

Case Study: Supplementing Program Analysis with Natural Language Analysis to Improve a Reverse Engineering Task

David Shepherd,

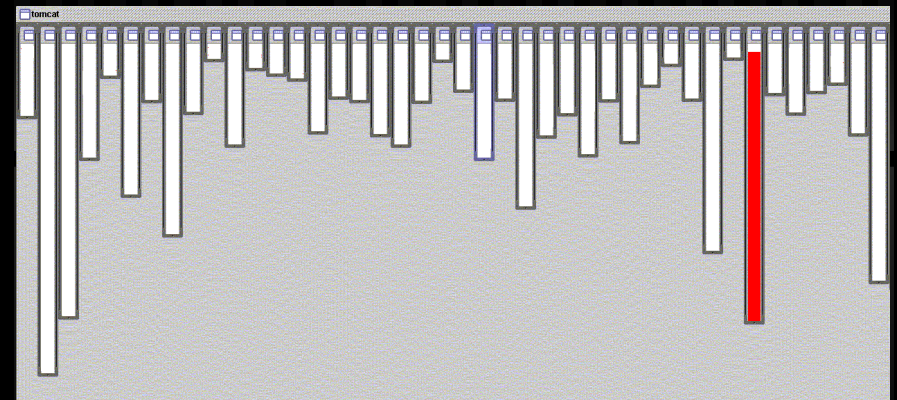
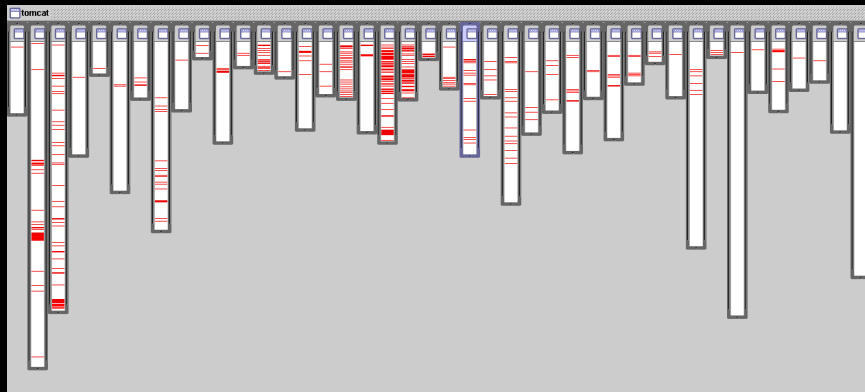


Lori Pollock, and K. Vijay-Shanker

University of Delaware

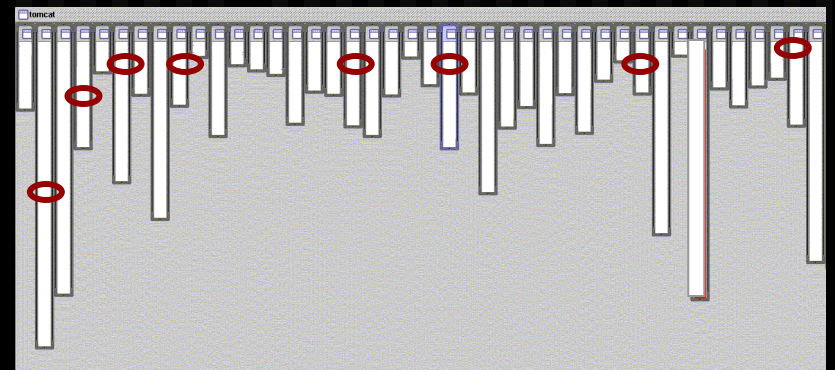
Consider the Reverse Eng Task...

Aspect-Oriented Programming, language support for:



Aspect Mining

Locates refactoring candidates



Classic Aspect Candidates

✓ Before/After Aspects [Mastering AspectJ, Aspectj.com]

```
Circle.setRadius(x){  
  r = x;  
  Display.update()  
}
```

```
Square.setSize(x){  
  size = x;  
  Display.update()  
}
```

Aspect Updater
After Shape.set*()
Call Display.update()

✓ Certain Design Patterns [Hanneman 02]

```
public MultiFigure(){  
  numFigs = n;....  
  addListener(Drawing)  
}
```

```
public Square(y){  
  side = y;  
  addListener(MultiFigure)  
}
```

Aspect Listener
After Figure.construct()
Add parent as listener

Current Aspect Mining Analyses

Traditional Program Analyses (TPA)

Fan-In Analysis

[Marin et al. WCRE 2005]



Method
Display.update()

Call Graph Fragment

Code Clone Analysis

[Shepherd SERP 04, Bruntink ICSM 04]

```
private int do_Snipe() {  
    Auctions.startBlocking();  
    AuctionServer as = m_auction.getServer();  
    m_auction.setLastStatus("Firing actual snipe.");  
    if(m_auction.isMultiSniped()) {
```

```
Auctions.startBlocking();  
AuctionServer as = m_auction.getServer();  
m_auction.setLastStatus("Preparing snipe.");
```

```
    }  
    MQF  
    // Log in  
    m_cj = as.getSignInCookie(null);  
    if (m_cj == null) {  
        // Alert somebody that we couldn't log in?  
        m_auction.setLastStatus("Pre-snipe login failed. Snipe-");
```

Traditional Program Analyses (TPA) can be effective

Consider Natural Language Analyses (NLA)

Scenario

✓ Class: DrawApplication

✓ Methods

✓ addViewChangeListener

✓ removeViewChangeListener

Opposite Verbs

Natural Language Analyses can be effective

Quiz: Is TPA Enough?

Which method(s) are good candidates?



Call Graph Fragments

Traditional Program Analyses (TPA) can be misleading

Quiz: Is NLA Enough?

Scenario A

- ✓ Method uses verb “add”
- ✓ Another method in the same class uses verb “remove”

Scenario B

- ✓ Method uses verb “add”
- ✓ Another method in the same class uses verb “remove”

Natural Language Analyses can be misleading

Summary: TPA & NLA Complementary

Using NLA and TPA together to **build a case** for refactoring



Fan-in is high

Opposite verb is used

```
Auctions.startBlocking()  
AuctionServer as = m_auc  
m_auction.setLastStatus()
```

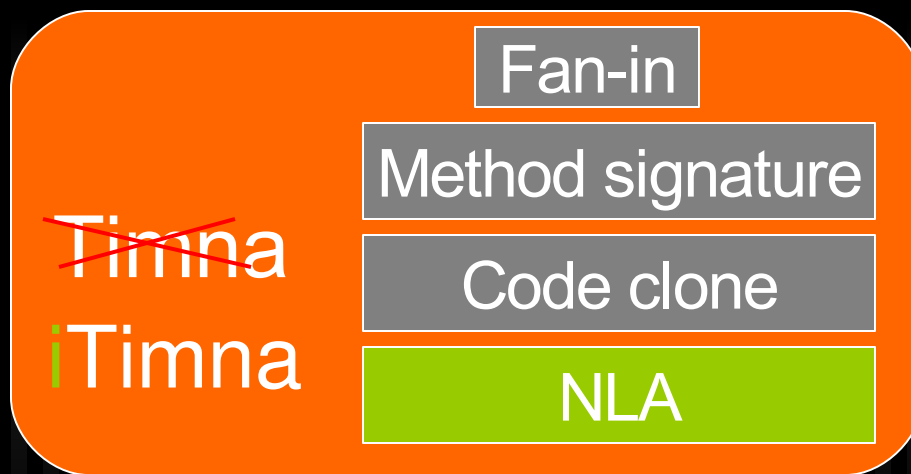
```
Auctions.startBlocking()  
AuctionServer as = m_auc  
m_auction.setLastStatus()
```

Called in a clone

Supplementing TPA with NLA makes a stronger case for refactoring

Talk Progression

- ✓ Aspect Mining Framework: **Timna**



- ✓ Integration of Natural Language in **iTimna**
- ✓ Evaluation

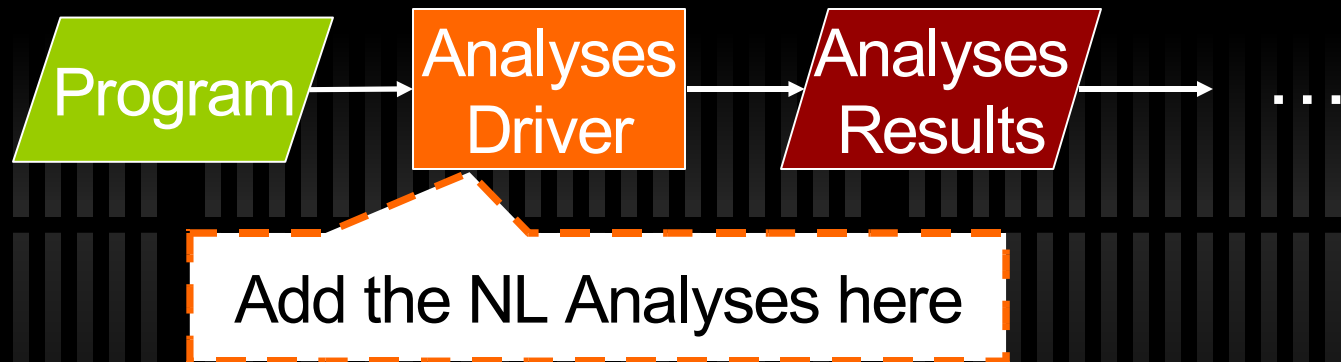
Timna: An Aspect Mining Framework

Timna uses

```
Method name | Fan-in | Code Clone | Is Void  
op If( Fan-in >3 and Is-Void)  
ad then the method is a refactoring candidate
```



iTimna: Adding NLA to Timna



TPAnalyses

- ✓ Code clone
- ✓ Signature analysis
- ✓ Callsite pairings

NLAnalyses

- ✓ Opposite verbs
- ✓ Observer words
- ✓ Constraint verbs

An Example NLA

- ✓ NLA: **Past-Tense Verb**
- ✓ Intuition: Past-Tense indicates **reactionary**
- ✓ Process:
 1. Extract verb [AOSD 06]
 2. Use morphological analysis to detect verb tense
- ✓ Examples:
 - ✓ `figureSelectionChanged()` \Rightarrow *changed* \Rightarrow past-tense
 - ✓ `setSelection` \Rightarrow *set* \Rightarrow not past-tense
 - ✓ `drawingInvalidated()` \Rightarrow *invalidated* \Rightarrow past-tense

Evaluation

✓ Question:

What effect does integrating NLA with TPA have on AM?

✓ Methodology:

Manually tag a program, Use AM tools, Check results

✓ Measures: Precision (quality) and Recall (completeness)

✓ Subject Program: JHotDraw (2,739 methods)

✓ AM Tools: Fan-in, Timna, iTimna

~~Tagger bias?~~

~~Overfitting?~~

~~Enough examples?~~

Quantitative Results



✓ Fan-in

Precision **37** Recall **2**

✓ Timna

Precision **62** Recall **60**

✓ iTimna

Precision **81** Recall **73**

Qualitative Results

NLA and Simple Program Information

```
private void endEdit() {  
    if (fURLTarget != null) {  
        setURL(fURLTarget, fTextField.getText());  
        fURLTarget = null;  
        fTextField.endOverlay();  
    }  
}
```

If (Void Return Type & Opposite Verb & No Parameter)

Then the method is a candidate

Qualitative Results

NLA and Simple Program Information

```
public void mouseReleased(MouseEvent e) {  
    if (isEnabled()) {  
        iState = iOldState;  
        repaint();  
        if (contains(e.getX(), e.getY())) {  
            fListener.paletteUserSelected(this);  
        }  
    }  
}
```

If (Not Void Return Type && Constraint Verb)

Then the method is a candidate

Related Work

- ✓ Mining Dynamic Information

[Breu ASE 04, Ceccato WCRE 04]

Single Analyses

- ✓ Mining Static Information

[Marin WCRE 04, Shepherd SERP 04, Bruntink ICSM 04]

- ✓ Combining/Comparing Approaches

Simple Combinations

- ✓ Sorts [Marin WCRE 2006]

Recommend

- ✓ Qualitative Comparison [Tourwe IWPC 05]

Combining

Conclusions



- ✓ Created **NLAs** for Aspect Mining
- ✓ Showed that TPA and NLA can **work together**
- ✓ Evaluated prototype, shows strong **promise**

Supplementing TPA with **NLA** improves the **effectiveness** of our aspect mining framework and warrants further evaluation