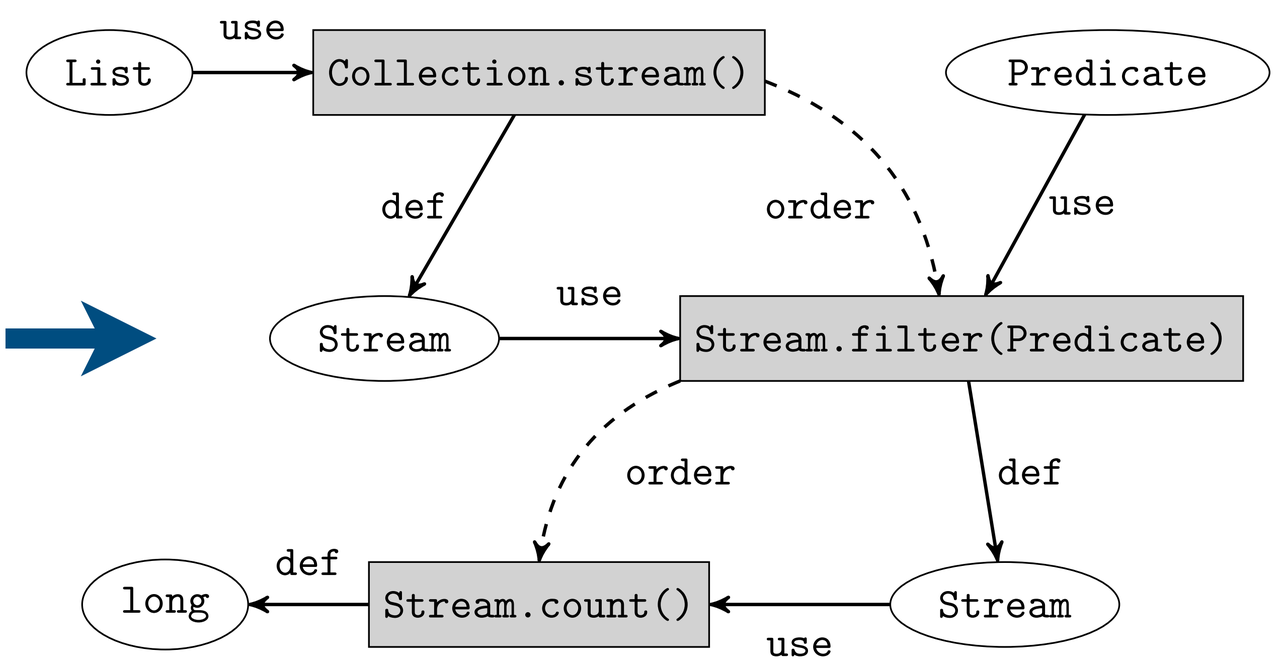
Corpus	GrouMiner*		BigGroum		
	Time (min)	# P	Time (min)	# Part.	# P
Corpus 1	0.95	3	10.4	3	4
Corpus 2	121.8	7	618.2	1	0

Goal: Understand library usages
to estimate modernisation effort

Today

Groums for OO languages

```
void my_method(List<String> lst) {
    return lst.stream()
        .filter(s -> s.startsWith("test_"))
        .count();
}
```

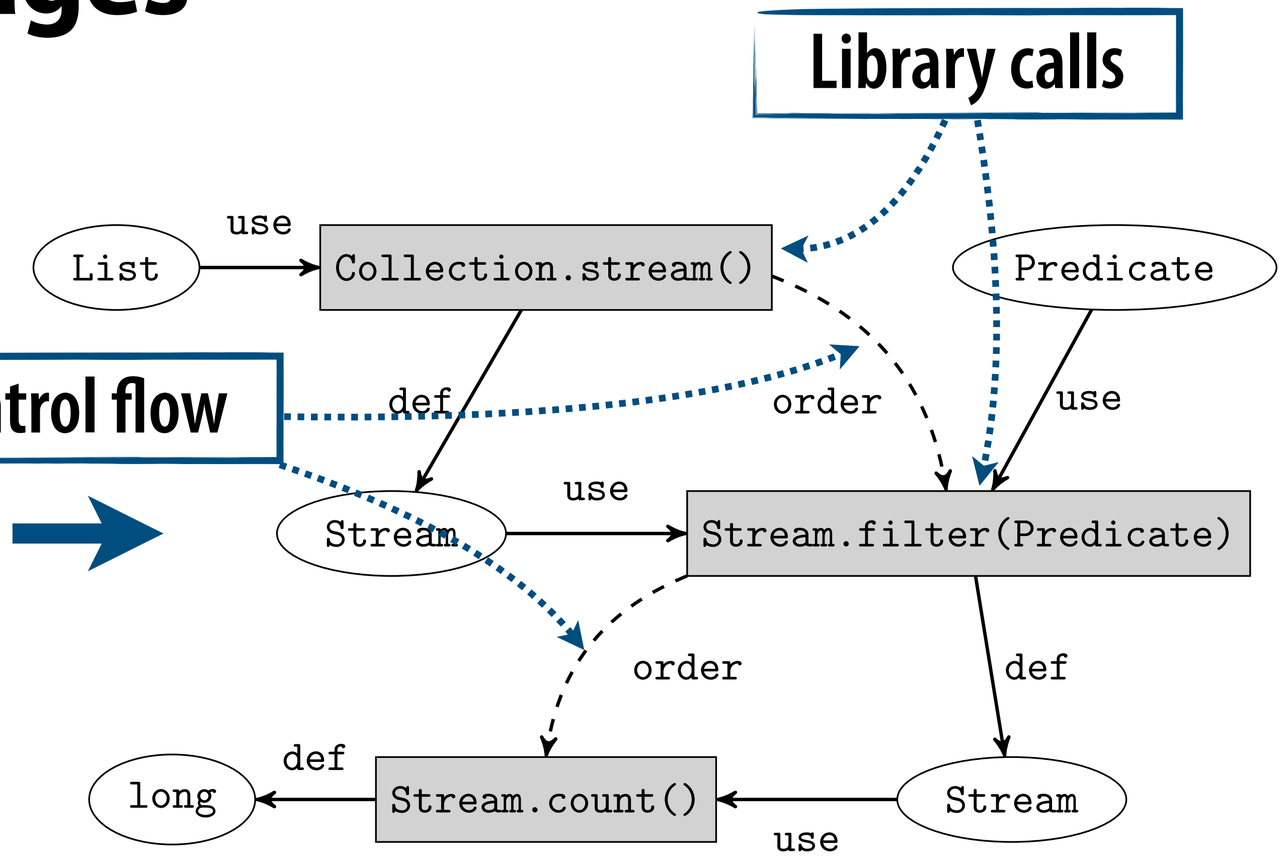


Java snippet using Collection API

Graph-based object usage model

Groums for OO languages

```
void my_method(List<String> lst) {
    return lst.stream()
        .filter(s -> s.startsWith("test_"))
        .count();
}
```

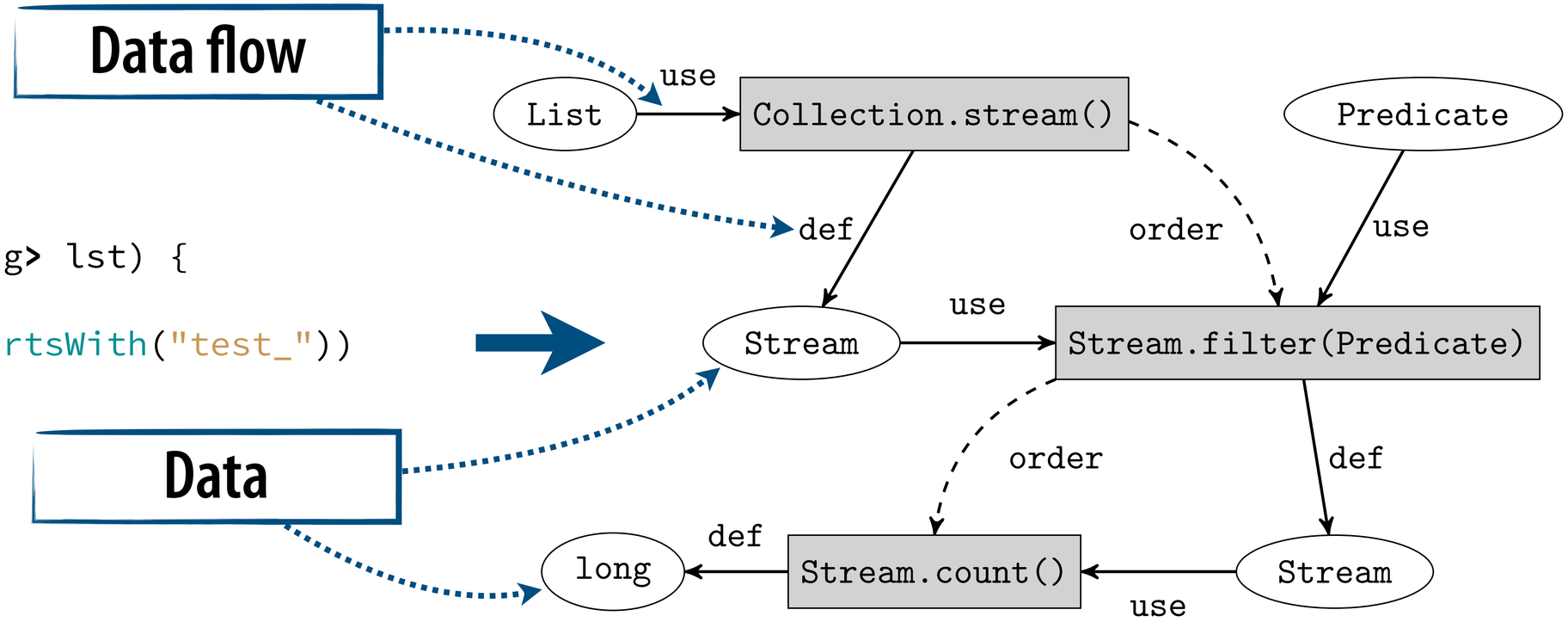


Java snippet using Collection API

Graph-based object usage model

Groums for OO languages

```
void my_method(List<String> lst) {
    return lst.stream()
        .filter(s -> s.startsWith("test_"))
        .count();
}
```



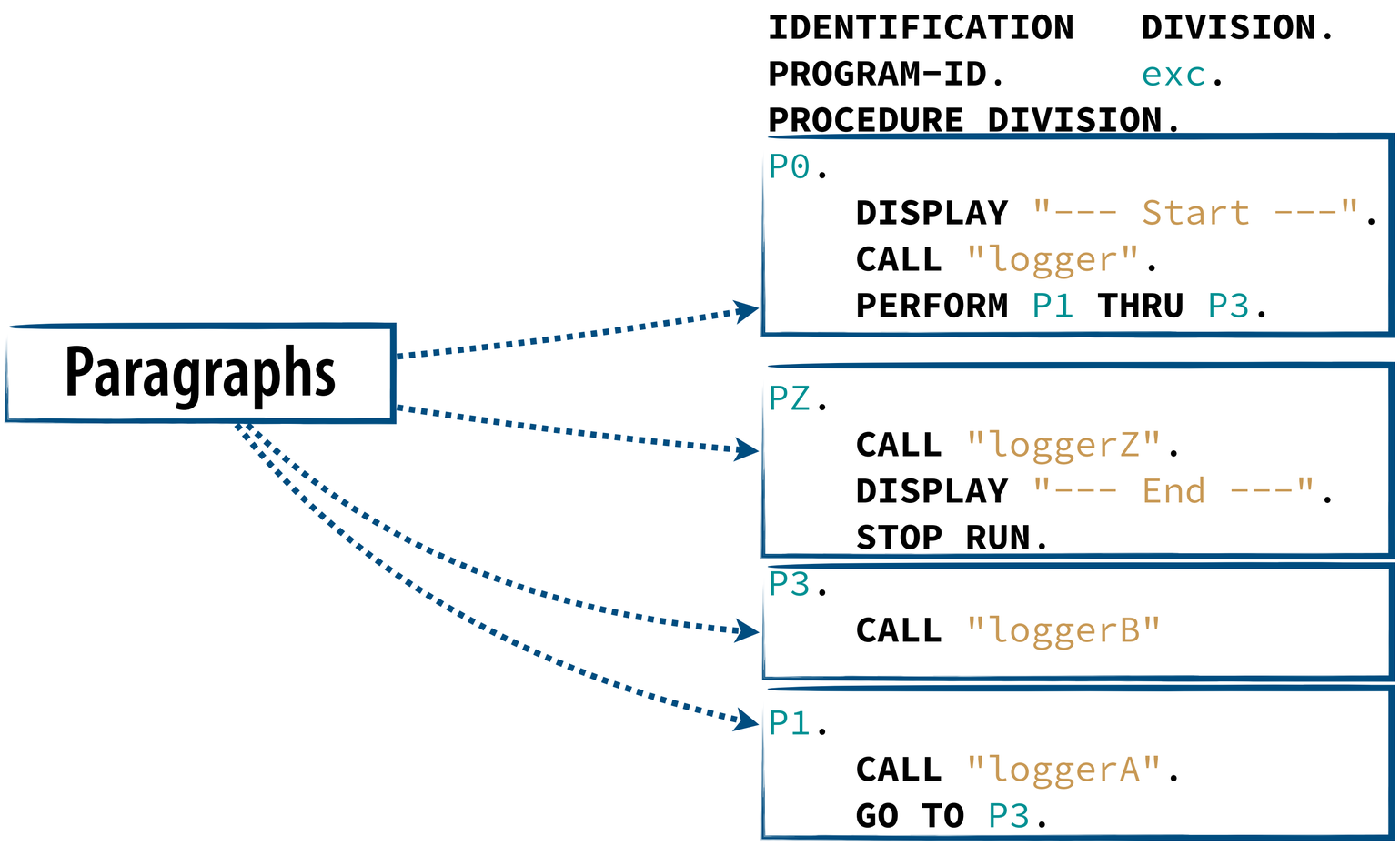
Java snippet using Collection API

Graph-based object usage model

COBOL Primer

```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.

PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"

P1.
    CALL "loggerA".
    GO TO P3.
```

COBOL Primer



COBOL Primer

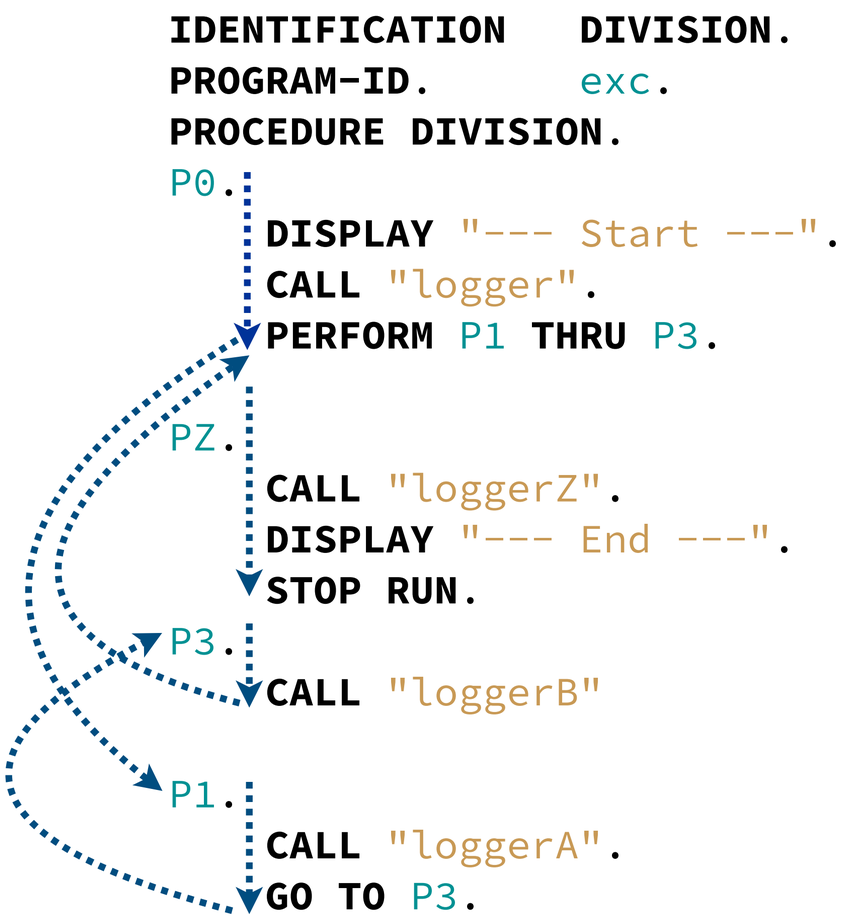
Call statement
Calls an external program.
“Library call”

```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.

PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"

P1.
    CALL "loggerA".
    GO TO P3.
```

Perform statement
Jump to `P1`, execute until `P3`, jump back

Go to statement
Jump to `P3` and continue execution

COBOL Primer



Jumping leads to complex control flow

Groums for COBOL

```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.

PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"

P1.
    CALL "loggerA".
    GO TO P3.
```

Single paragraph (often) too small

Groums for COBOL

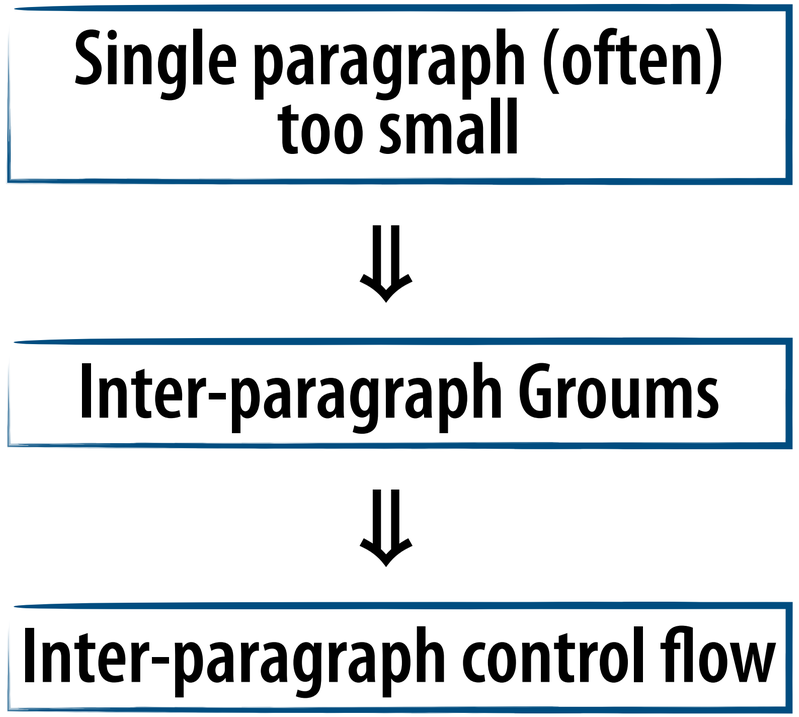
```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.

PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"

P1.
    CALL "loggerA".
    GO TO P3.
```

Single paragraph (often) too small

Inter-paragraph Groums

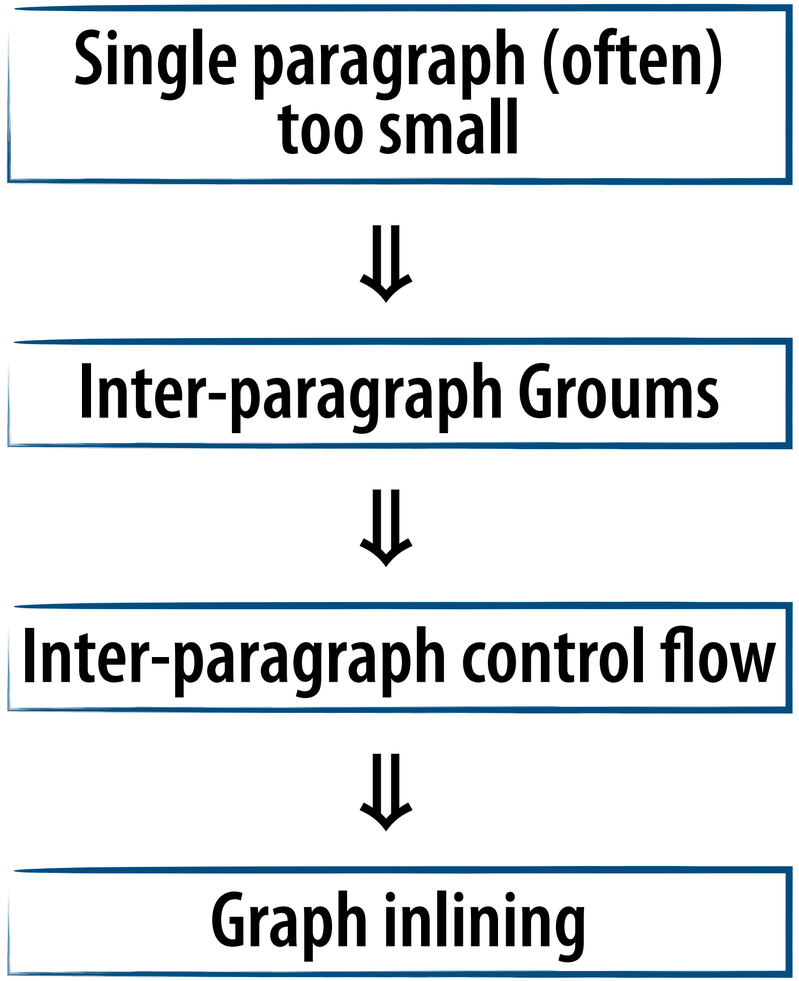
Groums for COBOL

```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.

PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"

P1.
    CALL "loggerA".
    GO TO P3.
```



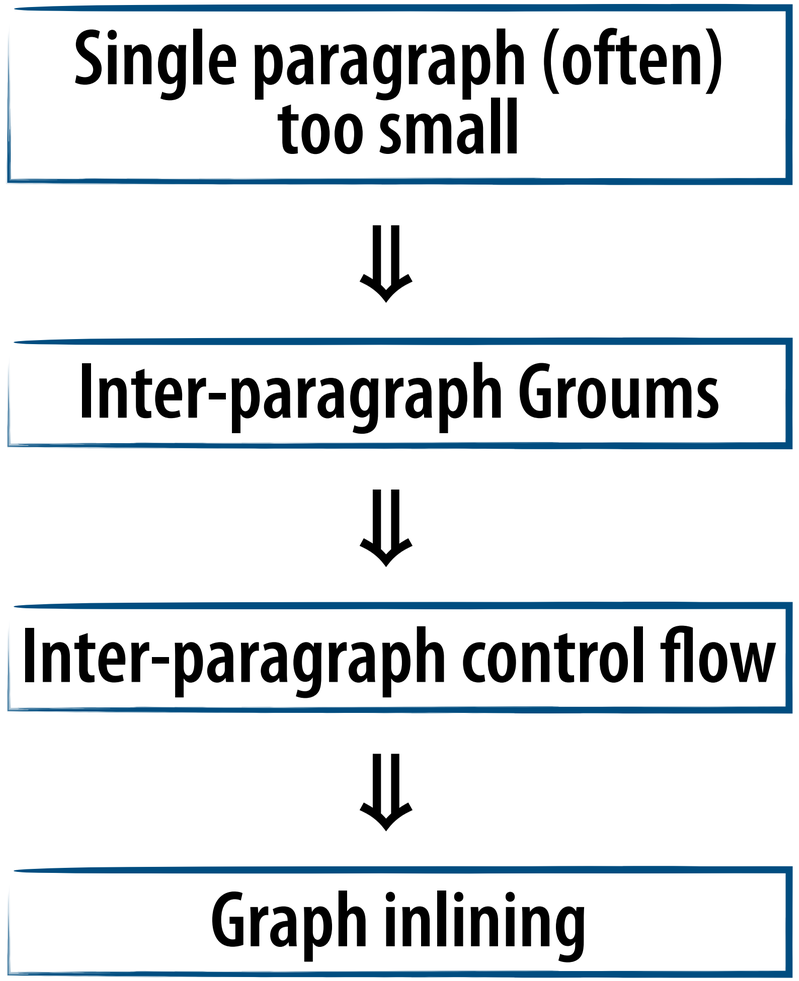
Groums for COBOL

```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.

PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"

P1.
    CALL "loggerA".
    GO TO P3.
```



Groums for COBOL

```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.

PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"

P1.
    CALL "loggerA".
    GO TO P3.
```



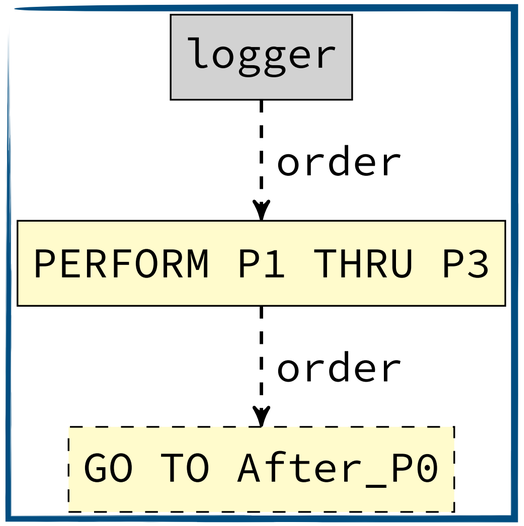
Other challenges
- Definition of library calls
- Absence of “def” edges
- Iteration through jumps
- Exit calls
- ...

Phase 1: Intermediate Groum Construction

```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.

PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"

P1.
    CALL "loggerA".
    GO TO P3.
```

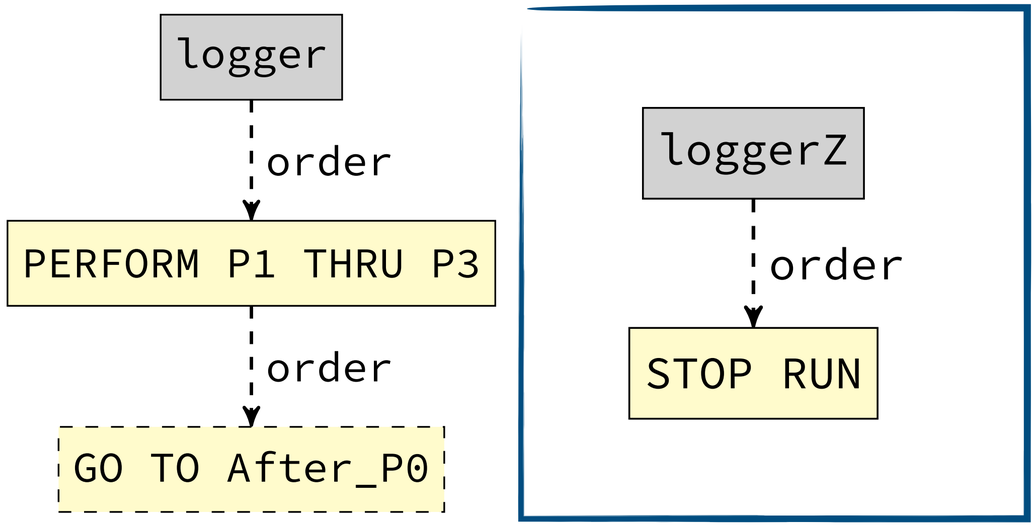
Phase 1: Intermediate Groum Construction

```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.

PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"

P1.
    CALL "loggerA".
    GO TO P3.
```



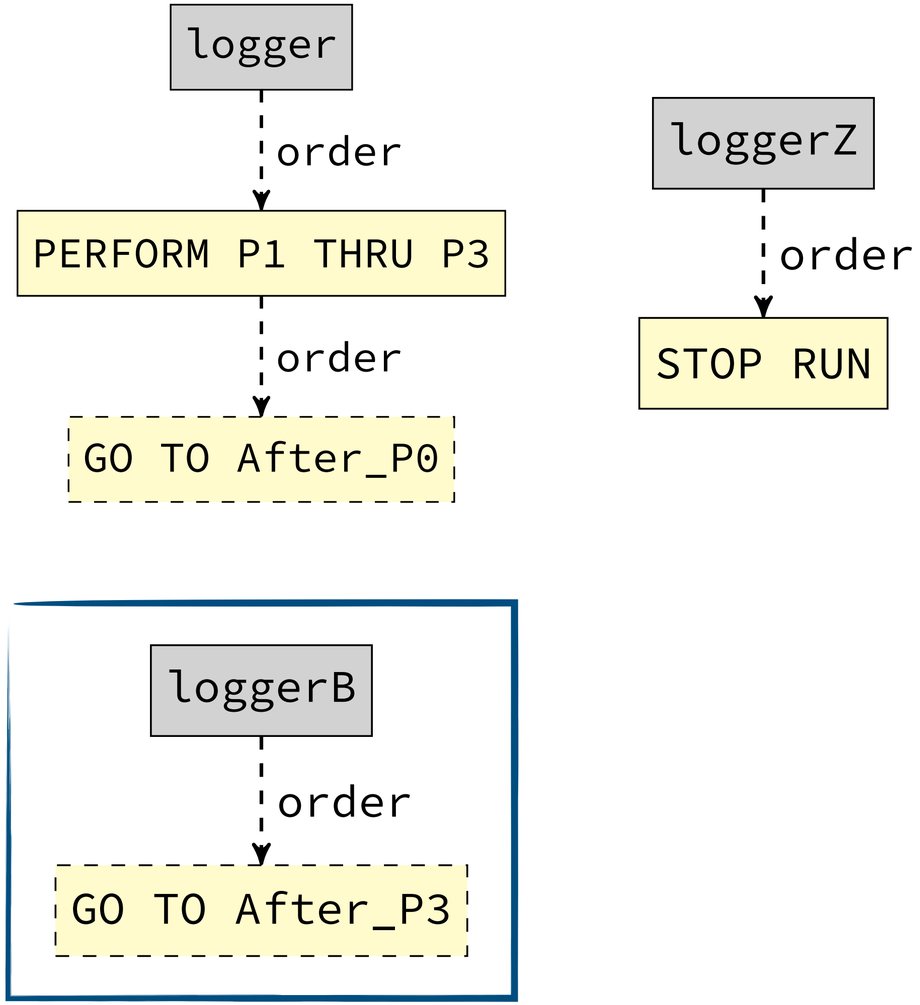
Phase 1: Intermediate Groum Construction

```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.

PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"

P1.
    CALL "loggerA".
    GO TO P3.
```



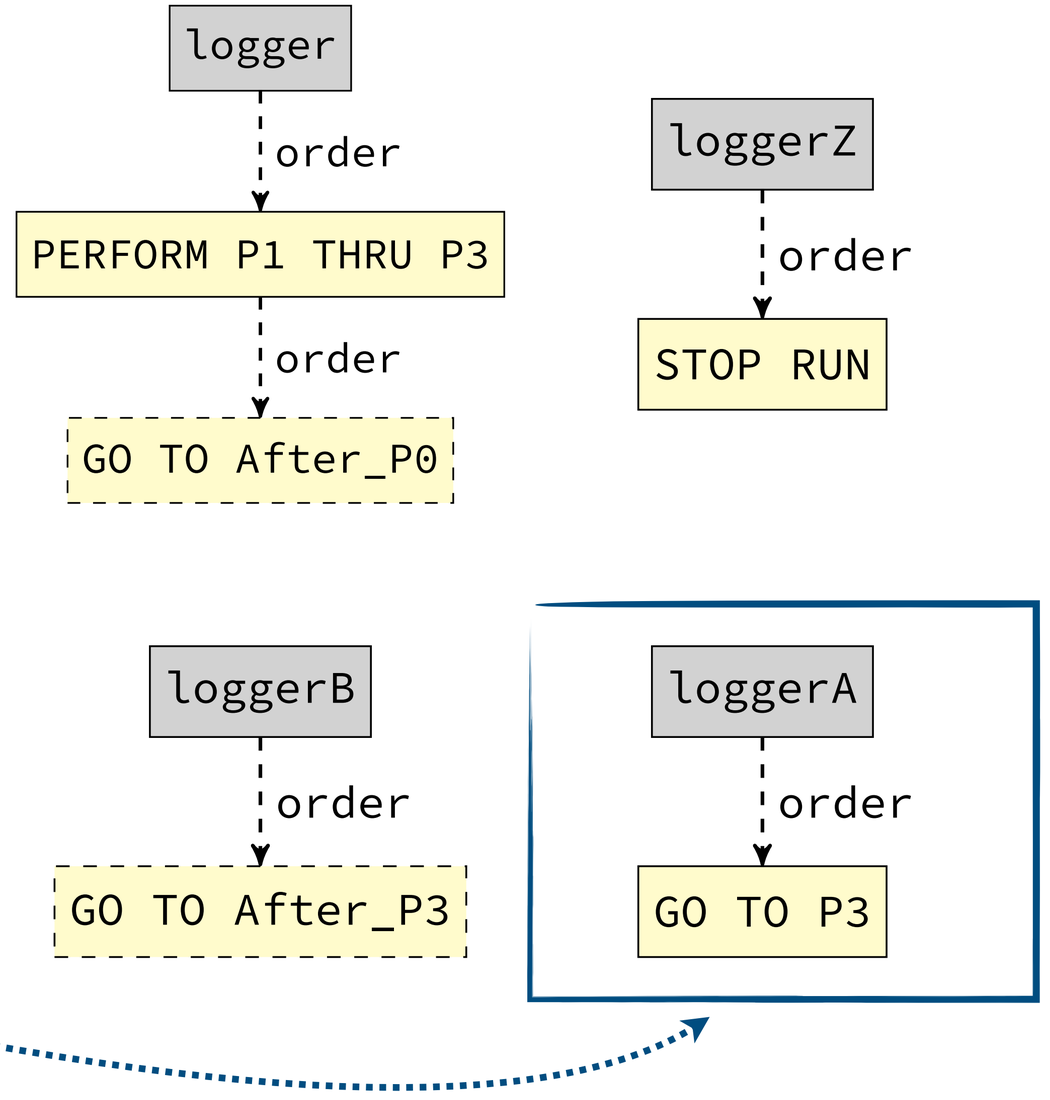
Phase 1: Intermediate Groum Construction

```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.

PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"

P1.
    CALL "loggerA".
    GO TO P3.
```



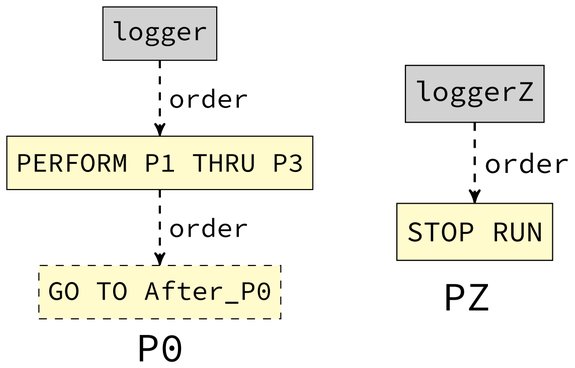
Phase 1: Intermediate Groum Construction

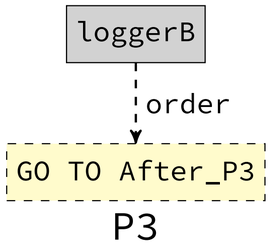
```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.

PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"

P1.
    CALL "loggerA".
    GO TO P3.
```

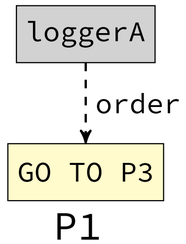


Phase 2: Groum Inlining

```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.
PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"
P1.
    CALL "loggerA".
    GO TO P3.
```





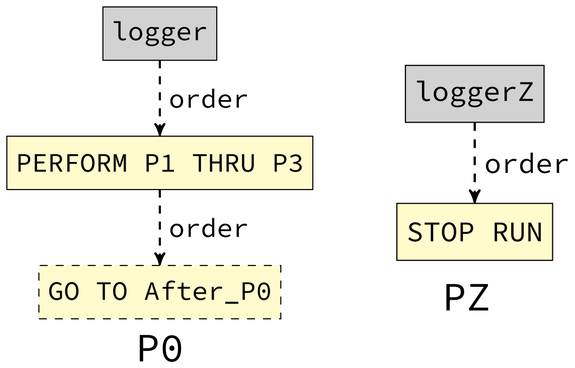


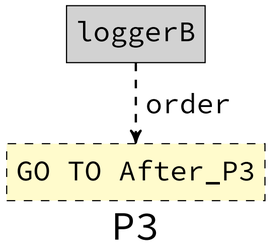
Jump Table

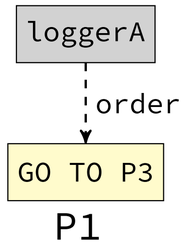
After_P0	PZ
After_P3	P0

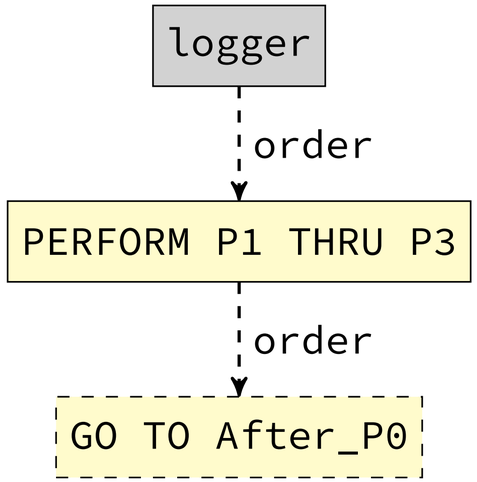
Phase 2: Groum Inlining

```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.
PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"
P1.
    CALL "loggerA".
    GO TO P3.
```







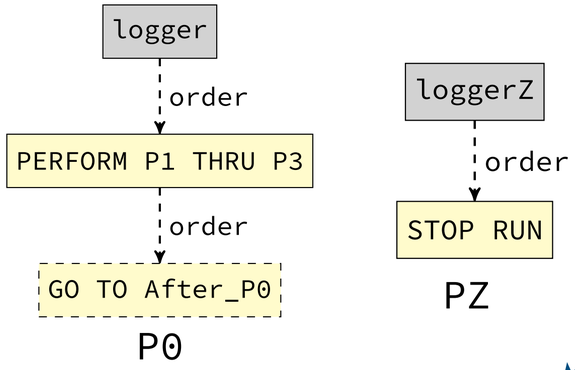


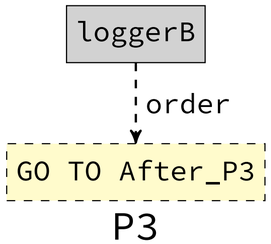
Jump Table

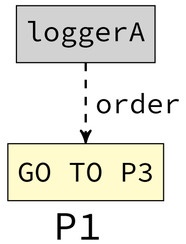
After_P0	PZ
After_P3	P0

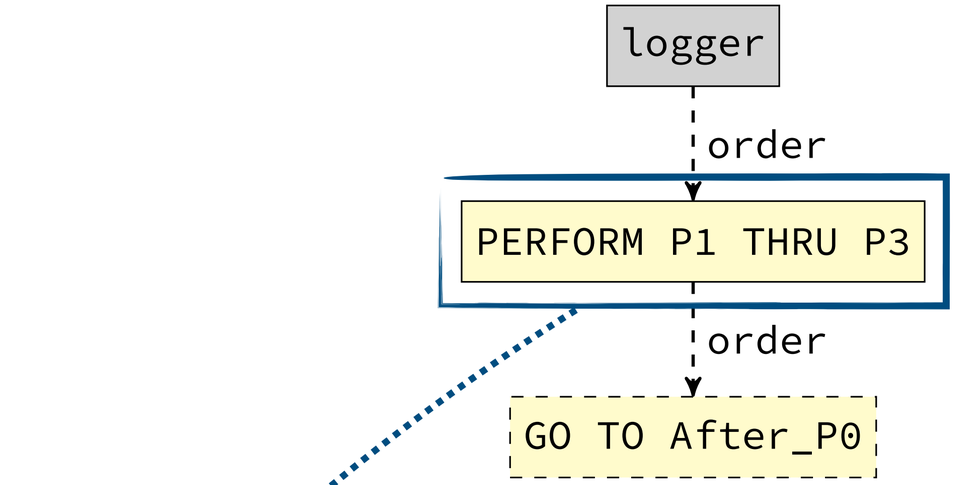
Phase 2: Groum Inlining

```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.
PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"
P1.
    CALL "loggerA".
    GO TO P3.
```









PERFORM

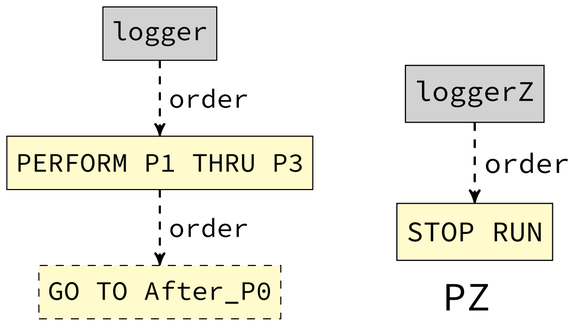
- Inline first paragraph
- Adjust last paragraph in jump table

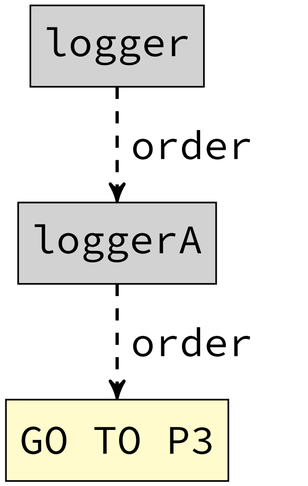
Jump Table

After_P0	PZ
After_P3	P0

Phase 2: Groum Inlining

```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.
PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"
P1.
    CALL "loggerA".
    GO TO P3.
```





GO TO After_P0

PERFORM

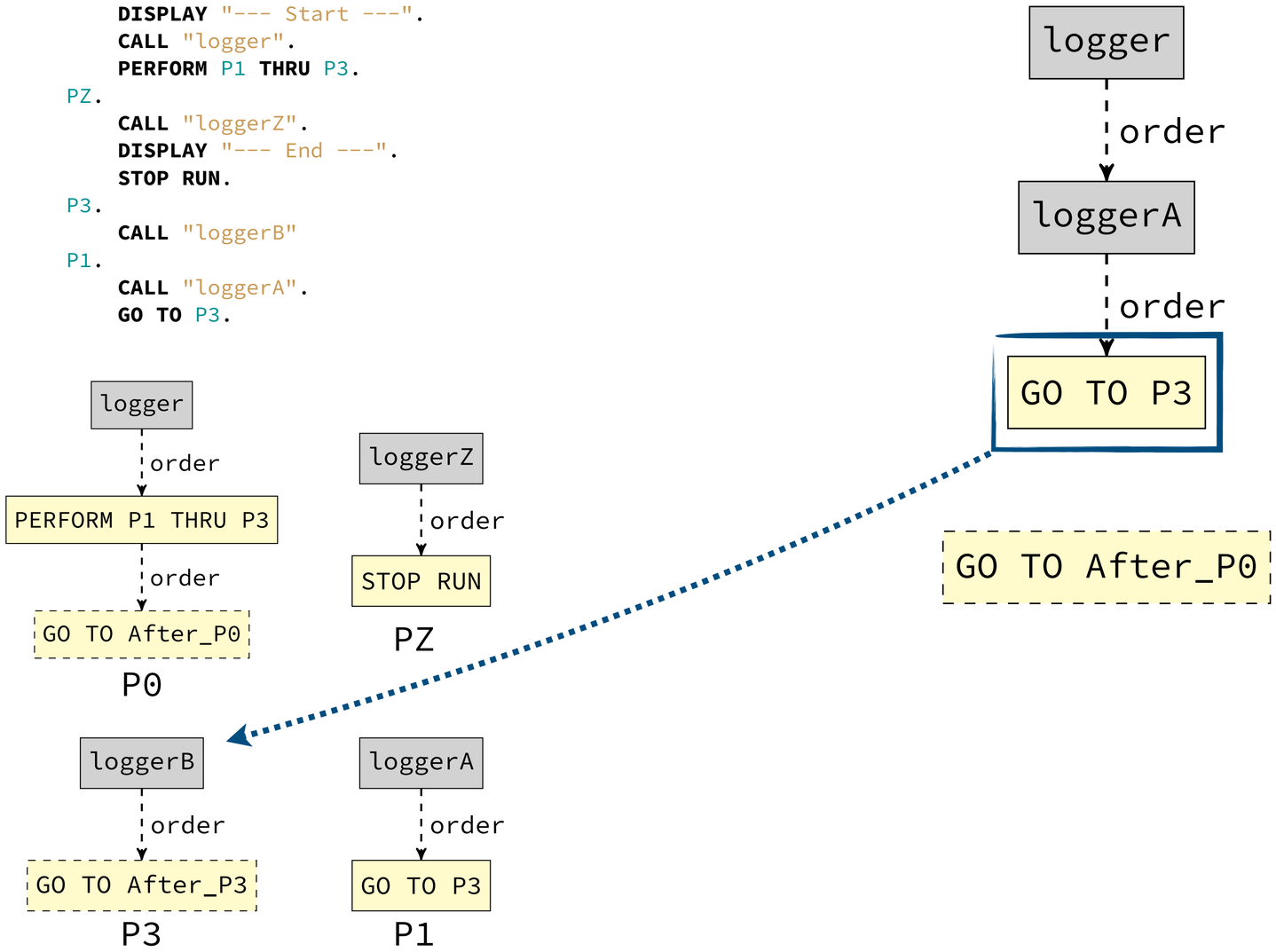
- Inline first paragraph
- Adjust last paragraph in jump table

Jump Table

After_P0	PZ
After_P3	GO TO After_P0

Phase 2: Groum Inlining

```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.
PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"
P1.
    CALL "loggerA".
    GO TO P3.
```



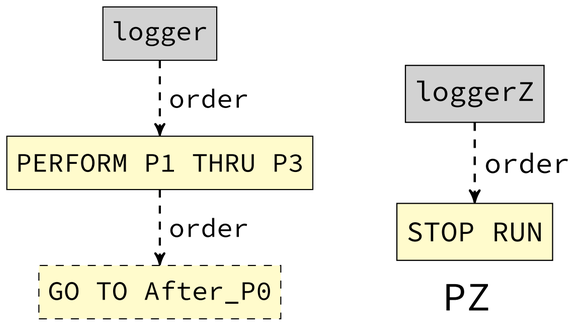
GO TO
Inline target paragraph

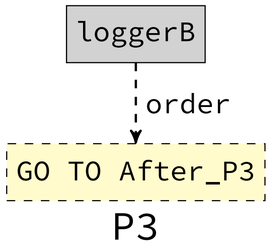
Jump Table

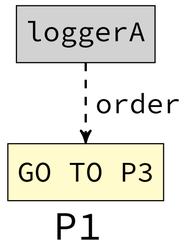
After_P0	PZ
After_P3	GO TO After_P0

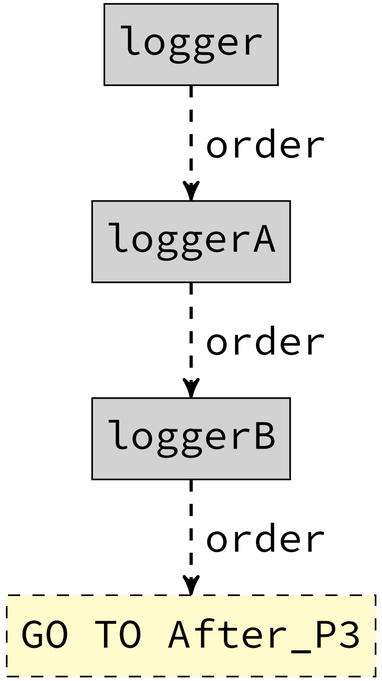
Phase 2: Groum Inlining

```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.
PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"
P1.
    CALL "loggerA".
    GO TO P3.
```









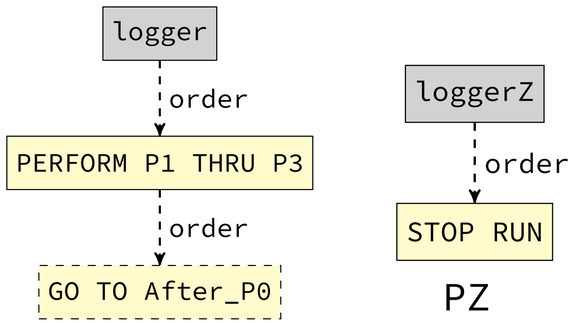
GO TO
Inline target paragraph

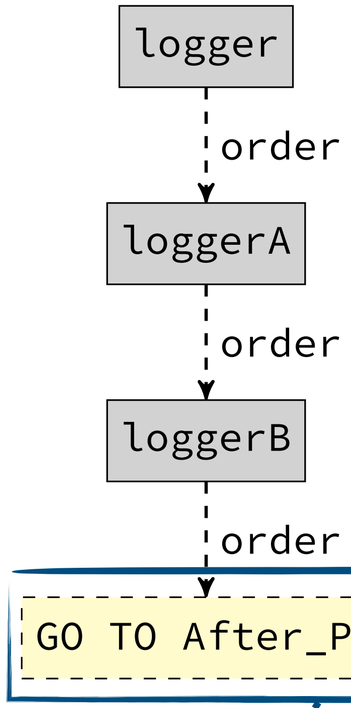
Jump Table

After_P0	PZ
After_P3	GO TO After_P0

Phase 2: Groum Inlining

```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.
PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"
P1.
    CALL "loggerA".
    GO TO P3.
```





Implicit jump after paragraph
Consult jump table

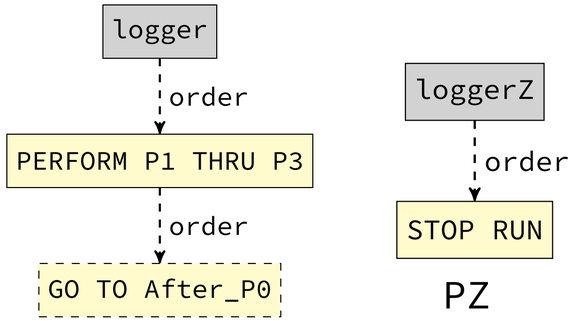
End of PERFORM
Reset jump table entry

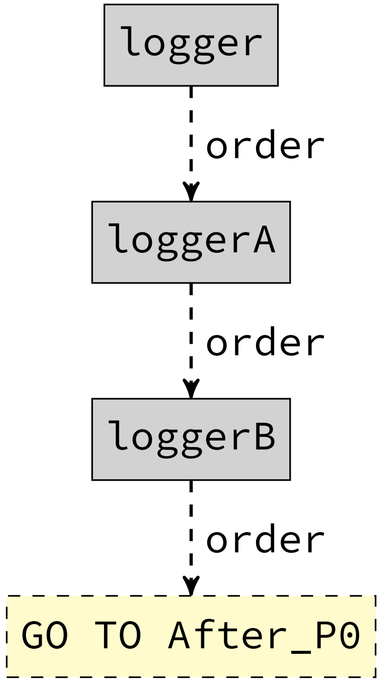
Jump Table

After_P0	PZ
After_P3	GO TO After_P0

Phase 2: Groum Inlining

```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.
PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"
P1.
    CALL "loggerA".
    GO TO P3.
```





Implicit jump after paragraph
Consult jump table

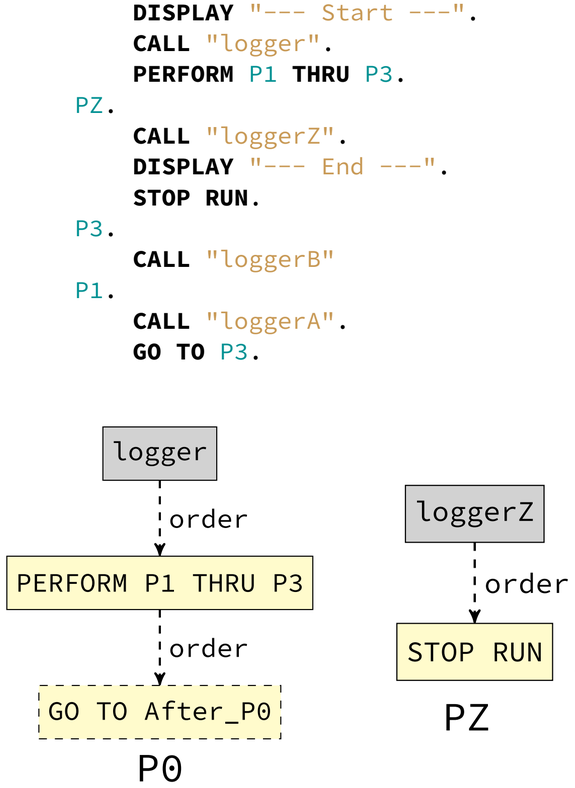
End of PERFORM
Reset jump table entry

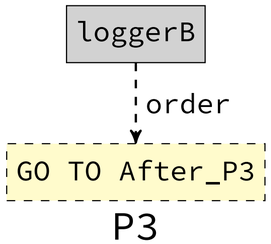
Jump Table

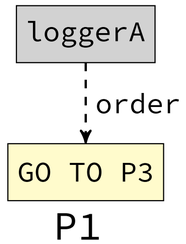
After_P0	PZ
After_P3	P0

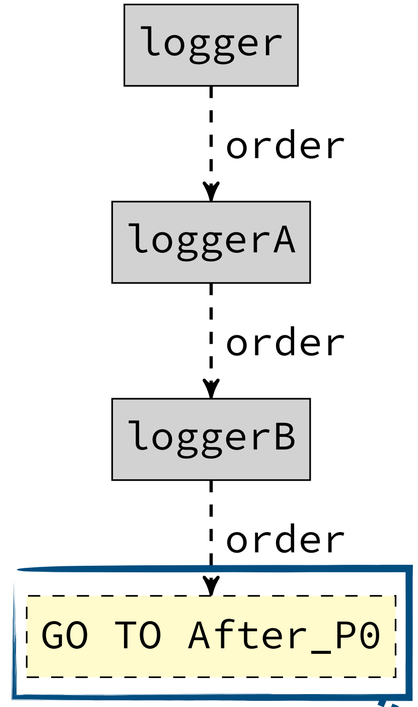
Phase 2: Groum Inlining

```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.
PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"
P1.
    CALL "loggerA".
    GO TO P3.
```









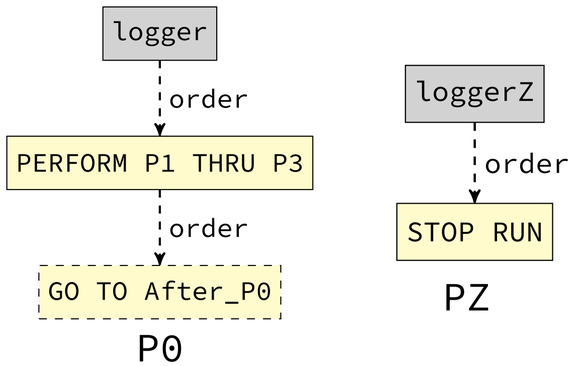
Implicit jump after paragraph
Consult jump table

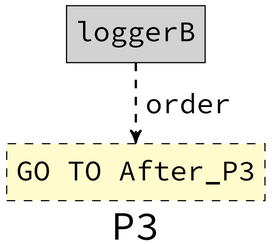
Jump Table

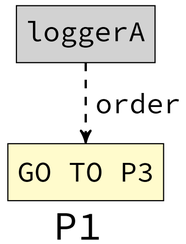
After_P0	PZ
After_P3	P0

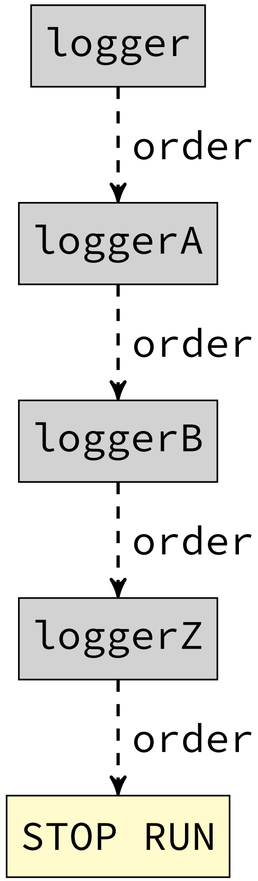
Phase 2: Groum Inlining

```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.
PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"
P1.
    CALL "loggerA".
    GO TO P3.
```









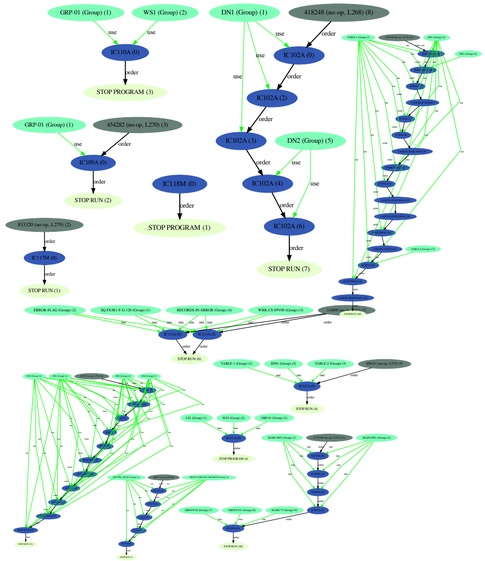
Implicit jump after paragraph
Consult jump table

Jump Table

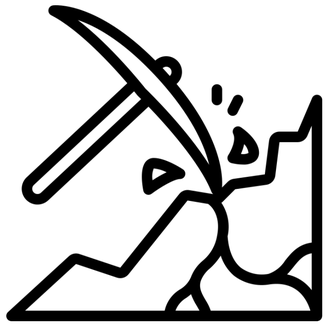
After_P0	PZ
After_P3	P0

Next Steps: Mining, Visualisation

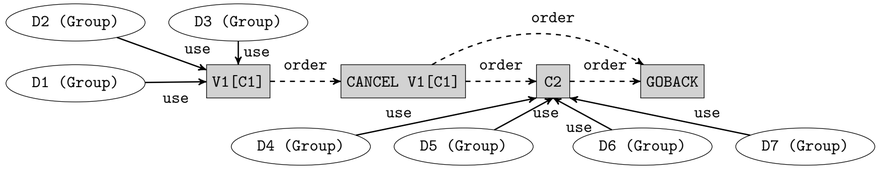
Groums

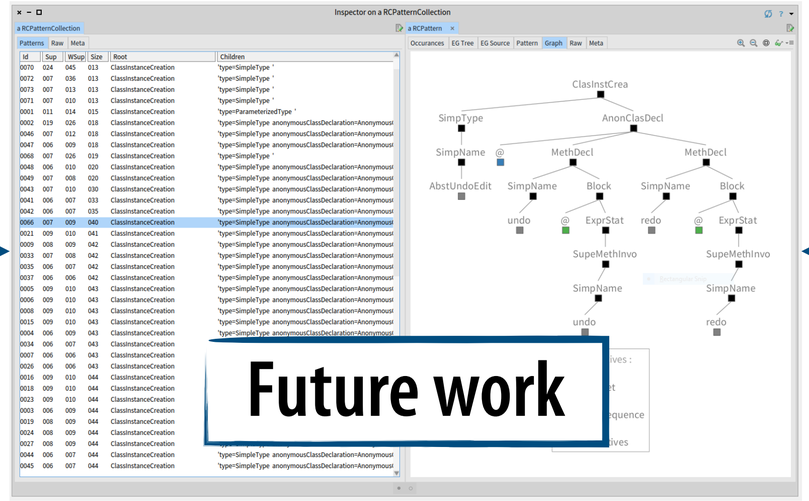


Frequent Subgraph Mining



Library Usage Patterns



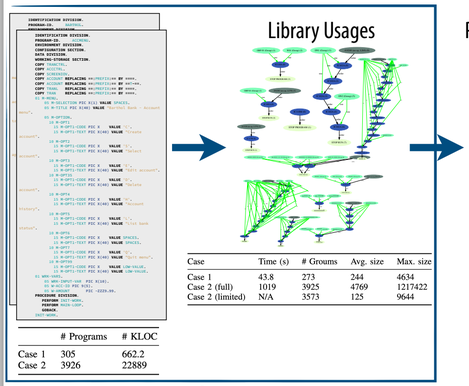


Future work

Visualisation Tool

Conclusion

COBOL Library Usage Pattern Mining



Library Usages

Pattern Mining

Library Usage Patterns

Case	Time (s)	# Grooms	Avg. size	Max. size
Case 1	43.8	273	244	4634
Case 2 (full)	1019	3925	4769	1217422
Case 2 (limited)	N/A	3573	125	9644

	# Programs	# KLOC
Case 1	305	662.2
Case 2	3926	22889

Today

Corpus	GrouMiner*		BigGroum		
	Time (min)	# P	Time (min)	# Part.	# P
Corpus 1	0.95	3	10.4	3	4
Corpus 2	121.8	7	618.2	1	0

Goal: Understand library usages
to estimate modernisation effort

Grooms for COBOL

```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.

PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"

P1.
    CALL "loggerA".
    GO TO P3.
```

Single paragraph (often) too small

⇓

Inter-paragraph Grooms

⇓

Inter-paragraph control flow

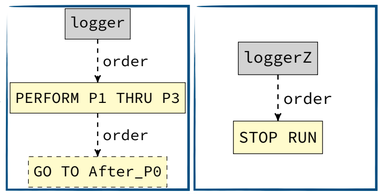
⇓

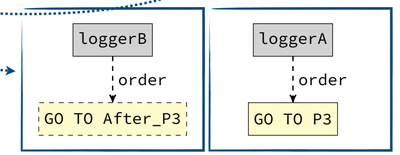
Graph inlining

Other challenges
- Definition of library calls
- Absence of “def” edges
- Iteration through jumps
- Exit calls
- ...

Phase 1: Intermediate Groum Construction

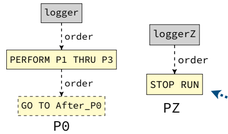
```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.
PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"
P1.
    CALL "loggerA".
    GO TO P3.
```

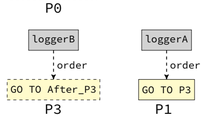


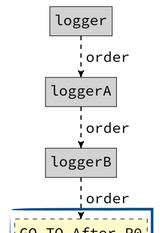


Phase 2: Groum Inlining

```
IDENTIFICATION   DIVISION.
PROGRAM-ID.      exc.
PROCEDURE DIVISION.
P0.
    DISPLAY "--- Start ---".
    CALL "logger".
    PERFORM P1 THRU P3.
PZ.
    CALL "loggerZ".
    DISPLAY "--- End ---".
    STOP RUN.
P3.
    CALL "loggerB"
P1.
    CALL "loggerA".
    GO TO P3.
```







Implicit jump after paragraph
Consult jump table

End of PERFORM
Reset jump table entry

Jump Table

After_P0	PZ
After_P3	P0