Sample Spaces and Feature Models: There and Back Again

Krzysztof Czarnecki\textsuperscript{1} Steven She\textsuperscript{1} Andrzej Wąsowski\textsuperscript{2}

\textsuperscript{1}University of Waterloo, Canada

\textsuperscript{2}IT University of Copenhagen, Denmark

Software Product Line Conference 2008
Overview

Feature Models

Sample Spaces

Feature Model with Soft Constraints
Overview

Feature Models

Sample Spaces

Feature Model with Soft Constraints

destroy encourages init
start encourages stop
stop encourages start
init encourages destroy

Sample Set of Configurations
Overview

Feature Models

Sample Spaces

Feature Model with Soft Constraints

destroy encourages init
start encourages stop
stop encourages start
init encourages destroy

Czarnecki, She, Wąsowski. Sample Spaces and Feature Models: There and Back Again.
Outline

1 Motivation

2 Probabilistic Feature Models
   - Semantics of Soft Constraints
   - Joint Probability Distributions

3 Configuration

4 Application: Feature Model Mining
   - Mining on Applets

5 Conclusions
Outline

1 Motivation

2 Probabilistic Feature Models
   - Semantics of Soft Constraints
   - Joint Probability Distributions

3 Configuration

4 Application: Feature Model Mining
   - Mining on Applets

5 Conclusions
Basic Feature Models

Feature Models...

- represent commonality and variability in a product line.
- describe a set of legal configurations.
- But... existing feature models cannot express preference among legal configurations.
Basic Feature Models

- represent commonality and variability in a product line.
- describe a set of legal configurations.
- But... existing feature models cannot express preference among legal configurations.
Basic Feature Models

Feature Models...
- represent commonality and variability in a product line.
- describe a set of legal configurations.
- But... existing feature models cannot express preference among legal configurations.
Basic Feature Models

Feature Models...

- represent commonality and variability in a product line.
- describe a set of legal configurations.
- But... existing feature models can not express preference among legal configurations.
Basic Feature Models

- represent commonality and variability in a product line.
- describe a set of legal configurations.
- But... existing feature models cannot express preference among legal configurations.
Feature Models...

- represent commonality and variability in a product line.
- describe a set of legal configurations.
- But... existing feature models can not express preference among legal configurations.
Basic Feature Models

Feature Models...

- represent commonality and variability in a product line.
- describe a set of legal configurations.
- But... existing feature models cannot express preference among legal configurations.
Feature Models...

- represent commonality and variability in a product line.
- describe a set of *legal configurations*.
- **But**... existing feature models can not express *preference* among legal configurations.
Probabilistic Feature Models (PFMs)

Probabilistic Feature Models add soft constraints.

...a constraint that should be satisfied by most configurations, but some may violate it.
Interactive Configuration

- gear
  - manual
  - automatic
- drive by wire
- for North America

Drive by wire → automatic

Automatic given gear [20%]
Automatic given North America [80%]
Interactive Configuration

Probabilistic view of the PFM.

drive by wire $\rightarrow$ automatic

automatic $\textit{given}$ gear $[20\%]$  
automatic $\textit{given}$ North America $[80\%]$
We begin by selecting **Car** and **Gear**.

- Drive by wire → automatic
- Automatic given gear [20%]
- Automatic given North America [80%]

**Czarnecki, She, Wąsowski. Sample Spaces and Feature Models: There and Back Again.**
Next, we select for North America.
Interactive Configuration

Now, we select automatic.

drive by wire → automatic
automatic given gear [20%]
automatic given North America [80%]
Outline

1 Motivation

2 Probabilistic Feature Models
   - Semantics of Soft Constraints
   - Joint Probability Distributions

3 Configuration

4 Application: Feature Model Mining
   - Mining on Applets

5 Conclusions
The semantics of a basic feature model... is defined as a conjunction of it’s *hard constraints* as a *propositional formula*.
Logical Components of a Basic Feature Model

Feature Dependencies

This formula denotes a set of legal configurations.
Logical Components of a Basic Feature Model

Feature Dependencies

- **Car**
  - **gear**
  - **drive by wire**
  - **for North America**

Mandatory

- **Car**
  - **gear**

Group

- **manual**
  - **automatic**
  - **false**

This formula denotes a set of legal configurations.
Logical Components of a Basic Feature Model

Feature Dependencies

Mandatory

Group

This formula denotes a set of legal configurations.
Logical Components of a Basic Feature Model

Feature Dependencies

- Car
- gear
- drive by wire
- for North America
- manual
- automatic

Mandatory

- Car
- gear

Group

- manual
- automatic
- false

\[ \land \text{Car} \]

Root

\[ \land \phi \]

Additional Constraints

This formula denotes a set of legal configurations.
Logical Components of a Basic Feature Model

Feature Dependencies

Mandatory

Group

This formula denotes a set of legal configurations.
A probabilistic feature model is...

A basic feature model + soft constraints

Components of a Car

- Gear
- Drive by wire
- For North America
- Manual
- Automatic

drive by wire → automatic
Components of a Probabilistic Feature Model

A probabilistic feature model is...

A basic feature model + soft constraints

---

- Car
- gear
- drive by wire
- for North America
- manual
- automatic

drive by wire → automatic +
Feature Modeling with Soft Constraints

- **On-by-default** if cond. probability between 80% and 100%.
Feature Modeling with Soft Constraints

Off-by-default if cond. probability between 0 and 50%.
Feature Modeling with Soft Constraints

- **Applet**
  - **must override**
    - **paint**
    - **start**
    - **init**
  - **stop**
  - **destroy**

**Constraints**:
- **destroy encourages init**
- **start encourages stop**
- **stop encourages start**
- **init encourages destroy**
Feature Modeling with Soft Constraints

Czarnecki, She, Wąsowski. *Sample Spaces and Feature Models: There and Back Again.*
Basic feature models... specify a set of legal configurations.

Probabilistic feature models... specify a set of legal joint probability distributions (JPDs).

A joint probability distribution... assigns a probability to each possible configuration.
Basic feature models...

specify a set of *legal configurations*.

Probabilistic feature models...

specify a set of *legal joint probability distributions* (JPDs).

A joint probability distribution...

assigns a probability to each *possible configuration*. 
Legal Configurations Compared with JPDs

Basic Feature Model

Czarnecki, She, Wąsowski. Sample Spaces and Feature Models: There and Back Again.
Legal Configurations Compared with JPDs

Basic Feature Model

Probabilistic Feature Model

Czarnecki, She, Wąsowski. Sample Spaces and Feature Models: There and Back Again.
Basic feature models...

specify a set of legal configurations.

Probabilistic feature models...

specify a set of legal joint probability distributions (JPDs).

A joint probability distribution...

assigns a probability to each possible configuration.
Under-specification in PFMs

An abstract PFM is *under-specified* and specifies a range of JPDs.

![Feature Model Diagram]

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>P(a, b, c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td>$p_1 \geq 0.0$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
<td>$p_2 \geq 0.0$</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
<td>$p_3 \geq 0.0$</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
<td>$p_4 \geq 0.0$</td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td>$p_{5...8} = 0$</td>
</tr>
</tbody>
</table>
Under-specification in PFM

An abstract PFM is under-specified and specifies a range of JPDs.

\[ P(a, b, c) = \begin{cases} 0.1 & p_1 > 0.0 \\ 0.25p_1 & p_2 = 0.25p_1 \\ 0.1 & p_3 > 0.0 \\ 1 - 1.25p_1 - p_3 & p_4 = 1 - 1.25p_1 - p_3 \\ 0 & p_{5\ldots8} = 0 \end{cases} \]

where \( 1.25p_1 + p_3 \leq 1 \)

\[ c \text{ given } b \text{ [80%]} \]
### Under-specification in PFM

A concrete PFM specifies a single JPD.

- **Motivation**
  - Probabilistic Feature Models
- **Configuration**
- **Feature Model Mining**
- **Conclusions**

---

A probabilistic feature model (PFM) is under-specified and specifies a range of joint probability distributions (JPDs).

An abstract PFM is under-specified and specifies a range of JPDs.

A concrete PFM specifies a single JPD.

<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>c</th>
<th>$P(a, b, c)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$p_1 = 0.1$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$p_2 = 0.025$</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>$p_3 = 0.1$</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>$p_4 = 0.775$</td>
</tr>
<tr>
<td>0</td>
<td>:</td>
<td>:</td>
<td>$p_{5\ldots8} = 0$</td>
</tr>
</tbody>
</table>

- $c$ given $b$ [80%]
- $b$ [12.5%]
Outline

1 Motivation

2 Probabilistic Feature Models
   - Semantics of Soft Constraints
   - Joint Probability Distributions

3 Configuration

4 Application: Feature Model Mining
   - Mining on Applets

5 Conclusions
Configuration

Requires a single concrete JPD.

- Abstract PFM need to be completed.
- *Entropy maximization.*

Probabilistic Inference.

- Relation with *Bayesian Networks*.
- *Most probable explanation* algorithms.
- Adaptive guidance given current state.
Configuration

Requires a single concrete JPD.

- Abstract PFMs need to be completed.
- Entropy maximization.

Probabilistic Inference.

- Relation with Bayesian Networks.
- Most probable explanation algorithms.
- Adaptive guidance given current state.
Outline

1 Motivation

2 Probabilistic Feature Models
   ■ Semantics of Soft Constraints
   ■ Joint Probability Distributions

3 Configuration

4 Application: Feature Model Mining
   ■ Mining on Applets

5 Conclusions
Feature Model Mining

Synthesize a feature model that is representative of the variability in a sample set of configurations.

<table>
<thead>
<tr>
<th>conf</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

Sample Set
Feature Model Mining

Synthesize a feature model that is representative of the variability in a sample set of configurations.

<table>
<thead>
<tr>
<th>conf</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Association Rules

- $c \Rightarrow a$ [100%]
- $\land b \Rightarrow a$ [100%]
- $\land d \Rightarrow c$ [100%]
- $\land a \Rightarrow b \lor c$ [100%]
- $\land \ldots$
- $\land a \Rightarrow c$ [75%]
Feature Model Mining

Synthesize a feature model that is representative of the variability in a sample set of configurations.

Sample Set

<table>
<thead>
<tr>
<th>conf</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

Association Rules

- \( c \Rightarrow a \ [100\%] \)
- \( \land b \Rightarrow a \ [100\%] \)
- \( \land d \Rightarrow c \ [100\%] \)
- \( \land a \Rightarrow b \lor c \ [100\%] \)
- \( \land \ldots \)
- \( \land a \Rightarrow c \ [75\%] \)

Feature Model

Czarnecki and Wąsowski 2007

Association Rule Mining

Feature Model Synthesis

Czarnecki, She, Wąsowski. Sample Spaces and Feature Models: There and Back Again.
Feature Model Mining on Applets

Applets

Sample Set

Construct sample set by analysing overridden methods in 64 applets:

destroy, paint, init, start and stop.
Case Study Results

Mined Feature Model

- Applet
  - paint [75%]
  - start [59%]
  - init [97%]
  - stop [53%]
  - destroy [42%]

- stop given start [84%]
- start given stop [97%]
- paint given destroy [88%]
- paint given stop [88%]
  - more...

Expert-specified Model

- Applet
  - must override
  - stop[off]
  - destroy[off]
  - paint[on]
  - start[off]
  - init[on]

- destroy encourages init
- start encourages stop
- stop encourages start
- init encourages destroy

Czarnecki, She, Wąsowski. Sample Spaces and Feature Models: There and Back Again.
Outline

1 Motivation

2 Probabilistic Feature Models
   - Semantics of Soft Constraints
   - Joint Probability Distributions

3 Configuration

4 Application: Feature Model Mining
   - Mining on Applets

5 Conclusions
Related Work

Probabilistic Feature Models.

- Soft Constraints  [Czarnecki 2000] [Wada, Suzuki and Oba 2007]
- Feature Models and fuzzy logic  [Robak, Pieczyński, 2003]
- i* goal models  [Giorgini et al., 2002]

Reverse-engineering models.

- Using concept analysis  [Loesch and Ploedereder, 2007]
- Identifying code differences  [Jepsen et. al., 2007]
Conclusions

Probabilistic Feature Models.

- Basic feature models extended with *soft constraints*.
- Specifies a set of joint probability distributions.
- Modeling, reverse-engineering, configuration.
Conclusions

Probabilistic Feature Models.

- Basic feature models extended with soft constraints.
- Specifies a set of joint probability distributions.
- Modeling, reverse-engineering, configuration.

Questions?