Combining Conceptual and Domain-Based Couplings to Detect Database and Code Dependencies

Malcom Gethers, Amir Aryani, Denys Poshyvanyk

UMBC

AN HONORS UNIVERSITY IN MARYLAND

RMIT University, Australia

WILLIAM & MARY

2012 IEEE 12th International Working Conference on Source Code Analysis and Manipulation

Common problems in software maintenance

Systems with legacy code, e.g., COBOL

Hybrid systems, e.g., Python and Java

Multi-tier systems

Inaccessible maintenance history

Despite the issues...

Perform impact analysis

Information suitable for domain experts

User Interface Components (UIC)

How?

Domain-based coupling

Conceptual coupling

Combination

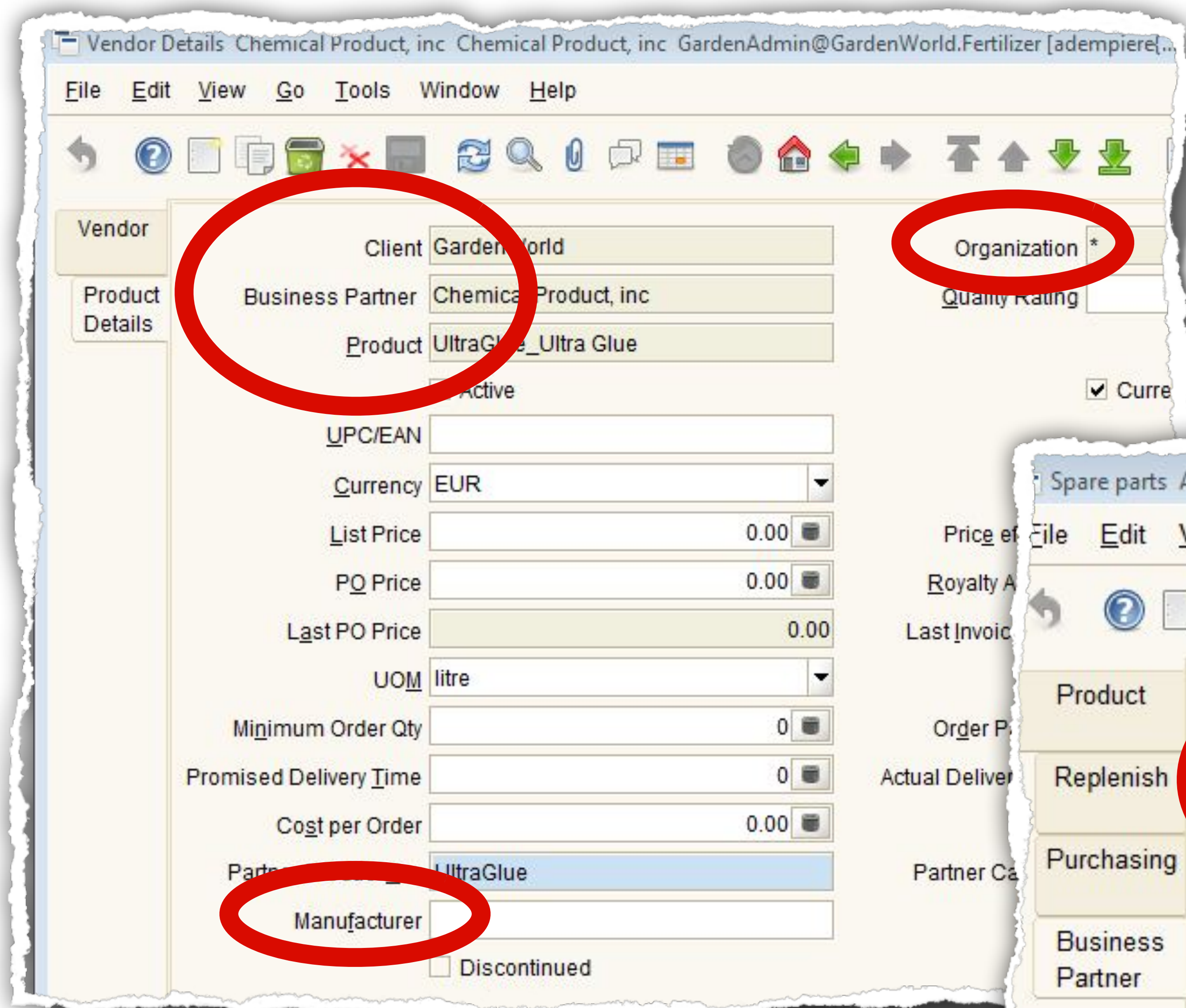
Motivations

Domain-based approach works without access to source code or design documents

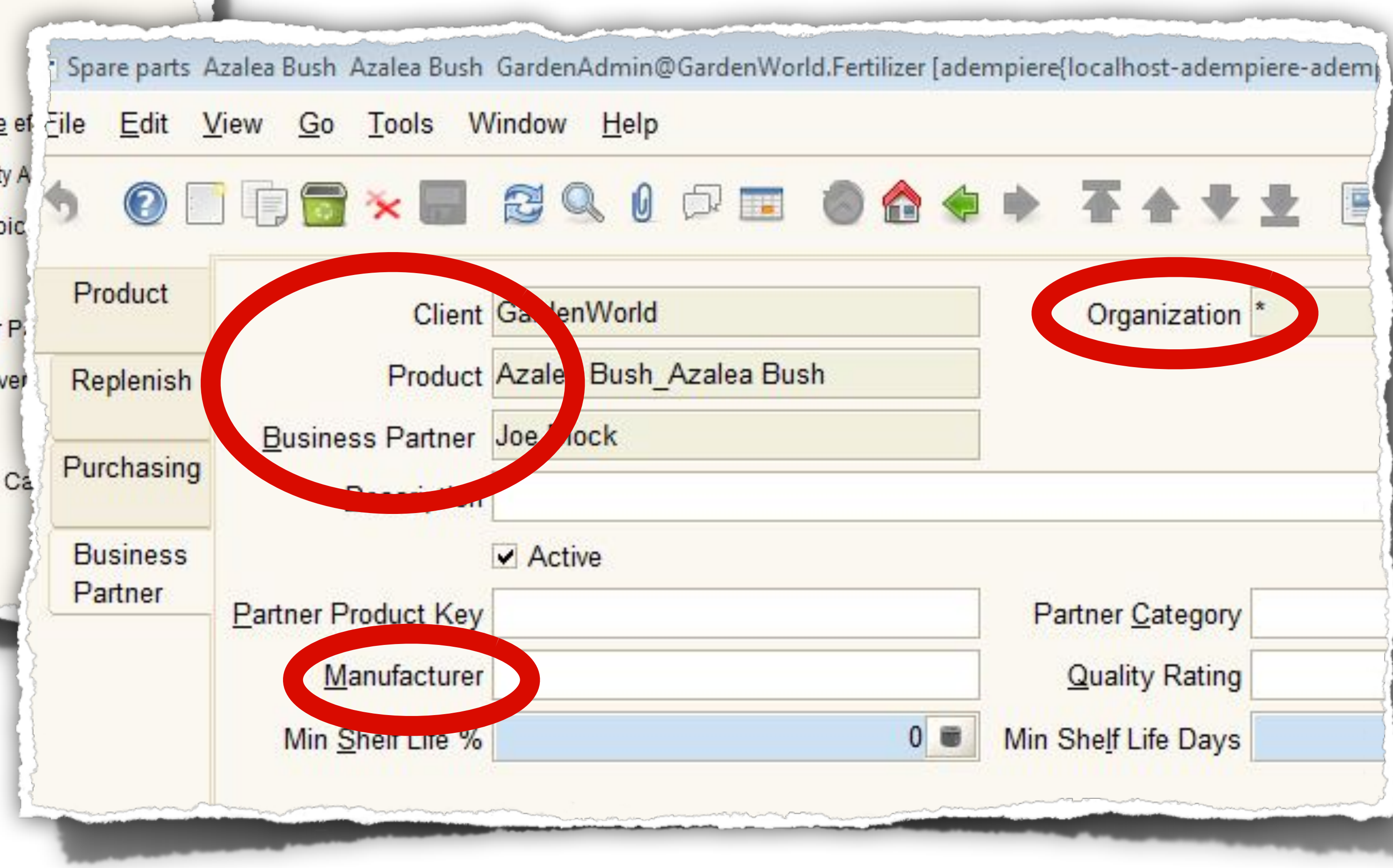
Conceptual coupling approach is language independent

The approaches complement each other

Example of UICs



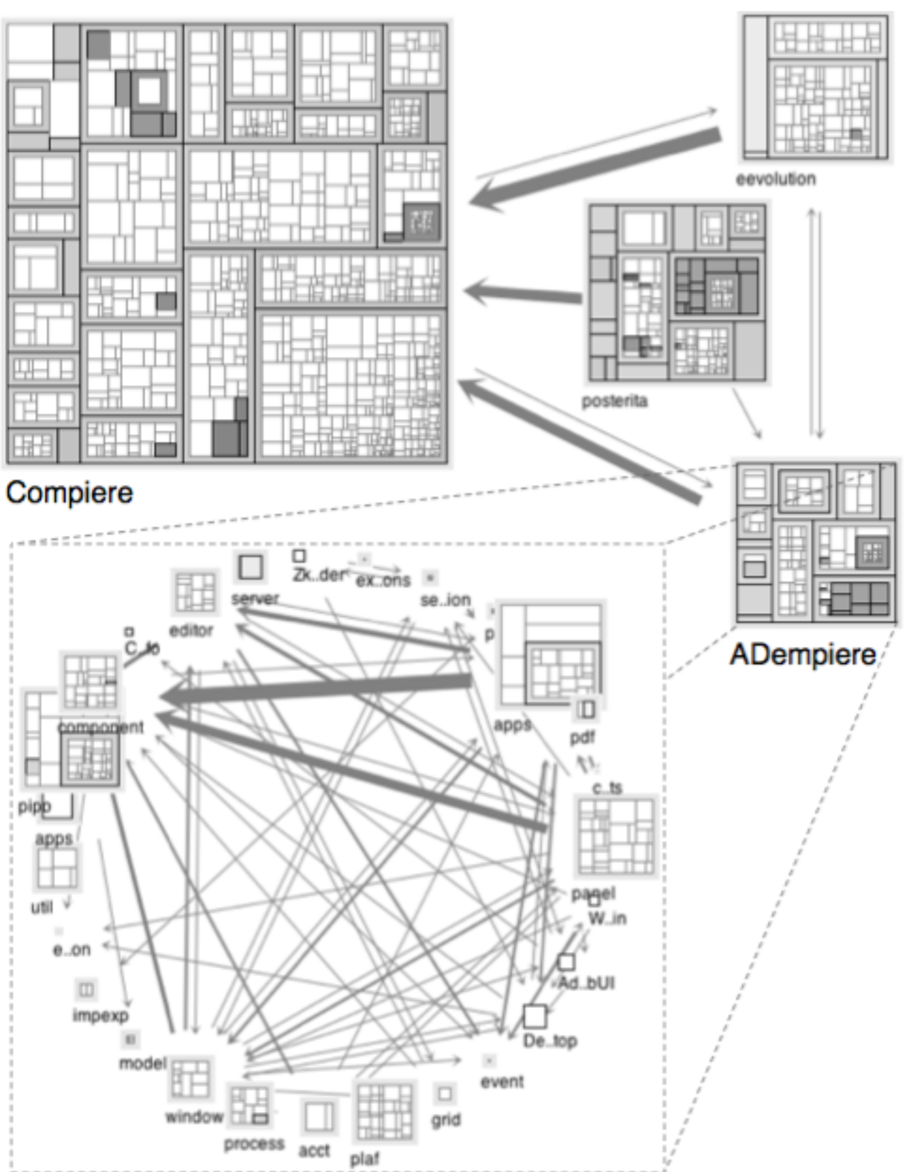
Domain variables



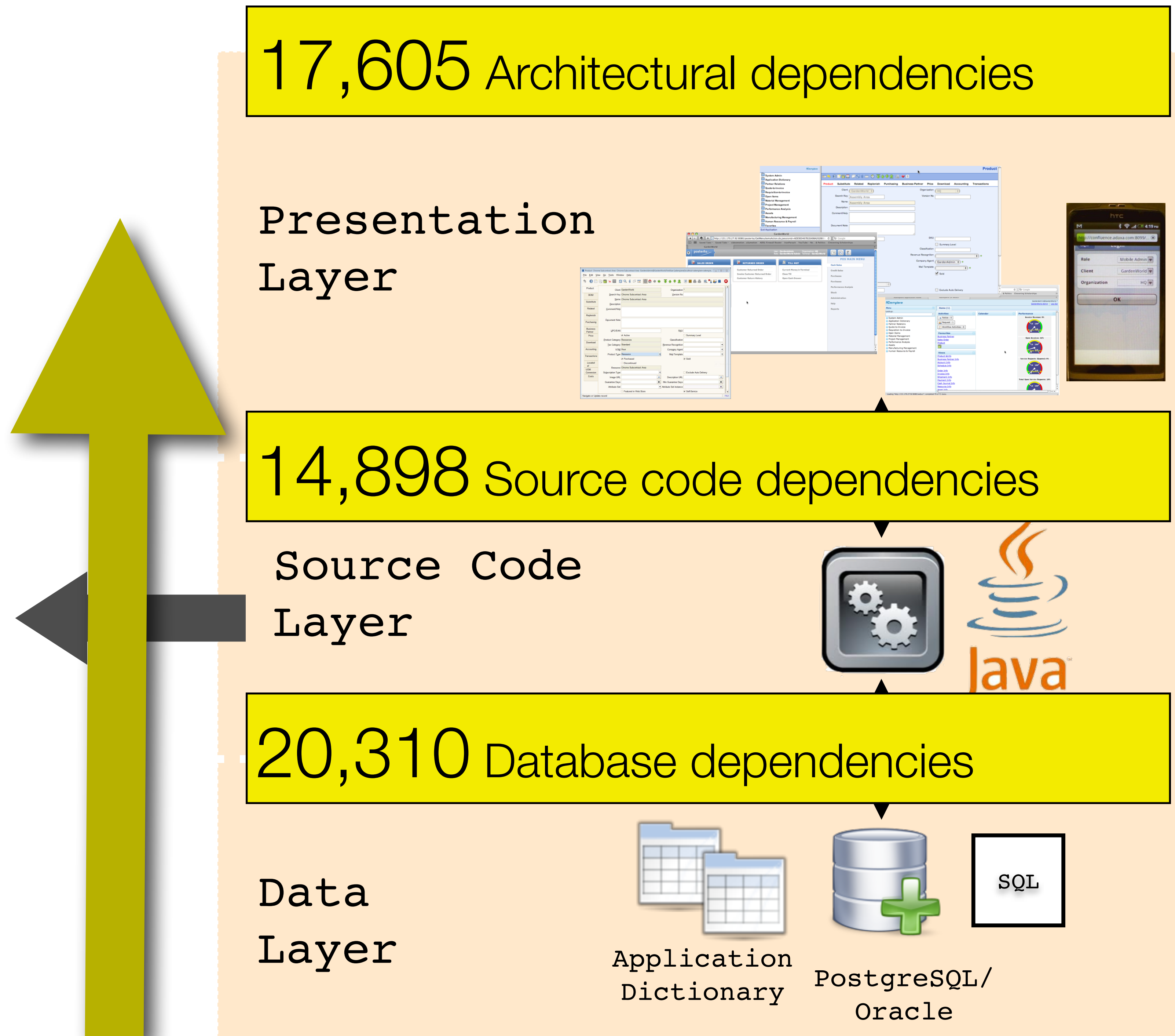
Case Study

- 120,111 times downloaded in 2011
- 3,531 Java Classes
- 2,569,854 lines of code
- Four distinct interfaces
- 347 screens

Dependencies



M. Lungu and M. Lanza, Softwarenaut, CSMR 2006



Case Study - Orthogonality

Do conceptual and domain-based coupling identify orthogonal dependencies?

Case Study - Orthogonality

	Architectural Dependencies (UICs)		
	CP 10	**CP 20**	**CP 30**
C (int) D	26%	26%	26%
C (diff) D	35%	38%	39%
D (diff) C	25%	25%	25%

C (int) D: Set intersection of correct dependencies identified by both conceptual and domain-based coupling

C (diff) D: Set difference of correct dependencies identified by conceptual and domain-based coupling

D (diff) C: Set difference of correct dependencies identified by conceptual and domain-based coupling

Case Study - Orthogonality

	Architectural Dependencies (UICs)		
	[illegible]	CP 20	CP 30
C (int) D	[illegible]	26%	26%
C (diff) D	[illegible]	[illegible]	39%
D (diff) C	[illegible]	[illegible]	[illegible]

Metrics are orthogonal!

C (int) D: Set intersection of correct dependencies identified by both conceptual and domain-based coupling

C (diff) D: Set difference of correct dependencies identified by conceptual and domain-based coupling

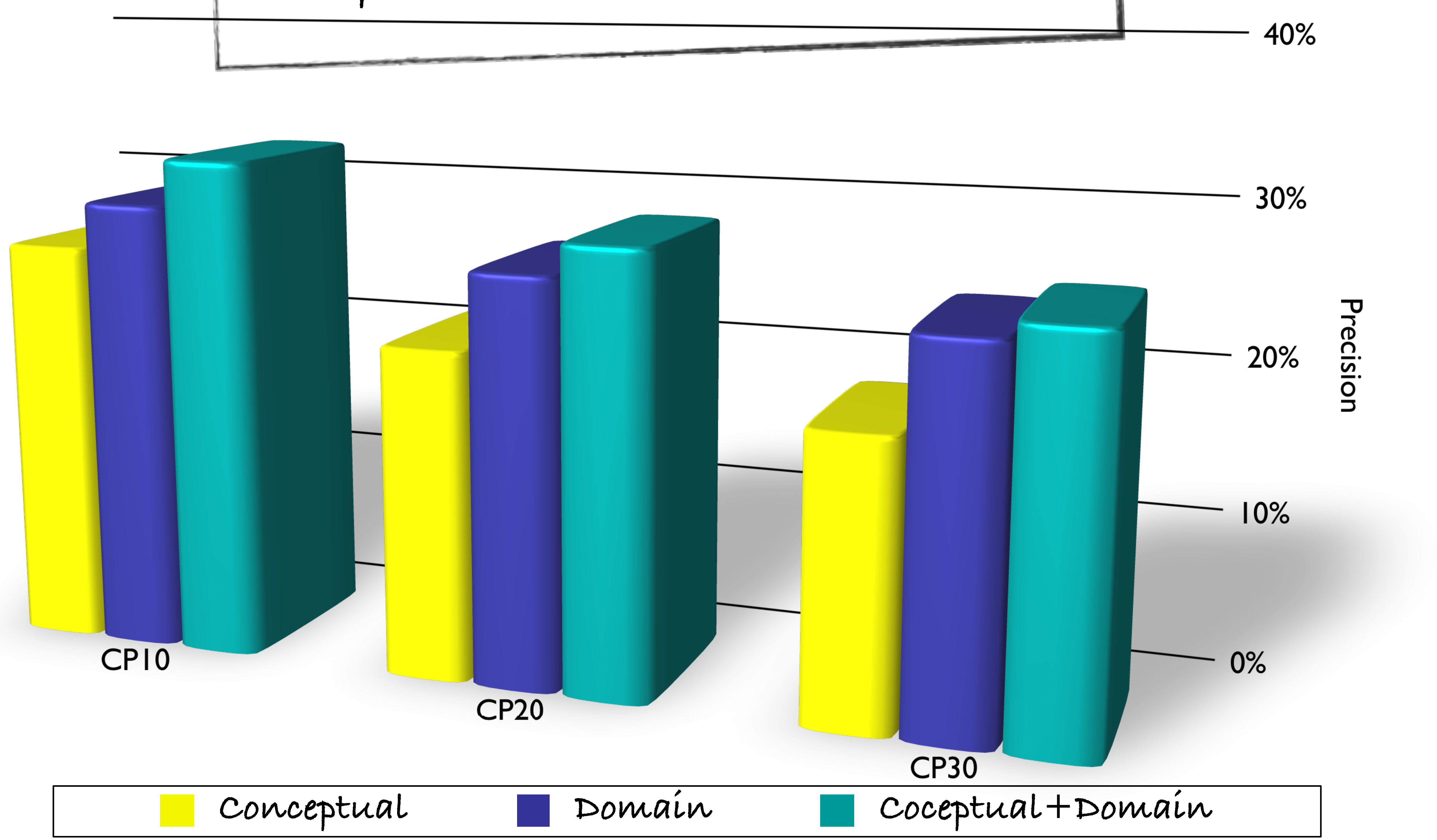
D (diff) C: Set difference of correct dependencies identified by conceptual and domain-based coupling

Case Study - Accuracy

Does combing conceptual and domain-based coupling improve the accuracy our ability to identify dependencies?

Case Study - Accuracy

Is it possible to improve the accuracy?



Case Study - Accuracy

The combination of conceptual and domain dependencies yields an improvement for identifying dependencies

Wilcoxon sign-ranked test indicates our findings are typically statistically significant

Case Study - Accuracy

...tion of conceptual and domain ... improvement for iden...

Wilcoxon sign-ran... ...gs are typically statistically sig...

The combination outperforms either individual technique.

Conclusion

Conceptual and domain-based coupling identify orthogonal sets of dependencies

Combining the metrics improves our ability to predict dependencies

Recall improvements of up to 7% over the baseline approach

Precision improvement up to 24% over the baseline approach

Thank You

SEMERU @ William and Mary

http://www.cs.wm.edu/semeru