# Variability Modeling in the Real: A Perspective from the Operating Systems Domain

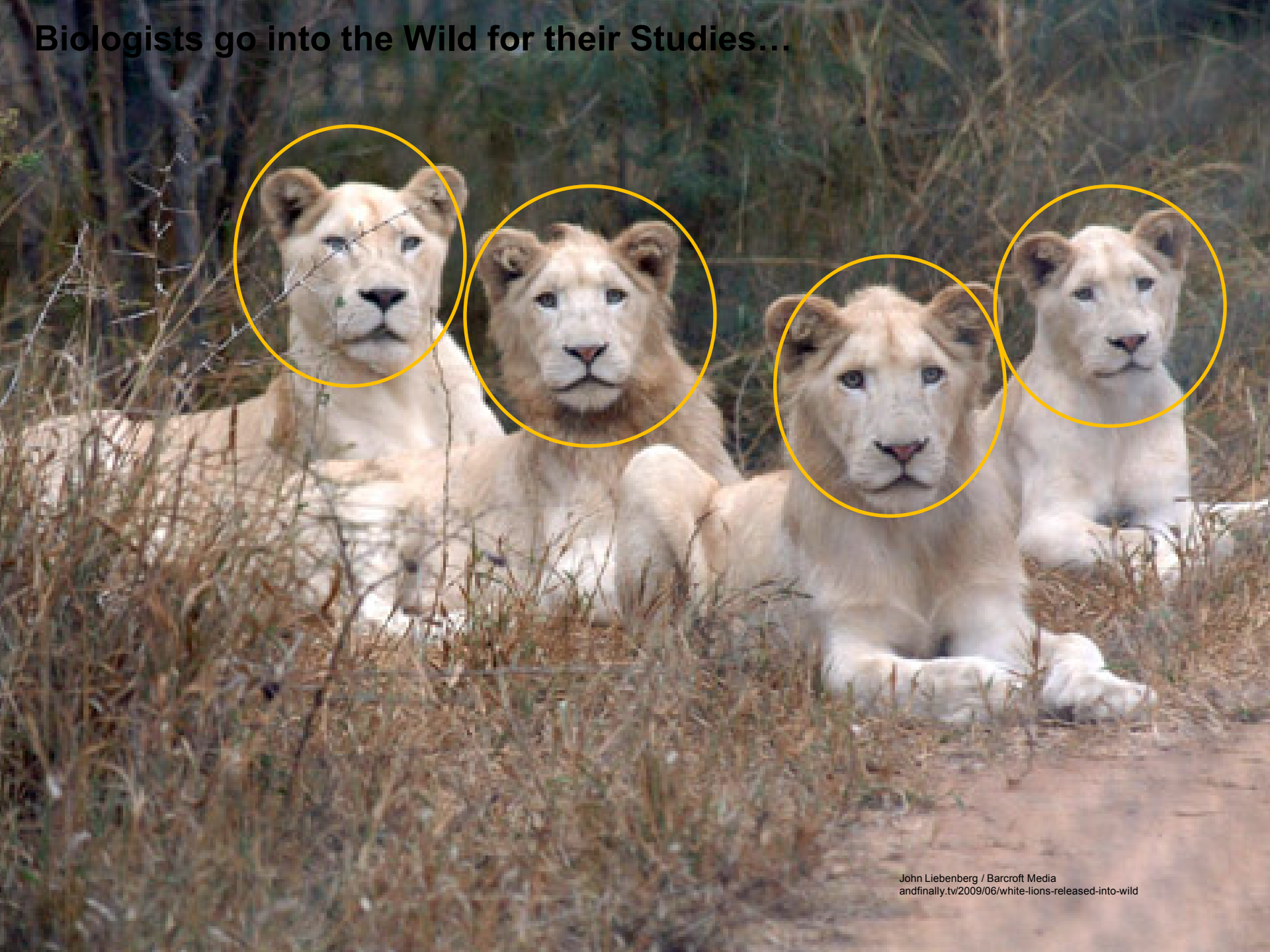
25th IEEE/ACM International Conference on Automated Software Engineering
ASE 2010
Antwerp, Belgium, Sept. 22nd, 2010

Thorsten Berger, Steven She, Rafael Lotufo,
Andrzej Wasowski, Krzysztof Czarnecki

# Biologists go into the Wild for their Studies…

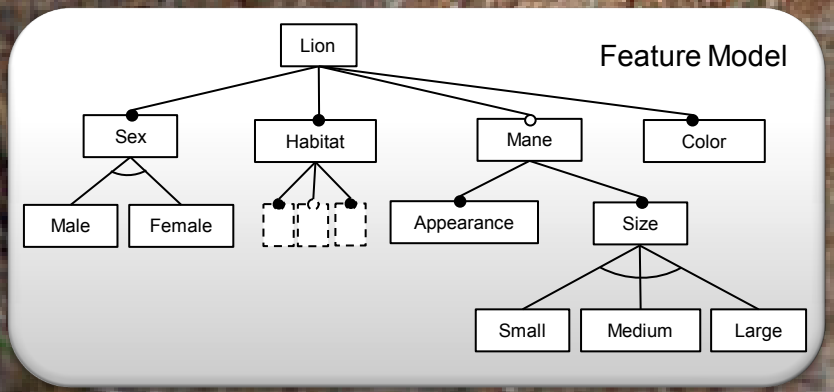
John Liebenberg / Barcroft Media
andfinally.tv/2009/06/white-lions-released-into-wild

# Biologists go into the Wild for their Studies…



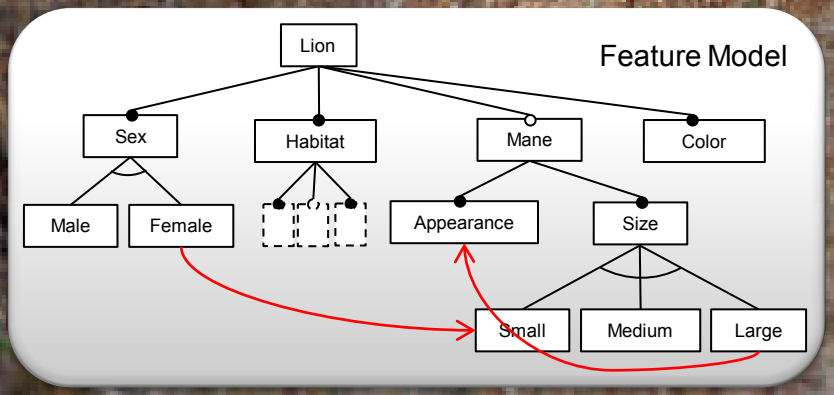
John Liebenberg / Barcroft Media
andfinally.tv/2009/06/white-lions-released-into-wild

# Biologists go into the Wild for their Studies…




John Liebenberg / Barcroft Media
andfinally.tv/2009/06/white-lions-released-into-wild

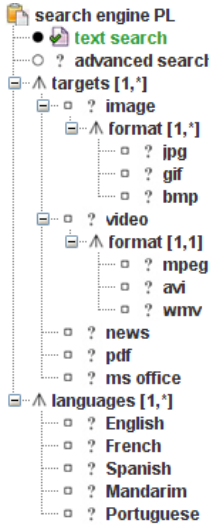
# Biologists go into the Wild for their Studies…



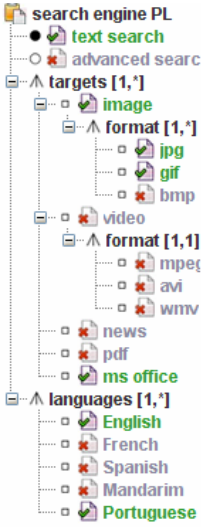

John Liebenberg / Barcroft Media
andfinally.tv/2009/06/white-lions-released-into-wild

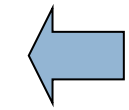
# Variability Modeling


**Variability Model**
(15810 possible configurations)

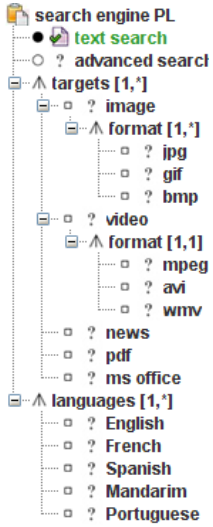


**Configuration #1**



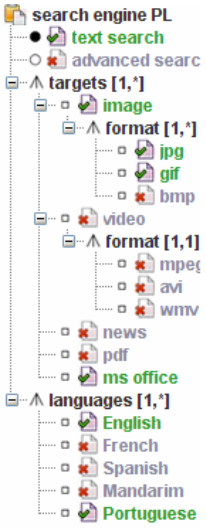



**Variability Model**
(15810 possible configurations)

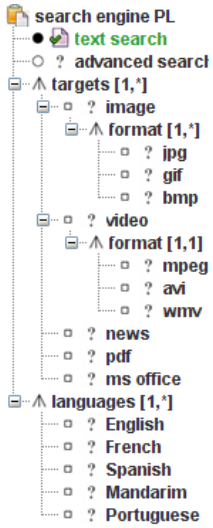




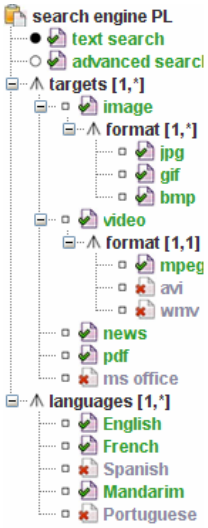

**Configuration #1**




**Variability Model**
(15810 possible configurations)



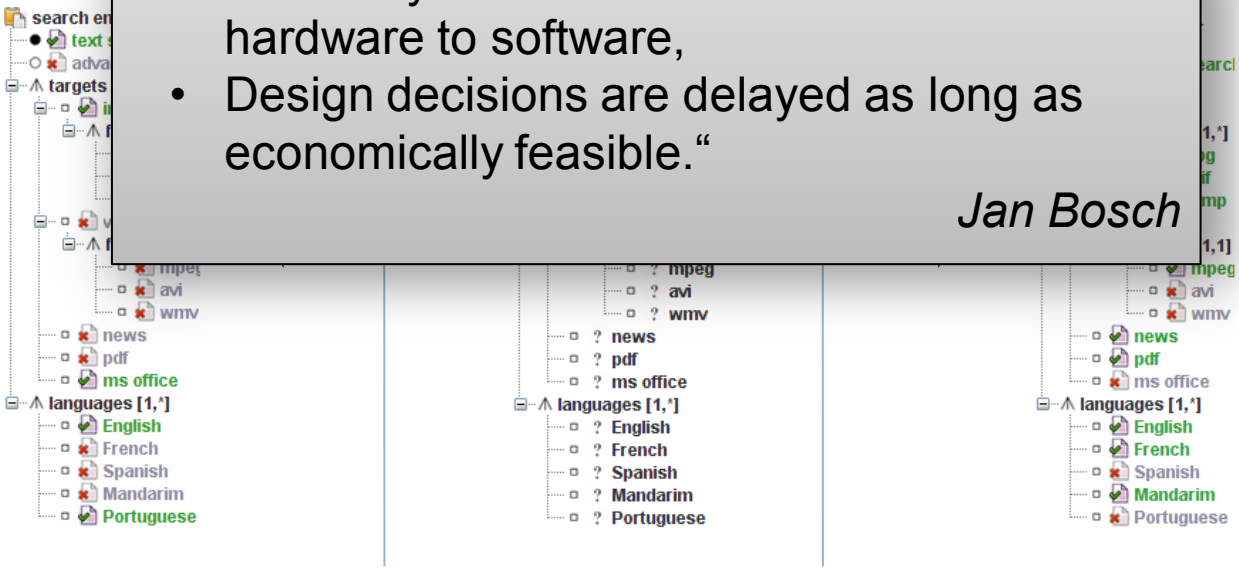

**Configuration #2**



"Variability needs in software are constantly increasing, because:

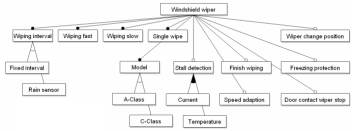
- Variability moves from mechanics and hardware to software,
- Design decisions are delayed as long as economically feasible."

*Jan Bosch*

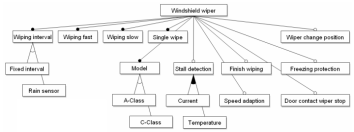



20 Years of Feature Models
Jubilee
Celebratee
Event
Hangover
Nice
Remarkable
Congratulations to
Kyo C. Kang
VaMoS 2010
Linz, Austria

## Other Domains




## Other Domains





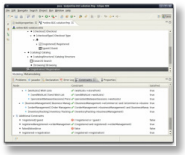
## Tools

pure·systems

BigLever Software, Inc.

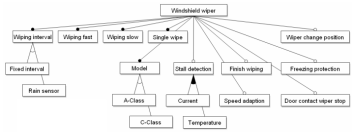
AHEAD
Feature Oriented Programming

FMP



IDE Feature

## Other Domains





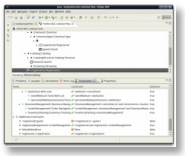
## Tools

pure-systems

BigLever Software, Inc.

AHEAD Feature Oriented Programming

FMP



FeatureIDE

## Standardization Efforts

OMG®

OBJECT MANAGEMENT GROUP

**Common Variability Language (CVL)**

eclipse

eclipse.org/proposals/feature-model

# NOW, WHAT‘S THE PROBLEM?

# What about Empirical Research?

- A lot of research on **variability**, but not on **real models**!

# What about Empirical Research?

- A lot of research on **variability**, but not on **real models**!
- Some **industry reports** available, but **not** the models!

# What about Empirical Research?

- A lot of research on **variability**, but not on **real models**!
- Some **industry reports** available, but **not** the models!
- **Assumptions** of real models used for **synthetic** examples!

# What about Empirical Research?

- A lot of research on **variability**, but not on **real models**!
- Some **industry reports** available, but **not** the models!
- **Assumptions** of real models used for **synthetic** examples!
- Literature studies:

# What about Empirical Research?

- A lot of research on **variability**, but not on **real models**!
- Some **industry reports** available, but **not** the models!
- **Assumptions** of real models used for **synthetic** examples!
- Literature studies:
  - *Hubaux et al. [VAMOS10]:*
    - **Only 2%** of reviewed papers (8 of 415) discuss **applications** of Feature Models in practice
    - Few details about their **usage** given

# What about Empirical Research?

- A lot of research on **variability**, but not on **real models**!
- Some **industry reports** available, but **not** the models!
- **Assumptions** of real models used for **synthetic** examples!
- Literature studies:
  - *Hubaux et al. [VAMOS10]:*
    - **Only 2%** of reviewed papers (8 of 415) discuss **applications** of Feature Models in practice
    - Few details about their **usage** given
  - *Chen et al. [SPLC09]:*
    - "There is **only little**, if any, **experimental** or detailed **comparative** analysis … of different VM approaches."
    - All VM approaches share **similar concepts**
    - Some sort of reference model needed for **model transformations**, **tools** and **future research**

# OUR STUDY…

# Three Research Questions

1. Can we provide quantitative and qualitative empirical evidence *whether* Feature Modeling concepts are used in real-world languages?

2. Are additional concepts needed?

3. Are the assumptions about real models in the literature correct?

# Three Research Questions

1. Can we provide quantitative and qualitative empirical evidence *whether* Feature Modeling concepts are used in real-world languages?

2. Are additional concepts needed?

3. Are the assumptions about real models in the literature correct?

# Three Research Questions

1. Can we provide quantitative and qualitative empirical evidence *whether* Feature Modeling concepts are used in real-world languages?

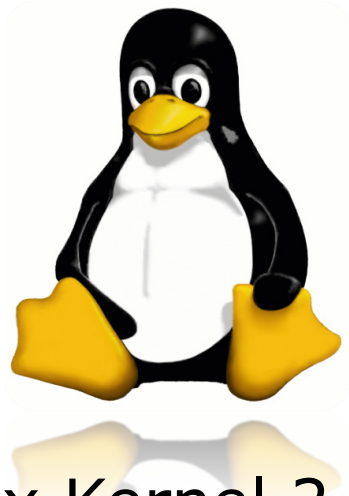
2. Are additional concepts needed?

3. Are the assumptions about real models in the literature correct?

We study

**CONCEPTS, SEMANTICS AND USAGE OF…**

## Kconfig language



Linux Kernel 2.6.32
(22 hardware architectures,
6.4 mio. SLOC)

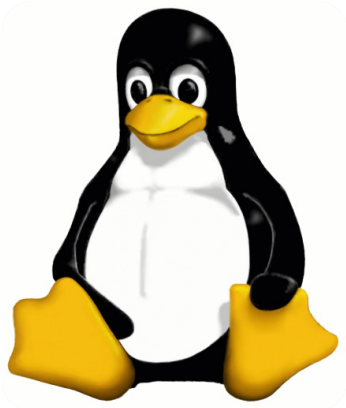
6320 Features (X86)

## CDL language

eCos 3.0
(embedded RTOS,
116 architectures,
~1 mio. SLOC)

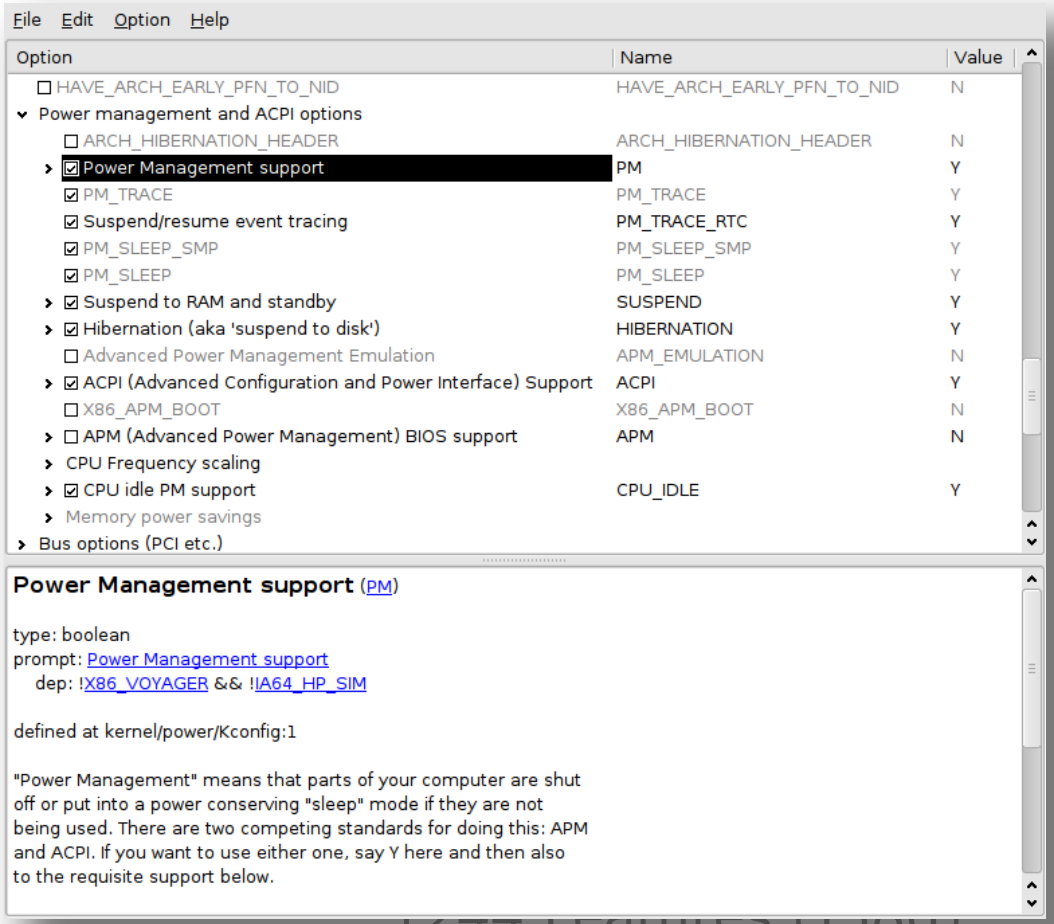
1244 Features (I386)

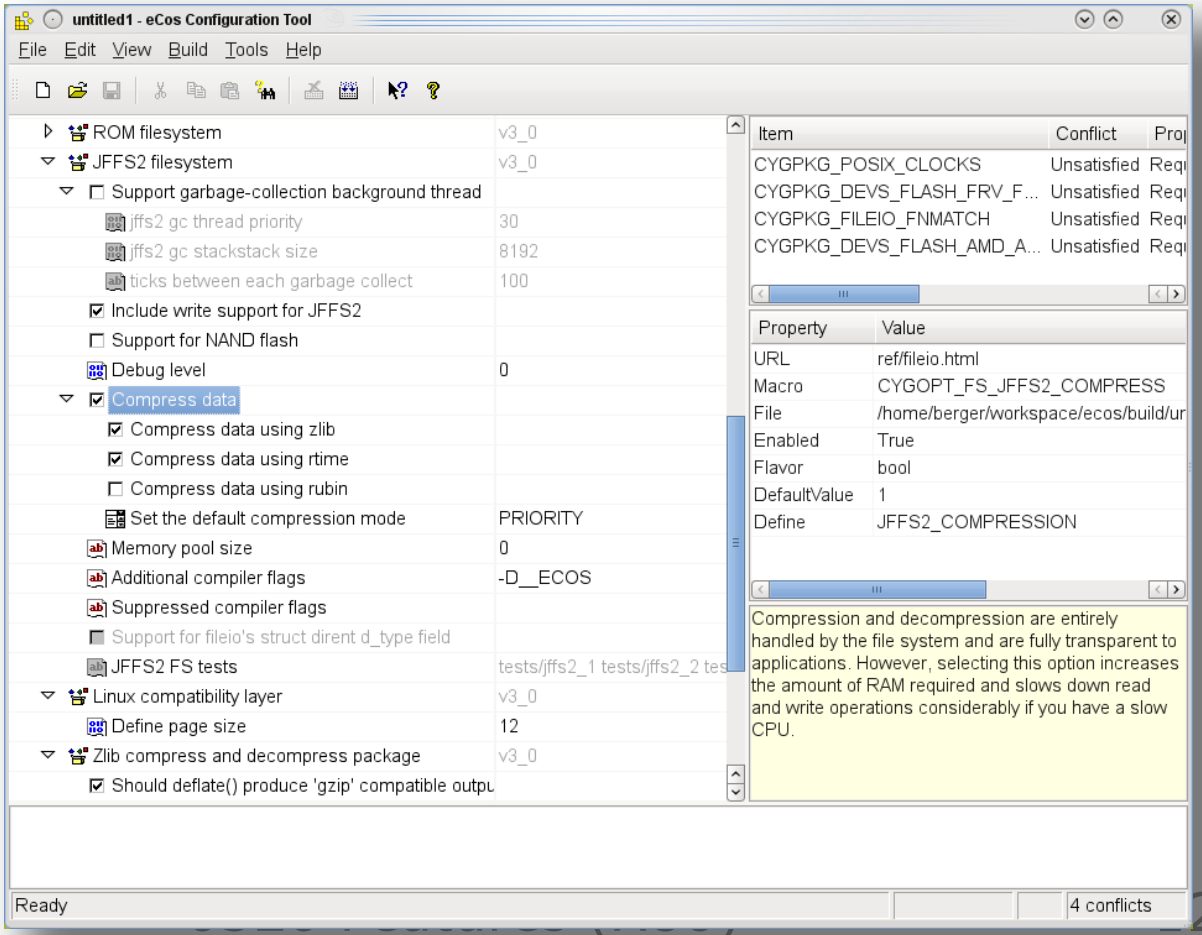
## Kconfig language



Linux Kernel 2.6.32
(22 hardware architectures,
6.4 mio. SLOC)

6320 Features (X86)






## CDL language

eCos 3.0
(embedded RTOS,
116 architectures,
~1 mio. SLOC)

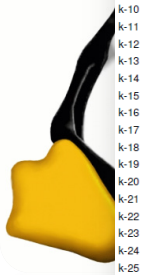
244 Features (I386)

## Kconfig language

```
menuconfig MISC_FILESYSTEMS
  bool "Miscellaneous filesystems"

 if MISC_FILESYSTEMS

  config JFFS2_FS
   tristate "Journalling Flash File System" if MTD
   select CRC32 if MTD




 config JFFS2_FS_DEBUG
  int "JFFS2 Debug level (0=quiet, 2=noisy)"
  depends on JFFS2_FS
  default 0
  range 0 2
  --- help ---
    Debug verbosity of ...


 config JFFS2_FS_WRITEBUFFER
  bool
  depends on JFFS2_FS
  default HAS_IOMEM


 config JFFS2_COMPRESS
  bool "Advanced compression options for JFFS2"
  depends on JFFS2_FS

 config JFFS2_ZLIB
  bool "Compress w/zlib..." if JFFS2_COMPRESS
  depends on JFFS2_FS
  select ZLIB_INFLATE
  default y

 choice
  prompt "Default compression" if JFFS2_COMPRESS
  default JFFS2_CMODE_PRIORITY
  depends on JFFS2_FS
  config JFFS2_CMODE_NONE
   bool "no compression"
  config JFFS2_CMODE_PRIORITY
   bool "priority"
  config JFFS2_CMODE_SIZE
   bool "size (EXPERIMENTAL)"
 endchoice
endif
```

## CDL language

```
cdl_component MISC_FILESYSTEMS {
 display   "Miscellaneous filesystems"
 flavor    none
}

cdl_package CYGPKG_FS_JFFS2 {
  display    "Journalling Flash File System"
  requires   CYGPKG_CRC
  implements CYGINT_IO_FILEIO
  parent     MISC_FILESYSTEMS
  active_if  MTD

  cdl_option CYGOPT_FS_JFFS2_DEBUG {
   display          "Debug level"
   flavor           data
   default_value    0
   legal_values     0 to 2
   define           CONFIG_JFFS2_FS_DEBUG
   description      "Debug verbosity of...."
  }

  cdl_option CYGOPT_FS_JFFS2_NAND {
   flavor           bool
   define           CONFIG_JFFS2_FS_WRITEBUFFER
   calculated       HAS_IOMEM
  }

  cdl_component CYGOPT_FS_JFFS2_COMPRESS {
   display          "Compress data"
   default_value    1

     cdl_option CYGOPT_FS_JFFS2_COMPRESS_ZLIB {
      display          "Compress data using zlib"
      requires         CYGPKG_COMPRESS_ZLIB
      default_value    1
     }

     cdl_option CYGOPT_FS_JFFS2_COMPRESS_CMODE {
      display          "Set the default compression mode"
      flavor           data
      default_value    { "PRIORITY" }
      legal_values     { "NONE" "PRIORITY" "SIZE" }
   }
  }
}
```



Linux K
(22 hardw
6.4

3.0
RTOS,
ctures,
LOC)

6320 F

res (I386)

What do we mean by variability model

# SEMANTICS?

- Configuration Space Semantics

- Configuration Space Semantics

Kconfig Model

```
menuconfig MISC_FILESYSTEMS
  bool "Miscellaneous filesystems"

if MISC_FILESYSTEMS

  config JFFS2_FS
   tristate "Journalling Flash FS" if MTD
   select CRC32 if MTD

  config JFFS2_FS_DEBUG
   int "JFFS2 Debug level (0=quiet,
2=noisy)"
   depends on JFFS2_FS
   default 0
   range 0 2
   --- help ---
    Debug verbosity of ...
```

$$[\![\cdot]\!]_{\text{kconfig}} : \text{Kconfig} \rightarrow \mathcal{P}(\text{Confs})$$

$C_1$={(JFFS2,y),(JFFS2_DEBUG,2),…}

$C_2$={(JFFS2,m),(JFFS2_DEBUG,0),…}

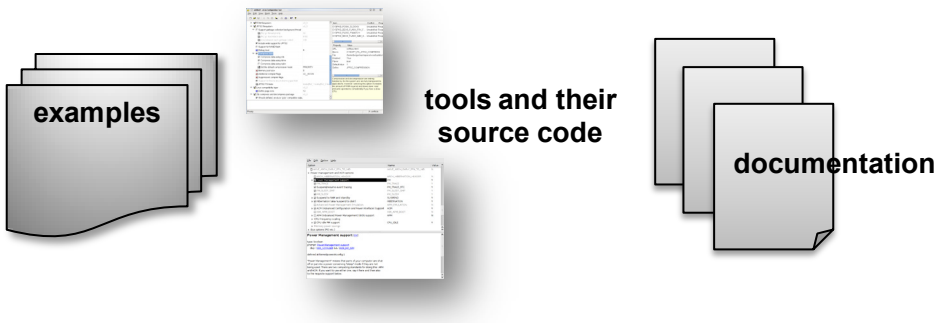
$C_n$={(JFFS2,n),(JFFS2_DEBUG,0),…}

# METHODOLOGY

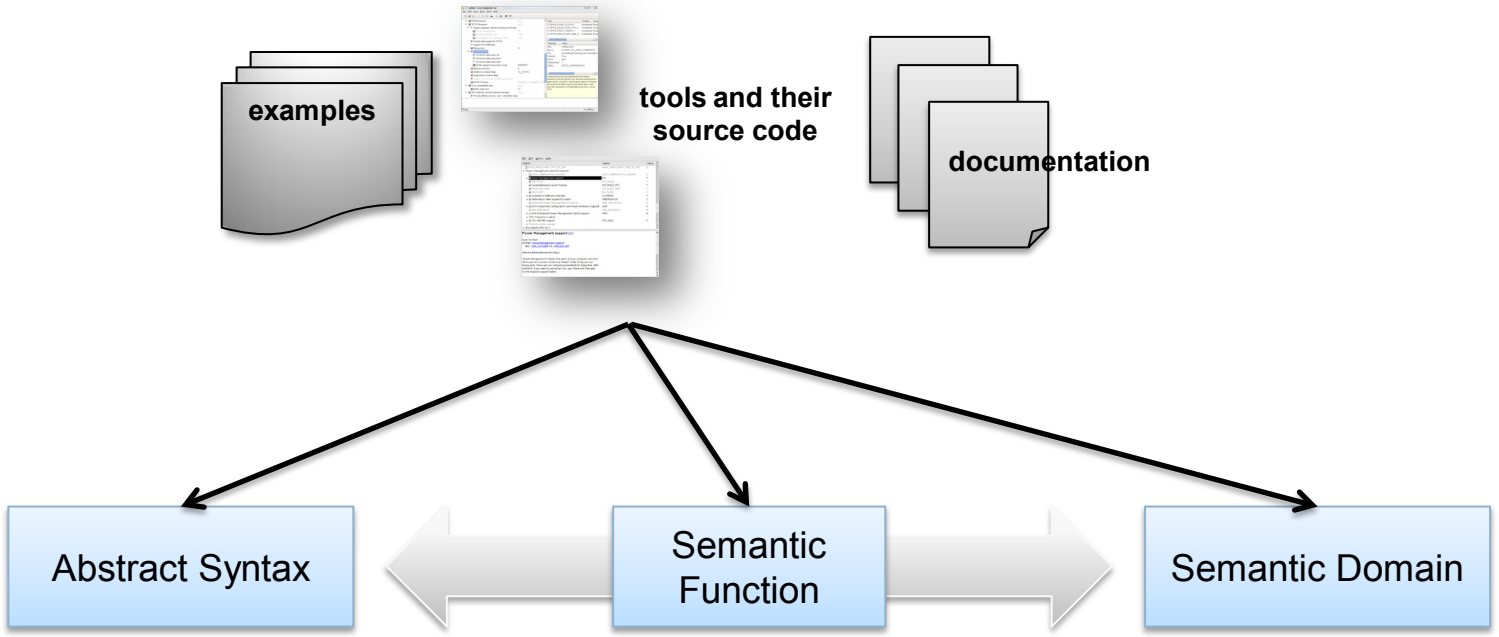
# Methodology - Design Space

- Reverse-engineered formal semantics (denotational style)

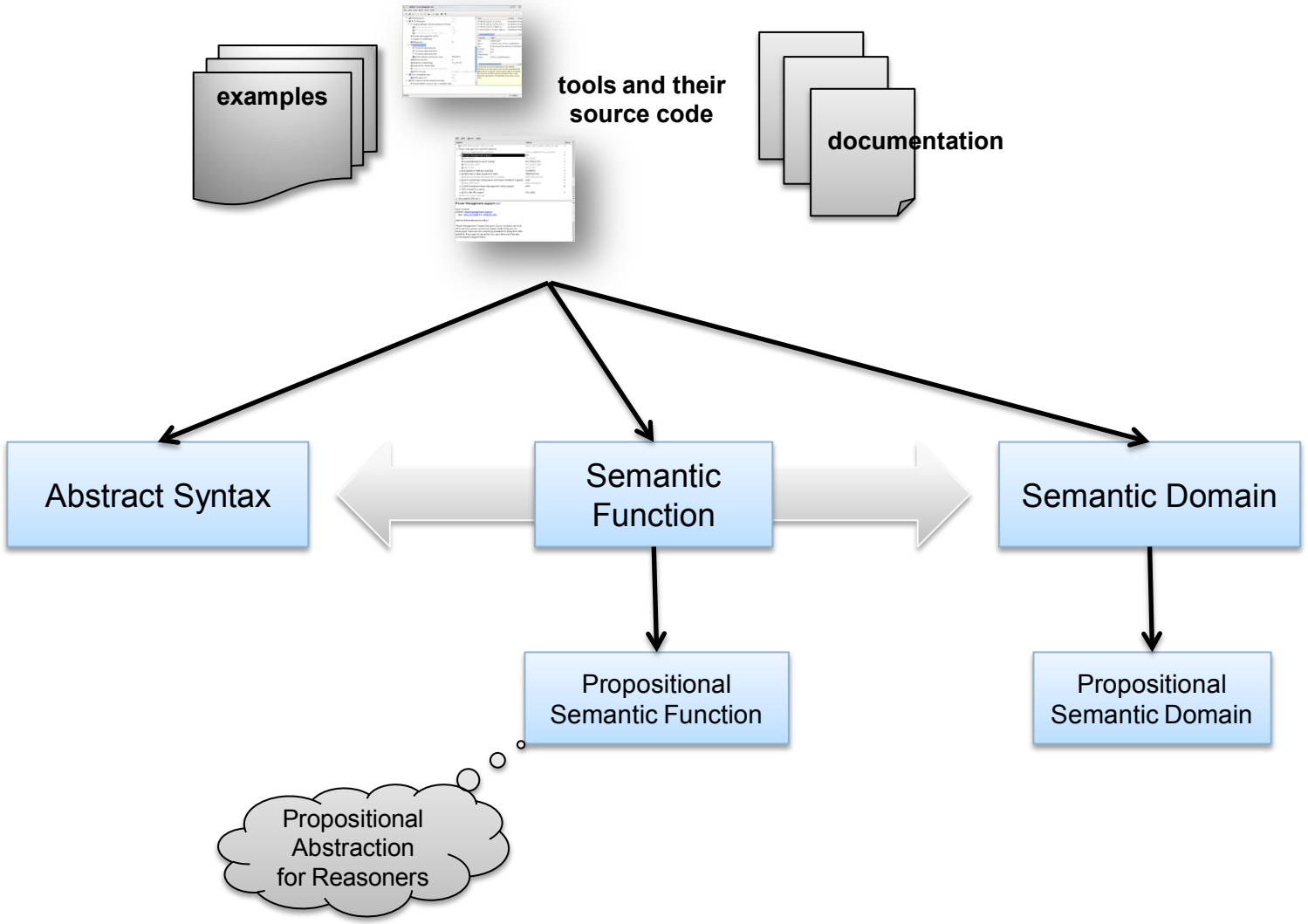
- Reverse-engineered formal semantics (denotational style)



- Reverse-engineered formal semantics (denotational style)

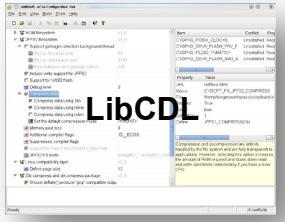


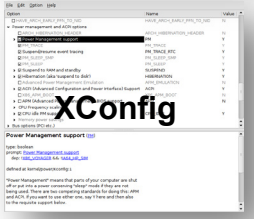
- Reverse-engineered formal semantics (denotational style)



# Methodology - Language Usage

extension of configurators

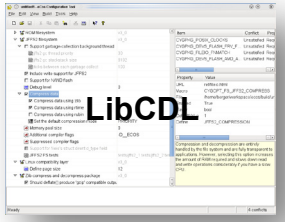


# Methodology - Language Usage

extension of configurators

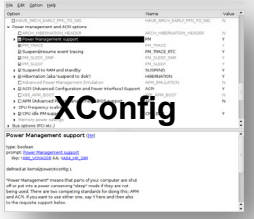
Linux and eCos models




model export

CDL

Intermediate Model




model export

Kconfig

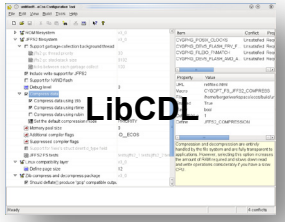
Intermediate Model

# Methodology - Language Usage

extension of configurators

Linux and eCos models

analysis infrastructure

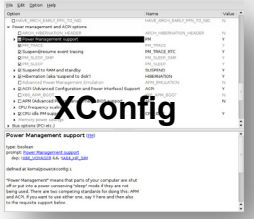



model export

CDL

Intermediate Model

parsing




model export

Kconfig

Intermediate Model

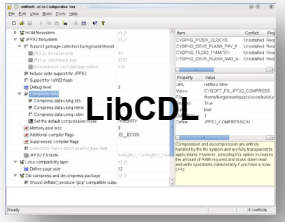
parsing

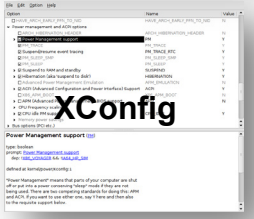
# Methodology - Language Usage

extension of configurators

Linux and eCos models

analysis infrastructure




model export

model export

CDL

Intermediate Model

Kconfig

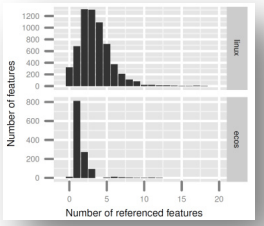
Intermediate Model

parsing

parsing

statistics

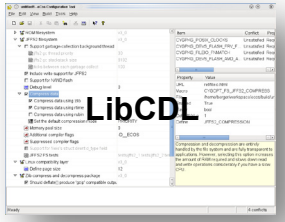
statistics

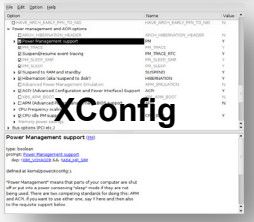


extension of configurators

Linux and eCos models

analysis infrastructure




model export

model export

CDL

Intermediate Model

Kconfig

Intermediate Model

parsing

parsing

transformation

statistics

statistics

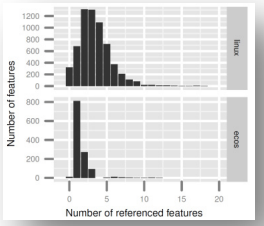
transformation

propositional abstraction

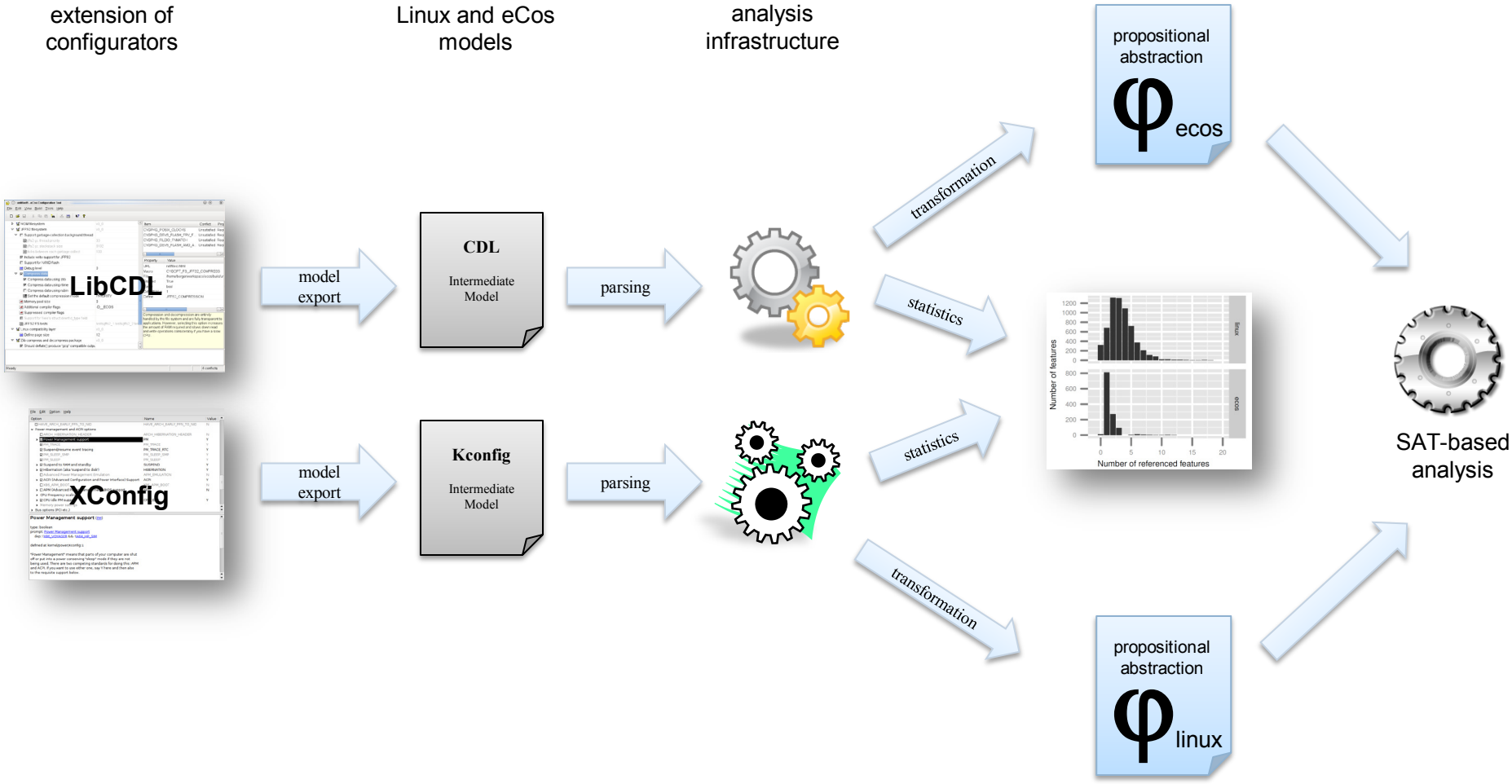
$\varphi_{ecos}$

propositional abstraction

$\varphi_{linux}$



# Methodology - Language Usage



# RESULTS

# Three Research Questions

1. Can we provide quantitative and qualitative empirical evidence *whether* Feature Modeling concepts are used in real-world languages?

2. Are additional concepts needed?

3. Are the assumptions about real models in the literature correct?

# Three Research Questions

1. Can we provide quantitative and qualitative empirical evidence *whether* Feature Modeling concepts are used in real-world languages?


2. Are additional concepts needed?

3. Are the assumptions about real models in the literature correct?

- Concepts:
  - Boolean (optional), Integer and String features (attributes)
  - Hierarchy
  - Group constraints
  - Cross-tree constraints

# Feature Modeling Concepts

- Concepts:
  - Boolean (optional), Integer and String features (attributes)
  - Hierarchy
  - Group constraints
  - Cross-tree constraints

- But intricate semantic interactions
  - Needed many iterations to get the formal semantics right
  - Often slightly different from what we expected

# Feature Modeling Concepts

- Concepts:
  - Boolean (optional), Integer and String features (attributes)
  - Hierarchy
  - Group constraints
  - Cross-tree constraints
- But intricate semantic interactions
  - Needed many iterations to get the formal semantics right
  - Often slightly different from what we expected
- Some (but minor) violations of feature modeling rules

- Concepts:
  - Boolean (optional), Integer and String features (attributes)
  - Hierarchy
  - Group constraints
  - Cross-tree constraints
- But intricate semantic interactions
  - Needed many iterations to get the formal semantics right
  - Often slightly different from what we expected
- Some (but minor) violations of feature modeling rules
- Languages benefit from being domain-specific

# Three Research Questions

1. Can we provide quantitative and qualitative empirical evidence *whether* Feature Modeling concepts are used in real-world languages?


2. Are additional concepts needed?

3. Are the assumptions about real models in the literature correct?

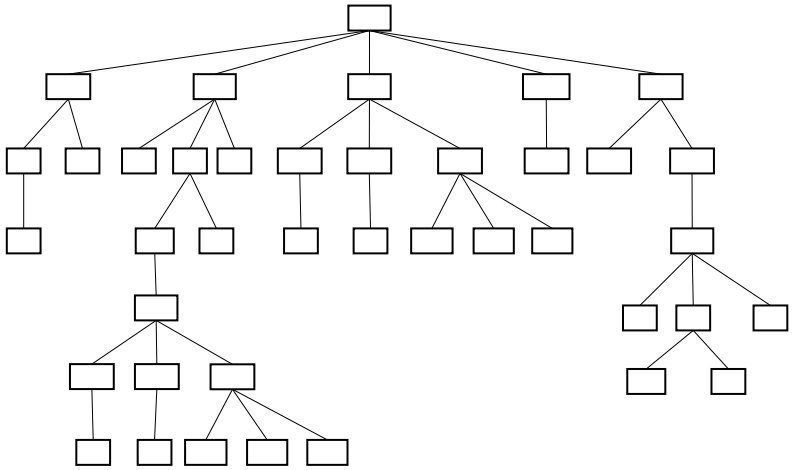
# Three Research Questions

1. Can we provide quantitative and qualitative empirical evidence *whether* Feature Modeling concepts are used in real-world languages?

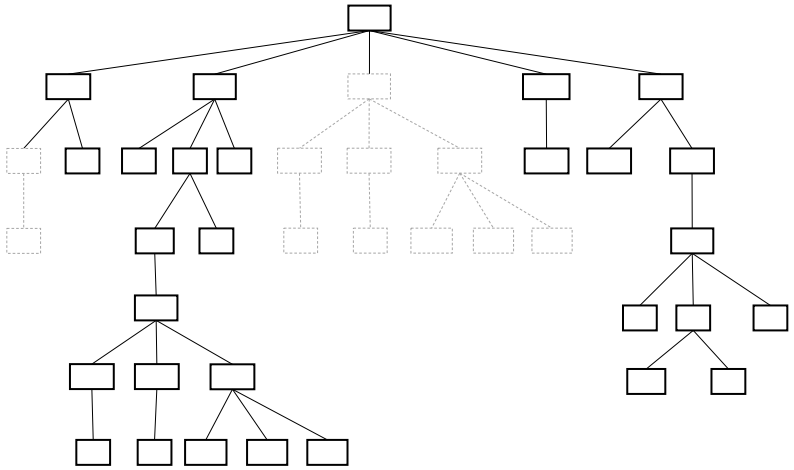

2. Are additional concepts needed?


3. Are the assumptions about real models in the literature correct?

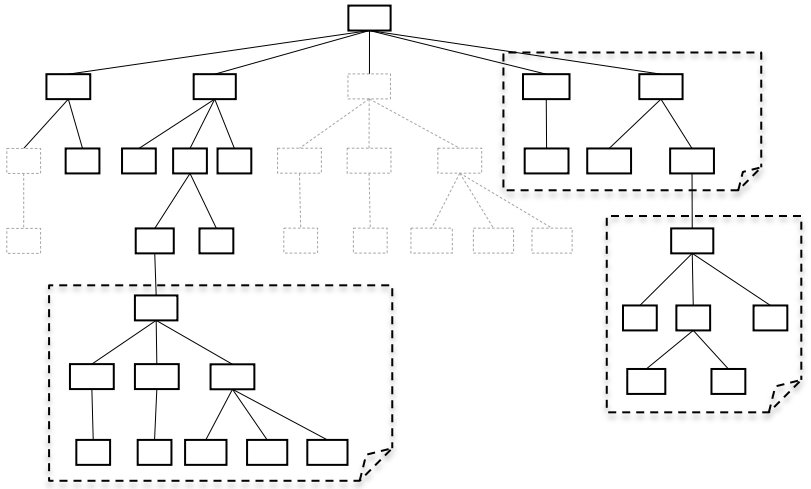
- Concepts for scalability



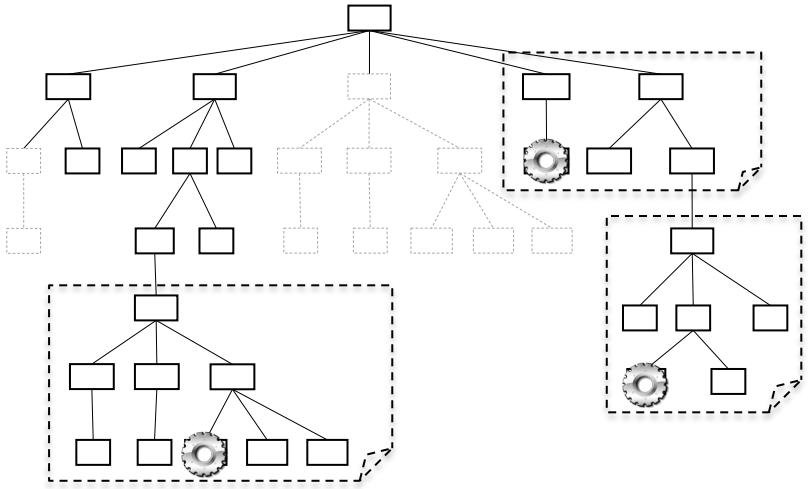
- Concepts for scalability
  - Visibility



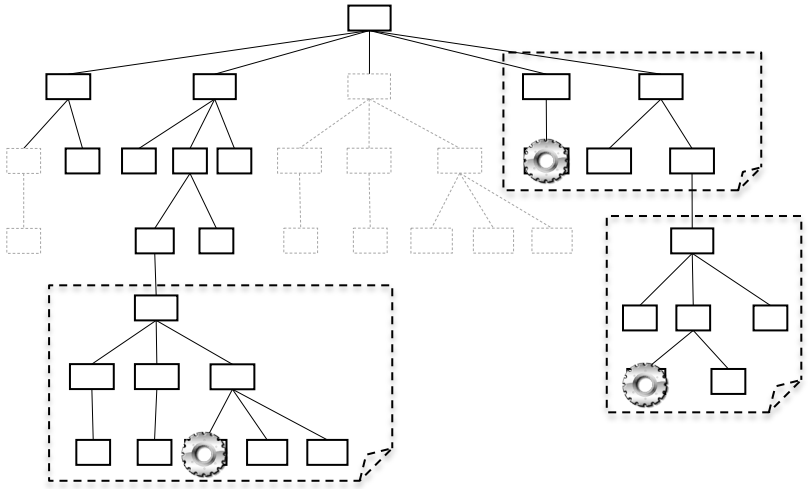
- Concepts for scalability
  - Visibility
  - Modularization



- Concepts for scalability
  - Visibility
  - Modularization
  - Derived defaults / derived features

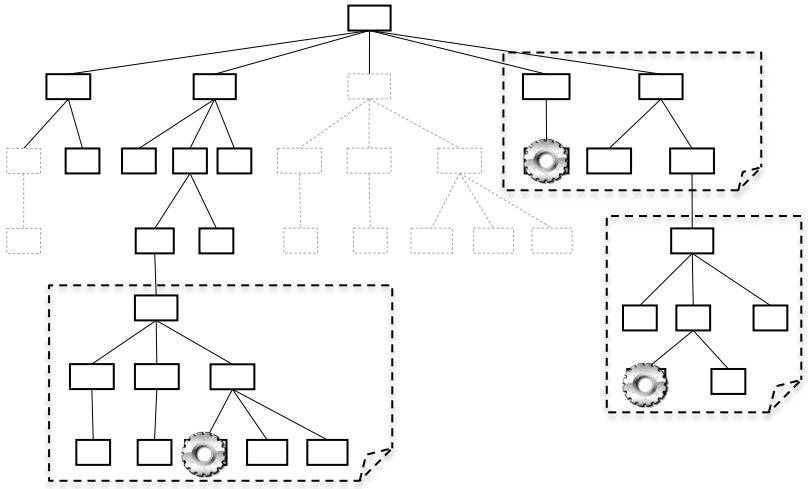


- Concepts for scalability
  - Visibility
  - Modularization
  - Derived defaults / derived features



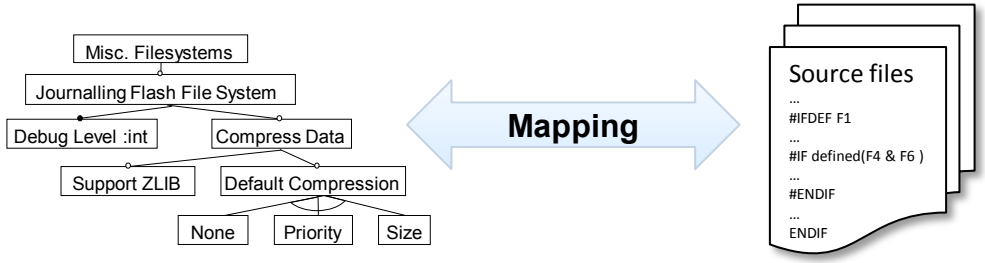
- Expressive constraints
  - Kconfig: Three-state logic (follows Kleene‘s rules)
  - CDL: Comparison, arithmetic and String operators

- Concepts for scalability
  - Visibility
  - Modularization
  - Derived defaults / derived features



- Expressive constraints
  - Kconfig: Three-state logic (follows Kleene‘s rules)
  - CDL: Comparison, arithmetic and String operators

- Code mappings / build specifications



# Three Research Questions

1. Can we provide quantitative and qualitative empirical evidence *whether* Feature Modeling concepts are used in real-world languages?


2. Are additional concepts needed?


3. Are the assumptions about real models in the literature correct?

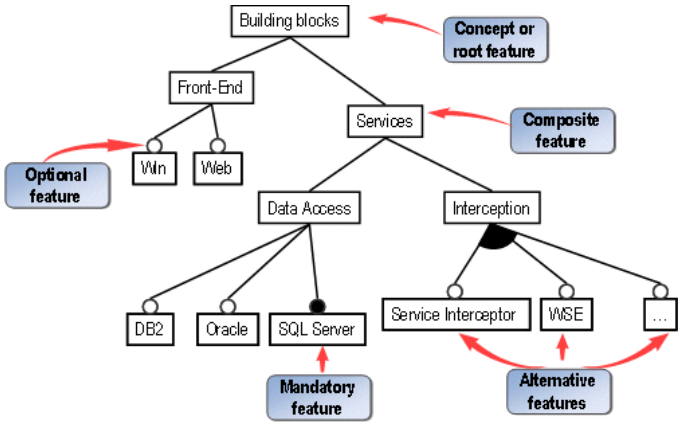
# Three Research Questions

1. Can we provide quantitative and qualitative empirical evidence *whether* Feature Modeling concepts are used in real-world languages?

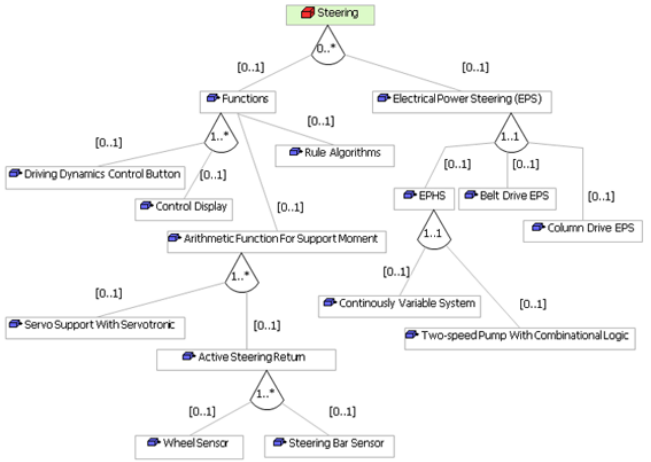

2. Are additional concepts needed?


3. Are the assumptions about real models in the literature correct?

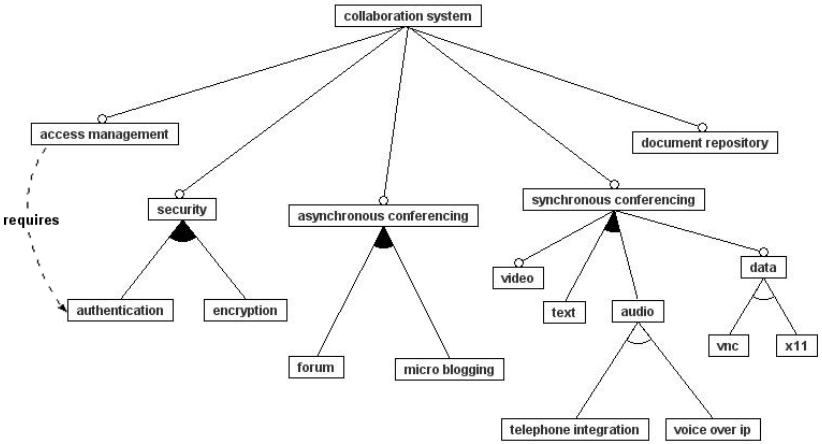
- We always see nicely balanced trees…




msdn.microsoft.com/en-us/library/aa925157.aspx




www.feasiple.de/description/bsp_ess_en.html




code.google.com/p/dslvariantmanagement/wiki/DemoShowCase

- But Linux and eCos models are very shallow!

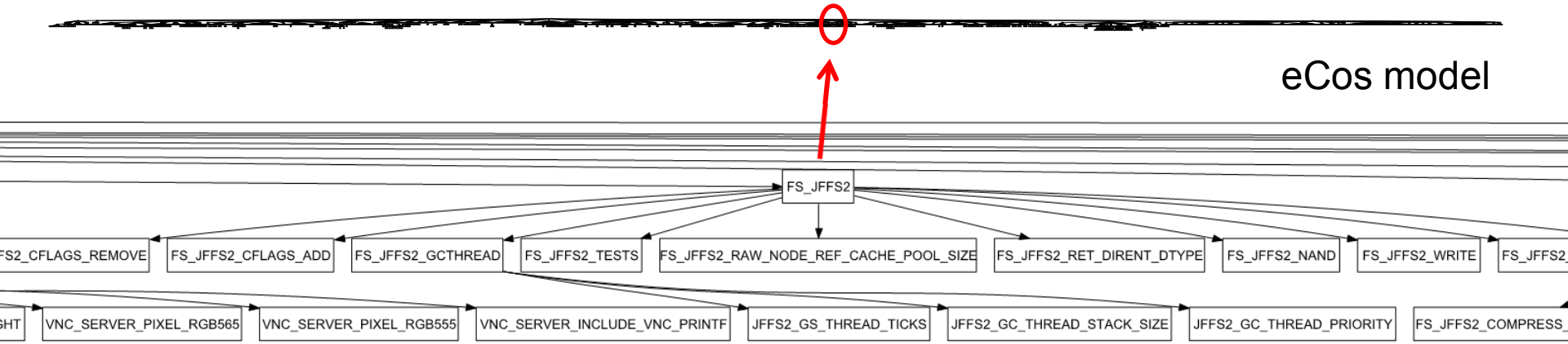
| | Avg. depth | Max. depth | Max. branching | No. of leaves |
|---|---|---|---|---|
| Linux | 4 | 8 | 158 | 5316 (84%) |
| eCos | 3 | 6 | 29 | 947 (76%) |

- But Linux and eCos models are very shallow!

eCos model

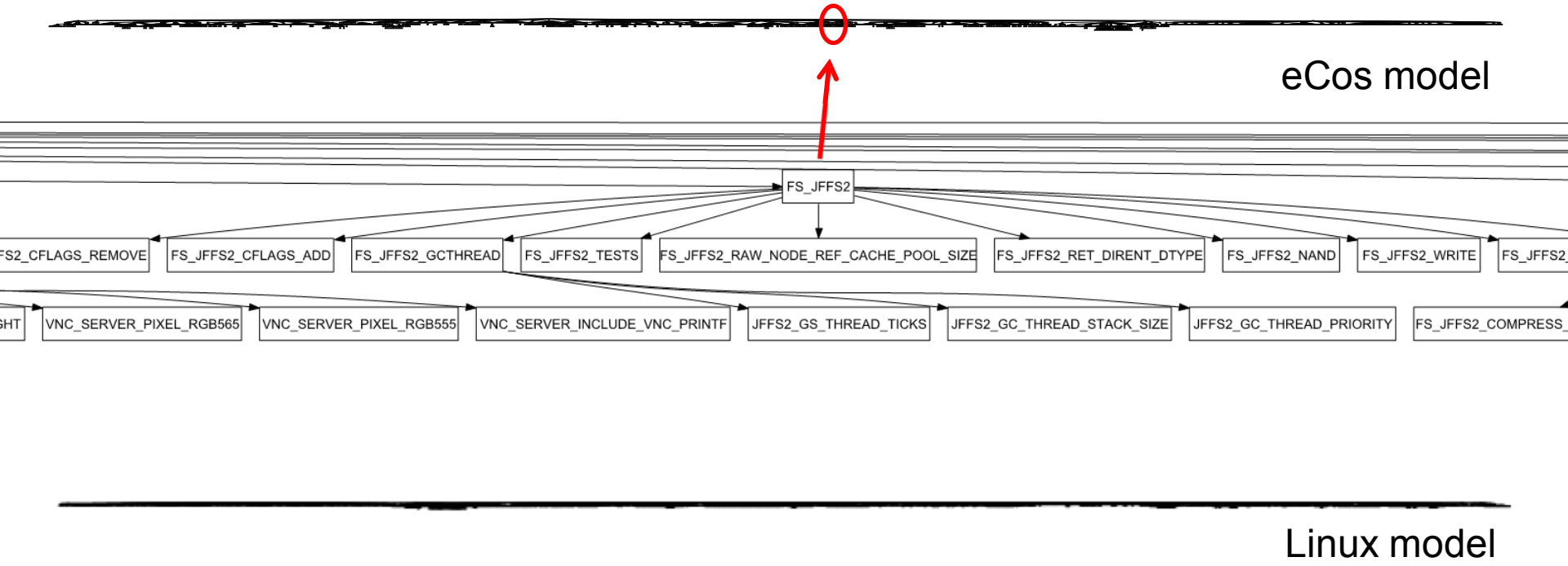
| | Avg. depth | Max. depth | Max. branching | No. of leaves |
|---|---|---|---|---|
| Linux | 4 | 8 | 158 | 5316 (84%) |
| eCos | 3 | 6 | 29 | 947 (76%) |

- But Linux and eCos models are very shallow!




| | Avg. depth | Max. depth | Max. branching | No. of leaves |
|---|---|---|---|---|
| Linux | 4 | 8 | 158 | 5316 (84%) |
| eCos | 3 | 6 | 29 | 947 (76%) |

- But Linux and eCos models are very shallow!



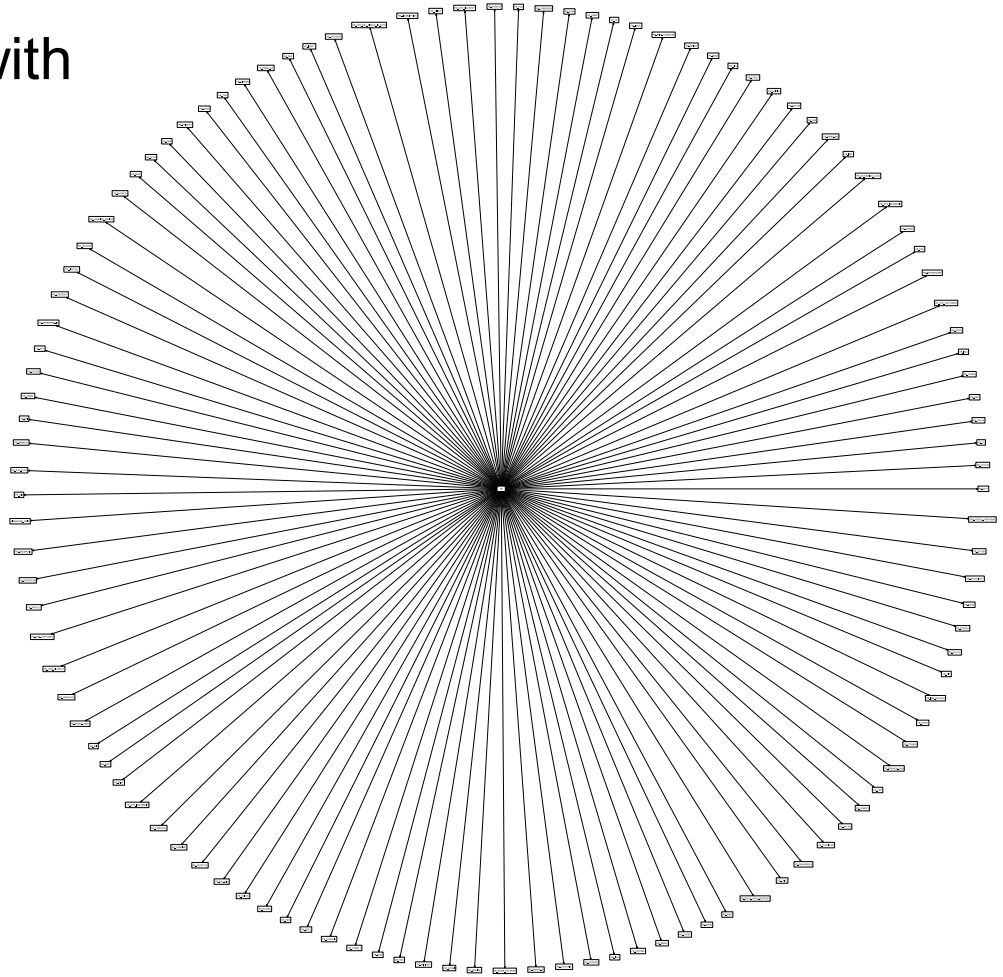

| | Avg. depth | Max. depth | Max. branching | No. of leaves |
|---|---|---|---|---|
| Linux | 4 | 8 | 158 | 5316 (84%) |
| eCos | 3 | 6 | 29 | 947 (76%) |

# Assumptions

- Branching factors with high variation!

| | Avg. depth | Max. depth | Max. branching | No. of leaves |
|---|---|---|---|---|
| Linux | 4 | 8 | 158 | 5316 (84%) |
| eCos | 3 | 6 | 29 | 947 (76%) |

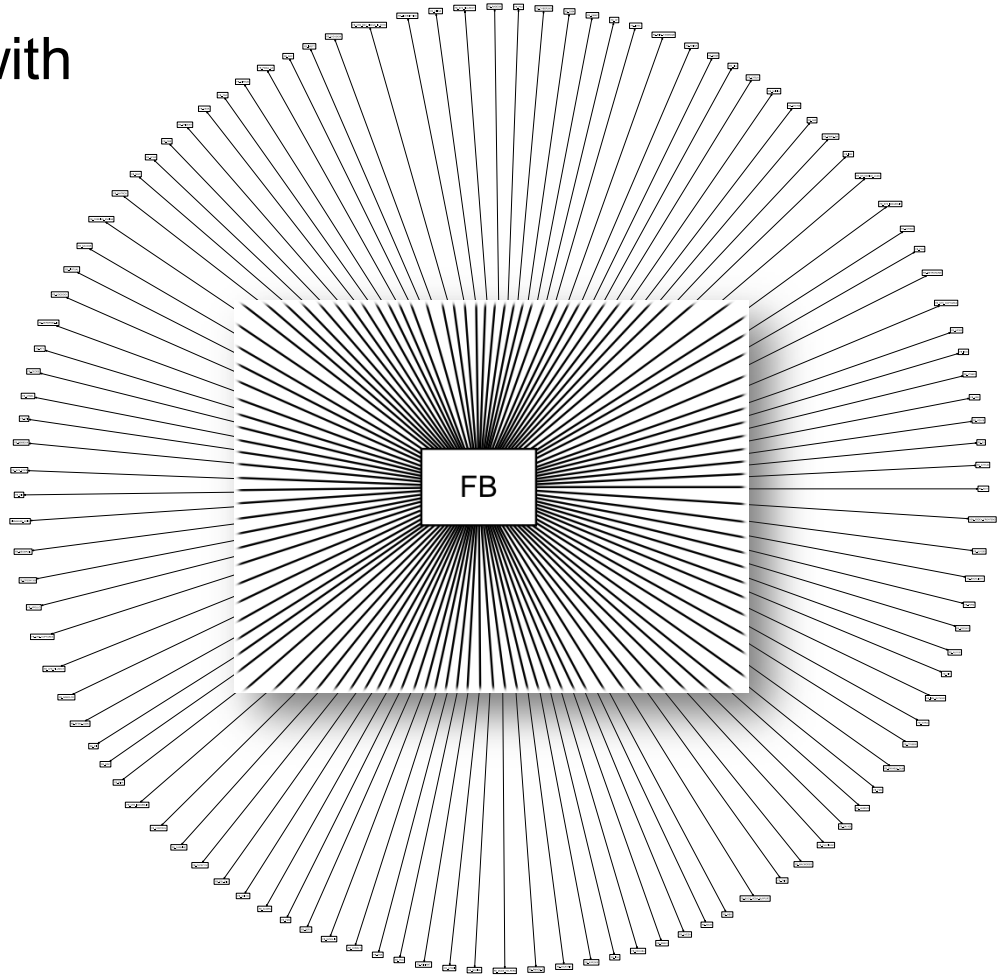
- Branching factors with high variation!



| | Avg. depth | Max. depth | Max. branching | No. of leaves |
|---|---|---|---|---|
| Linux | 4 | 8 | 158 | 5316 (84%) |
| eCos | 3 | 6 | 29 | 947 (76%) |

# Assumptions

- Branching factors with high variation!




| | Avg. depth | Max. depth | Max. branching | No. of leaves |
|---|---|---|---|---|
| Linux | 4 | 8 | 158 | 5316 (84%) |
| eCos | 3 | 6 | 29 | 947 (76%) |

# Some More Findings

- Cross-Tree Constraints ratio
  - Literature: 30-40%
  - Our models: 86%

# Some More Findings

- Cross-Tree Constraints ratio
  - Literature: 30-40%
  - Our models: 86%
- Very large constraints exist (Linux: 56 features, eCos: 21)

# Some More Findings

- Cross-Tree Constraints ratio
  - Literature: 30-40%
  - Our models: 86%
- Very large constraints exist (Linux: 56 features, eCos: 21)
- Very few group constraints (OR, XOR, MUTEX)

# Some More Findings

- Cross-Tree Constraints ratio
  - Literature: 30-40%
  - Our models: 86%
- Very large constraints exist (Linux: 56 features, eCos: 21)
- Very few group constraints (OR, XOR, MUTEX)
- Kconfig and CDL configurators
  - Configuration process is *Re-Configuration*!
  - Limited or no reasoning support
    - Kconfig relies on an imperative construct for choice propagation
    - CDL has an inference engine, which is correct but incomplete

# Some More Findings

- Cross-Tree Constraints ratio
  - Literature: 30-40%
  - Our models: 86%
- Very large constraints exist (Linux: 56 features, eCos: 21)
- Very few group constraints (OR, XOR, MUTEX)
- Kconfig and CDL configurators
  - Configuration process is *Re-Configuration*!
  - Limited or no reasoning support
    - Kconfig relies on an imperative construct for choice propagation
    - CDL has an inference engine, which is correct but incomplete
- Many more details in the paper!

# CONCLUSIONS

# Conclusions

- Empirical studies are fundamentally necessary in the VM field to guide future research and to provide requirements for tool developers.

# Conclusions

- Empirical studies are fundamentally necessary in the VM field to guide future research and to provide requirements for tool developers.

- The more ways we look at how real languages are designed and how models look like, the more confidence we have that we understand Software Product Lines.

# Conclusions

- Empirical studies are fundamentally necessary in the VM field to guide future research and to provide requirements for tool developers.

- The more ways we look at how real languages are designed and how models look like, the more confidence we have that we understand Software Product Lines.

- In studying the models in-depth, our findings have confirmed – and refuted – previous knowledge about variability languages and models.

# Conclusions

- Empirical studies are fundamentally necessary in the VM field to guide future research and to provide requirements for tool developers.
- The more ways we look at how real languages are designed and how models look like, the more confidence we have that we understand Software Product Lines.
- In studying the models in-depth, our findings have confirmed – and refuted – previous knowledge about variability languages and models.
- Understanding languages and extracting these models that were evolved over 10 years kept us PhD students and the professors busy for almost half a year!


# Variability Modeling in the Real: A Perspective from the Operating Systems Domain

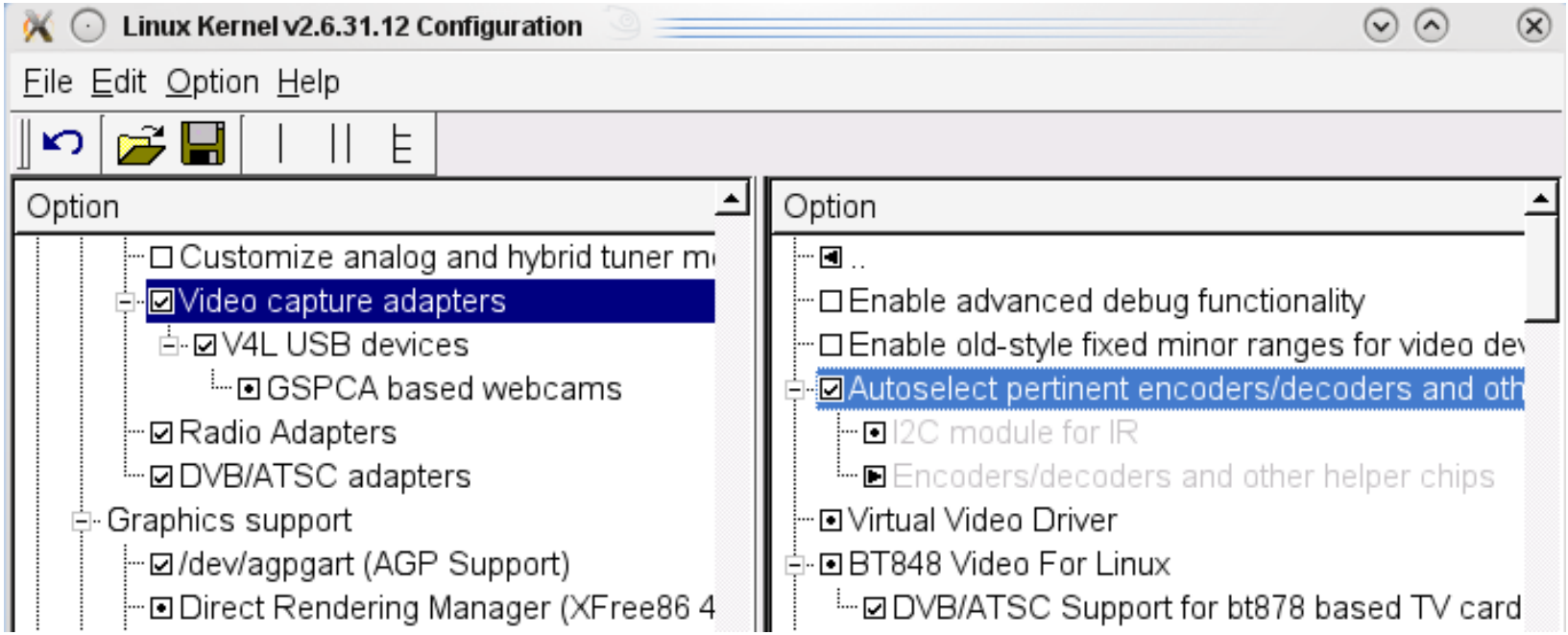
25th IEEE/ACM International Conference on Automated Software Engineering
ASE 2010
Antwerp, Belgium, Sept. 22nd, 2010

Thorsten Berger, Steven She, Rafael Lotufo,
Andrzej Wasowski, Krzysztof Czarnecki

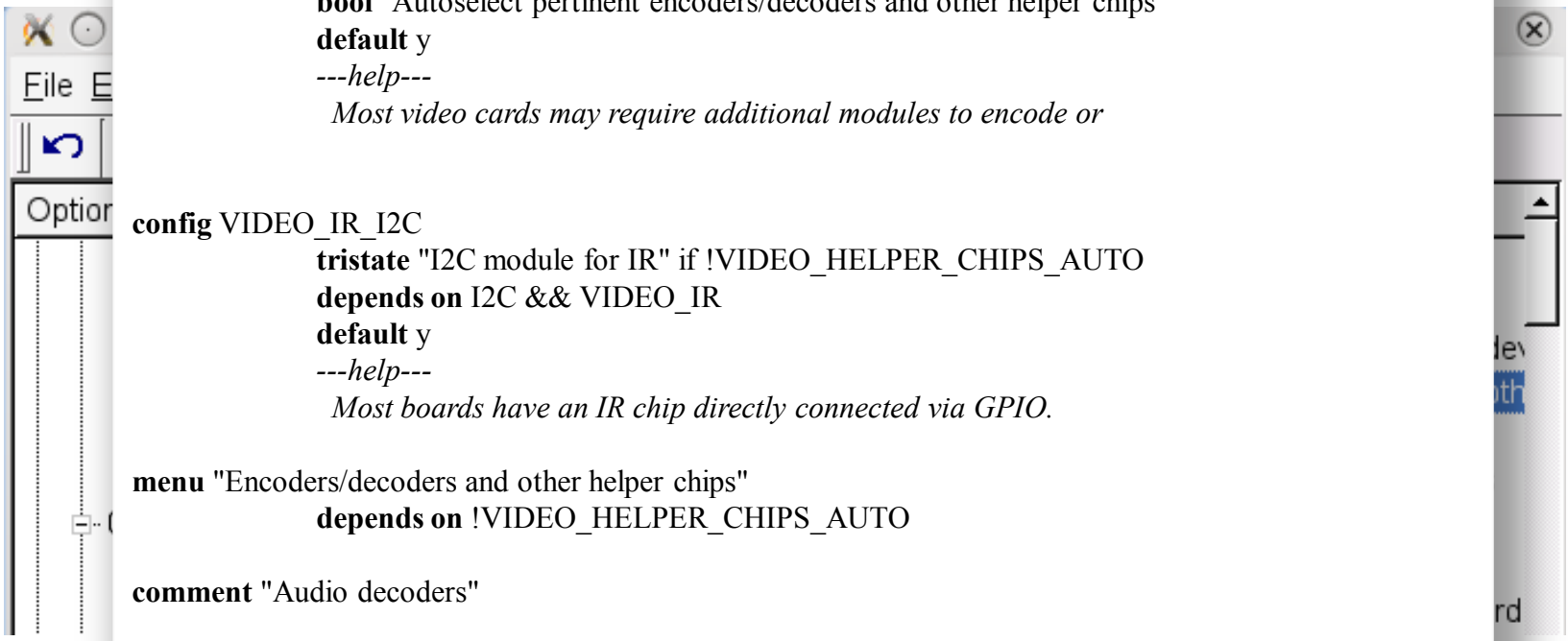
# APPENDIX

# EXAMPLES

- Children can exclude their parent



- Children can exclude their parent




```
config VIDEO_HELPER_CHIPS_AUTO
        bool "Autoselect pertinent encoders/decoders and other helper chips"
        default y
        ---help---
          Most video cards may require additional modules to encode or


config VIDEO_IR_I2C
        tristate "I2C module for IR" if !VIDEO_HELPER_CHIPS_AUTO
        depends on I2C && VIDEO_IR
        default y
        ---help---
          Most boards have an IR chip directly connected via GPIO.

menu "Encoders/decoders and other helper chips"
        depends on !VIDEO_HELPER_CHIPS_AUTO

comment "Audio decoders"

config VIDEO_TVAUDIO
        tristate "Simple audio decoder chips"
        depends on VIDEO_V4L2 && I2C
        ---help---
          Support for several audio decoder chips found on some bt8xx boards:
```

- Defaults can impose constraints in Kconfig

```
config DW_DMAC
        tristate "Synopsys DesignWare AHB DMA support"
        depends on AVR32
        select DMA_ENGINE
        default y if CPU_AT32AP7000
        ---help---
          Support the Synopsys DesignWare AHB DMA controller.  This
          can be integrated in chips such as the Atmel AT32ap7000.
```

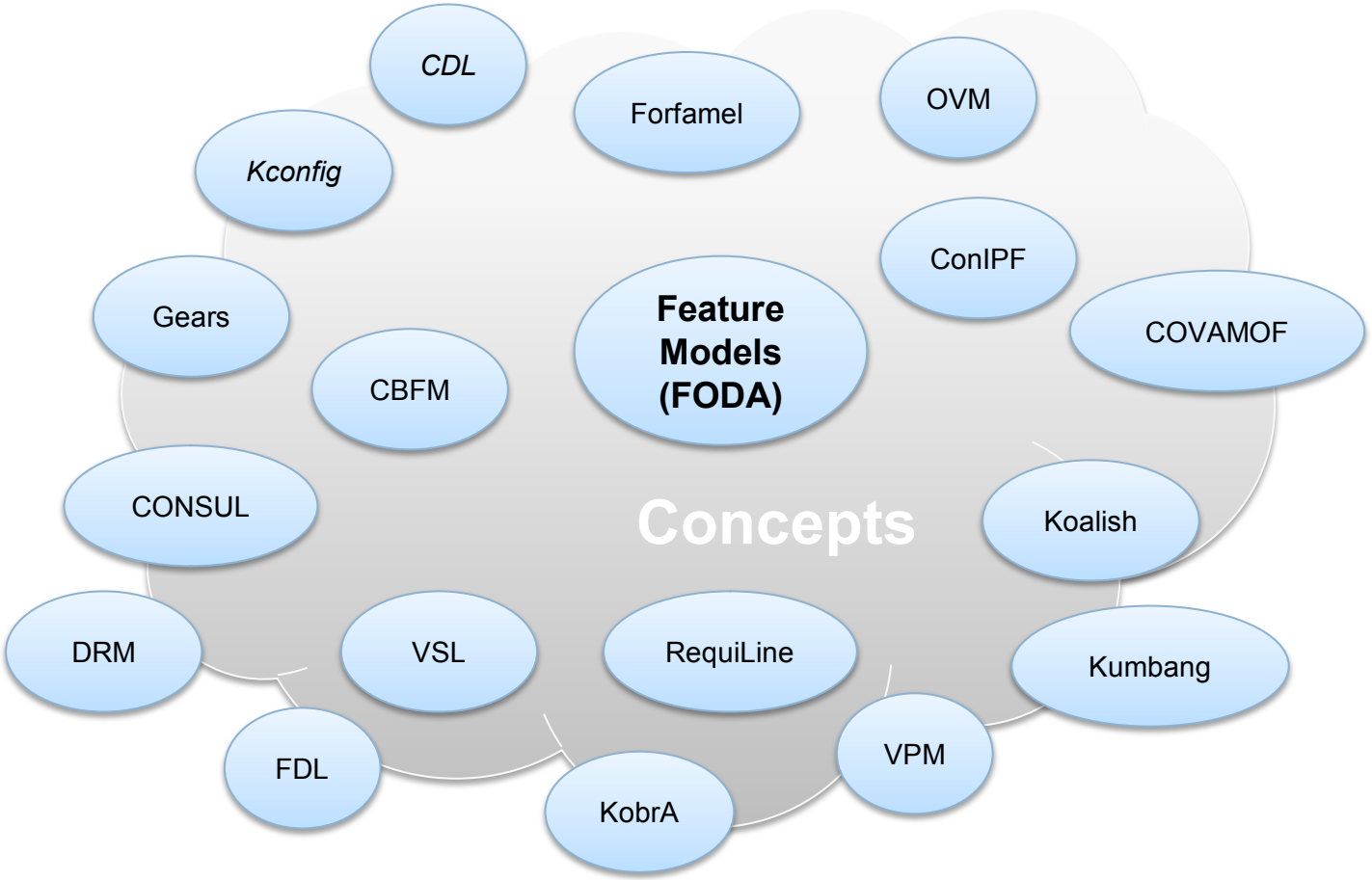
- We thought just
  - DW_DMAC ⟶ DMA_Engine Λ AVR32
- But instead also
  - !AVR32 Λ CPU_AT32AP7000 → DW_DMAC Λ DMA_ENGINE

- Computation of test cases

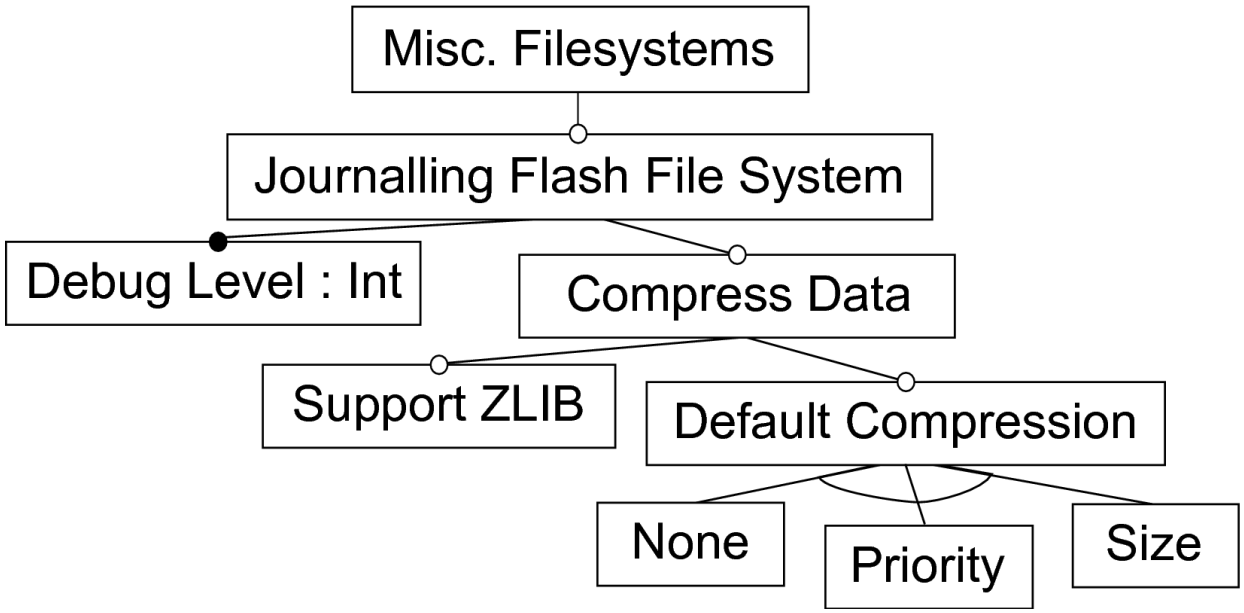
```
cdl_component CYGPKG_HAL_TESTS {
        display "Common HAL tests"
        flavor  data
        no_define
        calculated {  "tests/context tests/basic,,
                        . ((!CYGINT_HAL_TESTS_NO_CACHES) ? " tests/cache" : "")
                        . ((CYGPKG_HAL_BUILD_COMPILER_TESTS) ? " tests/cpp1 tests/vaargs" : "")
                        . ((!CYGVAR_KERNEL_COUNTERS_CLOCK) ?   " tests/intr" : "") }
        description   "
                        This option specifies the set of tests for the common HAL."
```

# VARIABILITY MODELING APPROACHES

# Variability Modeling Techniques



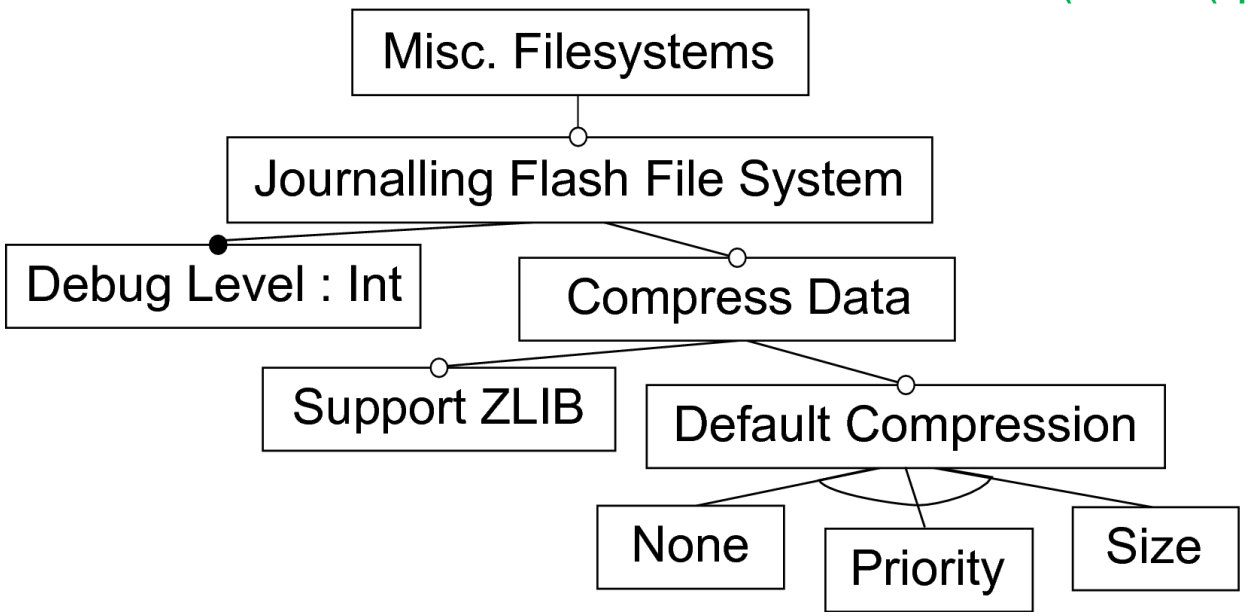
- JFFS2 filesystem




$$\text{Support ZLIB} \rightarrow \text{ZLIB Inflate}$$
$$\text{JFFS2} \rightarrow \text{CRC} \wedge \text{MTD}$$
$$0 \leq \text{Debug Level} \leq 2$$

- JFFS2 filesystem

**Features**
(Boolean (optional), String, Int)

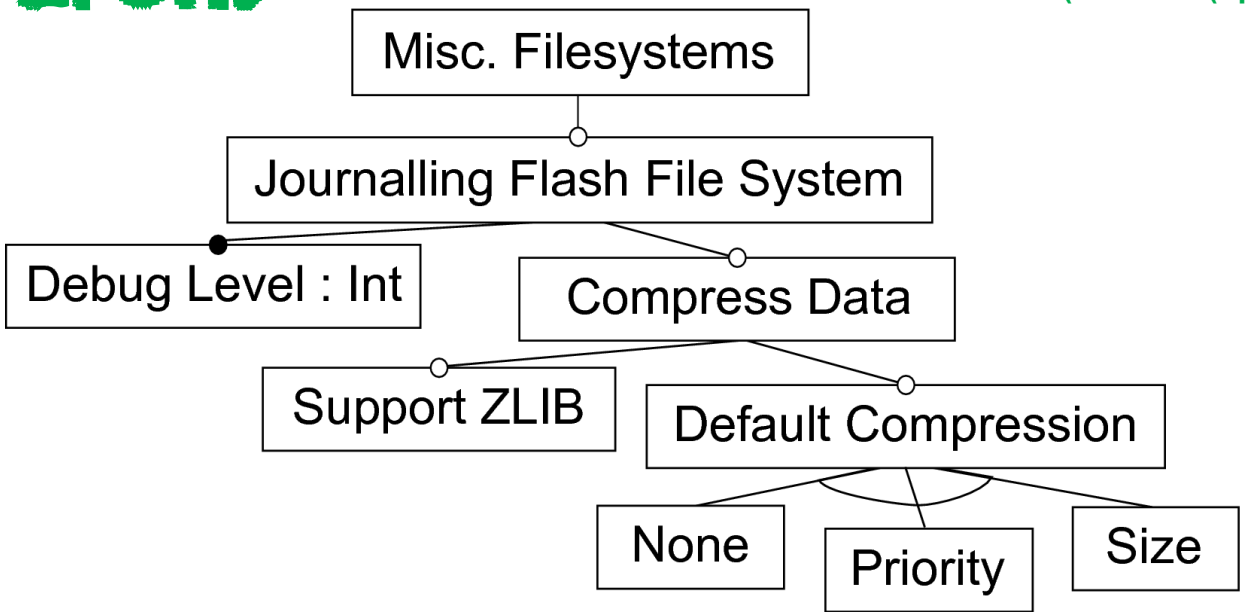



$$\text{Support ZLIB} \rightarrow \text{ZLIB Inflate}$$
$$\text{JFFS2} \rightarrow \text{CRC} \wedge \text{MTD}$$
$$0 \leq \text{Debug Level} \leq 2$$

- JFFS2 filesystem

Hierarchy

Features
(Boolean (optional), String, Int)

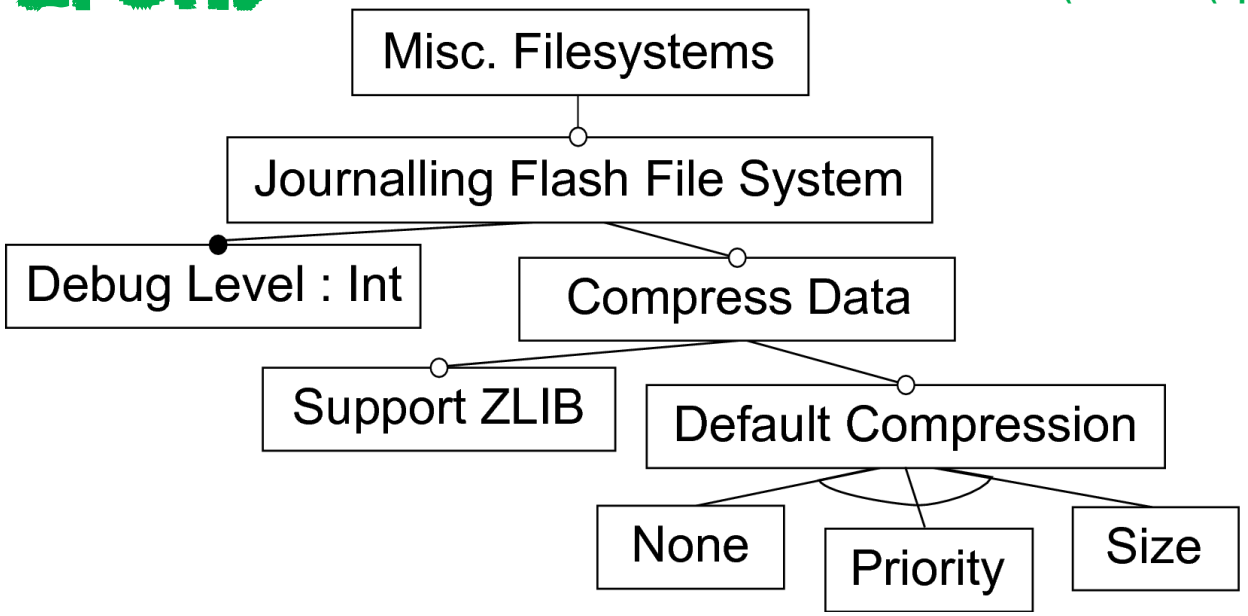



$$\text{Support ZLIB} \rightarrow \text{ZLIB Inflate}$$
$$\text{JFFS2} \rightarrow \text{CRC} \wedge \text{MTD}$$
$$0 \leq \text{Debug Level} \leq 2$$

- JFFS2 filesystem

**Hierarchy**

**Features**
(Boolean (optional), String, Int)



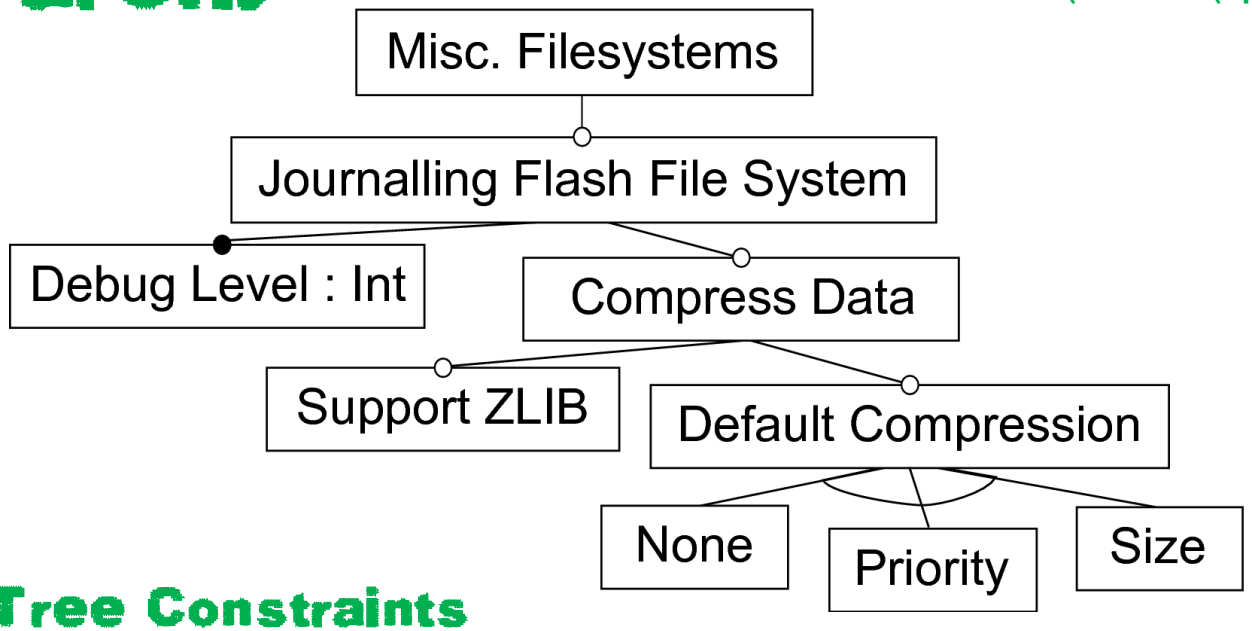

$$\text{Support ZLIB} \rightarrow \text{ZLIB Inflate}$$
$$\text{JFFS2} \rightarrow \text{CRC} \wedge \text{MTD}$$
$$0 \leq \text{Debug Level} \leq 2$$

**Group Constraints**

- JFFS2 filesystem

**Hierarchy**

**Features**
(Boolean (optional), String, Int)




**Cross-Tree Constraints**

$$\text{Support ZLIB} \rightarrow \text{ZLIB Inflate}$$
$$\text{JFFS2} \rightarrow \text{CRC} \wedge \text{MTD}$$
$$0 \leq \text{Debug Level} \leq 2$$

**Group Constraints**