

An empirical study on the efficiency of different design pattern representations in UML class diagrams

Bruno Henrique Oliveira
Pedro Victor Pinheiro

30/08/2010

Agenda

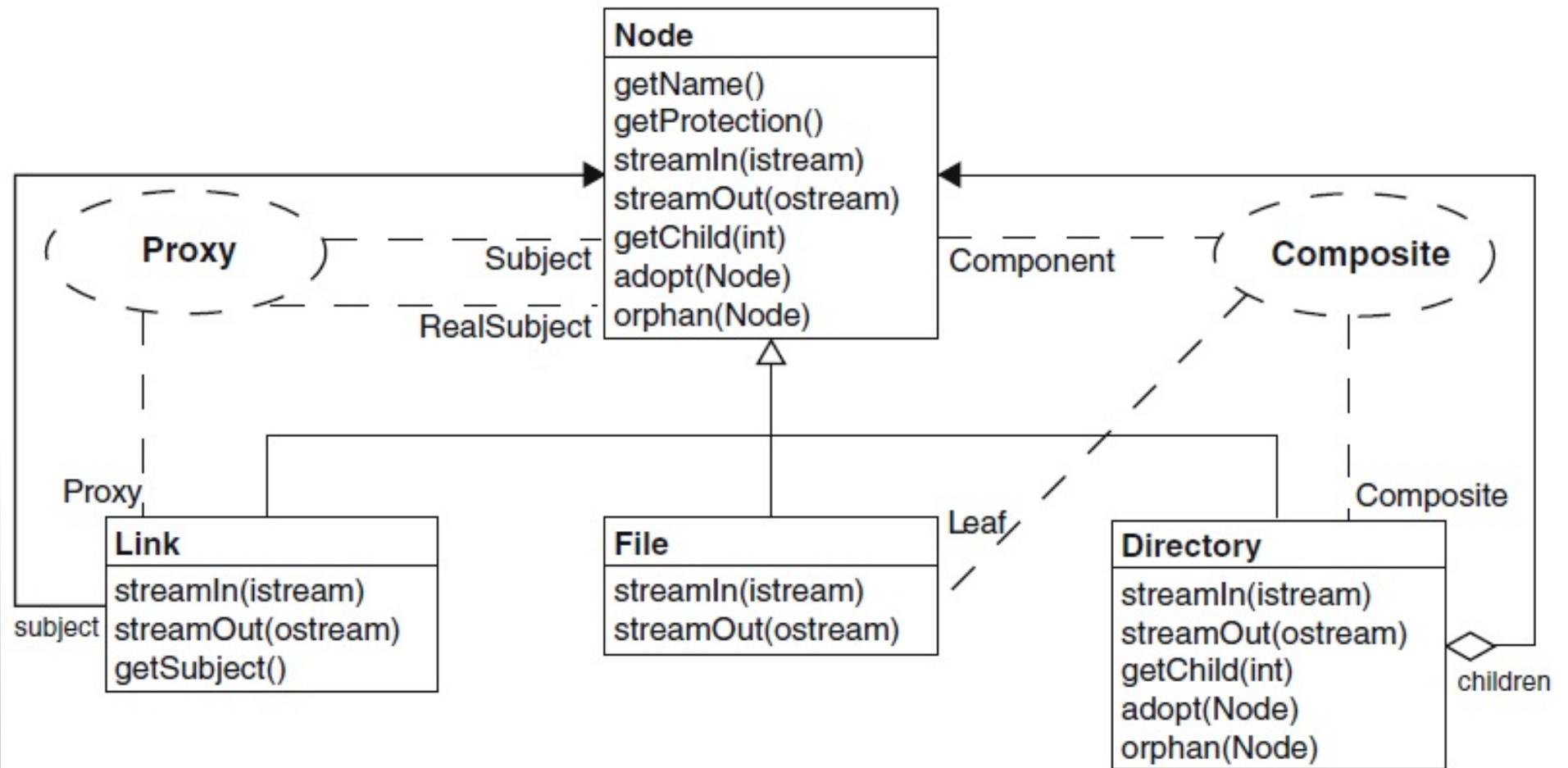
- Introdução
 - O Problema
 - Objetos
- O Experimento
 - Objetivo
 - Contexto
 - Hipóteses
 - Sujeitos
 - Padrões de Projeto
 - Variáveis
 - Projeto
 - Instrumentação
 - Ameaças à validade
 - Análise dos resultados
- Conclusão

Introdução

- Padrões de projeto
 - Soluções para problemas recorrentes ao se projetar programas orientados a objetos;
- Diagramas ajudam a:
 - construir uma representação mental da arquitetura do programa
 - desenvolver eficientemente o programa
 - manter eficientemente o programa

Introdução

- UML Colaboration notation
 - A representação mais comum usada para visualizar padrões de projeto



O Problema

- A UML Collaboration Notation possui limitações
- Várias representações foram propostas para suprir essas limitações
- Nenhum estudo empírico foi realizado para avaliar essas representações propostas

O Problema

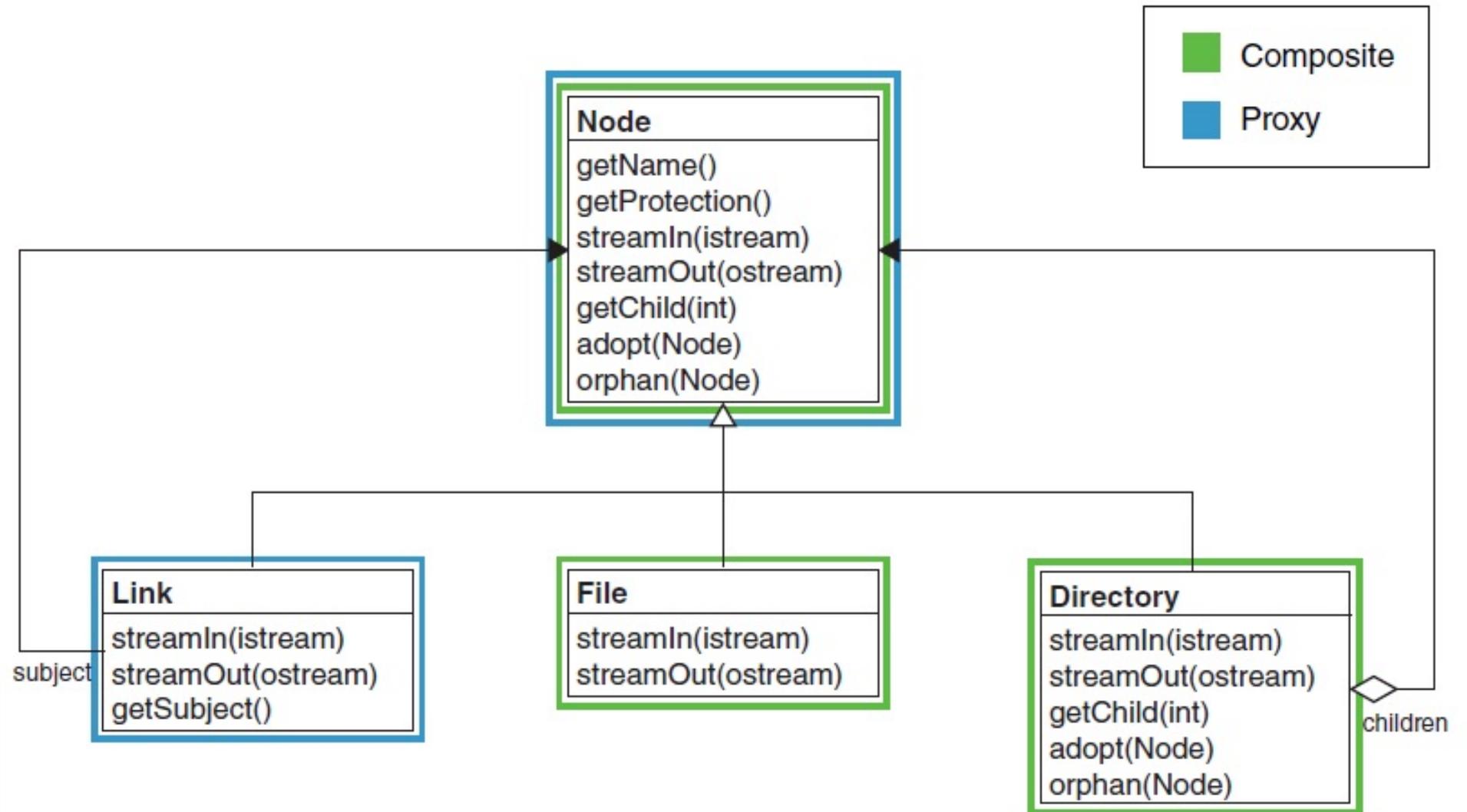
- As representações são divididas em:
 - Representações baseadas em não UML
 - Representações mono-diagrama
 - Representações multi-diagramas
 - Representações baseadas em UML
 - **Representações mono-diagrama (mais usada)**
 - Representações multi-diagramas

Tratamento

- Para realizar o experimento foram escolhidas as seguintes representações:
 - Diagramas de classe pattern-enhanced, rotulado de *Sc□hauer* (fortemente visual)
 - Diagramas UML stereotype-enhanced, rotulado de *Dong* (fortemente textual)
 - Notação “pattern:role”, rotulado de *Gamma* (visual e textual)

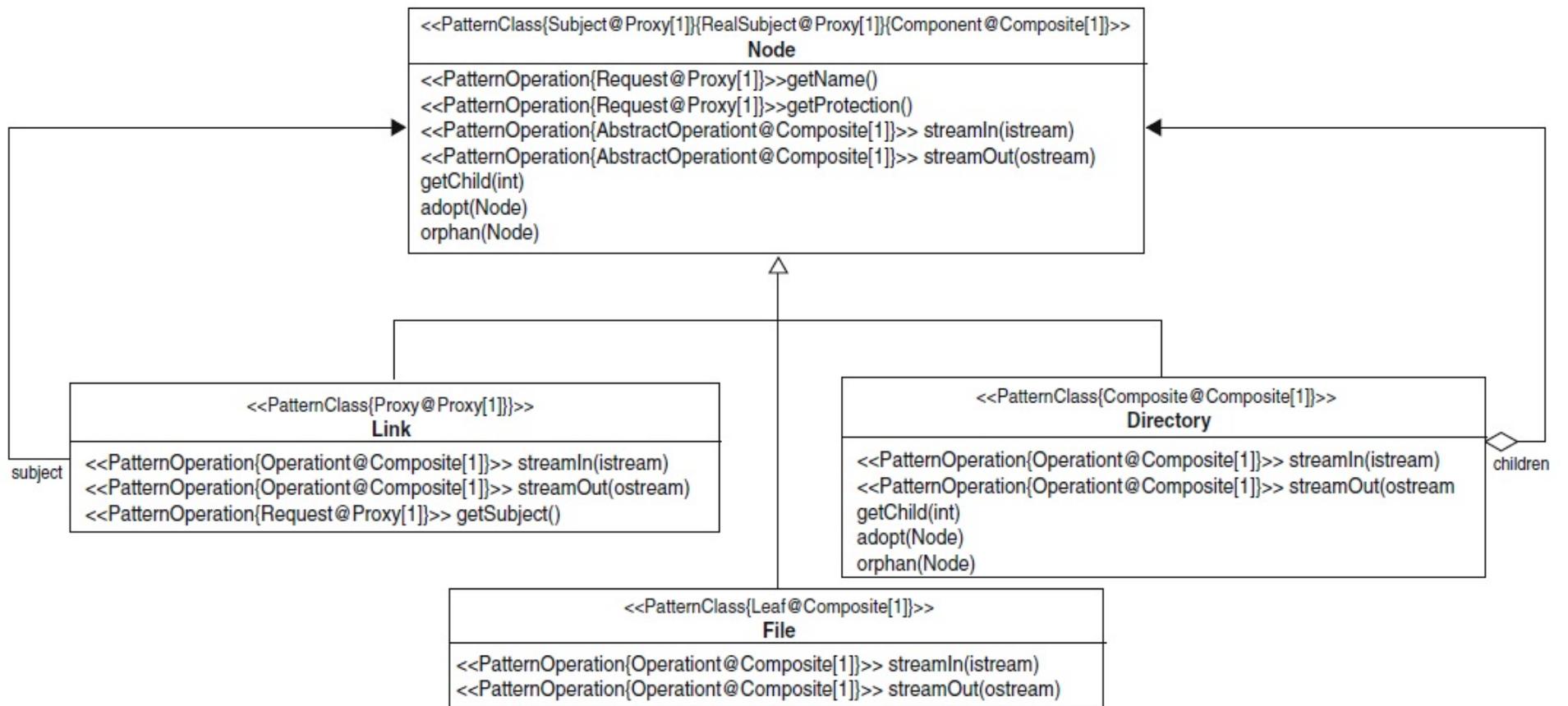
Tratamento

- Diagramas de classe pattern-enhanced, rotulado de *Sc hauer* (fortemente visual)



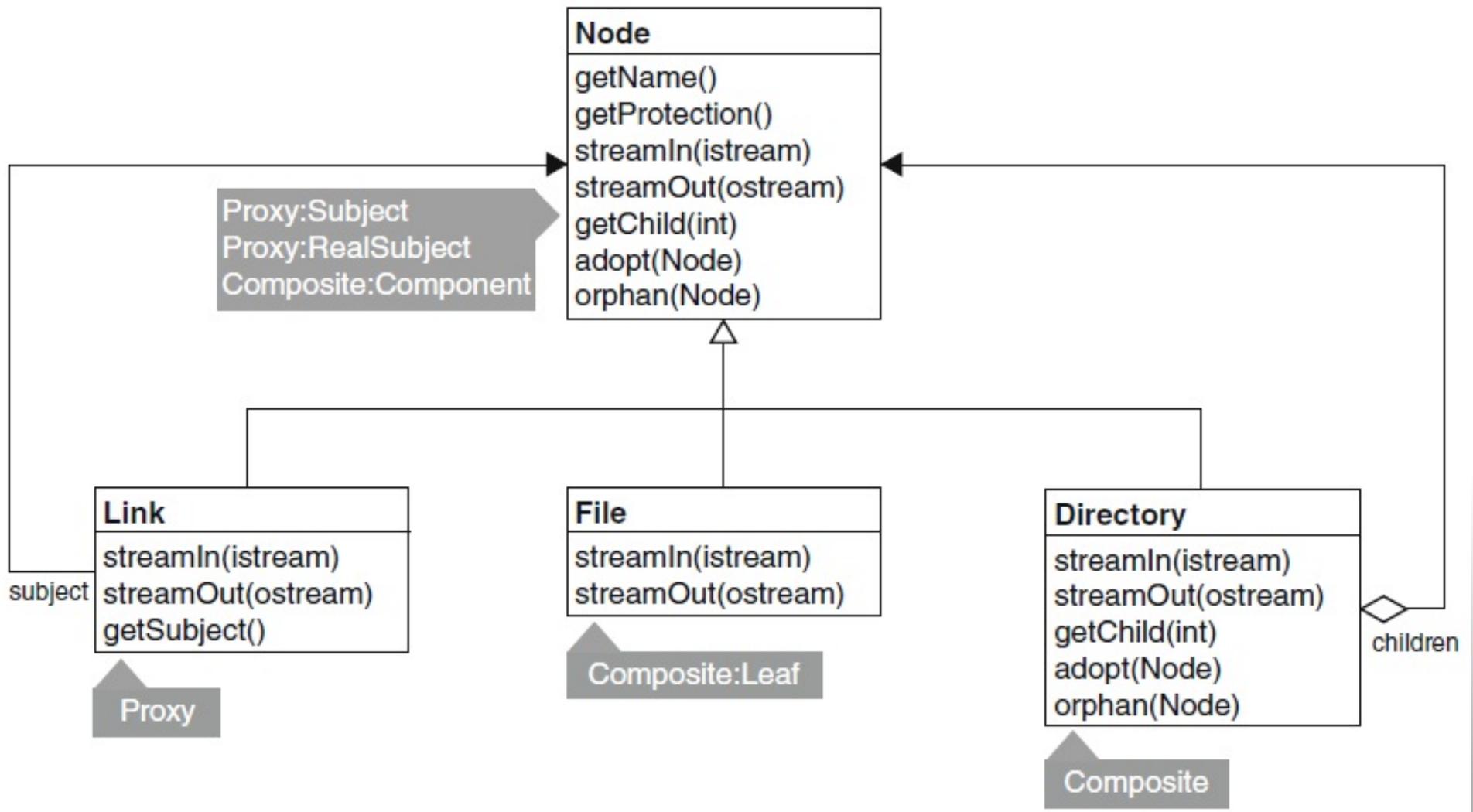
Tratamento

- Diagramas UML stereotype-enhanced, rotulado de *Dong* (fortemente textual)



Tratamento

- Notação “pattern:role”, rotulado de *Gamma* (visual e textual)



O Experimento

- O que será investigado:
 - Avaliar se ocorre ou não um aumento na performance dos desenvolvedores ao realizar tarefas de compreensão de padrões de projeto usando as representações Dong, Gamma e Schauer, quando comparadas à Representation UML.

O Experimento

- Tarefas de compreensão de padrões de projeto:
 - Participação - Identificar todas as classes que participam em um padrão de projeto
 - Composição - Identificar todos os padrões de projetos que uma classe participa
 - Papel - Identificar os papéis que uma classe participa em um padrão de projeto
- Essas tarefas são recorrentes na compreensão de programas e mantém curto os ensaios experimentais para garantir uma maior precisão dos dados registrados

Hipóteses

Hipóteses Nulas:

- H01: Não existe diferença na média dos esforços e precisão dos sujeitos usando a representação UML e os sujeitos usando a representação Dong.
- H02: Não existe diferença na média dos esforços e precisão dos sujeitos usando a representação UML e os sujeitos usando a representação Gamma.
- H03: Não existe diferença na média dos esforços e precisão dos sujeitos usando a representação UML e os sujeitos usando a representação Schauer.

Hipóteses

Hipóteses Alternativas:

- $H_{\alpha 3.1}$: A média dos esforços e precisão é superior para os sujeitos que usam a representação Dong do que os sujeitos que usam a representação UML.
- $H_{\alpha 3.2}$: A média dos esforços e precisão é inferior para os sujeitos que usam a representação Dong do que os sujeitos que usam a representação UML.
- $H_{\alpha 2.1}$: A média dos esforços e precisão é superior para os sujeitos que usam a representação Gamma do que os sujeitos que usam a representação UML.
- $H_{\alpha 2.2}$: A média dos esforços e precisão é inferior para os sujeitos que usam a representação Gamma do que os sujeitos que usam a representação UML.
- $H_{\alpha 1.1}$: A média dos esforços e precisão é superior para os sujeitos que usam a representação Schauer do que os sujeitos que usam a representação UML.
- $H_{\alpha 1.2}$: A média dos esforços e precisão é inferior para os sujeitos que usam a representação Schauer do que os sujeitos que usam a representação UML.

Sujeitos

- O experimento foi realizado com 24 sujeitos voluntários da pós-graduação da Universidade de Montreal
- Tinham experiência de pelo menos 2 anos em desenvolvimento usando diagramas de classe UML e padrões de projeto em seus estudos.
- A amostra foi considerada suficientemente heterogênea em termos de conhecimento de padrões de projeto

Sujeitos

- Os sujeitos foram divididos em grupos balanceados para simplificar e fortalecer a análise estatística dos dados coletados.

	15	40
<i>Dong</i>	$S_9, S_{11}, S_{12}, S_{21}, S_{23}, S_{24}$	$S_3, S_5, S_6, S_{15}, S_{17}, S_{18}$
<i>Gamma</i>	$S_6, S_8, S_{10}, S_{18}, S_{20}, S_{22}$	$S_2, S_4, S_{12}, S_{14}, S_{16}, S_{24}$
<i>Schauer</i>	$S_4, S_5, S_7, S_{16}, S_{17}, S_{19}$	$S_1, S_{10}, S_{11}, S_{13}, S_{22}, S_{23}$
<i>UML</i>	$S_1, S_2, S_3, S_{13}, S_{14}, S_{15}$	$S_7, S_8, S_9, S_{19}, S_{20}, S_{21}$

Padrões de Projeto

- Os padrões de projeto escolhidos foram:
 - Composite
 - Prototype
 - Template Method
 - State
 - Singleton
- São representantes dos padrões de projeto criacional, estrutural e comportamental.
- São claramente distinguíveis entre si
- Possuem em sua maioria um impacto subjetivo positivo nas características de qualidade do software (dispensabilidade, compreensibilidade e reusabilidade)

Variáveis

Independentes

- Representações (para 15 e 40 classes)
 - UML
 - Schauer (fortemente textual)
 - Gamma (visual e textual)
 - Dong (fortemente textual)

- Tarefas
 - Participação
 - Composição
 - Papel

Variáveis

Independentes

- Duas variáveis para mitigação de impacto no estudo:
 - Conhecimento no JHotDraw
 - Conhecimento em Padrões de Projeto

- Diagramas (Objeto Experimental)

Variáveis

Dependentes

As variáveis dependentes são escolhidas de acordo com as hipóteses formuladas e as variáveis independentes, baseadas na capacidade do sistema de rastreamento ocular.

Desempenho avaliado em termos de porcentagem de respostas corretas (CAP) e o esforço gasto pelos desenvolvedores para realizar as tarefas.

- AOG (*area of glance*)

Qualquer classe ou elemento da notação pertencente ao diagrama.

- AOI (*area of interest*)

Classe ou elemento relevante da notação que devem ter atenção especial por parte dos sujeitos envolvidos na hora de executar determinada tarefa.

Variáveis

Dependentes

AFD (Average Fixation Duration)

$$AFD = \frac{\sum_{i=1}^n (ET(F_i) - ST(F_i)) \text{ in AOG}}{n}$$

Onde,

$ET(F_i)$ = Encerramento da fixação (tempo)

$ST(F_i)$ = Início da fixação (tempo)

n = Numero de fixações

Quanto maior o AFD, significa que os usuários gastaram mais tempo para assimilar a informação representada.

Variáveis

Dependentes

ROAFT (Ratio of “On_target:All_target” Fixation Time)

$$ROAFT = \frac{\sum_{i=1}^n (ET(F_i) - ST(F_i)) \text{ in } AOI}{\sum_{j=1}^m (ET(F_j) - ST(F_j)) \text{ in } AOG}$$

Onde,

$ET(F_i)$ = Encerramento da fixação (tempo)

$ST(F_i)$ = Início da fixação (tempo)

n = Numero de fixações

Quanto menor a razão ROAFT, menor a eficiência.

Variáveis

Dependentes

ROAF (Ratio of “On_target:All_target” Fixations)

$$ROAF = \frac{\textit{Total Number of Fixations in AOI}}{\textit{Total Number of Fixations in AOG}}$$

Razões pequenas indicam baixa eficiência causadas pela necessidade de maior esforço para encontrar elementos necessários para realizar a tarefa.

Projeto

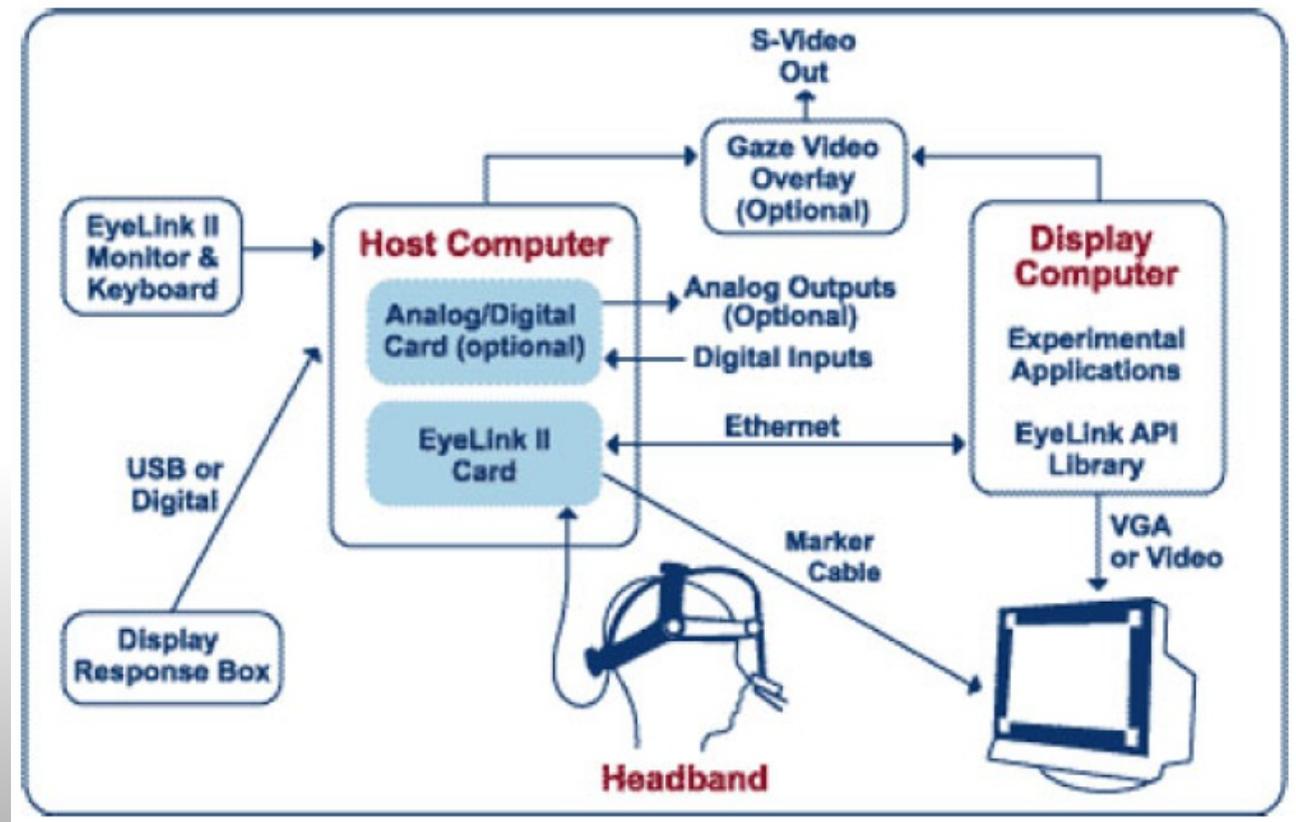
1. Inscrição pessoal, via email ou website (descrição do estudo).
2. Apresentação de um tutorial para introduzir as 4 representações utilizadas no estudo com alguns exemplos de diagramas para familiarizar os sujeitos com suas tarefas.
3. Apresentação do sistema de rastreamento ocular.
4. Acomodação do sujeito em um cadeira de dentista com um travesseiro para evitar movimentos com a cabeça.
5. Explicação da parte técnica do estudo (respostas devem ser dadas em voz alta).

Projeto

6. Calibração do sistema de rastreamento ocular.
7. Apresentação de tutorial básico do JHotDraw
8. Aquisição de dados (sem restrição de tempo).
9. Realização de pequena auto-avaliação sobre conhecimento dos sujeitos em padrões de projeto e JHotDraw. Questionário dado após experimento para não revelar o propósito do estudo
10. Ao final do experimento é dado um presente simbólico aos participantes. É solicitado que eles não compartilhem quaisquer informações com outros potenciais participantes até a data final do estudo.

Instrumentação

- Eye Tracking System (rastreador ocular)
- Formulários de pesquisa de conhecimento
- Computadores
- Website
- Videos
- Tutoriais



Ameaças à Validade

Internas

- Maturação
 - Ilustração das representações e tipos de tarefas a serem realizadas no tutorial
 - Apresentação das tarefas em ordem diferente para diferentes sujeitos
- Instrumentação
 - Minimização dos movimentos da cabeça (cadeira de dentista e travesseiro).
 - Compensação de desvio do movimento dos olhos.
- Difusão dos tratamentos
 - Solicitação de sigilo

Ameaças à Validade

Construção

- Mono-operação
 - Risco de utilização de apenas um sistema - O risco foi aceito para não elevar o grau de complexidade do estudo.
- Mono-método
 - Utilização de apenas um unico tipo de medida. Para mitigar esse risco foram utilizadas 4 variáveis dependentes, e os resultados foram comparadas entre si.
- Adivinhação de hipótese
 - Durante o experimento o objetivo do estudo não foi revelado para evitar que os sujeitos conhecessem as hipóteses.
- Apreensão
 - Sujeitos foram informados da ausência de risco na utilização do aparelho.
 - Garantia de anonimato nas respostas
 - Sem tempo limite para realizar as tarefas

Ameaças à Validade

Externas

- Interação de seleção e tratamento
 - Garantir que os sujeitos do estudo são amostras representativas de profissionais de software. Para isso foram selecionados estudantes já graduados com bom conhecimento em UML.
- Interação de configuração e tratamento
 - Garantir que o tamanho e complexidade do programa sejam significativos. Foi escolhido o programa JHotDraw pois utiliza extensivamente padrões de projeto. Foram utilizados diagramas com 15 classes (recomendado), e com 40 para analisar os mesmos pontos em programas mais complexos.

Ameaças à Validade

Conclusão

- Suposições violadas de testes estatísticos
 - Foram verificados e respeitados todos os pressupostos nos quais as análises foram baseadas.
- Confiabilidade das medidas
 - Foram utilizadas medidas bem documentadas.
 - Extremo cuidado na calibração do sistema de rastreamento ocular.
- Irrelevâncias aleatórias da configuração experimental
 - Ambiente silencioso
 - Teste preliminar (não contabilizados)
- Heretogenidade aleatória dos sujeitos
 - Seleção de sujeitos com níveis de conhecimentos variados sobre padrões de projeto

Análise dos resultados

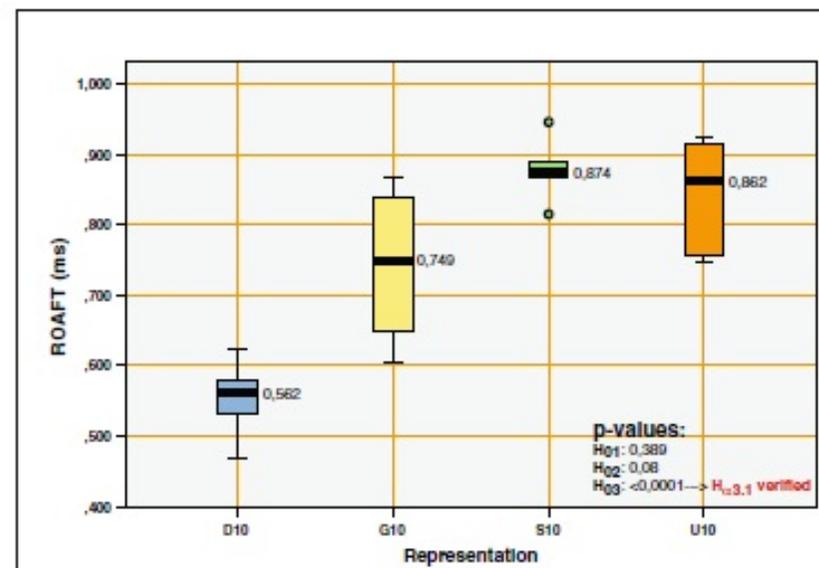
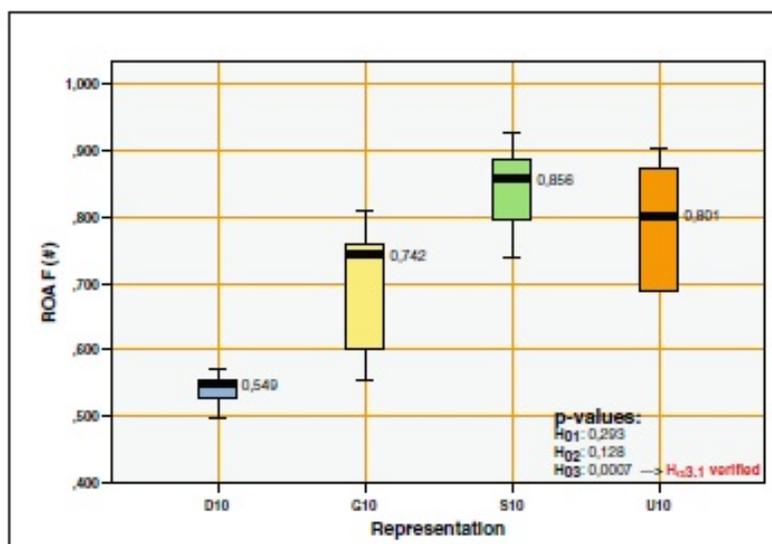
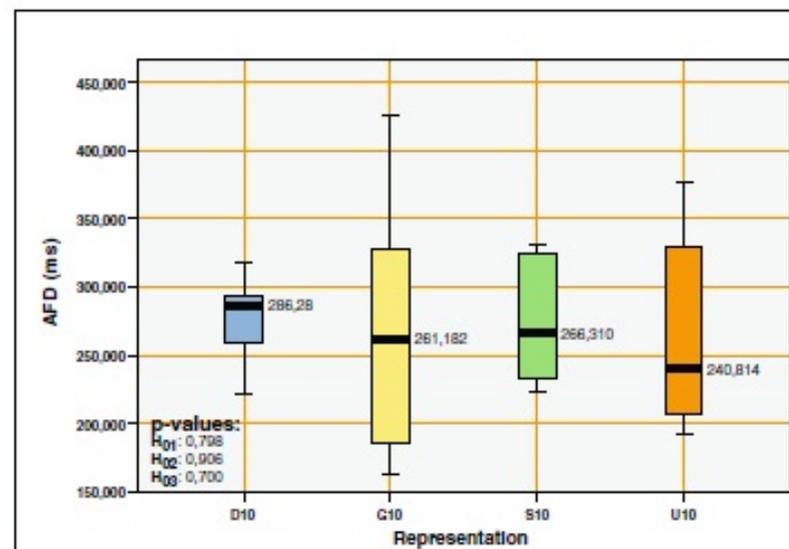
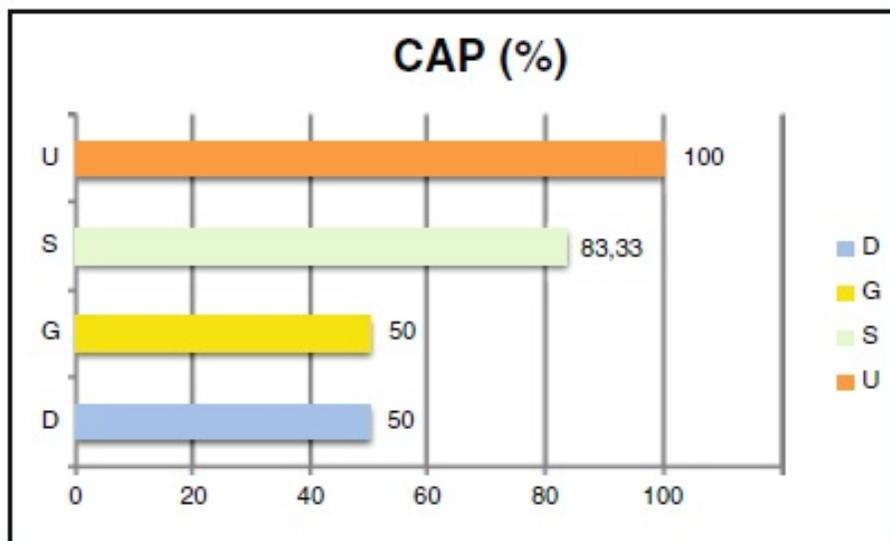
Participação

Diagrams	Perf.	Effort measures		
	CAP (%)	AFD (ms)	ROAF (#)	ROAFT (ms)
Dong₁₅	50.00	277.61	0.54	0.55
Gamma₁₅	60.00	270.60	0.70	0.74
Schauer₁₅	83.33	273.85	0.84	0.87
UML₁₅	100.00	264.66	0.79	0.84

Média das variáveis dependentes para as representações com 15 classes

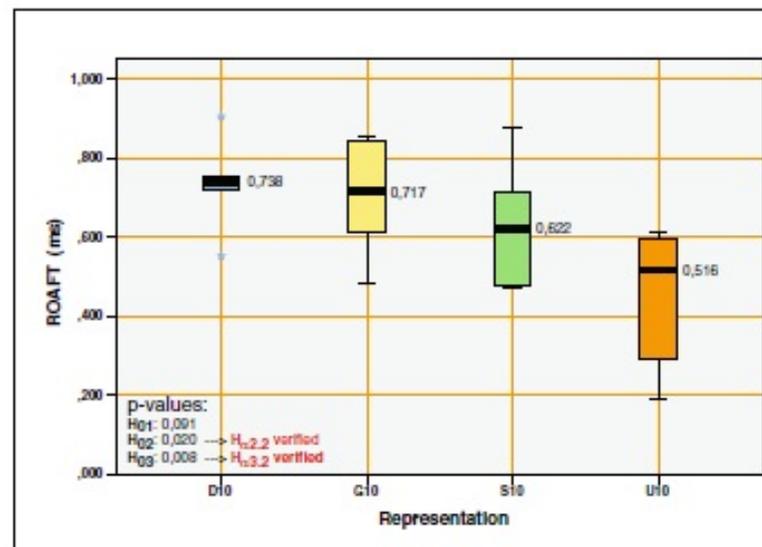
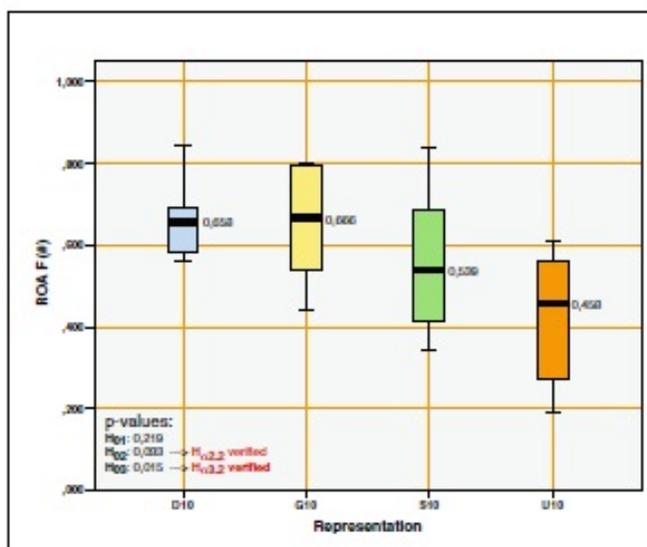
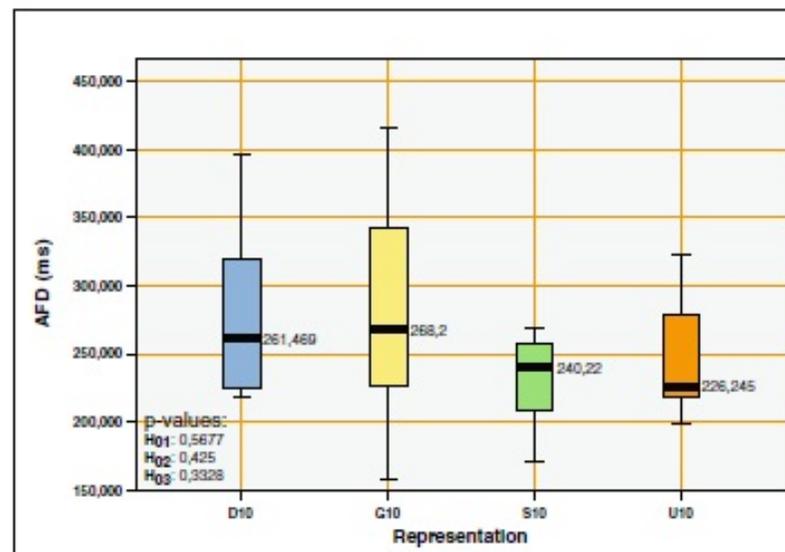
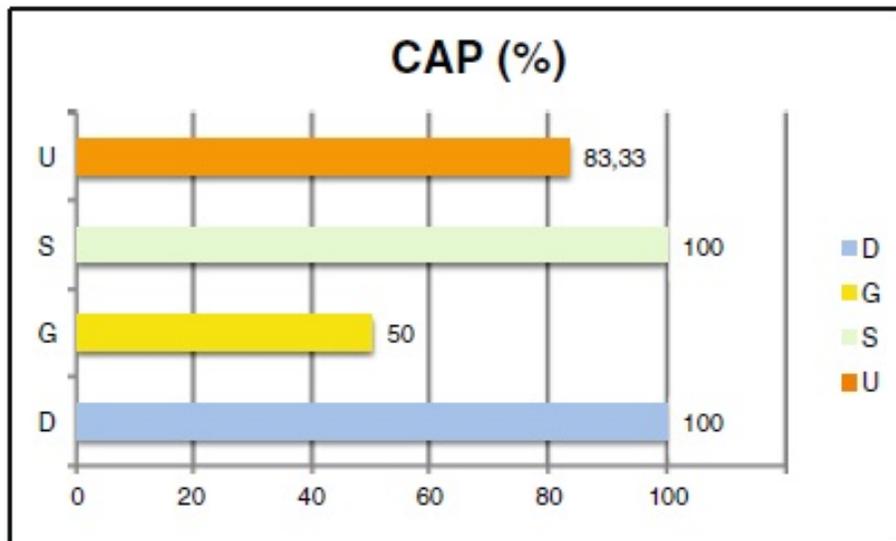
Análise dos resultados

Participação



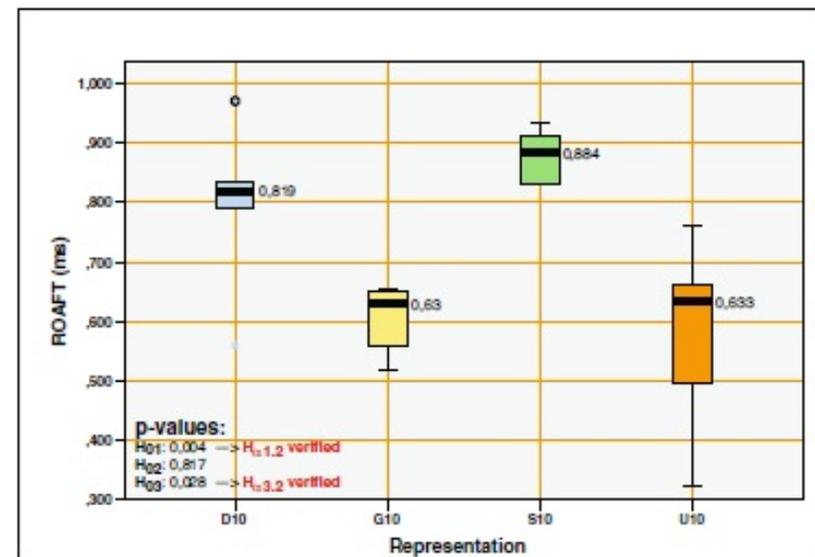
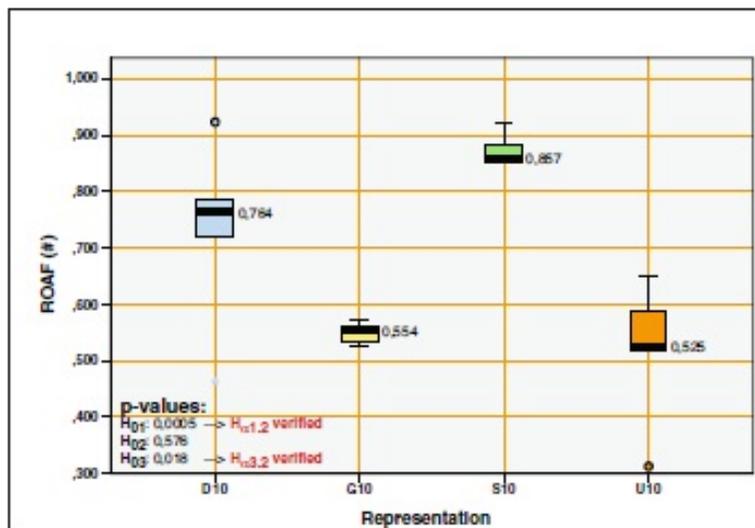
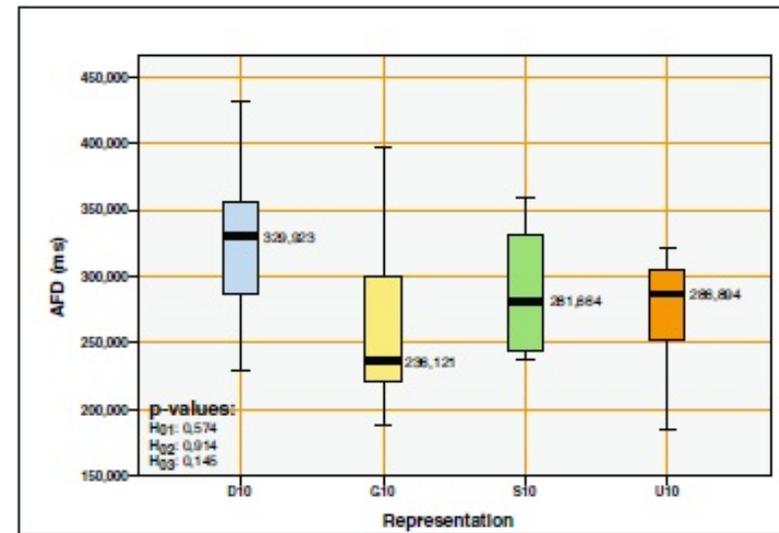
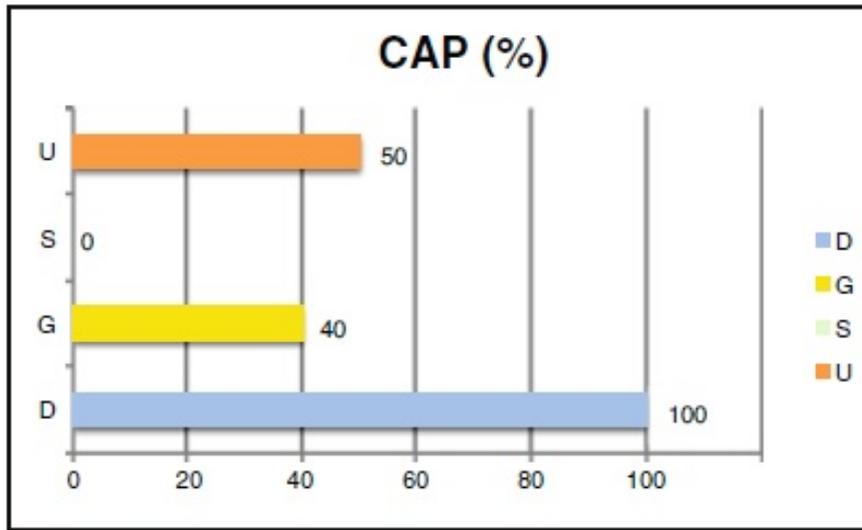
Análise dos resultados

Composição



Análise dos resultados

Papel



Conclusão

- A análise mostrou que diagramas UML melhorados com riqueza semântica são mais eficientes do que os diagramas de colaboração da UML para 15 classes.
- Os diagramas de colaboração da UML e os diagramas de classe melhorados são mais eficientes pra se encontrar as classes participantes em um padrão de projeto.
- Diagramas de 40 classes são difíceis de ler e portanto não se pode tirar conclusões desses diagramas.
- Fornece um *framework* para comparação dos diagramas e notações atuais e futuras.
- Todo material utilizado foi documentado com detalhes para que o experimento possa ser replicado.

Obrigado

Dúvidas?

bruno.ubatuba@gmail.com
pedrovictor.pinheiro@gmail.com