

レビュー効率化にむけた産学連携の取組み

国立大学法人 奈良先端科学技術大学院大学 情報科学研究科

森崎 修司

smrs@is.aist-nara.ac.jp

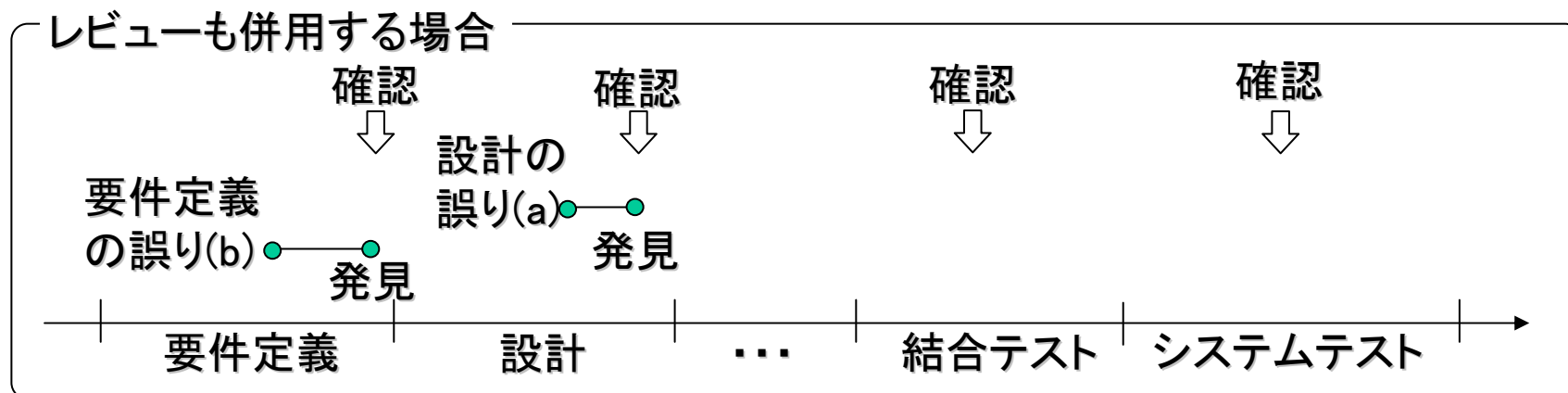
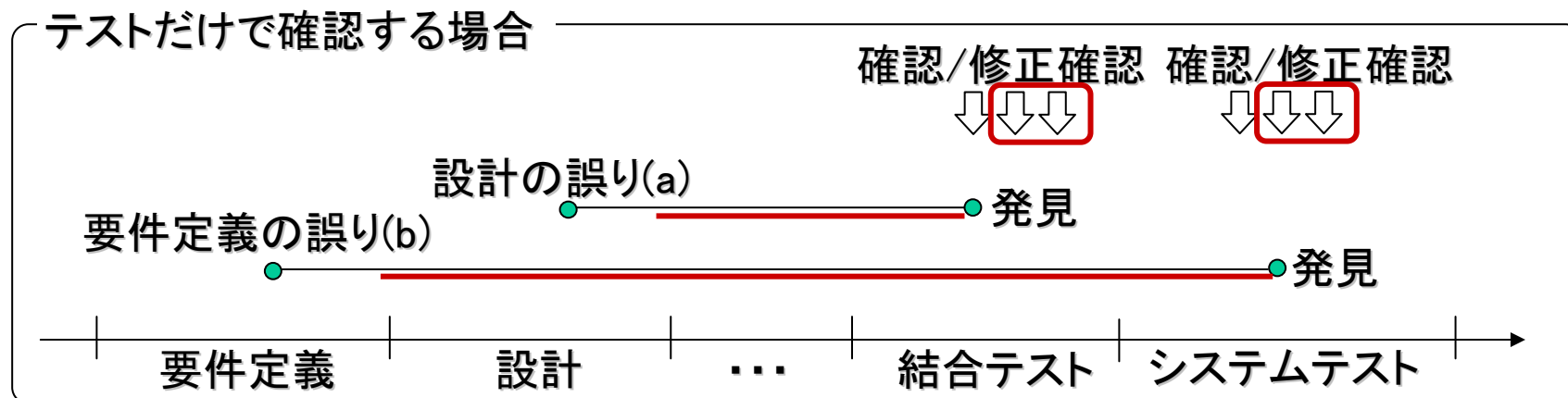
本スライドの作成にあたり、東芝デジタルメディアエンジニアリング 久保 匡氏、荻野 利彦氏、阪本 太志氏、東芝ソフトウェア技術センター山田 淳氏にご尽力いただきました。本資料に含まれる研究の一部は文部科学省「eSociety基盤ソフトウェアの総合開発」の委託に基づいて行われたものです。また、本資料に含まれる研究の一部は同省科学研究補助費(若手B:課題番号21700033)の助成を受けました。

アジェンダ

- レビューの課題
- 課題解決へ向けた取組み
 - 国際動向
 - 当学での取組み
- リスクベースドレビュー
 - 観点設定方法
 - 観点設定の効果
- 東芝デジタルメディアエンジニアリングとの共同研究事例
 - 観点設定
 - 実施結果

ソフトウェアレビューの課題

- ソフトウェアの規模増大に伴い、レビュー対象も増大
 - 網羅的なレビューを実施するのは現実的ではなくなっている。
- レビューの効果(理論上)



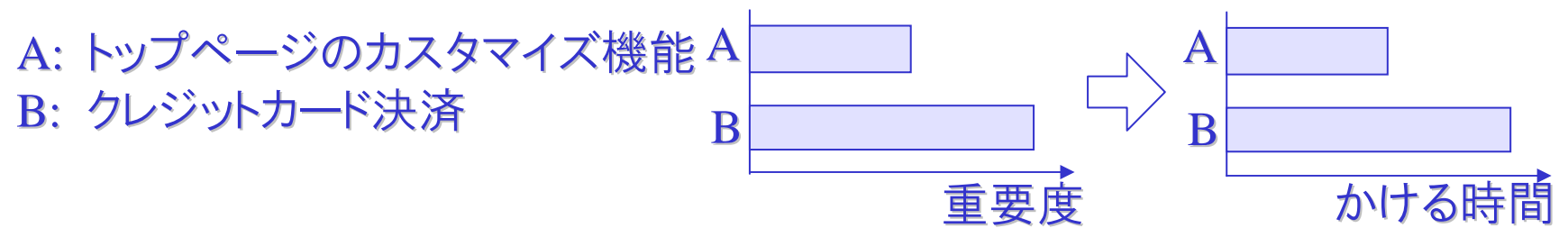
出典: 森崎 修司 “ソフトウェアインスペクションの動向” 情報処理学会誌 Vol. 50, No. 5, p. 377-384 ³



課題解決にむけた国際動向

- 時間管理型レビュー(優先順位つきレビュー)
ユースケースの重要度に応じてレビュー時間を配分

オンラインショッピングサイトでのユースケース例



出典: Kai Petersen, Kari Rönkkö, Claes Wohlin, “The impact of time controlled reading on software inspection effectiveness and efficiency: a controlled experiment <http://portal.acm.org/citation.cfm?id=1414004.1414029>

- 観点設定型(レビュー観点の設定)
 - エキスパートプログラマ: 性能、保守性
 - 初級プログラマ: 例外処理の存在確認
- 専門家によるフェーズつきレビュー
 - コーディングルール、リリース、セキュリティ

当学での取組み

- リスクベースドレビュー

- プロジェクトに応じたリスク設定

- レビュー開始前にリスク設定し、リスクの高い欠陥から指摘する。

- 修正確認テスト規模大をリスクとして設定

- 潜在的な修正確認テスト規模を意識したレビュー

- <http://thinkit.jp/article/896/1/>

- 田村 晃一, 亀井 靖高, 上野 秀剛, 森崎 修司, 松村 知子, 松本 健一, “見逃し欠陥の回帰テスト件数を考慮したコードレビュー手法,” 電子情報通信学会技術研究報告, Vol.108, No.173, pp.61-66, July 2008

- <http://se.naist.jp/~morisaki/publications/#withoutReview>

- ソフトウェアFMEA(故障モード影響解析)

- 山科 隆信, 森崎 修司, “大規模ソフトウェアの保守開発を対象とした故障モード影響解析(FMEA)適用の試み”, ユニシス技報, pp. 107-121, Vol. 28, No. 4, (Mar. 2009) http://www.unisys.co.jp/tec_info/tr99/99abs.htm#9908

- バグ票の修正工数の分析にもとづくレビュー観点設定

- 効果計測

- リスク設定により得られる効果と薄れる効果

- 支援環境

リスクベースドレビュー(観点設定型レビュー)

- プロジェクト計画時点でリスク定義をする。
 - プロジェクト、ソフトウェアのリスク
- 定義したリスクに関連するものを優先的に指摘する。
 - 例「前にデグレして顧客にかなり叱られている」
 - デグレードをリスクとして、デグレードのチェックを中心に

記録フォーマット		
1. リスク定義		
2. リスク記述		
3. 修正/確認コストの低減指標		
4. 指摘記録		
発見順番	時刻	指摘内容
1		

記録フォーマット				
1. リスク定義				
2. リスク記述				
3. 修正/確認コストの低減指標				
4. 指摘記録				
発見順番	時刻	指摘内容	(1) リスク低減	(2) 修正/確認コストの低減
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
(1) リスク低減の度合い	定性指標	A	リスク低減に効果があった	
		B	リスク低減にはほとんど効果がなかった	
(2) 修正/確認コストの低減	定量指標	人時	省略できたと推測される修正コスト	
		件	省略できたと推測されるテスト件数	
	定性指標	A	後工程で発見されたときよりも修正工数の削減ができた。	
		B	後工程で発見されたときと修正工数はほとんど変わらない。	

<http://se.aist-nara.ac.jp/html/review/formats.html>で書式と記載例を公開しています



過去のバージョンのバグ票からレビュー観点設定

- 派生開発、保守開発が前提
- 観点設定
 - 修正工数の概要把握
相関ルール分析に基づき、修正工数が大きな不具合の傾向を把握する。
 - 概要での情報にもとづき、バグ票を精査
レビューの観点としてふさわしいものを選択する。
- レビュー実施
 - 観点を設定してレビューする。
- 効果計測と改善
 - レビュー実施後
指摘欠陥を見逃した場合の修正コストを推測する。
 - テスト実施後
 - 観点にあてはまるものの見逃された欠陥の分析
 - 観点到設定していなかった不具合の修正コストを計測

不具合管理表から修正工数に関する相関ルールを抽出

- (再現性=低) かつ (機能=A) かつ (テスト種別=異常系)
 → 修正工数(平均 18.6人時 標準偏差 3.21)
 出現頻度: 3/8, 全体平均の1.86倍

再現性	機能	テスト種別	修正工数(人時)
低	A	異常系	20
常に	B	正常系	3
高	B	正常系	4
常に	X	異常系	2
低	A	異常系	15
高	X	異常系	10
低	A	異常系	21
低	A	正常系	5

前提部を含む不具合の修正工数の平均: 18.6

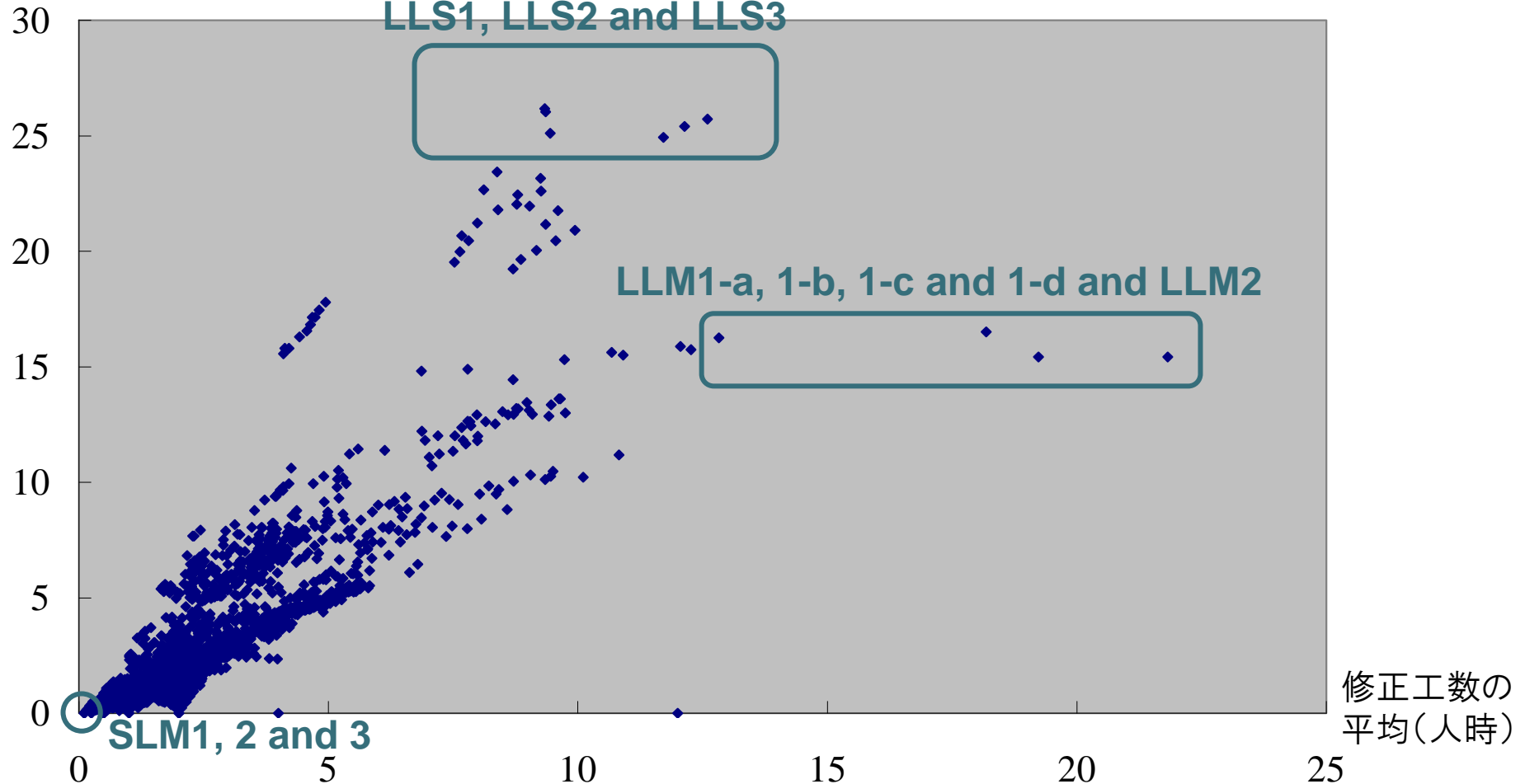
修正工数の全体平均: 10.0

基準化平均 = $18.6 / 10.0 = 1.86$

もし相関ルールが事実を表してれば、機能Aの異常系のレアケースを重点的にレビューする。

抽出ルール分布(出現頻度 0.5%以上の17,000ルールを抽出)

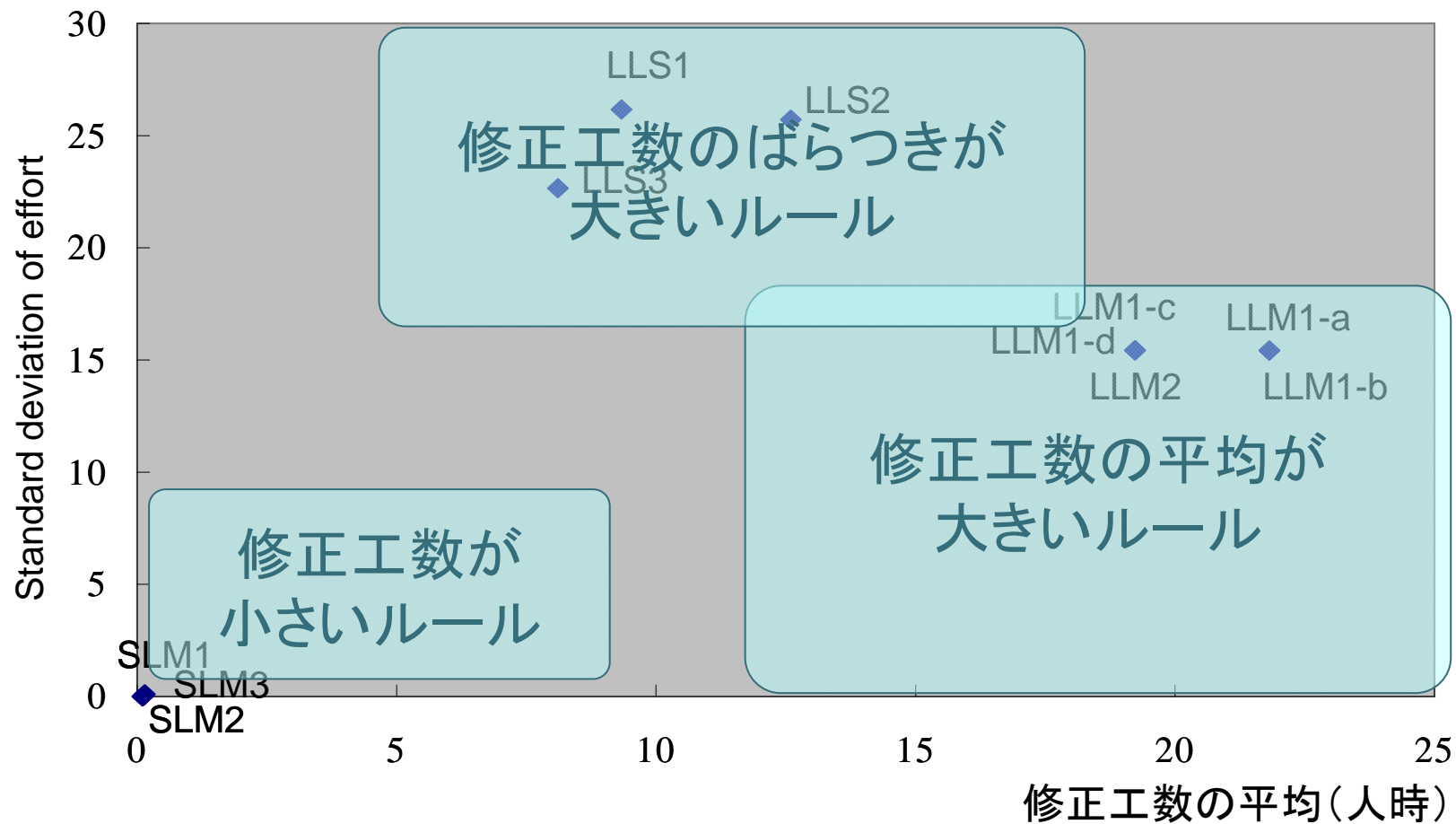
修正工数の標準偏差



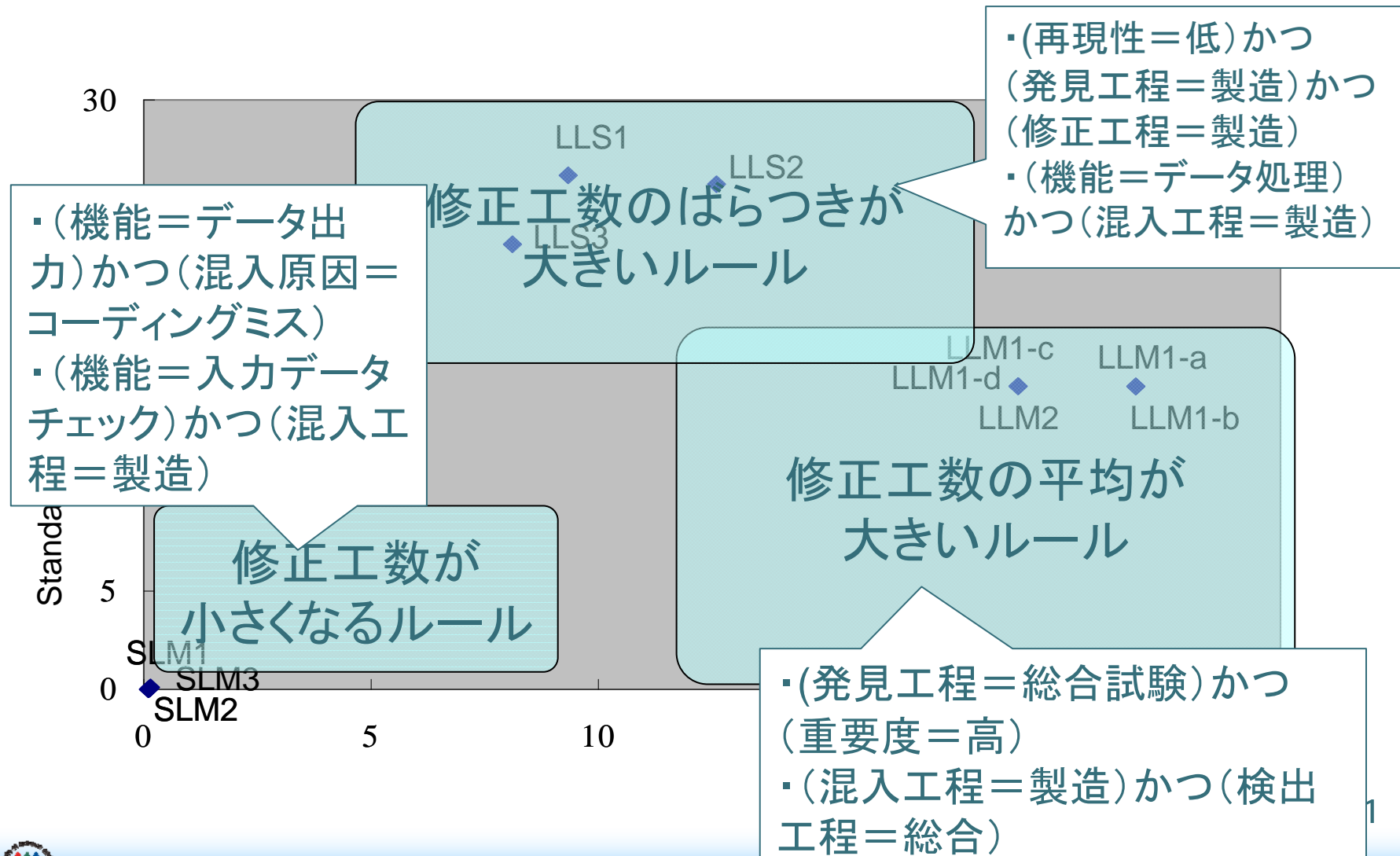
出典: 森崎、門田、玉田、松村、松本: “Defect Data Analysis Based on Extended Association Rule Mining”, Proceedings of International Workshop on Mining Software Repository, pp.17-24. <http://se.naist.jp/~morisaki/publications/#i-200705>

抽出ルール傾向分析

修正工数の標準偏差



抽出ルールの特徴分析



レビュー観点絞込みの評価

- 目的
 - レビュー観点の設定による指摘内容の変化を明らかにする。
 - (A) 観点設定なし、(B) 設定する、(C) 設定し指摘の都度確認、で比較
- 対象ドキュメント
 - Webアプリケーションの外部仕様書
 - 日本語で記述
- レビューア
 - 普段ソフトウェア開発に従事する実務者
 - 4～5名で6グループ
- 演習(セミナーの一環)として実施
 - 実施時間: 40分

観点の有無による欠陥指摘の違い

- A. リスク設定なし

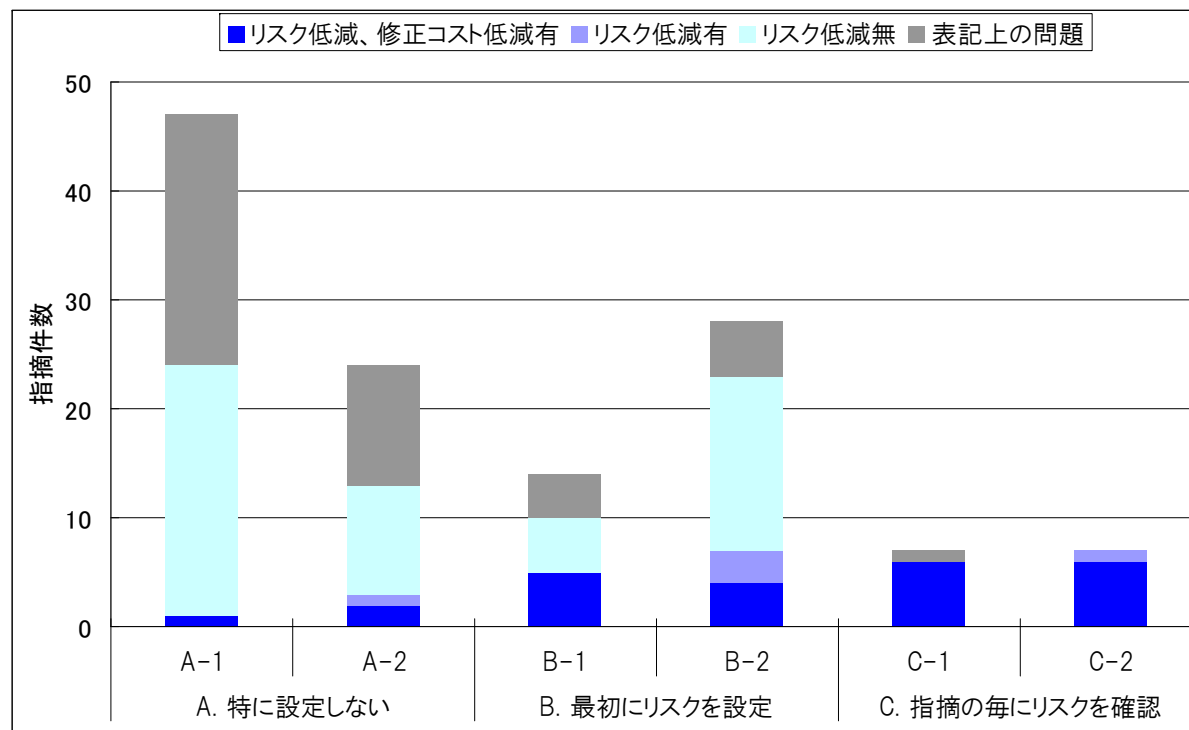
ID	指摘内容	リスク低減	修正コスト低減
...
3	ページ番号が1ページ目だけ欠けている	なし	なし
4	章の番号が違う(2. 3 4.のようにドットがないものがある)	なし	なし
...

- B, C. リスク設定あり (リスク: セキュリティ)

ID	指摘内容	リスク低減	修正コスト低減
...
5	本人確認用IDが不正に使用される可能性	あり	あり
6	個人情報登録前にどこまで公開するか同意を得る画面が必要	あり	あり
...

レビュー観点絞込みによる評価結果

- 指摘件数: グループA > グループB > グループC
- リスク低減指摘件数: グループC > グループB > グループA
- リスク低減につながら、見逃されてテストで修正されるよりも修正コストが小さくなる指摘: グループC > グループB > グループA



アジェンダ

- レビューの課題
- 課題解決へ向けた取組み
 - 国際動向
 - 当学での取組み
- リスクベースドレビュー
 - 観点設定方法
 - 観点設定の効果
- 東芝デジタルメディアエンジニアリングとの共同研究事例
 - 観点設定
 - 実施結果

DMEにおけるレビューの問題点

- レビューア自身の経験や知識に頼ったレビューを実施しており、レビュー効率はレビューアの質に大きく左右される。
- レビュー結果のデータ収集を行っているが、有効に活用されていない。
 - 指摘目標件数、レビュー時間、など
- 必ずしも正しいレビュー手法に則ったレビューが実施されているとは限らない。
 - レビュー対象物の事前配布、事前説明の実施、レビューアの選定
- 不具合情報の集計は行っているが、有効な分析手法が見出せず、レビューへのフィードバックが上手くいっていない。

対象とした不具合管理票

- プロジェクト
 - デジタルメディア処理ソフトウェア開発
(デジタルテレビの機能を模したメディア処理等を行う、組込み応用ソフトウェア開発)
 - C/C++で650KLOC(流用込み)、約3.5年
 - 派生開発
- 不具合
 - 発見フェーズ 単体～総合試験
 - 分析対象不具合件数 約350件
 - バグ以外の要求事項等は含まれない。

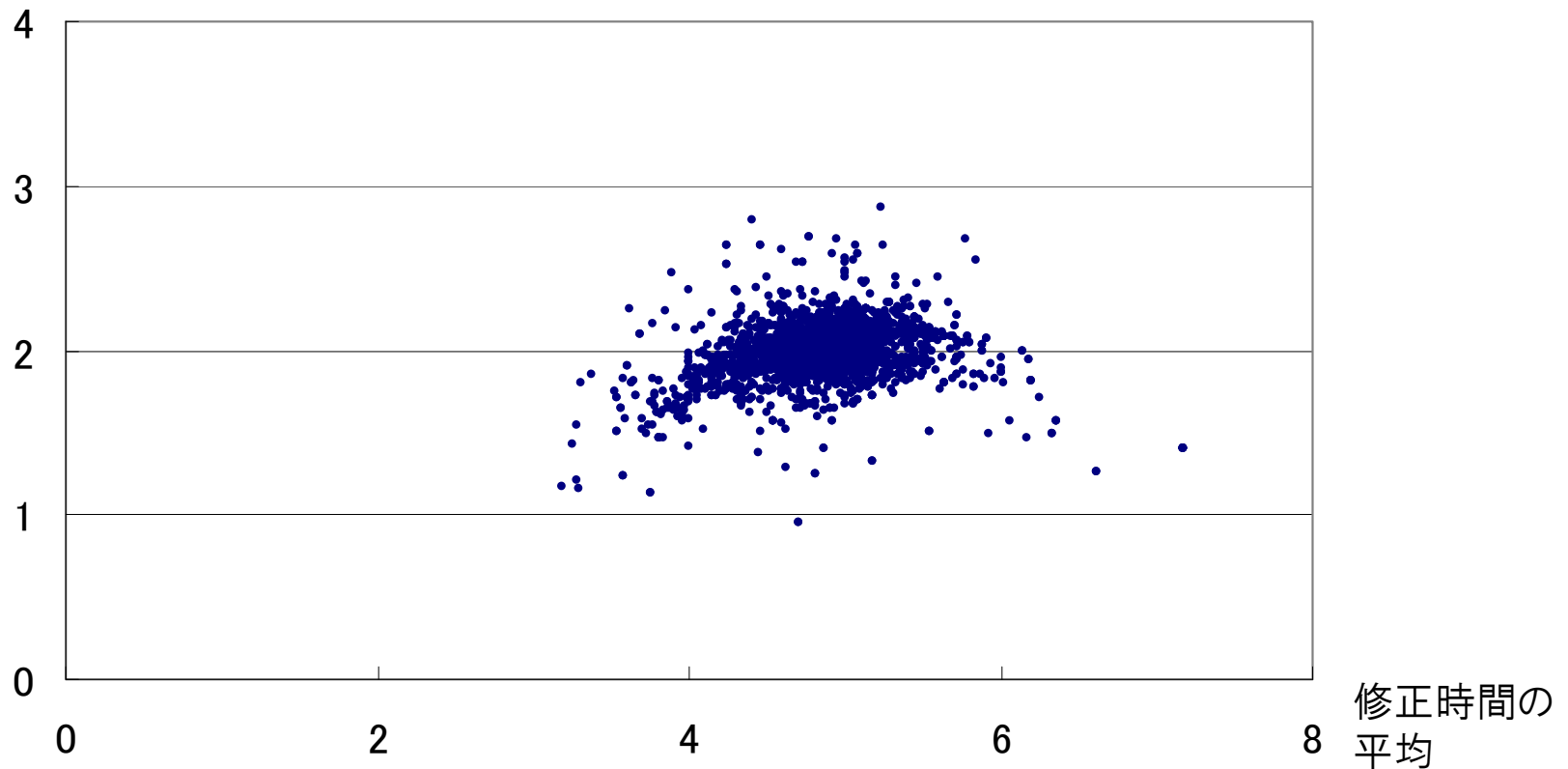
対象とした不具合管理票の項目

変数名	値
修正時間	修正に必要となった時間(手動で入力)
解析時間	不具合の原因箇所特定までの時間
担当者	担当者名(18名)
サブシステム カテゴリ	5種類(ミドルウェア、アプリケーション等)
サブシステム名	53種類(ミドルウェア、機能名等)
原因種別	設計ミス、設計漏れ、調査漏れ、調査ミス、外部設計漏れ、外部設計ミス、内部設計漏れ、内部設計ミス、コーディング漏れ、コーディングミス
修正期間	解決日と発生日の差(休日含む)
検査種別	正常系、異常系、限界系、その他
発生種別	新規バグ、内在バグ、既知バグ、その他

DMEとの共同研究事例

抽出ルール分布(修正期間にもとづく)

修正時間の標準偏差



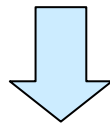
注目したルール

- 出現頻度、修正期間の平均、標準偏差が大きいものに着目

ルール	出現頻度	修正期間平均 (対全体平均)	修正期間標準偏差 (対全体標準偏差)
(サブシステムカテゴリ=MW(P)) かつ(原因種別=内部設計ミス)	3.8%	6.62 (1.26)	1.26 (0.73)
(担当者=担当-015)かつ (検査種別=異常系)	3.5%	6.33 (1.32)	1.50 (0.73)
(原因種別=内部設計漏れ)	10.2%	5.68 (1.18)	2.09 (0.62)
(サブシステムカテゴリ=MW(F)) かつ(検査種別=異常系)	4.4%	5.67 (1.18)	2.29 (1.02)
(サブシステム名=FW-00 MW(P)-001)	14.0%	5.62 (1.17)	2.11 (1.04)

相関ルールを元にバグ票を精査した結果

- 注目したルールに当てはまるバグの現象説明欄等からレビューの観点として設定できる項目を導出

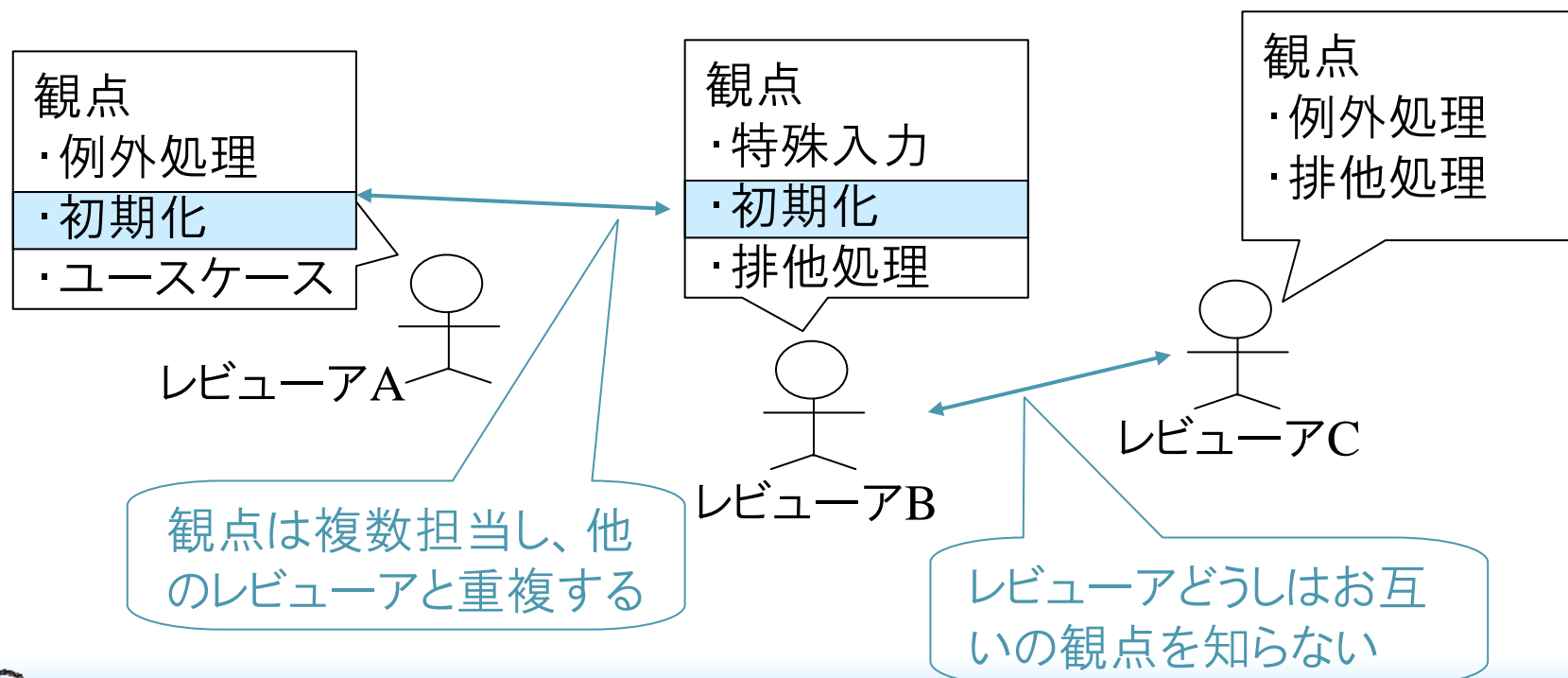


設計レビューの観点として適当なものを選択 (表中色塗りのセル)

分類	件数
例外処理	20
パラメータ初期化/クリア	8
特殊入力データ/パラメータ	8
排他処理	5
内部シーケンス	5
ユースケース	4
アルゴリズム	3
バッファオーバーフロー	2
データ領域管理	2
状態遷移	2
条件分岐	2
IF(戻り値)	2
操作ミス	1

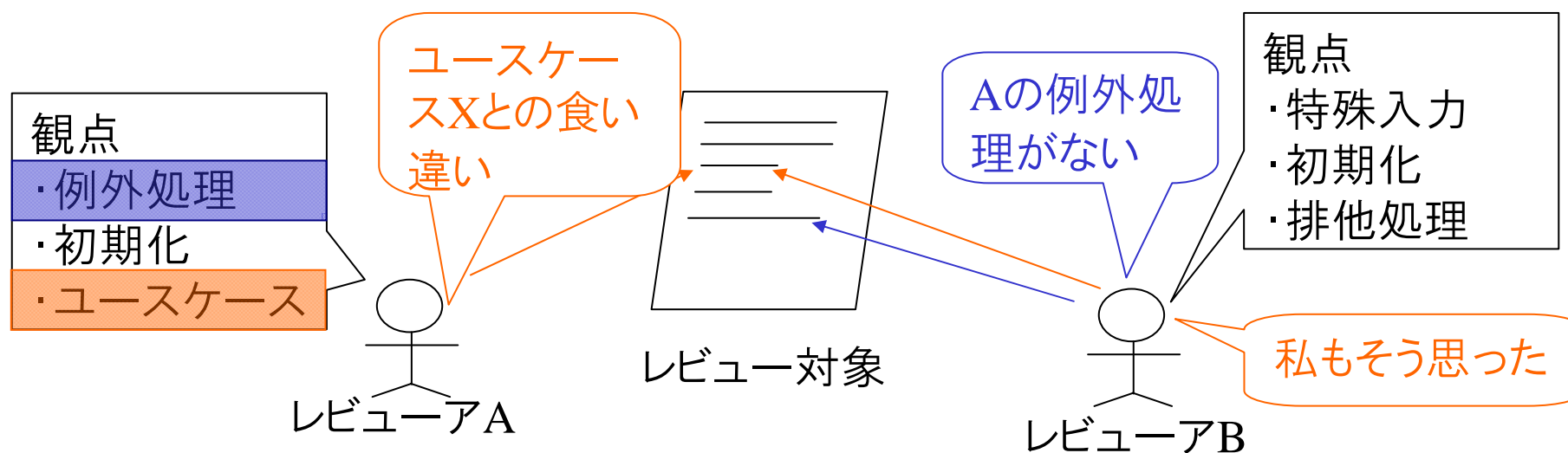
試行概要: レビューアへの観点割当て

- レビューア毎に異なる観点を割り当てる。観点は毎回同じ人に同じものを割り当てる。
 - 1人2~3観点(欠員の場合の考慮)
 - 観点の割り当てはランダム
- 他のレビューアの観点は伝えない。観点の設定の有無による違いを見るため。



試行概要: レビューアへの指示

- 割り当てられた「観点」を中心にレビュー対象をチェックする。
 - 割り当てられた観点以外の指摘も可
 - 他のレビューアの指摘時に「同様の指摘をしようと思った」場合、その旨を報告する。
- 観点が適用されていない場合、発見できた指摘か否かを判断するため



試行概要

- 試行1
 - 設計書作成者が内容を説明しながら、レビューアが気づいた点を指摘
 - 事前準備はなく、会議での読み合わせのみ
 - 対象
 - 外部設計書(サブシステムA)
 - 内部設計書(サブシステムB)
 - レビューア
 - プロジェクトメンバ4名 + 有識者 1名
 - 外部設計書、内部設計書ともに同一メンバ
- 試行2
 - レビューア: 3名(試行1と同一メンバ、特に対象に関わりが深いメンバ)
 - 対象: 外部設計書(サブシステムA)
 - 指摘会議のみ

試行結果1

- レビュー対象
 - 外部設計書(サブシステムA)
 - 内部設計書(サブシステムB)
- レビュー観点設定による効果は見られなかった。
 - 各レビューアに設定した観点とは無関係の指摘がほとんど
 - 仕様に詳しい一部のレビューアからの指摘が中心

レビュー対象	外部設計書 (サブシステムA)	内部設計書 (サブシステムB)
レビュー工数	6人時	7.5人時
全指摘件数	11件	6件
欠陥レベル中～高の指摘	6件	4件
観点を適用したと思われる指摘	1件	1件
予測修正工数	7人日	9人日

試行結果2 – 外部設計書(サブシステムA)

- レビュー方法
 - レビュー対象を事前配布し、各レビューアがチェック
 - レビュー時間は試行1の外部設計書と同様
 - レビュー結果
 - 観点からの指摘が多く見られた。
 - 欠陥レベル中以上の指摘率が高い。
- ⇒ 欠陥レベルの高い問題が見つかりやすく、観点適用の効果も高い。

	試行結果1	試行結果2
レビュー工数	6人時	3人時
人数	5名	3名
全指摘件数	11件	9件
欠陥レベル中～高の指摘	6件	8件
観点を適用したと思われる指摘	1件	6件
予測修正工数	7人日	16人日

考察

- レビュー実施方法

- レビュー当日の初見の状態では、観点からの指摘を探すことが難しい。
⇒ 事前にレビュー対象物の概要説明を実施して、レビューアに知識を与えた上でレビューを実施した方が効果が高い。
⇒ レビュー対象の事前配布 & 事前チェックも効果的と思われる。

- 設計書の記述レベル

- 設計仕様書は記載内容の専門性が高すぎる場合、専門知識保有者以外からの発言が挙がらない。
⇒ レビューメンバは専門知識保有者のみで構成する方が効率的。
この場合、「観点」に関する指摘は事前レビューで実施。
- 「観点」に関する内容が記載されていなければ指摘すらできない。
⇒ 逆に、観点を意識して設計書を作成することでレビュー効率が期待される。

- 試行方法

- 試行1が試行2に対する学習効果となっている可能性があるため、更なる検討を重ねる。

実施手順

- きっかけ
 - レビュー技法の講演 (2008/12)
(東芝 ソフトウェア技術センターとの共同研究契約に基づく)
- 準備
 - 講演内容にもとづく適用提案をDMEより、いただく。(2009/4)
 - レビュー観点の抽出にバグ票の修正工数分析を提案(打合せ 1 2009/6)
- レビュー観点抽出
 - 相関ルール抽出ツールの貸与(2009/7)
 - ルール抽出(DME) (2009/7)
 - ルールの解釈例の説明(打合せ 2 2009/7)
 - ルールの解釈とそれに基づく観点設定(DME)
- レビュー実施(DME) (2009/8)

まとめ

- レビューの課題
 - 規模増大
 - Low impact指摘
- 国際動向と当学での取組み
 - 時間管理型、観点設定型、専門家によるフェーズつきレビュー
- リスクベースドレビュー
 - 不具合管理票からの観点設定
 - 観点設定による効果の調査
- 東芝デジタルメディアエンジニアリングとの共同研究
 - 相関ルール分析に基づく過去のバージョンのバグの修正工数分析
 - 修正工数にもとづくレビュー観点設定
 - 観点の設定、事前の読込みにより効果が期待できる。