

Untangling Source Code Changes Using Program Slicing



Ward Muylaert (@wardmuylaert) and Coen De Roover (@oniroi)

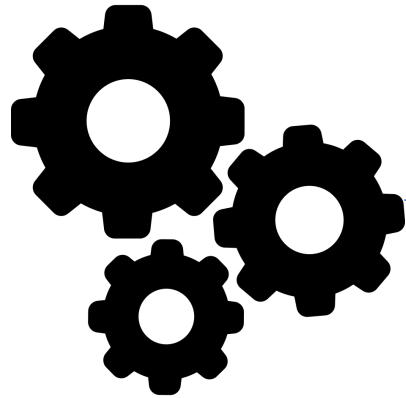
BENEVOL 2017.

Untangling Source Code Changes Using Program Slicing



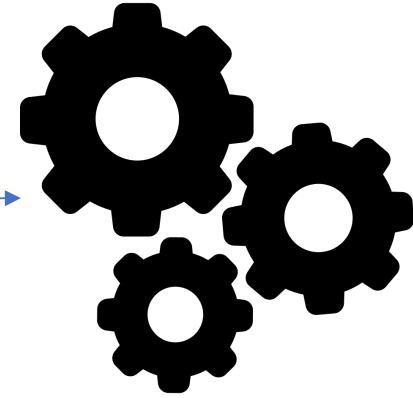
Ward Muylaert (@wardmuylaert) and Coen De Roover (@oniroi)

BENEVOL 2017.



Version x

```
6 arch/arm64/kernel/cpu_ops.c
@@ -31,13 +31,13 @@ extern const struct cpu_operations cpu_psci_ops;
31
32 const struct cpu_operations *cpu_ops[NR_CPUS]
33   __ro_after_init;
34
35 -static const struct cpu_operations
36   *dt_supported_cpu_ops[] __initconst = {
37   &smp_spin_table_ops,
38   &cpu_psci_ops,
39   NULL,
40 };
41
42 -static const struct cpu_operations
43   *acpi_supported_cpu_ops[] __initconst = {
44   #ifdef CONFIG_ARM64_ACPI_PARKING_PROTOCOL
45   &acpi_parking_protocol_ops,
46   #endif
47 };
48
49 +static const struct cpu_operations *const
50   dt_supported_cpu_ops[] __initconst = {
51   &smp_spin_table_ops,
52   &cpu_psci_ops,
53   NULL,
54 };
55
56 +static const struct cpu_operations *const
57   acpi_supported_cpu_ops[] __initconst = {
58   #ifdef CONFIG_ARM64_ACPI_PARKING_PROTOCOL
59   &acpi_parking_protocol_ops,
60   #endif
61 };
62
```



Version y

6 arch/arm64/kernel/cpu_ops.c View

@@ -31,13 +31,13 @@ extern const struct cpu_operations cpu_psci_ops;

31 const struct cpu_operations *cpu_ops[NR_CPUS]

32 __ro_after_init;

33

34 -static const struct cpu_operations

35 #dt_supported_cpu_ops[] __initconst = {

36 &cpu_spin_table_ops,

37 &cpu_psci_ops,

38 NULL,

39 };

40

41 -static const struct cpu_operations

42 #acpi_supported_cpu_ops[] __initconst = {

43 #ifdef CONFIG_ARM64_ACPI_PARKING_PROTOCOL

44 &acpi_parking_protocol_ops,

45 #endif

31

32 const struct cpu_operations *cpu_ops[NR_CPUS]

33 __ro_after_init;

34

35 +static const struct cpu_operations *const

36 dt_supported_cpu_ops[] __initconst = {

37 &cpu_spin_table_ops,

38 &cpu_psci_ops,

39 NULL,

40 };

41

42 +static const struct cpu_operations *const

43 acpi_supported_cpu_ops[] __initconst = {

44 #ifdef CONFIG_ARM64_ACPI_PARKING_PROTOCOL

45 &acpi_parking_protocol_ops,

46 #endif

6 arch/arm64/kernel/cpu_ops.c View

@@ -31,13 +31,13 @@ extern const struct cpu_operations cpu_psci_ops;

31 const struct cpu_operations *cpu_ops[NR_CPUS]

32 __ro_after_init;

33

34 -static const struct cpu_operations

35 #dt_supported_cpu_ops[] __initconst = {

36 &cpu_spin_table_ops,

37 &cpu_psci_ops,

38 NULL,

39 };

40

41 -static const struct cpu_operations

42 #acpi_supported_cpu_ops[] __initconst = {

43 #ifdef CONFIG_ARM64_ACPI_PARKING_PROTOCOL

44 &acpi_parking_protocol_ops,

45 #endif

31

32 const struct cpu_operations *cpu_ops[NR_CPUS]

33 __ro_after_init;

34

35 +static const struct cpu_operations *const

36 dt_supported_cpu_ops[] __initconst = {

37 &cpu_spin_table_ops,

38 &cpu_psci_ops,

39 NULL,

40 };

41

42 +static const struct cpu_operations *const

43 acpi_supported_cpu_ops[] __initconst = {

44 #ifdef CONFIG_ARM64_ACPI_PARKING_PROTOCOL

45 &acpi_parking_protocol_ops,

46 #endif



Fix issue #31415

ward committed 7 days ago ✓



Add timetraveling feature

ward committed 7 days ago ✓



Refactor benevol presentation

ward committed 7 days ago ✓


```
6 arch/arm64/kernel/cpu_ops.c
@@ -31,13 +31,13 @@ extern const struct cpu_operations cpu_psci_ops;
31
32 const struct cpu_operations *cpu_ops[NR_CPUS]
__ro_after_init;
33
34 -static const struct cpu_operations
+static const struct cpu_operations *const
*dt_supported_cpu_ops[] __initconst = {
dt_supported_cpu_ops[] __initconst = {
35     &cpu_spin_table_ops,
36     &cpu_psci_ops,
37     NULL,
38 };
39
40 -static const struct cpu_operations
+static const struct cpu_operations *const
*acpi_supported_cpu_ops[] __initconst = {
acpi_supported_cpu_ops[] __initconst = {
41     #ifdef CONFIG_ARM64_ACPI_PARKING_PROTOCOL
42     &acpi_parking_protocol_ops,
43     #endif
44 }
```

```
6 arch/arm64/kernel/cpu_ops.c
@@ -31,13 +31,13 @@ extern const struct cpu_operations cpu_psci_ops;
31
32 const struct cpu_operations *cpu_ops[NR_CPUS]
__ro_after_init;
33
34 -static const struct cpu_operations
+static const struct cpu_operations *const
*dt_supported_cpu_ops[] __initconst = {
dt_supported_cpu_ops[] __initconst = {
35     &cpu_spin_table_ops,
36     &cpu_psci_ops,
37     NULL,
38 };
39
40 -static const struct cpu_operations
+static const struct cpu_operations *const
*acpi_supported_cpu_ops[] __initconst = {
acpi_supported_cpu_ops[] __initconst = {
41     #ifdef CONFIG_ARM64_ACPI_PARKING_PROTOCOL
42     &acpi_parking_protocol_ops,
43     #endif
44 }
```



Fix issue #31415

ward committed 7 days ago ✓



Add timetraveling feature

ward committed 7 days ago ✓



Refactor benevol presentation

ward committed 7 days ago ✓



Fix issue #271828 and some refactoring to do so

bad-ward committed 7 days ago ✓



Add timetravel, remove locations

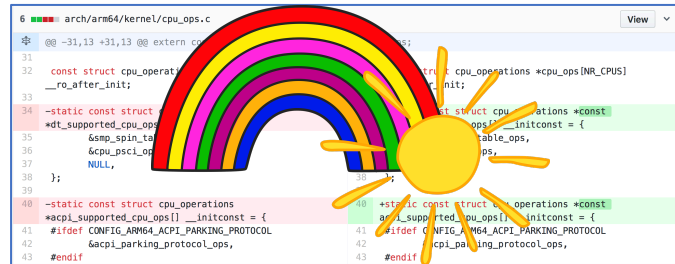
bad-ward committed 7 days ago ✓



Change 50 different things for procrastination

bad-ward committed 7 days ago ✓

Atomic commits



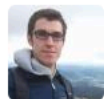
Fix issue #31415

ward committed 7 days ago ✓



Add timetraveling feature

ward committed 7 days ago ✓



Refactor benevol presentation

ward committed 7 days ago ✓



Fix issue #271828 and some refactoring to do so

bad-ward committed 7 days ago ✓



Add timetravel, remove locations

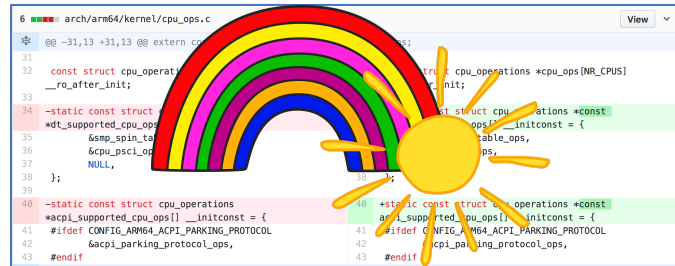
bad-ward committed 7 days ago ✓



Change 50 different things for procrastination

bad-ward committed 7 days ago ✓

Atomic commits



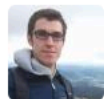
Fix issue #31415

ward committed 7 days ago ✓



Add timetraveling feature

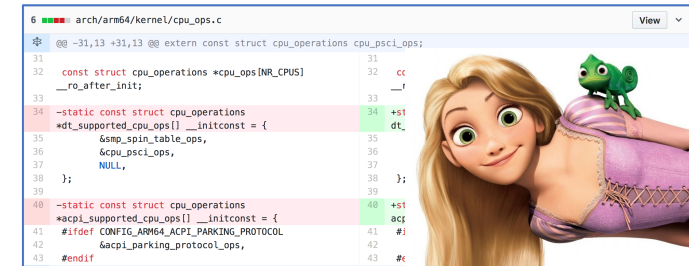
ward committed 7 days ago ✓



Refactor benevol presentation

ward committed 7 days ago ✓

Tangled commits



Fix issue #271828 and some refactoring to do so

bad-ward committed 7 days ago ✓



Add timetravel, remove locations

bad-ward committed 7 days ago ✓

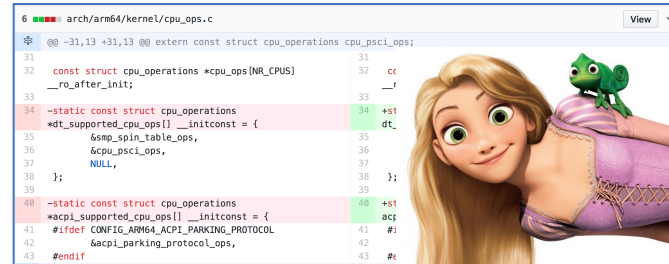


Change 50 different things for procrastination

bad-ward committed 7 days ago ✓

Integrate

Revert

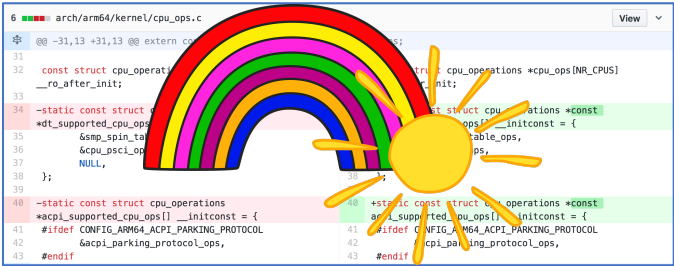
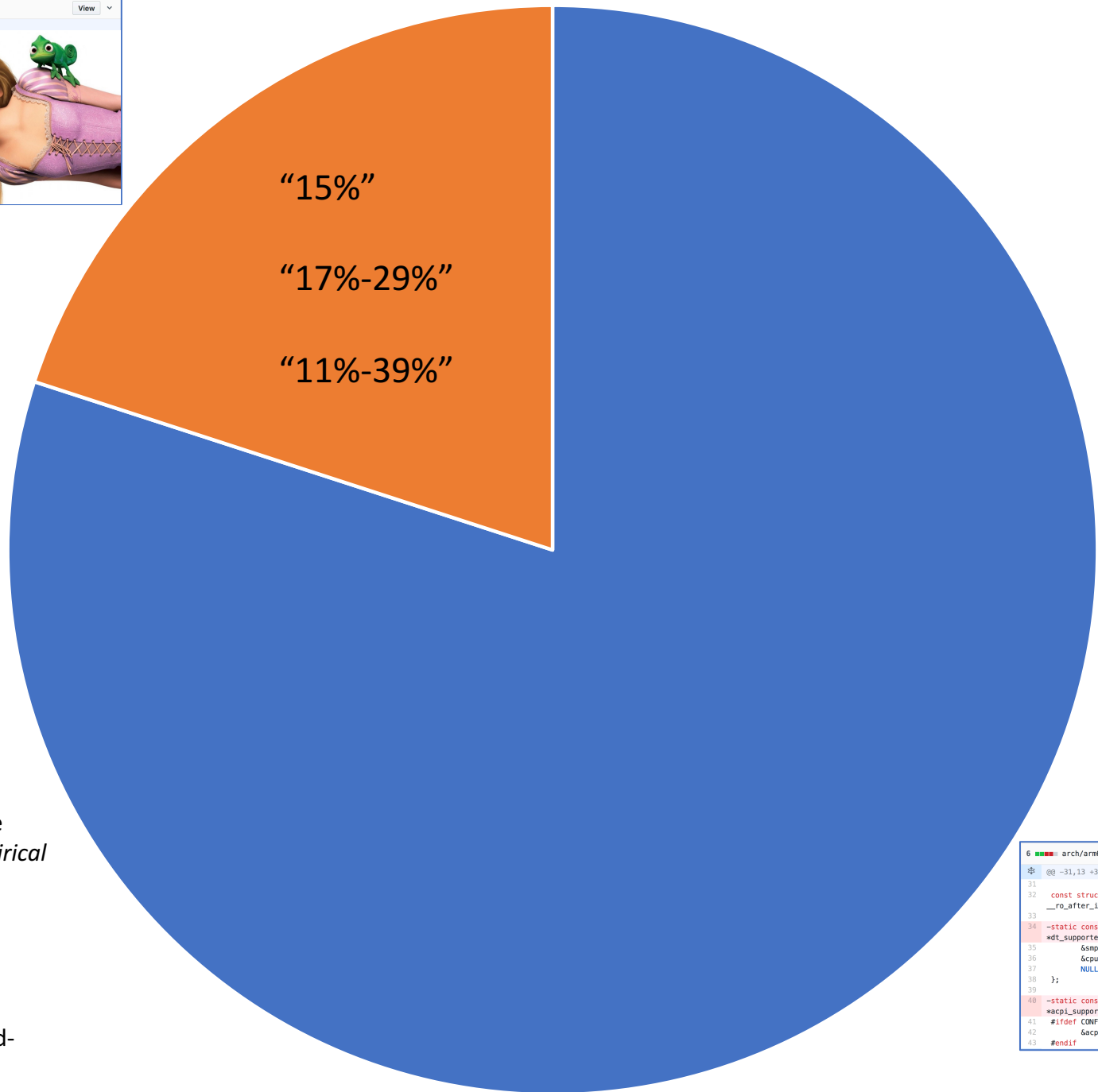
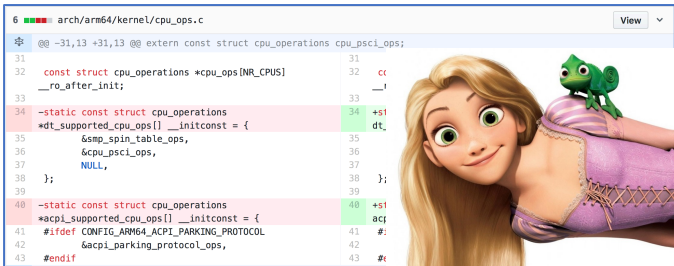


The central window displays a code editor with C code for CPU operations. The code is organized into two main sections, each starting with a line number (31 and 40). The first section defines a constant structure for CPU operations, including a pointer to a table of supported CPU operations. The second section defines a static constant structure for CPU operations, including a pointer to a table of supported CPU operations. A small image of Rapunzel from Disney's 'Tangled' is visible on the right side of the code editor.

```
6 arch/arm64/kernel/cpu_ops.c
31 @@ -31,13 +31,13 @@ extern const struct cpu_operations cpu_psci_ops;
32 const struct cpu_operations *cpu_ops[NR_CPUS]
33 __ro_after_init;
34 -static const struct cpu_operations
35 #dt_supported_cpu_ops[] __initconst = {
36     &mp_spin_table_ops,
37     &cpu_psci_ops,
38     NULL,
39 };
40 -static const struct cpu_operations
41 #acpi_supported_cpu_ops[] __initconst = {
42     #ifdef CONFIG_ARM64_ACPI_PARKING_PROTOCOL
43     &acpi_parking_protocol_ops,
44 #endif
45 };
46
```

Research

Understand

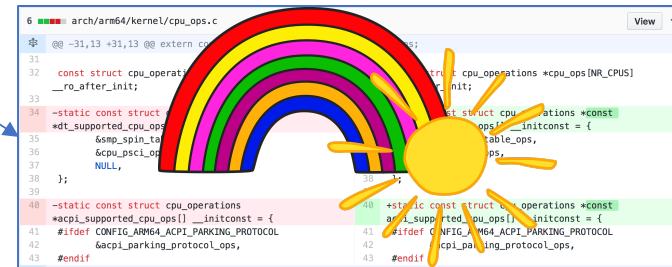
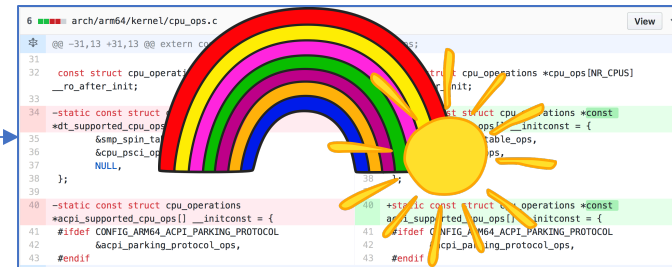
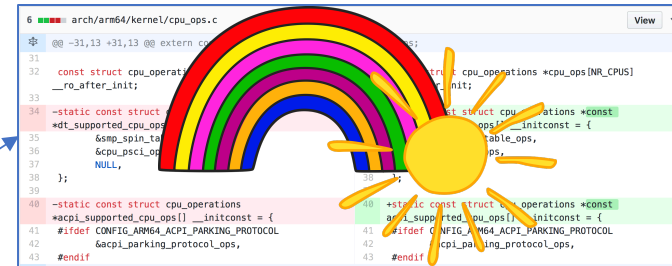
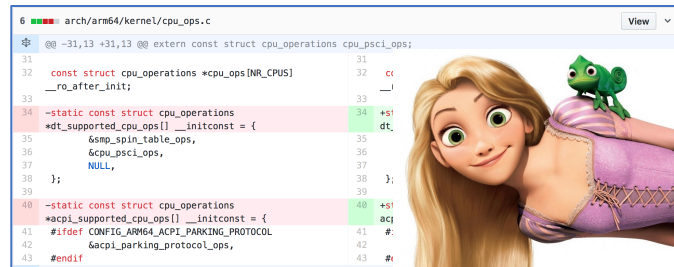


K. Herzig et al., “The impact of tangled code changes on defect prediction models,” *Empirical Software Engineering*, 2015.

Y. Tao et al., “Partitioning Composite Code Changes to Facilitate Code Review,” in *MSR*, 2015.

H. A. Nguyen et al., “Filtering Noise in Mixed-Purpose Fixing Commits to Improve Defect Prediction and Localization,” in *ISSRE*, 2013.

Goal

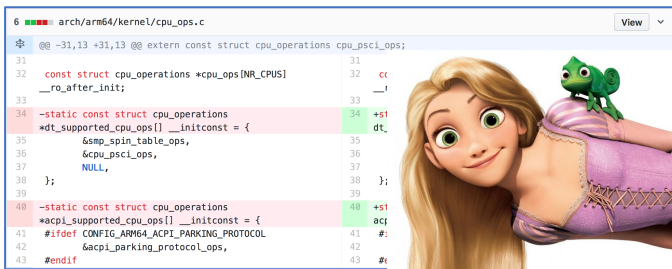


Hypothesis

*Related changes are part of the same program slice.
A commit may be untangled using the program slices.*

Hypothesis

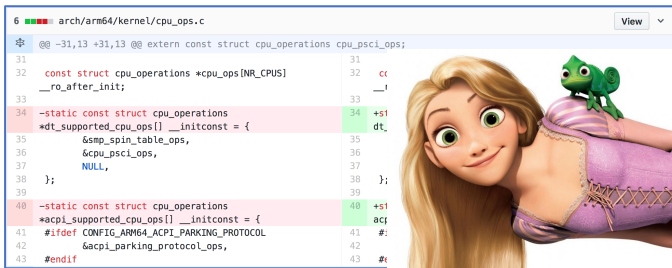
*Related changes are part of the same program slice.
A commit may be untangled using the program slices.*



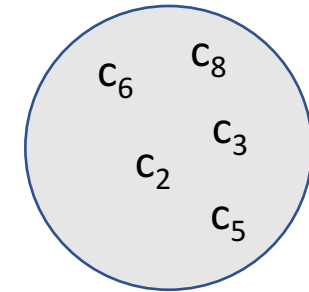
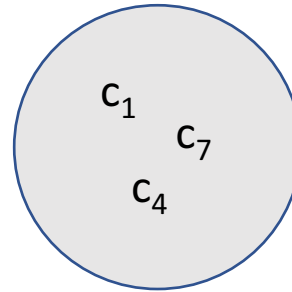
$C_1 C_2 C_3 C_4 C_5 C_6 C_7 C_8$

Hypothesis

*Related changes are part of the same program slice.
A commit may be untangled using the program slices.*



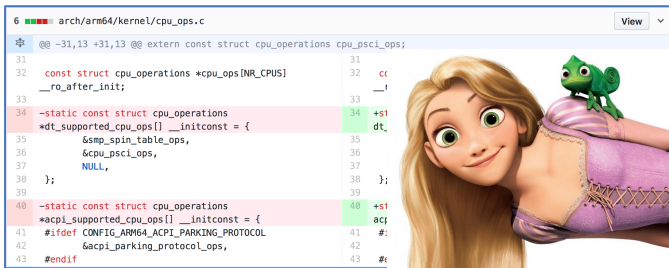
```
6 arch/arm64/kernel/cpu_ops.c
@@ -31,13 +31,13 @@ extern const struct cpu_operations cpu_psci_ops;
31
32 const struct cpu_operations *cpu_ops[NR_CPUS]
33 __ro_after_init;
34
35 -static const struct cpu_operations
36 *dt_supported_cpu_ops[] __initconst = {
37     &mpu_spin_table_ops,
38     &cpu_psci_ops,
39     NULL,
40 };
41
42 -static const struct cpu_operations
43 *acpi_supported_cpu_ops[] __initconst = {
44     #ifdef CONFIG_ARM64_ACPI_PARKING_PROTOCOL
45     &acpi_parking_protocol_ops,
46     #endif
47     NULL,
48 };
49
```



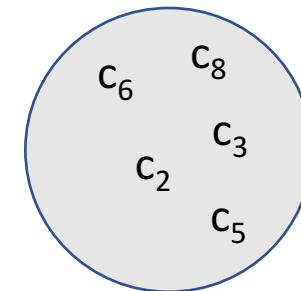
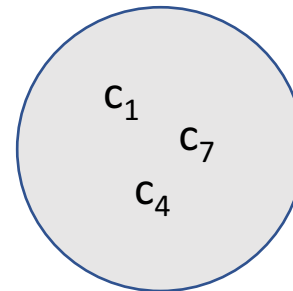
C_1 C_2 C_3 C_4 C_5 C_6 C_7 C_8

Hypothesis

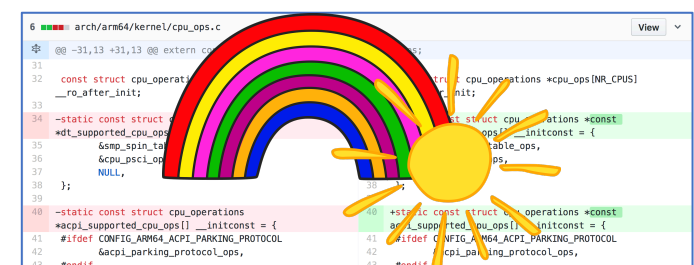
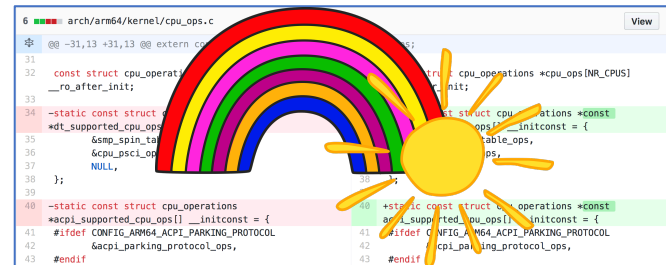
*Related changes are part of the same program slice.
A commit may be untangled using the program slices.*



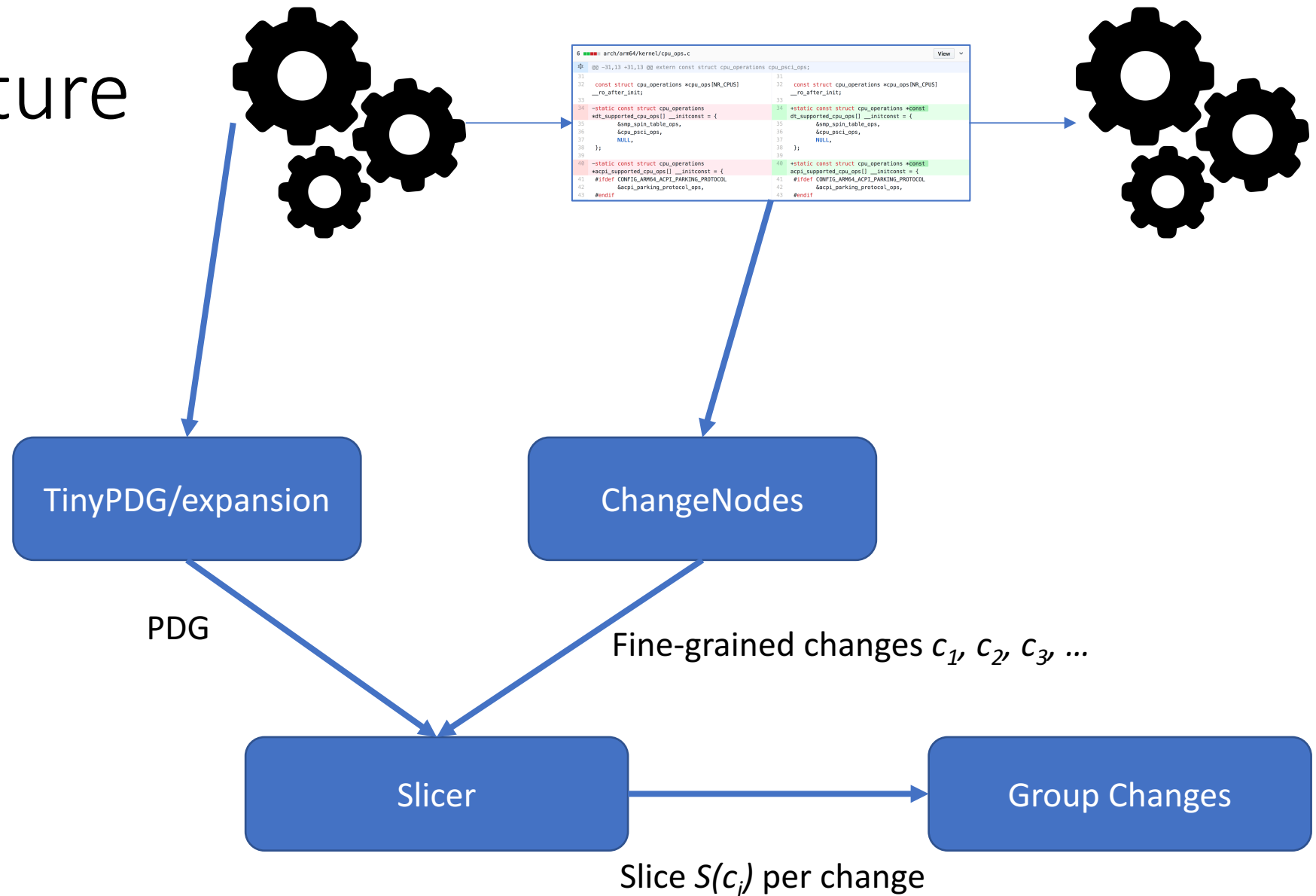
```
6 arch/arm64/kernel/cpu_ops.c
@@ -31,13 +31,13 @@ extern const struct cpu_operations cpu_psci_ops;
31
32 const struct cpu_operations *cpu_ops[NR_CPUS]
33 __ro_after_init;
34
35 -static const struct cpu_operations
36 *dt_supported_cpu_ops[] __initconst = {
37     &dt_spin_table_ops,
38     &cpu_psci_ops,
39     NULL,
40 };
41
42 -static const struct cpu_operations
43 *acpi_supported_cpu_ops[] __initconst = {
44     &acpi_parking_protocol_ops,
45     NULL,
46 };
47
48 #ifdef CONFIG_ARM64_ACPI_PARKING_PROTOCOL
49 #include <acpi_parking_protocol.h>
50 #endif
```

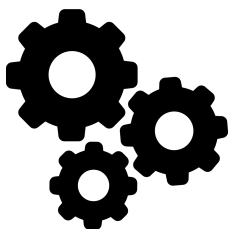


$C_1 C_2 C_3 C_4 C_5 C_6 C_7 C_8$



Architecture





TinyPDG

- Java
- Eclipse AST
- Intra-procedural

TinyPDG fork WIP

- Java
- Eclipse AST
- Inter-procedural

PDG


```
6 arch/arm64/kernel/cpu_ops.c
@@ -31,13 +31,13 @@ extern const struct cpu_operations cpu_psci_ops;
31 const struct cpu_operations *cpu_ops[NR_CPUS]
32 __ro_after_init;
33
34 -static const struct cpu_operations
35 +static const struct cpu_operations *const
36 #dt_supported_cpu_ops[] __initconst = {
37     &smp_spin_table_ops,
38     &cpu_psci_ops,
39     NULL,
40 };
41
42 -static const struct cpu_operations
43 +static const struct cpu_operations *const
44 #acpi_supported_cpu_ops[] __initconst = {
45     #ifdef CONFIG_ARM64_ACPI_PARKING_PROTOCOL
46     &acpi_parking_protocol_ops,
47     #endif
48 };
49
50 #endif
```

ChangeNodes

- Java
- Eclipse AST



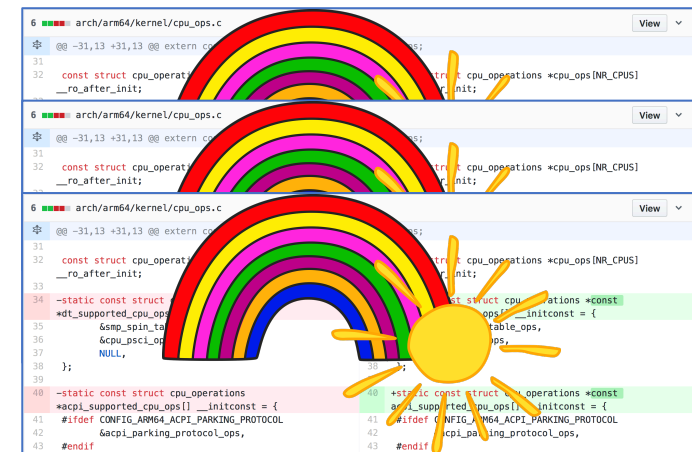
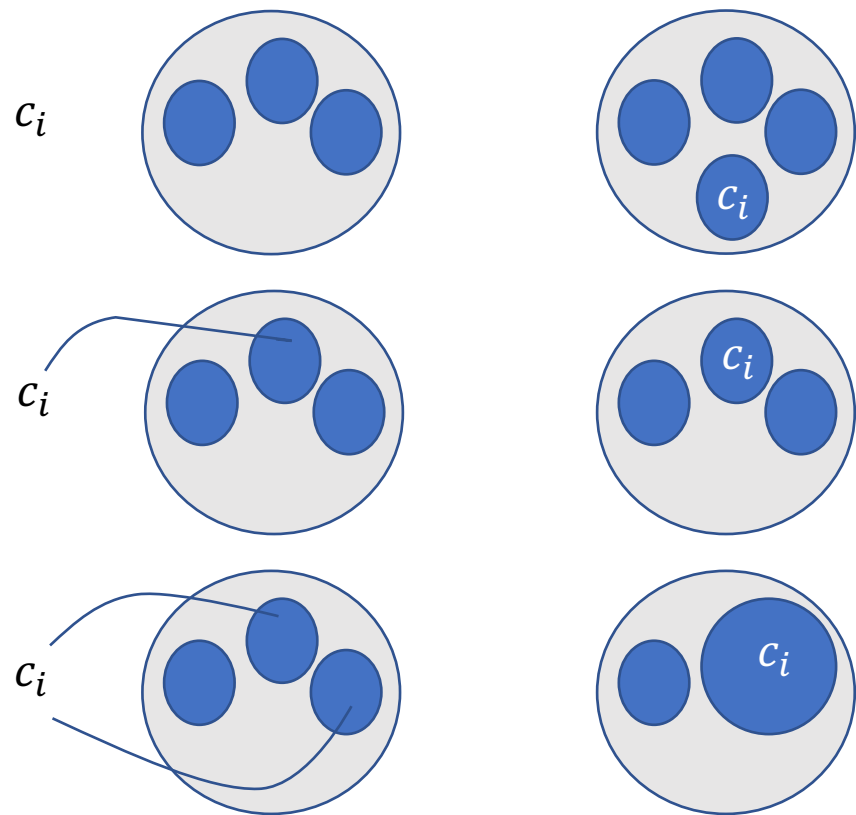
Fine-grained changes c_1, c_2, c_3, \dots

- Insert
- Update
- Move
- Delete



$$c_i \in S(c_j) \vee c_j \in S(c_i)$$

Slices $S(c_1), S(c_2), S(c_3), \dots$



Evaluation Dataset

ArgoUML,
GWT, Jaxen,
JRuby, Xstream

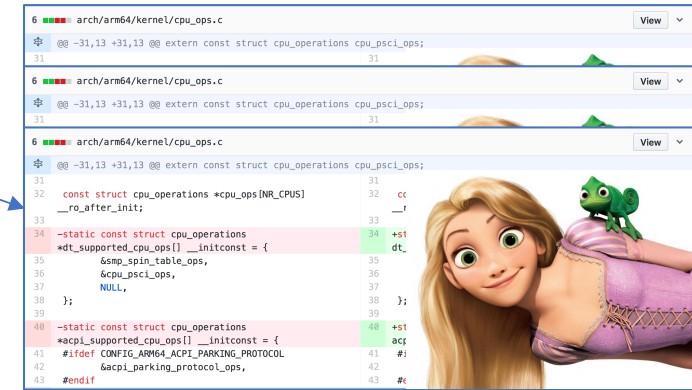
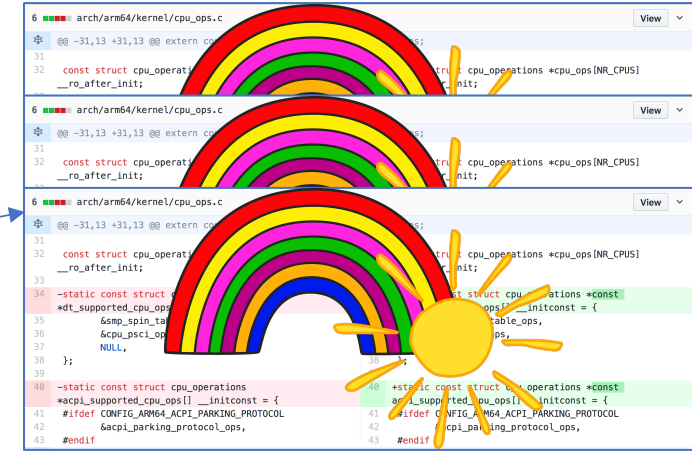


Table III

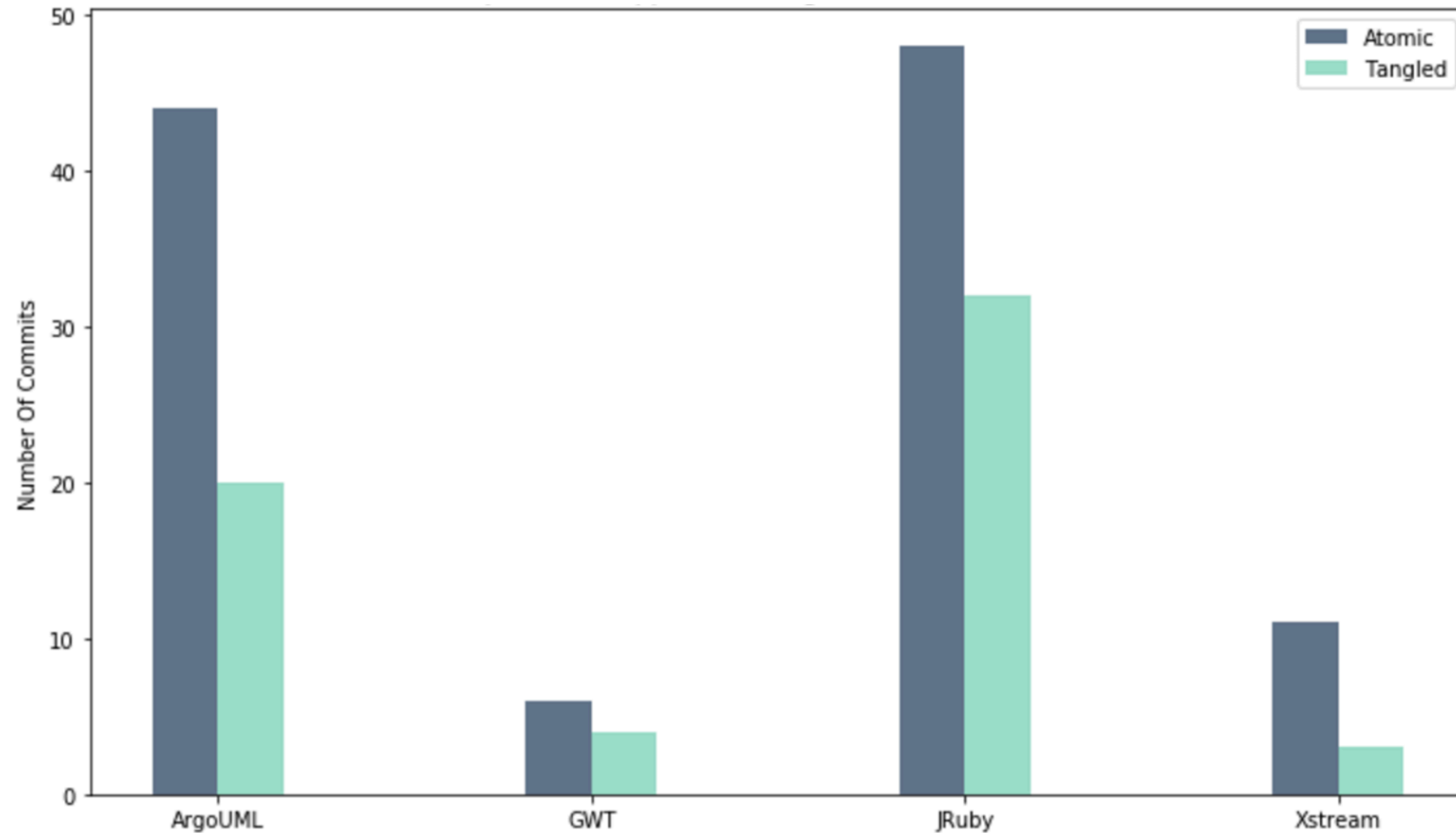
DETAILS OF PROJECTS USED DURING EXPERIMENTS.

	ArgoUML	Google Web Toolkit	Jaxen	JRuby	XStream
lines of code	164,851	266,115	20,997	101,799	22,021
#history months	150	54	114	105	90
#developers	50	120	20	67	12
Total #change sets	16481	5326	1353	11134	1756
#bug fixes	2,945	809	105	2,977	312
#atomic bug fixes	125 (4.2%)	44 (5.4%)	32 (30.5%)	200 (6.7%)	40 (12.8%)

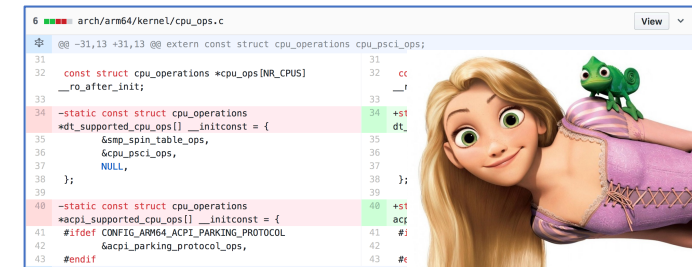
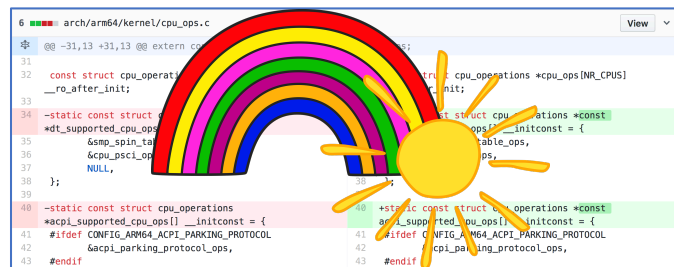
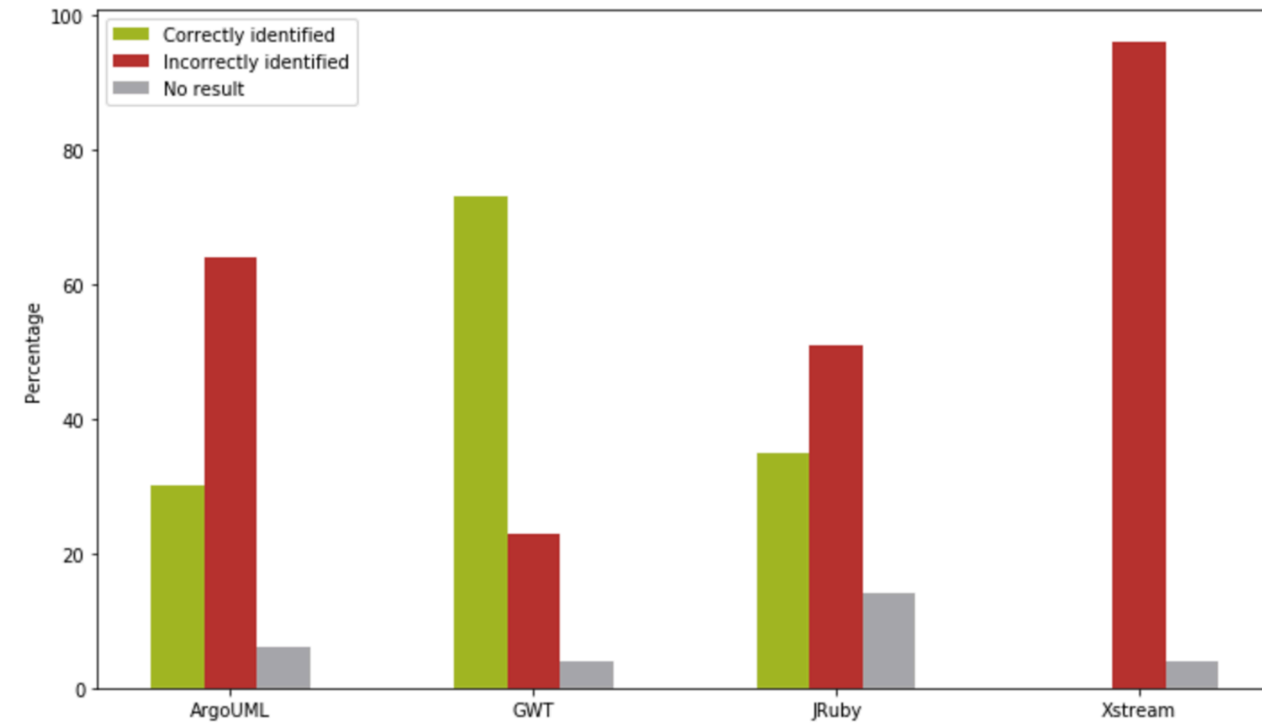
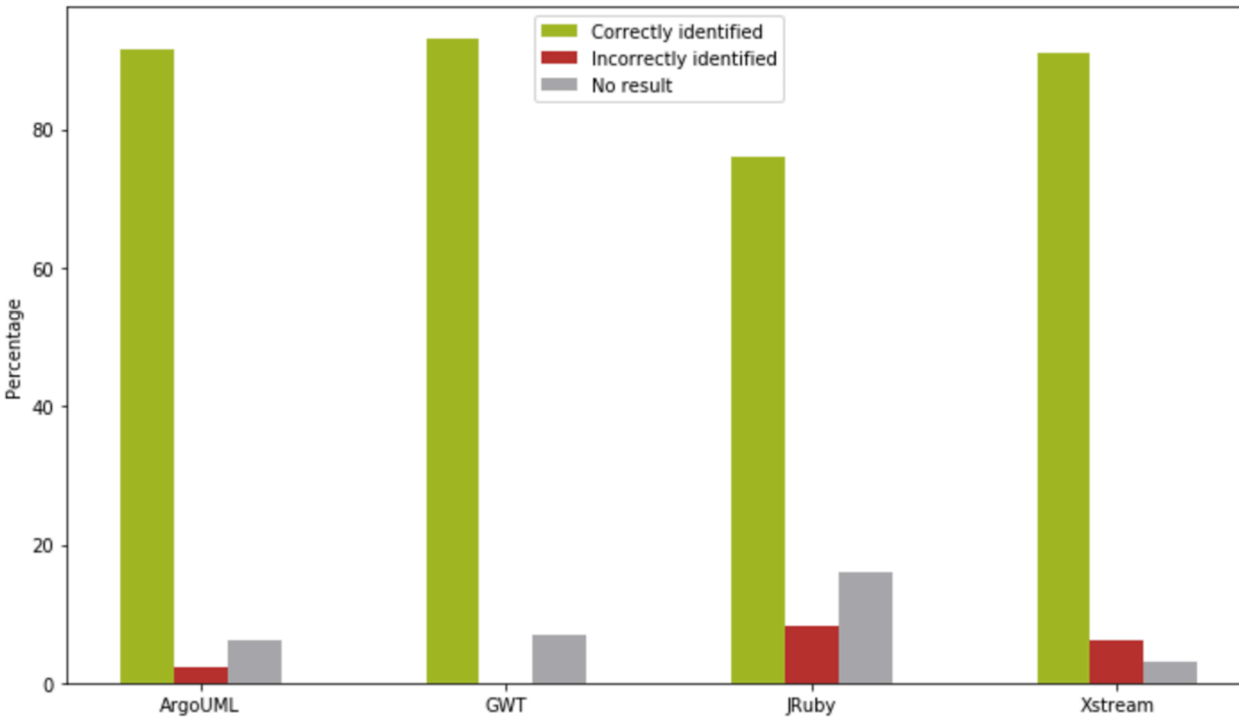
K. Herzig et al., “The impact of tangled code changes on defect prediction models,” *Empirical Software Engineering*, 2015.

K. Herzig et al., “Untangling Changes,” Saarland University, 2011. [Tech Report]

Preliminary Intra-procedural Results



Preliminary Intra-procedural Results

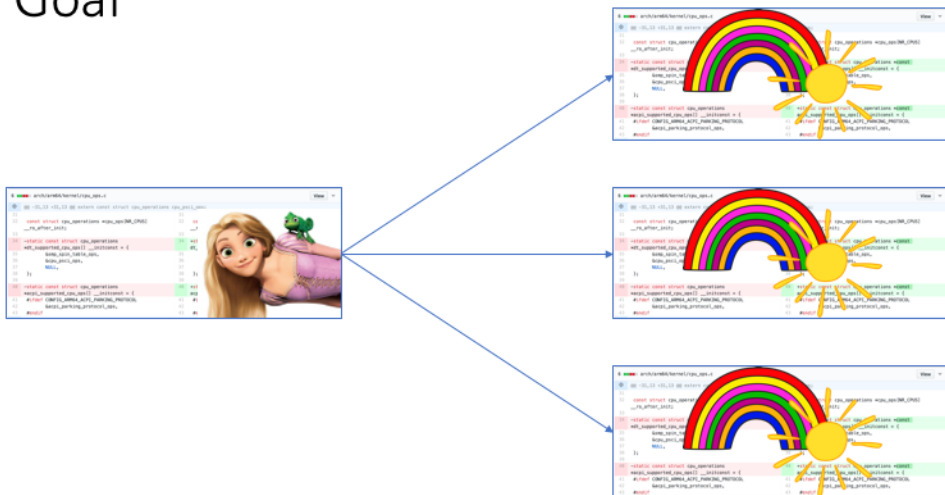


Current and Future Work

- Inter-procedural dependence graphs and slicing
- Perform full evaluation

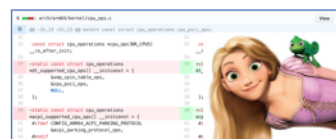


Goal

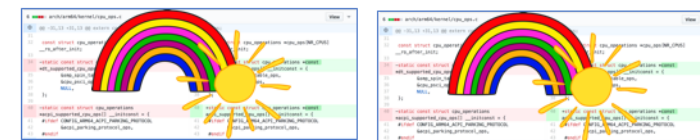


Hypothesis

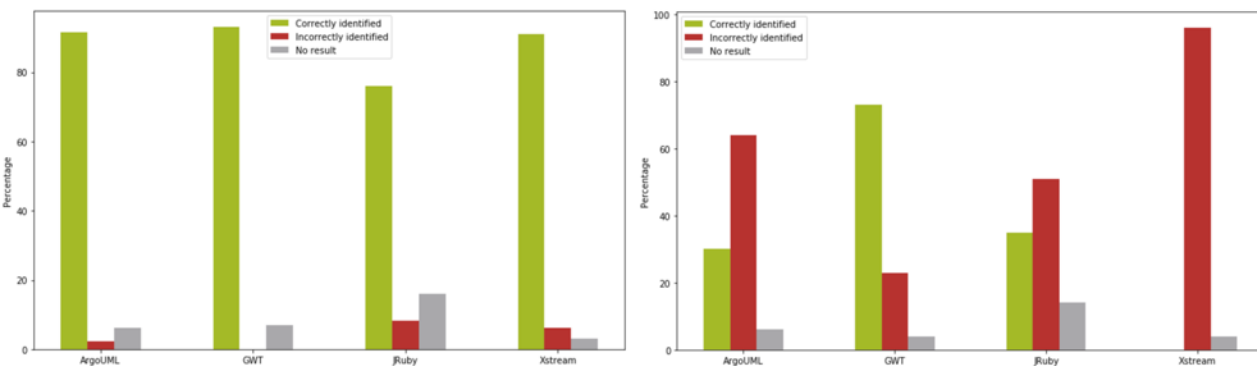
*Related changes are part of the same program slice.
A commit may be untangled using the program slices.*



$c_1 c_2 c_3 c_4 c_5 c_6 c_7 c_8$



Preliminary Intra-procedural Results

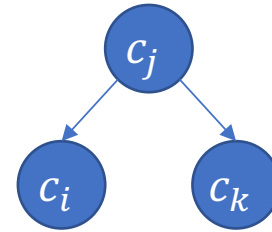


Current and Future Work

- Inter-procedural dependence graphs and slicing
- Perform full evaluation

Backup slide: reflexive, symmetric, transitive

$$c_i R c_j \Leftrightarrow c_i \in S(c_j) \vee c_j \in S(c_i)$$



$$\begin{array}{l} c_i R c_j \\ c_j R c_k \\ \neg(c_i R c_k) \end{array}$$

$$c_i R' c_j \Leftrightarrow \exists c_{t_1}, \dots, c_{t_n} : c_i R c_{t_1}, \dots, c_{t_n} R c_j$$

$$\begin{array}{l} c_i R' c_j \\ c_j R' c_k \\ c_i R' c_k \end{array}$$