André Constantino da Silva

“Interação Multimodal em Ambientes de EaD: proposta de arquitetura e impactos”

CAMPINAS
2014

Universidade Estadual de Campinas
Instituto de Computação

André Constantino da Silva

"Interação Multimodal em Ambientes de EaD: proposta de arquitetura e impactos"

Orientador(a): **Profa. Dra. Heloísa Vieira da Rocha**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciência da Computação do Instituto de Computação da Universidade Estadual de Campinas para obtenção do título de Doutor em Ciência da Computação.

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA TESE DEFENDIDA POR ANDRÉ CONSTANTINO DA SILVA, SOB ORIENTAÇÃO DE PROFA. DRA. HELOÍSA VIEIRA DA ROCHA.

Assinatura do Orientador(a)

CAMPINAS
2014

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Matemática, Estatística e Computação Científica
Maria Fabiana Bezerra Muller - CRB 8/6162

Si38i

Silva, André Constantino da, 1981-
Interação multimodal em ambientes de EaD : proposta de arquitetura e impactos / André Constantino da Silva. – Campinas, SP : [s.n.], 2014.

Orientador: Heloísa Vieira da Rocha.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Computação.

1. Interação homem-máquina. 2. Ensino a distância. I. Rocha, Heloísa Vieira da,1954-. II. Universidade Estadual de Campinas. Instituto de Computação. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Multimodal interaction on e-Learning Environments : proposal for architecture and impacts
Palavras-chave em inglês:
Computer-human interaction
Distance education
Área de concentração: Ciência da Computação
Titulação: Doutor em Ciência da Computação
Banca examinadora:
Heloísa Vieira da Rocha [Orientador]
Renata Pontin de Mattos Fortes
Joice Lee Otsuka
André Santanchè
Fernanda Maria Pereira Freire
Data de defesa: 14-03-2014
Programa de Pós-Graduação: Ciência da Computação

TERMO DE APROVAÇÃO

Defesa de Tese de Doutorado em Ciência da Computação, apresentada pelo Doutorando André Constantino da Silva, aprovada em 14 de março de 2014, pela Banca examinadora composta pelos professores doutores:

Prof^a^. Dr^a^. Renata Pontin de Mattos Fortes
ICMC / USP

Prof^a^. Dr^a^. Joice Lee Otsuka
UAB / UFSCar

Prof. Dr. André Santanchè
IC / UNICAMP

Prof^a^. Dr^a^. Fernanda Maria Pereira Freire
NIED / UNICAMP

Prof^a^. Dr^a^. Heloísa Vieira da Rocha
IC / UNICAMP

Instituto de Computação
Universidade Estadual de Campinas

Interação Multimodal em Ambientes de EaD: proposta de arquitetura e impactos

André Constantino da Silva[1]

14 de março de 2014

Banca Examinadora:

- Profa. Dra. Heloísa Vieira da Rocha (*Orientadora*)
  Instituto de Computação (IC) - UNICAMP
- Prof. Dr. André Santanchè
  Instituto de Computação (IC) - UNICAMP
- Dra. Fernanda Maria Pereira Freire
  Núcleo de Informática Aplicada à Educação (NIED) - UNICAMP
- Profa. Dra. Renata Pontin de Mattos Fortes
  Instituto de Ciências Matemáticas e de Computação (ICMC) – USP
- Profa. Dra. Joice Lee Otsuka
  Secretaria Geral de Educação a Distância (SEaD) – UFSCar
- Prof. Dr. Celmar Guimarães da Silva (*Suplente*)
  Faculdade de Tecnologia (FT) – UNICAMP
- Dra. Maria Cecília Martins (*Suplente*)
  Núcleo de Informática Aplicada à Educação (NIED) - UNICAMP
- Delano Medeiros Beder (*Suplente*)
  Departamento de Computação (DC) – UFSCar

[1]Esta tese de doutorado foi desenvolvida com suporte financeiro da CAPES e do CNPq (141423/2011-0).

" Numa folha qualquer
Eu desenho um sol amarelo
E com cinco ou seis retas
É fácil fazer um castelo...".

Trecho da música Aquarela de autoria de Toquinho lançada em 1983.

Ao meu pai (in memoriam), ausente desta etapa de minha vida. Te amo.

Agradecimentos

À Deus, que me possibilitou alcançar esta meta, abençoando-me em todos os momentos.

Aos meus pais, Antônio e Bernadete, e aos meus irmãos, Antônio e Rafael, que, apesar da distância e da minha ausência em muitos momentos de suas vidas, me incentivaram a continuar com os estudos.

À toda minha família, que compreendeu o motivo de minha ausência e que sempre demonstrou seu orgulho por mim.

À minha orientadora, professora e amiga Profa. Dra. Heloísa Vieira da Rocha, a quem sou triplamente agradecido. Primeiramente, por acreditar no meu potencial e abrir a oportunidade de realizar o doutoramento. Em segundo, por me dar a liberdade de desenvolver este trabalho seguindo o meu "faro" de pesquisador, podendo, não somente aprender a realizar uma pesquisa, mas também a me aprimorar como pesquisador. E, também, por ser um modelo exemplar de professora e de pessoa. Muito do que sou como professor e pessoa devo à ti.

À colega de pesquisa e amiga Fernanda Maria Pereira Freire, pelos ótimos momentos em reuniões e intervalos para um café, sempre com muitos rabiscos no papel, e pelos conselhos e apoio. Apesar de não possuir vínculo formal neste trabalho, sua presença foi fundamental para a sua realização.

À Ciça, pelo carinho e apoio tanto no trabalho quanto no concurso para professor no IFSP.

Aos meus colegas de orientação, Ariel e Ricardo, que, ao longo desses seis anos, realizamos diversos trabalhos juntos.

Ao professor Dr. Rodolfo J. de Azevedo, por aceitar minha participação no projeto *Tablet PC Technology at UNICAMP*.

Aos meus amigos do e-Labora Delano, Joice, Celmar, Eduardo, Alex, Glauber entre outros, pela troca de experiências sobre desenvolvimento de software, ambientes de EaD e tecnologias.

À Vânia, ao Delano e à Joice (novamente) que me possibilitaram vivenciar a educação a distância por meio de tutoria na UAB.

Ao Fábio, Glauber, Elaine, Cristian, Tatá, Raoni, Márcio, Jefferson, Leonardo e Rodrigo, dentre outros amigos, que fiz na UNICAMP e em Campinas. Vocês sempre poderão contar comigo, obrigado pelo apoio.

Aos professores e funcionários do Instituto de Computação e, em especial, do Núcleo de Informática Aplicada a Educação (NIED), pela atenção e colaboração, oferecendo um espaço rico em discussões nas quais pude desenvolver minhas ideias, não somente as relacionadas a este trabalho, mas a prática docente.

Ao NIED e ao Instituto de Computação por disponibilizarem os equipamentos utilizados neste trabalho.

À CAPES, ao CNPq, ao Instituto de Computação, à UNICAMP e ao IFSP pelo apoio financeiro por meio de bolsas e auxílios viagem, os quais possibilitaram a divulgação dos resultados do meu trabalho a pesquisadores de diversas partes do Brasil e do mundo.

Resumo

Ambientes de Educação a Distância (EaD) proveem ferramentas para apoiar atividades de ensino e aprendizagem de cursos presenciais, semipresenciais e a distância, utilizando a infraestrutura da Web para disponibilizar suas funcionalidades aos usuários. As ferramentas existentes nesses ambientes oferecem a manipulação de mídias ricas, como imagem, áudio e vídeo, possibilitando a criação e visualização de conteúdo pelos usuários; cujas interfaces de usuário foram projetadas para ter boa usabilidade em computadores convencionais, equipados com os dispositivos de entrada teclado e mouse e como dispositivo de saída um monitor de tamanho médio com alta resolução. Por utilizarem a Web como meio de acesso, os ambientes estão suscetíveis ao acesso por uma variedade de dispositivos, inclusive os móveis que possibilitam utilizar as funcionalidades do ambiente a qualquer hora em qualquer lugar. Ressalta-se que os dispositivos móveis são equipados com telas sensíveis ao toque e/ou caneta, modalidades de entrada não existentes em um computador convencional. Portanto, os ambientes de EaD necessitam suportar diversas modalidades de interação, que serão utilizadas pelos usuários considerando o seu objetivo, o dispositivo, o contexto físico e social. Projetar uma interface específica para cada dispositivo não é uma tarefa trivial e resulta em um alto número de linhas de código que devem ser mantidas; assim propomos a adoção da multimodalidade. A abordagem de pesquisa adotada contempla: (i) revisão da literatura sobre conceitos fundamentais da área de Interação Humano-Computador e multimodalidade; (ii) investigação dos problemas de interação que ocorrem ao acessar ambientes de EaD por modalidades que não foram consideradas durante a especificação de sua interface; (iii) desenvolvimento de uma ferramenta que aproveite os benefícios das modalidades de entrada caneta e multitoque; (iv) elaboração de uma arquitetura para ambientes de EaD com interação multimodal. Percebeu-se que o impacto da multimodalidade pode ir além de possibilitar acesso por diversos periféricos de interação, podendo surgir novas funcionalidades que podem apoiar a produção de conteúdos que antes era difícil, ou impossível, de serem produzidos nos ambientes.

Palavras-chave: Interação Humano-Computador, Informática Aplicada a Educação, Ambientes de EaD, Interação Multimodal, Usabilidade, Mobilidade e Dispositivos Móveis.

Abstract

e-Learning Environments provide tools to support teaching and learning activities in classroom, blended or on-line courses, by using the infrastructure of the Web to provide its functionality to users. The tools in these environments offer the handling of rich media such as image, audio and video, allowing the creation and visualization of content; their user interfaces were designed to have good usability using a desktop computer with keyboard and mouse as input devices and a high-resolution medium-size display as output devices. By using the Web as a means of access, these environments are susceptible to access by a variety of devices, including mobile ones that enable users to use the environment anytime and anywhere. It is emphasized that the mobile devices are equipped with touch and/or pen input modalities, two modalities that do not available on desktop computers. Therefore, e-learning environments need to support different modes of interaction, which will be used by the users considering their goal, the device, and the physical and social context. Designing a particular interface for each device is not a trivial task and results in a high number of code lines that must be maintained, and then we propose the adoption of multimodality. The research approach adopted contemplates: (i) literature review about fundamental concepts of Human-Computer Interaction and multimodality; (ii) investigation of the interaction problems that occur when accessing e-learning environments with modalities that have not been considered in the design process; (iii) development of a tool that take the benefits of pen and multitouch input modalities; (iv) proposal for an architecture for distance education environments with multimodal interaction. We noticed that the impact of multimodality can go beyond enabling access for various peripherals and may emerge new functionalities that support the production of content that were difficult before, or impossible, to be produced in the environments.

Keywords: Human-Computer Interaction, Informatics Applied on Education, e-Learning Environments, Multimodal Interaction, Usability, Mobility and Mobile Devices.

Lista de Figuras[1]

[1] Devido a esta tese conter artigos publicados e consequentemente possuir numeração própria de figuras e tabelas, decidiu-se nesta listagem acrescentar entre colchetes o número do capítulo que possui a tabela ou figura, e usar letras para distinguir entre os artigos quando mais de um artigo compor o capítulo.

Lista de Tabelas

Lista de Abreviaturas e Siglas

ACM	*Association for Computing Machinery*
API	*Application Programming Interface*
ASR	*Automatic Speech Recognition*
ATM	*Automatic Teller Machine*
BRETAM	*Breakthrough, Replication, Empiricism, Theory, Automation, Maturity model*
CARE	*Complementarity, Assignment, Redundancy and Equivalence*
CCXML	*Voice Browser Call Control*
DOM	*Document Object Model*
EaD	Educação a distância
FAME	*Framework for Adaptive Multimodal Environments*
GUI	*Graphical User Interface*
ICARE	*Interaction Complementarity Assignment, Redundancy and Equivalence*
IHC	Interação Humano-Computador
InkML	*Ink Markup Language*
ISO	*International Organization for Standardization*
LCMS	*Learning Content Management Systems*
LMS	*Learning Management Systems*
MCD	Memória de Curta Duração
MIA	Memória da Imagem Auditiva
MIV	Memória da Imagem Visual
MLD	Memória de Longa Duração
MMI	*Multimodal Interaction*
MMI-WG	*Multimodal Interaction Working Group*
MMWA	*MultiModal Web Approach*
MPIH	Modelo de Processador de Informação Humano
MT	Memória de Trabalho
MWUIs	*Multimodal Web User Interfaces*

NUI	*Natural User Interface*
PC	Processador Cognitivo
PM	Processador Motor
PP	Processador Perceptual
SCXML	*State Chart Extensive Markup Language*
SIGCHI	*Special Interest Group on Computer-Human Interaction*
SMIL	*Synchronized Multimedia Integration Language*
SVG	*Scalable Vector Graphics*
UIDL	*User Interface Description Language*
UsiXML	*User Interface eXtensible Markup Language*
TIC	Tecnologia da Informação e Comunicação
VoiceXML	*Voice Extensible Markup Language*
W3C	*World Wide Web Consortium*
XHTML	*eXtensible HyperText Markup Language*
XML	*eXtensible Markup Language*

Sumário

Capítulo 1

Introdução

Este Capítulo apresenta o contexto geral do trabalho desenvolvido, a área de estudo o qual se vincula, bem como definições para os principais conceitos adotados ao longo do seu desenvolvimento.

1.1 Contextualização

A Interação Humano-Computador (IHC) é “a disciplina preocupada com o design, avaliação e implementação de sistemas computacionais interativos para uso humano e com o estudo dos principais fenômenos ao redor deles” (ACM SIGCHI, 1992, p. 6; Rocha e Baranauskas, 2003). Segundo Barbosa e da Silva (2010) a área de IHC está “interessada na qualidade de uso desses sistemas e no seu impacto na vida dos seus usuários”, e complementam:

> “De acordo com Hewett e seus colegas (1992), os objetos de estudo de IHC podem ser agrupados em cinco tópicos inter-relacionados: a natureza da interação humano-computador; o uso de sistemas interativos situado em contexto; características humanas; arquitetura de sistemas computacionais e da interface com usuários; e processos de desenvolvimento preocupados com o uso” (Barbosa e da Silva, 2010, p. 10).

Segundo a IHC, projetar uma interface, ou seja, fazer um *design* de interface, é uma atividade que envolve conhecer: (i) a tecnologia e suas restrições; (ii) os usuários, seus conhecimentos, suas preferências e limitações; (iii) a área da aplicação e as tarefas que serão realizadas por meio do dispositivo computacional; (iv) o contexto de uso, além de ferramentas adequadas para apoiar o processo de desenvolvimento da interface de usuário. Conforme Barbosa e da Silva (2010) afirmam “Conhecer essas tecnologias e dispositivos é fundamental para sermos capazes de propor, comparar, avaliar e tomar decisões sobre formas alternativas de interação com sistemas computacionais” (p. 10). Apesar da variedade de conhecimentos necessários para se desenvolver uma interface de usuário, Shneiderman (2006) descreve que “o desafio dos projetistas é compreender com mais profundidade o que o usuário deseja.”

Ainda segundo Barbosa e da Silva (2010) a “IHC se beneficia de conhecimentos e métodos de outras áreas fora da Computação para conhecer melhor os fenômenos envolvidos no uso de sistemas computacionais interativos” (p. 12). A IHC é, portanto, uma área multidisciplinar, envolvendo pesquisadores das áreas de Ciência da Computação, Psicologia, Ergonomia e Fatores Humanos, Engenharia, *Design*, Semiótica, Etinografia, Sociologia e Linguística, cujo foco é o ser humano interagindo com a tecnologia. Sustentada por essas áreas, a IHC se desenvolveu em relação a vários aspectos: paradigmas de interação, *hardware* e as suas possibilidade de aplicação, compreensão de métodos para o desenvolvimento e avaliação das interfaces de usuário, compreensão do impacto da adoção de um sistema computadorizado, papel social da tecnologia, entre outros. Mas é necessário compreender algo que também evolui: o *hardware*, as áreas de aplicação e, principalmente, os anseios das pessoas em relação ao que podem realizar com um dispositivo computacional. Se antes projetava-se interfaces para um número conhecido de usuários, hoje projeta-se para milhões de pessoas. Se antes projetava-se interfaces para apoiarem o trabalho, hoje a computação é aplicada em vários nichos diferentes: educação, saúde, governo, entretenimento (Shneiderman, 2006).

Assim, muitos desafios ainda persistem e envolvem a variedade de usuários, de hardware e de contextos de uso:

> "... os profissionais da "nova informática" ainda terão muito trabalho pela frente. Terão de dar resposta à difícil questão: Como tornar as tecnologias de informação e comunicação utilizáveis por todos os usuários?
>
> O *design* para usuários frequentes e experientes é bastante difícil, mas o *design* para uma ampla plateia de iniciantes é um desafio muito maior. Passar de uma mala-direta com o cadastro de cem engenheiros de software para cem mil professores escolares e para cem milhões de eleitores registrados demandará inspiração e transpiração." (Shneiderman, 2006, p. 31-32)
>
> "Manter uma ampla gama de *hardware*, *software* e redes não é tarefa fácil. O trabalho é ainda mais desafiador quando se considera as necessidades de conciliar os antigos *hardware* e *software* e apresentar novos dispositivos e ambientes." (Shneiderman, 2006, p. 58)

Preece *et al.* (1994) já apontavam esses problemas: como dar conta da rápida evolução tecnológica? Como garantir que os design ofereçam uma boa IHC ao mesmo tempo que exploram o potencial e funcionalidade da nova tecnologia?

Mas o que seria "a nova tecnologia" nos tempos atuais? Kugler (2008), apoiado em relatórios do grupo Gartner[2], descreve que um dos grandes desafios para a área de Tecnologia da Informação e Comunicação (TIC) nos próximos 25 anos são as interfaces naturais e não-táteis e a tradução automática da fala. Esse desafio, somado à tendência em trocar gradualmente o mouse por interfaces alternativas emergentes que trabalham com reconhecimento facial, movimento e gestos (Hayes, 2008) trazem novos desafios para a área de IHC, conforme já apontavam Preece *et al.* (1994). E o impacto pode ir além do uso do equipamento pelo homem:

> "As novas tecnologias da informática incluiriam *displays* do tamanho de uma parede, aplicados para *palmtops*, minis sensores médicos parecidos

---

[2] O grupo Gartner oferece consultoria e realiza pesquisa na área de tecnologia de informação, oferecendo informações sobre tendências e aconselhamento a clientes desde 1979. Mais informações em http://www.gartner.com.

> com joias e computadores sensíveis ao tato que mudam experiências sensoriais e formas de pensar." (Shneiderman, 2006, p. 21)

Assim, projetar um software cujo tempo de vida percorrerá décadas exige projetar uma interface que também se adeque às evoluções da época e possibilite boas experiências ao usuário com o passar dos anos. Logo é necessário ir além de possibilitar o acesso às ferramentas de TIC usando um dispositivo qualquer. É necessário considerar que o número de periféricos está aumentando ao longo dos anos e possibilitando que usuários interajam com as aplicações de forma diferente da habitual.

É, portanto, pensando na adaptação à evolução que este trabalho de tese se insere. Pode-se pensar que não existem aplicações que perdurem tanto tempo, mas é suficiente pensar nos Sistemas Operacionais para quebrar esse argumento. Os Sistemas Operacionais também são *software* que possuem interfaces de usuário e que se adequam às mudanças de tecnologia e de preferência dos usuários. Por exemplo, o sistema operacional Windows, desde o seu lançamento, sofreu diversas mudanças. Ao longo do tempo "ditou" como deveria ser uma interface gráfica para computadores de mesa (*desktops*) e, inclusive, seu modelo foi migrado para um dispositivo móvel, o *palmtop* e o sistema operacional Windows CE. Por não considerar as particularidades do dispositivo (a pequena tela e o modo de entrada de dados por caneta) e, com o surgimento de outros sistemas operacionais para dispositivos móveis projetados para uso em telas sensíveis ao toque, o padrão adotado foi baseado nesses outros sistemas operacionais. Em 2012, com o lançamento da versão 8, a interface do SO Windows foi aproximada das interfaces dos sistemas operacionais de dispositivos móveis que se tornaram padrão. Motivo: ter um modelo de interface de sistema operacional semelhante para qualquer dispositivo.

Essa não foi a primeira tentativa da Microsoft em produzir um sistema operacional cuja interface apresente características semelhantes entre vários dispositivos. Desde a versão XP, a Microsoft modifica gradualmente o seu sistema operacional para atender aos usuários de Tablet PC, dispositivo equipado com caneta eletrônica. Inicialmente, isso foi feito com pequenas modificações relacionadas ao mapeamento das ações da caneta com ações do mouse (movimentar o cursor, clicar com o botão esquerdo e com o botão direito) e alguns aplicativos que estimulam o uso da caneta (notadamente o Windows Journal, software desenvolvido para tornar o computador um caderno digital). Com o passar dos anos, a Microsoft foi incluindo modificações no seu sistema operacional para se adaptar ao uso da caneta. Este é o caso das caixas de seleção no Windows Explorer (software que possibilita a navegação entre diretórios e a visualização dos arquivos em uma unidade de armazenamento), pois se o usuário está utilizando a caneta com a mão não poderá utilizá-la para pressionar a tecla CTRL ou SHIFT do teclado (teclas utilizadas para indicar ao navegador de arquivos que mais de um arquivo deve ser selecionado). Atualmente o Tablet PC possui, além da interação a caneta, interação por multitoque, o que torna ainda mais desafiadora a adaptação da interface para essa variedade de periféricos.

Pode-se pensar que essa complexidade está relacionada somente aos computadores pessoais. Entretanto, a Internet é utilizada por uma variedade de dispositivos, tornando-se assim um espaço de interação de múltiplas tecnologias. Na rede encontram-se usuários de desktops e *laptops* (em sua maioria), mas o número de usuários de dispositivos móveis de pequeno tamanho está aumentando gradativamente. A Internet, forjada para os computadores convencionais, será acessada cada vez mais por usuários de dispositivos não convencionais, como celulares inteligentes (*smartphones*), *console* de jogos e tablets, sobre os quais não podemos dizer muito a respeito das suas características em relação aos periféricos de interação. Mas a expectativa dos usuários é que as páginas Web deverão ter seu conteúdo acessível por esses dispositivos e a mesma experiência (senão melhor) deverá ser oferecida.

O processo de *design* de interfaces de usuário nesse contexto se torna uma tarefa tão difícil, para não dizer mais difícil, do que a situação atual. As aplicações precisam evoluir nesse novo cenário e se adequarem ao *hardware* que o usuário está utilizando, garantindo a experiência por ele esperada e o surpreendendo de forma positiva. Tomemos por exemplo o Tablet PC, no qual a caneta, um de seus dispositivos de entrada, é utilizada como um dispositivo de apontamento durante a navegação de páginas Web. Ao ser usada apenas como um dispositivo de apontamento não é possível usá-la para desenhar ou escrever um texto manuscrito, uso original da caneta fora do mundo digital. Nesse caso, para obter o potencial da caneta no mundo digital, é necessário resgatar o propósito para o qual ela foi concebida, ou seja, escrever. Mas é importante manter uma consistência para tornar a interface mais fácil de ser utilizada, conforme afirma Shneiderman:

> “O *software* não deveria ser projetado de forma que os usuários pudessem operar o mesmo programa de calendário em um *palmtop*, *laptop* ou computador do tamanho de uma parede?” (Shneiderman, 2006, p. 58)

Refletindo a respeito dessa frase de Shneiderman, embora a consistência de uso deva ser mantida para tornar o uso da aplicação mais fácil, não podemos operar da mesma forma em dispositivos que utilizam periféricos de interação muito diferentes. A caneta deverá ser usada da mesma forma que o teclado? Se assim o afirmarmos, não usaremos a caneta para desenhar, e negaremos a evolução tecnológica, restringindo-nos ao que já foi criado. Mas, certamente, um nível de consistência deverá ser adotado por ser um dos princípios fundamentais de design da interação.

1.2 Conceitos

Ao longo deste estudo adotamos um conjunto de conceitos cujos significados são apresentados a seguir.

Como campo de estudo propomos as tecnologias na educação, mais especificamente, os ambientes de EaD, tipo de *software* que se encaixa na categoria de TICs, os quais, acreditamos, podem se beneficiar da diversidade de dispositivos. Os ambientes de EaD sofreram várias mudanças tecnológicas moldadas pelos avanços tecnológicos e necessidades de seus usuários nas últimas duas décadas, seja em relação à interface de usuário, seja na disponibilização de ferramentas.

Considerando as várias definições para e-Learning (do inglês *eletronic learning,* que significa aprendizado eletrônico) encontradas na literatura, trazemos a de Lima e Capitão (2003) que discutem o significado desse conceito e sua abrangência, afirmando que não existe consenso entre autores.

> "Na prática, o e-Learning é qualquer experiência de aprendizagem distribuída via Internet, Intranet, Extranet, CD ou DVD-ROM, pois o fundamental do e-Learning não é a tecnologia mas sim a forma de ensinar. Embora o e-Learning combine tecnologia e pedagogia, o importante é a experiência vivida pelo aluno na aprendizagem. Além disso, nem todo tipo de conteúdo requer interacção social" (Lima e Capitão, 2003, p. 38)[3]

O e-Learning se refere ao uso da tecnologia eletrônica, mais precisamente, as ferramentas de TIC na educação, independente se é aplicado em cursos formais ou informais; presenciais, semipresenciais ou a distância. Lima e Capitão também esclarecem a diferença entre ensino a distância e e-Learning:

> "... o e-Learning tem uma abrangência um pouco mais restrita que o ensino a distância porque não abrange os cursos por correspondência, as cassetas de áudio e de vídeo, a televisão, e outras tecnologias restritas à 'distância'. Por isso, poderá afirmar que e-Learning é uma forma de ensino a distância, mas ensino a distância não é necessariamente e-Learning." (Lima e Capitão, 2003, p. 37)

Neste trabalho, adotaremos uma definição de e-Learning que representa qualquer tipo de aprendizagem que envolva uma rede física de dados como a Internet, Intranet ou Extranet, usada para a distribuição de conteúdos e para a interação entre os participantes. Nosso propósito é fundamentar o nome das aplicações computacionais que apoiam esse tipo de aprendizagem.

Ambientes de EaD são aplicações computacionais que agregam ferramentas para apoiar a realização de atividades de ensino e aprendizagem por meio da plataforma Web, e são também conhecidos como Ambientes Virtuais de Aprendizagem, Plataformas de e-Learning, Ambientes On-line de Educação[4].

Na literatura costuma-se utilizar o termo LMS (do inglês *Learning Management Systems* que significa Sistemas de Gestão de Aprendizagem) para os sistemas cujo objetivo principal é automatizar aspectos administrativos do processo educacional, como a inscrição de alunos e registro de desempenhos; enquanto o termo LCMS (do inglês *Learning Content Management Systems* que significa Sistemas de Gestão de Conteúdo de Aprendizagem) é reservado para indicar gestão de conteúdos de aprendizagem envolvendo a criação, a catalogação, o armazenamento, a combinação e a sua distribuição (Lima e Capitão, 2003). Nesta tese, trataremos Ambientes de EaD como a composição desses dois sistemas: LMS e LCMS, pois os Ambientes de EaD empregados neste trabalho apresentam características de ambos.

Além das funcionalidades, a usabilidade e a acessibilidade são dois atributos chave para possibilitar um ensino inclusivo ao se adotar ambientes de EaD. Nielsen (1993) define o conceito de usabilidade por meio da decomposição do conceito "aceitabilidade geral de um sistema", dividindo-o em *aceitabilidade social* e *aceitabilidade prática*. Um dos conceitos que

---

[3] Trecho escrito em Português de Portugal.

[4] Embora tenha se discutido anteriormente a diferença entre e-Learning e EaD, preferiu-se adotar o termo "Ambiente de EaD" para representar esse tipo de sistema pois ele é mais empregado na comunidade científica brasileira.

compõe a *aceitabilidade prática* é o "Usefulness", que se refere ao fato de o sistema poder ser usado para atingir um determinado objetivo. Esse conceito é composto pela *utilidade* e *usabilidade*. A usabilidade é a combinação dos elementos: facilidade de aprendizado, eficiência, facilidade de lembrar, probabilidade de o usuário cometer poucos erros e a satisfação do usuário.

Segundo a norma ISO 9241-210 (1999) a usabilidade pode ser compreendida como "a eficácia, eficiência e a satisfação com a qual usuários específicos alcançam metas específicas em ambientes particulares". A eficácia é definida como a acurácia e a completude com a qual usuários específicos podem atingir metas específicas; enquanto a eficiência está relacionada aos recursos utilizados em relação à acurácia e à completude das metas atingidas; a satisfação, por sua vez, está relacionada ao conforto e à aceitação do sistema de trabalho pelos seus usuários e outras pessoas que são afetadas pelo uso de um sistema.

Acreditamos que a importância da usabilidade nos ambientes de EaD se deve ao fato de a interface de usuário não prejudicar as atividades de ensino e aprendizagem que são desenvolvidas por seu intermédio, sejam elas: a leitura de um texto, a elaboração de um relatório, a troca de mensagens eletrônicas.

A acessibilidade dos ambientes de EaD se justifica pelo fato de a educação ser um direito do ser humano. A acessibilidade pode ser compreendida como a "capacidade de acessar" e se beneficiar de algum sistema ou entidade. Segundo a W3C (2005), a acessibilidade da Web significa que pessoas com deficiência podem usar a Web beneficiando também outros usuários.

Com a popularização dos dispositivos móveis, equipamentos dotados de autonomia e leveza, o que possibilita que sejam facilmente transportados, e das redes sem fio, começam a surgir o uso desses dispositivos com propósitos educacionais. Sung *et al.* (2005) definem m-Learning (do inglês *mobile learning,* que significa aprendizagem móvel) como qualquer tipo de educação que acontece quando o aprendiz não tem uma localização geográfica fixa ou pré-determinada; ou aquela que acontece quando o aprendiz aproveita as oportunidades de aprendizado oferecidas pelas tecnologias móveis. Os dispositivos móveis permitem ao aluno estudar em qualquer lugar (*anywhere*) e a qualquer hora (*anytime*).

Pode-se dizer que e-Learning, por se tratar do uso de qualquer tecnologia eletrônica na educação, engloba também o m-Learning, embora Lima e Capitão suponham que "Num futuro próximo, face ao desenvolvimento e proliferação das tecnologias móveis sem fio, o termo e-Learning será provavelmente substituído por m-Learning." (Lima e Capitão, 2003, p. 68).

Ainda assim, pensando na tecnologia dos computadores convencionais e na dos dispositivos móveis, acreditamos em uma complementação de uso dessas tecnologias. Dependendo da tarefa a ser realizada, utilizar um determinado *hardware* será mais rápido, eficiente e/ou oferecerá uma experiência mais gratificante aos usuários. É com esse pensamento que desenvolvemos o presente trabalho, no qual as diversas tecnologias são utilizadas em prol das experiências dos usuários e do planejamento e realização das atividades de ensino e aprendizado.

O termo *Modalidade* será usado para definir o modo como uma entrada do usuário ou uma saída do sistema são expressas. Nigay e Coutaz (1995) definem modalidade como um método de interação que um agente pode usar para alcançar uma meta ressaltando que uma modalidade pode ser especificada em termos gerais como "usando fala" ou em termos mais específicos como "usando microfone". Por exemplo, pode-se usar a modalidade de entrada à caneta, que envolve os periféricos de entrada caneta, tela sensível à caneta, e a mão do usuário utilizada para segurar a caneta.

Neste trabalho exploramos as modalidades de entrada de dados baseada em caneta e toque. Existem várias pesquisas, inclusive no campo educacional, que aplicam o uso de um periférico semelhante à caneta, chamado de *stylus*, para interagir com aplicações computacionais ao invés do tradicional emprego do mouse e teclado para entrada de dados: a computação baseada em caneta (do inglês *Pen-based computing*). Segundo Kurtenbach (2010, p. 14, tradução nossa):

> "… a caneta pode remover a exigência de habilidade para digitar quando se opera um computador. Ao invés de digitar, o usuário simplesmente poderia escrever ou desenhar, e o computador poderia reconhecer e agir após receber esta entrada. O raciocínio foi que, suportando essa expressão "natural", a computação poderia ser acessível a qualquer um, usável em uma variedade ampla de tarefas da entrada de receitas de uma avó, até matemáticos resolvendo problemas com o auxílio de um computador."

As interfaces de usuário para computação baseada em caneta podem ser implementadas de vários modos, destacam-se: (i) usar a caneta como um periférico de entrada para apontamento, no qual as ações da caneta são mapeadas para ações das interfaces gráficas, como mover o ponteiro, acionar elemento de interface, selecionar e arrastar; (ii) utilizar manipulação direta e reconhecimento de escrita manuscrita ou gestos.

Outro periférico que tem se popularizado são as telas sensíveis ao toque, principalmente em dispositivos móveis. Existem várias tecnologias para possibilitar a interação por toque, com características de tempo de resposta diferentes ou possibilidade de entrada de apenas um ou vários toques simultâneos (multitoque).

Refletindo a respeito do uso da caneta e do toque, Kurtenbach (2010, p. 20, tradução nossa) acredita que:

> "Pesquisadores e projetistas têm realizado a observação chave que as interações usando a caneta são distintamente diferentes das interações a toque e podem ser exploradas em diferentes e interessantes formas".

E, reafirmando que o mouse e o teclado são os periféricos mais utilizados para interagir com o computador, Kurtenbach acredita que "a caneta continuará essencial para algumas tarefas, como rascunhar e entrar com uma ideia livre de forma, e será nessas aplicações que ela encontrará sucesso" (2010, p. 20, tradução nossa).

É possível empregar a movimentação da caneta ou do dedo na tela e atribuir-lhe um significado: a interação por gestos. Gestos podem ser gerados a partir de qualquer movimento do corpo, embora comumente relacionados com as mãos e a face. Gestos produzidos a partir de movimentos da caneta são chamados de *pen-based gestures,* enquanto que gestos gerados a partir de movimentos dos dedos são chamados de *touch gestures*, sendo que no caso do gesto envolver mais de um dedo é chamado de *multi touch gesture*.

Esclarecido o contexto do presente trabalho, bem como as definições dos principais conceitos, passamos à apresentação do seu escopo teórico (Capítulo 2). Na sequência, o Capítulo 3 apresenta a tese, propriamente dita, e a forma de condução do trabalho, cujos resultados foram divulgados no meio científico conforme o seu avanço. Assim, os capítulos numerados de 4 a 8 são compostos por artigos publicados em anais de eventos e periódicos científicos. O Capítulo 9 apresenta uma síntese das contribuições resultantes deste trabalho, suas limitações e aponta trabalhos futuros.

Capítulo 2

Escopo Teórico

2.1 Introdução

Como descrito no Capítulo 1, a Interação Humano-Computador (IHC) é uma disciplina que se preocupa com o projeto, a avaliação e a implementação de sistemas computacionais interativos para uso humano, bem como o estudo dos principais fenômenos ao redor deles (ACM SIGCHI, 1992, p. 6). A IHC é uma área multidisciplinar que envolve as disciplinas Psicologia Cognitiva, Psicologia Social e Organizacional, Ergonomia, Ciência da Computação, Inteligência Artificial, Linguística, Filosofia, Sociologia, Antropologia, Engenharia e Design (Preece *et al*., 1994). Um dos conceitos da área é a usabilidade, definida por Nielsen (1993) como a combinação dos elementos facilidade de aprendizado, eficiência, facilidade de lembrar, probabilidade de o usuário cometer poucos erros e a satisfação do usuário.

Este Capítulo apresenta o escopo teórico do trabalho. Na Seção 2.2 é descrito de forma geral o processo de *design* de uma interface de usuário. Na Seção 2.3 apresentamos conceitos da literatura para modalidades de interação e na Seção 2.3 a multimodalidade. A Seção 2.4, que encerra o capítulo, discute a adaptação de interfaces de usuário.

2.2 Design e Avaliação de Interfaces de Usuário

Seguindo conhecimentos da Interação Humano-Computador (Hix e Hartson, 1993; Shneiderman e Plaisant, 2004; Sharp *et al*., 2007; Preece *et al*, 1994; Nielsen, 1993; Rocha e Baranauskas, 2003), ao se projetar uma interface de usuário, deve-se considerar as características de usabilidade para satisfazer as tarefas que o usuário deseja ao interagir com o sistema (Paternò, 1999). Norman e Draper (1986) consideram o modelo mental como uma representação dinâmica sobre qualquer sistema ou objeto, que evolui na mente de um sujeito. O objetivo de um modelo mental é permitir à pessoa entender e antecipar o comportamento de um sistema físico sendo-lhe possível executar mentalmente o seu modelo (Norman e Draper, 1986).

Desenvolver interface de usuário com uma boa usabilidade é uma atividade difícil, exige que esse requisito seja considerado desde o início do processo de desenvolvimento para que não se torne extremamente oneroso. Nesse sentido, a literatura apresenta abordagens para a elaboração de interfaces de usuário que têm em comum a realização repetitiva de duas etapas: o design e a avaliação (Hix e Hartson, 1993; Shneiderman e Plaisant, 2004; Sharp *et al*., 2007; Preece *et al*., 1994; Nielsen, 1993). Teoricamente, na medida em que o designer realiza um número de repetições de design-avaliação, a interface de usuário tende a se aproximar do modelo mental do usuário.

Para facilitar a realização do design e minimizar o número de iterações de design-avaliação, a literatura dispõe de guidelines e heurísticas, isto é, regras escritas textualmente que, se seguidas, auxiliam a melhorar a usabilidade do sistema (Nielsen, 1993; Shneiderman e Plaisant, 2004).

Para realizar a avaliação, a literatura apresenta alguns métodos, que podem envolver somente especialistas em interfaces, os chamados métodos heurísticos, ou envolver os usuários, os denominados métodos empíricos.

Antes de realizar o design da interface de usuário, também é recomendado conhecer para quem se está desenvolvendo o sistema, como o sistema pode auxiliá-los no desempenho de suas atividades e escolher os recursos tecnológicos que serão utilizados para interagir com o sistema. A escolha desses recursos pode se basear nos dispositivos computacionais que serão utilizados pelo usuário, no conhecimento dos usuários a respeito dos dispositivos a serem utilizados e seus recursos de interação.

2.3 Modalidades de Interação

Modalidade é o termo empregado para definir o modo como uma entrada do usuário e uma saída do sistema são expressas. O modo de comunicação se refere ao canal de comunicação usado por duas entidades diferentes que interagem (Alty e McCartney, 1991). Nigay e Coutaz (1995) definem modalidade como um método de interação que um agente pode usar para alcançar uma meta. Segundo os autores uma modalidade pode ser especificada em termos gerais como “usando fala” ou em termos mais específicos como “usando microfone”. Bernsen (2008) e, posteriormente, Bernsen e Dykjær (2010) discutem que o termo modalidade não é muito informativo e não pode ser reduzido somente às modalidades sensoriais humanas, como a visão e a audição: “A modalidade, ou, mais explicitamente, uma modalidade de representação da informação, é uma forma de representar a informação em algum *medium*. Assim, uma modalidade é definida por seu *midium* e sua particular ‘forma’ da representação” (Bernsen, 2008, p. 7, tradução nossa). E fazem uma relação entre modalidade e o *medium*:

> “... humanos usam várias e diferentes modalidades para representar informação no mesmo *midium*. Nós usamos, por exemplo, *medium* gráfico (tudo o que é visível) para representar texto, imagens, expressões faciais, gestos e muito mais. Existem diferentes modalidades representadas no mesmo *medium*” (Bernsen e Dykjær, 2010, p. 71, tradução nossa).

Considerando esta definição, Bernsen (2008) comenta que os sistemas com Interface Gráfica de Usuário (do inglês *Graphical User Interface*, ou simplesmente, GUI) são multimodais: entrada de dados tátil e saída de dados gráfica, e ambos envolvem uma variedade de modalidades individualmente diferentes. Usando esse argumento, Bernsen apresenta uma posição bem diferente da posição de Oviatt e Cohen (2000 *apud* Bernsen, 2008, p. 6) que descrevem que “Sistemas multimodais são radicalmente diferentes do padrão GUI”. Para Bernsen e Dykjær (2010, p. 69, tradução nossa) “Interação baseada em GUI é um paradigma de interação multimodal útil que, por historicamente preceder os demais, é muito bem explorado, e mais familiar à maioria das pessoas, quando comparado a outras interações multimodais". Portanto, esses autores consideram GUI como uma modalidade de entrada e saída, como descreveremos mais adiante.

Bernsen (2008) também apresenta uma taxonomia de modalidades elementares que podem ser usadas para construir interfaces multimodais e possibilitar a interação multimodal. Sua base são as mídias gráficas, acústicas e táteis que são atualmente dominantes na troca de informação com sistemas computacionais interativos. Pode-se citar as modalidades texto digitado, texto manuscrito, gestos e movimento corporal.

Bernsen (2008) afirma que duas modalidades não são equivalentes, pois as modalidades diferem uma da outra em relação aos pontos fortes e fracos de expressividade e também em sua relação aos sistemas perceptual, cognitivo e emocional do ser humano.

Várias modalidades tornaram-se objeto de pesquisa nas últimas décadas, entre elas a fala (reconhecimento e sintetização), a escrita manuscrita (reconhecimento) e a interação com dedos por meio de toques e gestos. Assim, surgiram áreas de pesquisa dentro da Ciência da Computação que visam possibilitar a entrada e a saída de dados por uma modalidade, bem como explorar suas vantagens, como é o caso da Computação Baseada em Caneta, a interação por toque e por gestos, modalidades de estudo deste trabalho.

Computação Baseada em Caneta (do inglês *Pen-based Computing* também conhecido como *Pen Computing*) se refere à interação com uma interface de usuário computacional feita por meio de uma caneta ao invés de periféricos como teclado ou mouse.

A caneta pode ser utilizada como um dispositivo de apontamento. Neste caso o seu uso tem finalidade semelhante à de um mouse (mover um cursor para interagir com elementos na tela). Mas, caneta e mouse podem se diferenciar, caso o dispositivo utilizado pelo usuário seja uma tela sensível à caneta. Assim, o apontamento realizado com a caneta é diretamente sobre a tela (a ponta da caneta deve ser fixada sobre o objeto com o qual se deseja interagir), enquanto o mouse é um dispositivo de apontamento indireto (o ponteiro é movido conforme o usuário move o mouse, mas o mouse não fica sobre a tela). É, então, necessário que o usuário faça um esforço cognitivo para relacionar o movimento da mão com o resultado que provocará no ponteiro. O ponteiro se moverá na direção e com a intensidade que o mouse for se movimentado, mas o espaço percorrido na tela será diferente do espaço realizado com o mouse, embora proporcional.

A caneta também pode ser utilizada para produzir tinta eletrônica (do inglês *electronic ink*), um conjunto de pontos sequenciais gerado pelo rastro da caneta na tela e que representa o rastro de tinta deixado por uma caneta comum. A manipulação de objetos digitais usando recursos que correspondem ao mundo físico de forma e funcionalidade semelhantes é chamada de Manipulação Direta, resumida por Shneiderman e Plaisant (2004) por três princípios básicos: (i) Representação contínua do objeto de interesse; (ii) Ações físicas (clicar, arrastar, etc.) ao invés de sintaxe complexa; (iii) Operações incrementais reversíveis, cujo impacto no objeto de interesse é imediatamente visível. Shneiderman e Plaisant descrevem que a manipulação direta resulta em menor comprometimento de recursos cognitivos.

Isso significa que as interfaces de usuário para computação baseada em caneta podem ser implementadas de várias formas, mas as vantagens do uso da caneta aparecem de forma mais expressiva quando a aplicação é desenvolvida considerando manipulação direta, reconhecimento de escrita ou de gestos realizados com a caneta. Contextos em que papel e caneta estão presentes podem se beneficiar do estilo de interação baseado em caneta, oferecendo melhores experiências de uso. Dentre os contextos de estudo da computação baseada em caneta, destaca-se o educacional. Backon (2006) descreve que o teclado possibilita rápida digitação e possibilita a elaboração de textos estruturados em tópicos, mas destaca a importância da caneta nesta atividade:

> "o uso da caneta eletrônica é importante pois permite registrar e organizar pensamentos, ideias, problemas e soluções de forma clara e precisa, tornando o ensino e o aprendizado mais flexíveis. Além disso, a caneta eletrônica possibilita o registro natural do processo de desenvolvimento do pensamento e raciocínio em todas as suas etapas. Por exemplo, na resolução de cálculos grandes e complexos, é possível que sejam registrados os passos feitos e também desenhados diagramas que explicam a resolução do problema." (Backon, 2006, p. 10, tradução nossa)

Exemplo desse tipo de aplicação foi apresentado por Caceffo, da Rocha e de Azevedo (2011) que propõem uma ferramenta educacional baseada em caneta eletrônica com o propósito de construir um ambiente colaborativo fundamentado no modelo de aprendizado ativo. A ferramenta proposta pelos autores possibilita que professor e alunos compartilhem anotações realizadas por meio da caneta eletrônica em conjunto de slides disponibilizado pelo professor. A ferramenta apresenta funcionalidades para agrupar as várias submissões enviadas pelos alunos, possibilitando que o professor ofereça um *feedback* mais rápido aos alunos. O agrupamento realizado pela ferramenta considera a localização espacial das tintas eletrônicas das respostas dos alunos com a localização espacial das tintas eletrônicas que compõem o gabarito elaborado pelo professor[5].

Uma outra modalidade que se tornou comum é o uso de toque para interagir com a interface de usuário. Trata-se de uma modalidade do tipo tátil que emprega sensores para identificar uma ou mais regiões de uma superfície na qual um ou mais dedos entraram em contato. Telas sensíveis ao toque se tornaram comuns principalmente em dispositivos móveis.

Na manipulação direta em telas sensíveis ao toque, o ato de tocar (do inglês *tap*) realizado por um dedo pode acionar uma ação no objeto desenhado na região em que o toque foi realizado. É possível ter uma sequência contínua de toques sobre uma região, por exemplo, um toque duplo (do inglês *double tap*) que, semelhante ao clicar duplo do mouse, pode ser utilizado para acionar uma funcionalidade do objeto; um toque mantendo o dedo pressionado (do inglês *tap and hold*), semelhante ao clicar com o botão do mouse e deixá-lo pressionado.

Existe também a possibilidade de interagir utilizando mais de um dedo, a interação multitoque (do inglês *multitouch*). O *Natural User Interface Group* (Teiche *et al.,* 2009), em tradução literal, Grupo de Interface de Usuário Natural, descreve que:

> "Multitoque denota um conjunto de técnicas de interação que permite a usuários de computador controlar aplicações gráficas usando vários

[5] O Capítulo 6 apresenta trabalhos relacionados ao emprego da computação baseada em caneta no contexto educacional.

> dedos. Dispositivos multitoque consistem de uma tela sensível ao toque (ex. Monitor de computador, mesa, parede) ou *touchpad*, bem como software para reconhecimento de múltiplos pontos de toque simultâneos, em oposição às telas de toque comuns (ex. touchpad de computador, ATM), que reconhecem somente um ponto de toque por vez.” (Teiche *et al.*, 2009, p. 2, tradução nossa)

A sequência temporal de dados de entrada oriundos de um dispositivo como a tela sensível ao toque podem ser interpretados conjuntamente, originando um gesto. Karpov *et al.* (2008, p. 156, tradução nossa) definem gestos como “um expressivo, significante movimento do corpo, isto é, movimento físico dos dedos, mãos, braços, cabeça, face ou corpo que possui a intenção de passar informação ou interagir com o ambiente”. Gestos podem ser dotados de informação, como os que compõem uma linguagem de sinais ou quando se utiliza um gesto para complementar uma fala. Karpov *et al*. dividem a tecnologia de reconhecimento de gestos em três categorias:

1. Baseado em caneta: envolve o reconhecimento de gestos de periféricos de entrada 2D, como, por exemplo, uma caneta ou uma tela sensível ao toque;

2. Baseado em rastreamento de hardware: envolve o reconhecimento de gestos por meio do rastreamento da posição da cabeça, íris dos olhos (*eye gaze*), mãos ou corpo do usuário com dados coletados por meio de periféricos especiais, como luvas, capacetes ou roupas equipados com sensores como acelerômetros e giroscópio;

3. Baseado em visão: envolve o reconhecimento de gestos por meio de rastreamento da posição da cabeça, íris dos olhos, mãos ou corpo do usuário mas sem exigir periféricos especiais, apenas empregando câmeras que capturam imagens e algoritmos que as interpretam.

Norman e Nielsen (2010), ao analisarem as interfaces gestuais dos dispositivos móveis, alertam que as companhias estão ignorando as convenções estabelecidas e os princípios fundamentais de design da interação. Esses princípios, independentes de tecnologia, são: visibilidade (uma forma de interação adotada, exemplo um gesto, em um objeto deve estar presente em toda a aplicação quando o objeto está na tela), feedback (sistema oferecer uma resposta a ação do usuário e a resposta retornada pelo sistema deve ser a resposta esperada pelos usuários), consistência (em relação aos objetos de interação da aplicação, da plataforma e entre as plataformas), operações não-destrutivas (por isso é importante dispor a funcionalidade “desfazer”), *discoverability* (operações que podem ser descobertas por meio da exploração de menus), escalabilidade (as operações devem funcionar em telas pequenas, médias ou grandes), confiança (as operações devem funcionar e os eventos não podem ocorrer randomicamente). Os pesquisadores enfatizam que as interfaces gestuais devem considerar esses princípios para que haja uma boa usabilidade das aplicações.

Relacionando o mouse, os gestos, a interface GUI e a Interface de Usuário Natural, Teiche *et al.* (2009) destacam que:

> “O mouse e a interface GUI foram umas das principais razões para a grande penetração dos computadores na sociedade. Entretanto a técnica de interação é indireta e baseada em reconhecimento. A Interface de Usuário Natural com as telas multitoque é intuitiva, contextualizada e

> evocativa. A mudança da GUI para interface baseada em Gestos continuará a tornar os computadores parte integrante, mas discreta, do nosso estilo de vida." (Teiche *et al.*, 2009, p. 30, tradução nossa)

2.4 Multimodalidade na Interação Humano-Computador

Os designers não necessitam se limitar a uma única modalidade. Podem escolher várias modalidades que, se utilizadas em conjunto, possibilitam uma maior flexibilidade do sistema, além de outros benefícios. As interfaces com essa característica são chamadas de interfaces multimodais e os sistemas com interfaces multimodais são chamados de sistemas de interação multimodal.

O trabalho seminal sobre interfaces multimodais explora o uso conjunto de fala e gestos por apontamento (Bolt, 1980). Petejan (1984) explora o uso de fala e reconhecimento de movimentos labiais; Neal e Shapiro (1991) exploram o uso de fala e de gestos. Koons, Sparrell e Thorisson (1993) estudam as modalidades fala e rastreamento de olho; Cohen *et al*. (1997) estudam a fala em conjunto com o mouse e, Bangalore e Johnston (2009), as modalidades fala e reconhecimento de gestos registrados por tinta eletrônica. Embora a maioria desses trabalhos envolva a modalidade fala, ela não precisa ser necessariamente a modalidade primária de um sistema multimodal (Oviatt, 1999).

Mayes (1992) define um sistema de interação multimodal como aquele que tem a capacidade de se comunicar com um usuário por meio de diferentes modos de comunicação, utilizando mais do que uma modalidade, além de fornecer e extrair significados automaticamente.

Segundo Oviatt (2003), interfaces multimodais processam dois ou mais modos combinados de entrada de dados do usuário (como a fala, caneta, toque, gestos manuais, rastreamento do olho, movimentos do corpo e da cabeça) de maneira coordenada com a saída de dados do sistema multimídia.

Bernsen (2008) define um sistema de interação multimodal como "um sistema que usa no mínimo duas modalidades diferentes para entrada ou para saída" (p. 9, tradução nossa) e sistemas de interação unimodal como "sistemas que usam a mesma modalidade para entrada e saída" (p. 9, tradução nossa).

Lalanne *et al*. (2009) afirmam que sistemas de interação multimodal ou, simplesmente, sistemas multimodais, permitem aos usuários interagir com computadores por meio de várias modalidades de entrada de dados (ex. fala, gestos, movimentação de olho) e canais de saída (ex. texto, gráficos, som, avatares, fala sintética).

Segundo Bangalore e Johnston (2009), as interfaces multimodais possibilitam que as entradas do usuário e as saída do sistema sejam expressas no(s) modo(s) em que ela(s) melhor se ajusta(m) levando em conta tarefa, as preferências do usuário, o ambiente físico e social da interação.

O fornecimento de múltiplas formas de interação permite melhorar a acessibilidade, uma vez que uma interface multimodal pode ser utilizada por uma pessoa com deficiência por meio dos recursos de interação que consegue manipular, além de tornar o sistema mais flexível e com maior comodidade (Lalanne *et al*., 2009).

Para Bernsen (2008) a principal questão a respeito da multimodalidade é criar alguma coisa nova pois, "quando modalidades são combinadas, obtêm-se novas e emergentes propriedades de representações que não poderiam ser consideradas pelas modalidades individualmente" (p. 25, tradução nossa). Dumas, Lalanne e Oviatt (2009) propõem uma arquitetura genérica para sistemas multimodais, apresentado na Figura 2.1, que facilita a compreensão dos principais componentes de um sistema multimodal.

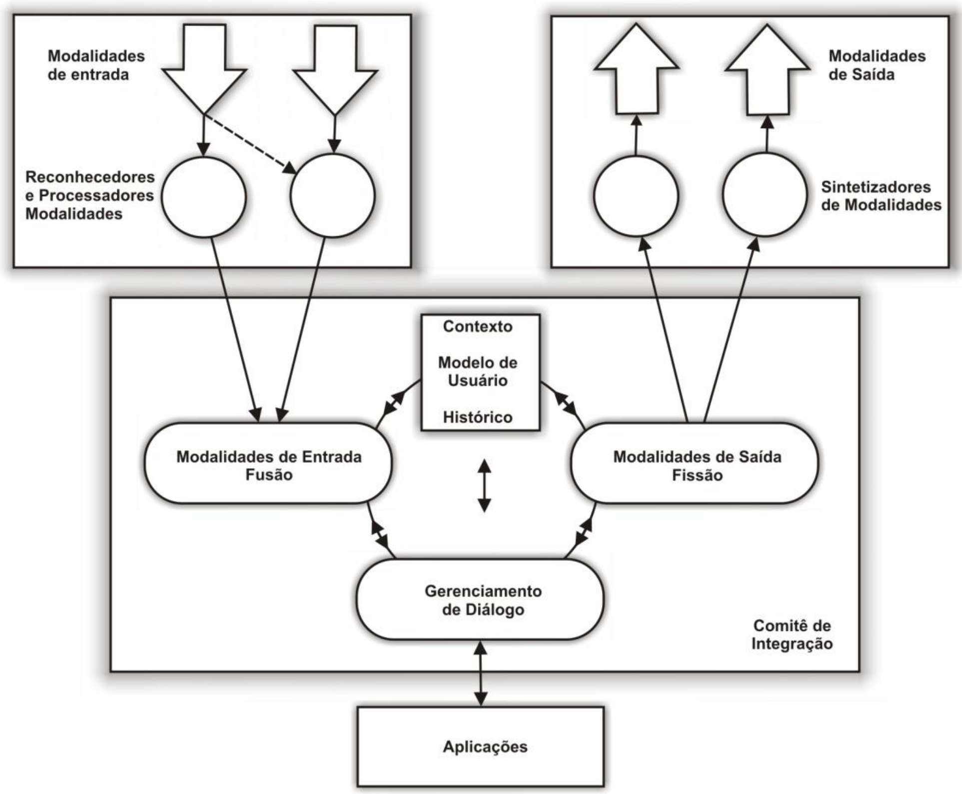


Figura 2.1 – Arquitetura para Sistemas Multimodais traduzida de Dumas, Lalanne e Oviatt (2009, p. 11, tradução nossa).

As modalidades de entrada são percebidas por meio de reconhecedores que enviam dados para as máquinas de fusão. As máquinas de fusão dão um significado comum para os dados de entrada e comunicam ao gerenciador de diálogo. O gerenciador de diálogo é responsável por identificar o estado do diálogo, a transição a realizar, a ação a comunicar para a aplicação e/ou mensagens a retornar por meio dos componentes de fissão. A máquina de fissão retorna a mensagem para o usuário por meio da modalidade mais adequada ou da combinação de modalidades, dependendo do perfil do usuário e do contexto de uso. O contexto de uso, por ser responsável por identificar a localização, o contexto e o perfil do usuário, comunica qualquer modificação do ambiente para os outros três componentes, de modo que eles possam adaptar as suas interpretações.

Os sistemas multimodais devem considerar todas as entradas que foram realizadas pelo usuário para identificar e realizar a ação solicitada. Mecanismos que interpretam os diversos fluxos de entrada, cujo significado pode variar de acordo com o contexto, a tarefa, o usuário e o tempo, são chamados de máquinas de fusão (Dumas, Lalanne e Oviatt, 2009).

Pelo fato de constituir parte central de sistemas multimodais existem na literatura técnicas para máquinas de fusão (Koons, Sparrell e Thorisson, 1993; Cohen *et al.*, 1997; Krahnstoever *et al.*, 2002; Flippo, Krebs e Marsic, 2003; Portillo, García e Carredano, 2006; Bangalore e Johnston, 2009) e abordagens para implementação de máquinas de fusão (Bouchet, Nigay e Ganille, 2004; Mendonça *et al.*, 2009).

Por meio da análise de trabalhos sobre máquinas de fusão, e usando o modelo BRETAM – *Breakthrough, Replication, Empiricism, Theory, Automation and Maturity* (Gaines, 1991 *apud* Lalanne *et al.*, 2009) para determinar a maturidade na área, os autores comentam que as ferramentas de software e modelos de arquitetura para fusão na interação multimodal estão na fase de Maturidade (última fase do modelo). Segundo o modelo BRETAM, quando um conceito está nessa fase significa que as teorias foram assimiladas e são usadas rotineiramente sem questionamentos.

Diante desta constatação, Lalanne *et al.* (2009) comentam que os próximos passos são a criação de procedimentos de avaliação e de engenharia das máquinas de fusão. Os procedimentos de avaliação precisam levar em consideração não somente a interação e a experiência do usuário mas, também, a eficiência e a eficácia das máquinas de fusão, identificando as vantagens e desvantagens dos diferentes mecanismos. Além disso, as máquinas de fusão precisam ser estudadas sob a ótica da engenharia para torná-las mais genéricas, sendo necessário desenvolver ferramentas para:

1) possibilitar ajustamentos finos da fusão pelo designer e pelos usuários finais;

2) simular e configurar rapidamente as máquinas de fusão para um sistema particular.

Seguindo essa direção, os trabalhos de Dumas, Ingold e Lalanne (2009) e Lalanne et *al*. (2009) realizam comparações entre as máquinas de fusão existentes.

Os mecanismos de fissão, responsáveis por determinar a modalidade mais relevante ou o conjunto de modalidades mais relevantes para retornar a mensagem ao usuário, são partes importantes de sistemas multimodais, conforme apresentado na Figura 2.1. Para determinar as modalidades adequadas para retorno da mensagem, os mecanismos de fissão utilizam informações do contexto de uso e do perfil do usuário (Dumas, Lalanne e Oviatt, 2009).

Segundo Foster (2002), os mecanismos de fissão devem realizar três tarefas:

1) seleção e estruturação do conteúdo: o conteúdo a ser incluído na apresentação deve ser selecionado e organizado em uma estrutura;

2) seleção da modalidade: a saída específica que será liberada em cada modalidade disponível deve ser especificada;

3) coordenação da saída: as saídas de cada canal devem ser coordenadas para que as formas de saída sejam apresentadas coerentemente.

Descrevemos até o momento definições de sistemas multimodais e os descrevemos sob uma perspectiva arquitetural. Na próxima Seção descrevemos resultados e observações encontradas na literatura a respeito do uso de sistemas com interação multimodal.

2.4.1 Impactos dos Estudos da Interação Multimodal

No campo de estudo de uso de interfaces multimodais, Dumas, Lalanne e Oviatt (2009) resumem os resultados de algumas pesquisas na área de psicologia cognitiva, que podem ser úteis ao se projetar interfaces multimodais:

• Humanos são capazes de processar parcialmente modalidades independentes devido à ativação de diferentes partes do cérebro humano para processar informação;

• Humanos tendem a produzir o padrão de interação interpessoal durante a interação multimodal com o sistema;

• A performance dos humanos é aumentada quando interagem multimodalmente devido ao funcionamento da percepção, da comunicação e da memória. Segundo Grant e Greenberg (2001 *apud* Dumas, Lalanne e Oviatt, 2009) quando processamos informação visual e auditiva durante uma palestra, somos capazes de extrair uma taxa maior de inteligibilidade léxica;

• A performance humana pode ser melhorada quando interage com duas modalidades que podem ser coprocessadas em partes separadas do cérebro;

• A performance para a realização de duas tarefas simultâneas que requerem o uso de dois domínios perceptuais (ou seja, uma tarefa visual e outra verbal) é próxima da eficiência da performance das tarefas individuais;

• Quando uma pessoa tenta executar duas tarefas simultâneas que usam o mesmo domínio perceptual, a performance é menor do que quando realizadas individualmente.

Chitarro (2010) discute a disposição de multimodalidades em dispositivos móveis considerando os aspectos perceptivo, motor, social, cognitivo considerando a distração e a carga na memória de trabalho. A partir de seus estudos, o autor sugere algumas guidelines:

• Projetar e avaliar a interface priorizando a atenção do usuário, performance, estresse e carga de trabalho cognitivo;

• Projetar a interface para minimizar a atenção e não somente mover a mesma informação para um outro canal de modalidade;

• Enviar informação suficiente pelos canais escolhidos de tal modo que ela possa ser fácil e rapidamente processada pelos sentidos;

• Escolher modalidades apropriadas. Não é suficiente combinar duas modalidades que funcionam bem isoladamente, é necessário testar o efeito da combinação;

• Explorar a sensibilidade ao contexto para minimizar as exigências de atenção e carga cognitiva nos diferentes níveis:

- Mostrar somente a informação relevante para a tarefa;

- Escolher a melhor modalidade ou combinação de modalidade segundo a tarefa e o contexto;

- Mostrar funções que podem ser úteis ou desejadas pelo usuário na situação em que ele se encontra.

Reeves *et al*. (2004) sugerem as seguintes guidelines:

• Os sistemas devem ser projetados para um amplo número de usuários e contextos de uso;

• Questões de privacidade e segurança devem ser consideradas com cuidado;

• Modalidades devem ser integradas de modo compatível com as preferências e capacidades do usuário;

• Os sistemas devem se adaptar facilmente aos diferentes contextos, perfis de usuários e necessidades da aplicação;

• Integrar as modalidades de modo complementar, possibilitando melhorar a robustez do sistema e oferecer ao usuário o controle para selecionar a modalidade.

No contexto de ensino e aprendizagem, Moreno e Mayer (2007) se apoiam na teoria cognitiva-afetiva de aprendizagem com mídias na tentativa de responder à pergunta "O que são ambientes de aprendizagem multimodais interativos e como eles podem ser projetados para promover o aprendizado do estudante?". A partir de suas pesquisas, definem cinco princípios para o design do material instrucional para ambientes de aprendizagem multimodal interativo: atividade guiada, reflexão, feedback, controle e pré-treinamento.

A teoria cognitiva-afetiva de aprendizagem com mídias expande a teoria cognitiva de aprendizagem multimídia para mídias tais como ambientes de aprendizagem com realidade virtual baseados em agentes e baseados em casos que podem apresentar materiais instrucionais que vão além de palavras e figuras (Moreno e Mayer, 2007). Essa teoria se baseia nas seguintes hipóteses sugeridas pela pesquisa cognitiva e motivacional, citadas por Moreno e Mayer (2007):

1. Humanos têm canais separados para processar diferentes modalidades de informação;

2. Somente algumas peças de informação podem ser ativamente processadas em um momento na memória de trabalho para cada canal;

3. Aprendizado significativo ocorre quando o aluno utiliza esforço consciente no processo cognitivo: selecionar, organizar, integrar novas informações com conhecimentos existentes;

4. Memória de longa duração consiste de uma estrutura dinâmica e evolutiva que engloba a memória de experiências passadas e a memória para conhecimentos de domínio geral;

5. Fatores motivacionais medeiam o aprendizado melhorando ou piorando o engajamento cognitivo;

6. Fatores metacognitivos medeiam o aprendizado pela regulação dos processos cognitivos e os seus efeitos;

7. Diferenças no conhecimento anterior e nas habilidades dos alunos podem afetar o quanto podem aprender com a mídia.

Segundo a teoria cognitiva-afetiva de aprendizagem com mídias, a mídia instrucional pode consistir de explicações verbais por meio da combinação de palavras faladas e escritas e representação de conhecimento não-verbal, tais como figuras e sons. Para que o aprendizado significativo ocorra os estudantes precisam, primeiro, observar e selecionar informação verbal e não-verbal relevante para processamento posterior na memória de trabalho. Portanto, Moreno e Mayer (2007) dizem que os ambientes de aprendizagem mais eficazes são os que combinam representação verbal e não-verbal do conhecimento por meio de uma forma de apresentação que mistura as modalidades e que deve-se examinar cuidadosamente a relação entre a demanda cognitiva imposta pelo ambiente de aprendizagem e o desfecho de aprendizado desejado.

2.4.2 Construção de Sistemas de Interação Multimodal

Os benefícios dos sistemas multimodais relatados tornam atraente a implementação desse tipo de sistema, mas o seu desenvolvimento é uma tarefa muito complexa (Dumas, Lalanne e Oviatt, 2009). Para apoiar o desenvolvimento, alguns trabalhos são encontrados na literatura, como os frameworks ICARE (Bouchet, Nigay e Ganille, 2004) e o FAME (Duarte e Carriço, 2006), e as abordagens de Stanciulescu *et al*. (2005), Talarico Neto *et al*. (2009).

Bouchet, Nigay e Ganille (2004) propõem um framework baseado em componentes, chamado ICARE (*Interaction CARE – Complementarity, Assignment, Redundancy and Equivalence*), para o desenvolvimento rápido de interfaces multimodais. O framework define dois tipos de componentes de software reutilizáveis: os componentes elementares que possibilitam a definição de modalidades e os componentes de composição que descrevem mecanismos de fusão de dados. Esses componentes são integrados à arquitetura do sistema para possibilitar a chamada de métodos que realizarão o comando. Os critérios para a invocação dos métodos são a complementaridade ou redundância dos dados e o tempo.

Os componentes de composição são baseados em 4 propriedades oriundas da relação entre os dados das diversas modalidades: Complementaridade, Atribuição, Redundância e Equivalência (*Complementarity, Assignment, Redundancy, Equivalence*), são chamadas de propriedades CARE.

O framework ICARE é um bom exemplo de encapsulamento do mecanismo de fusão, que se beneficia da reusabilidade, manutenibilidade e evolução, devido à adoção da orientação a componentes, impactando na diminuição de custos de desenvolvimento por meio de reuso de componentes. Conforme Bouchet, Nigay e Ganille afirmam, o propósito do framework está no tratamento de dados de entrada, embora afirmem que a abordagem é válida para a saída de dados, mas não demonstrem essa aplicação. Os autores acreditam que a entrada e a saída das modalidades de interação estão fortemente ligadas, ou seja, a escolha de uma modalidade de entrada pode ter impacto na escolha das modalidades de saída, indicando, como trabalhos futuros, o estudo dessa interligação.

Duarte e Carriço (2006) propõem o framework FAME (*Framework for Adaptive Multimodal Environments*), baseado em modelos que propõem uma arquitetura e um conjunto de guidelines para o processo de design de sistemas multimodais. As guidelines envolvem a identificação de variáveis de adaptação relacionadas ao ambiente físico, usuário, plataforma e dispositivos; a elaboração de templates para o modelo de interação; a definição das operações de fusão e fissão; e a definição de regras de adaptação de componentes.

O FAME propõe uma arquitetura que utiliza um conjunto de modelos para descrever atributos e comportamentos do usuário, da plataforma e do dispositivo, do ambiente e da interação e são alterados por módulos definidos na arquitetura para atender às modificações que ocorrem durante a interação.

Horchani *et al*. (2007) propõem um componente que define, a partir de um conjunto de possibilidades de resposta para a requisição do usuário, o conteúdo a ser exibido para o usuário na apresentação multimodal. Chamado de Componente para Estratégia de Diálogo (*Dialog Strategy Component*), esse componente utiliza várias informações como disponibilidade de uma modalidade e preferências do usuário, bem como regras para definir o comportamento do sistema em relação às estratégias de diálogo e às estratégias de apresentação. Assim, esse componente deve receber informação suficiente para gerar a melhor saída multimodal e não se basear somente na resposta que o sistema deve emitir.

Os autores implementaram o Componente Estratégico de Diálogo usando regras especificadas durante o design e definindo mecanismos que selecionam o diálogo e a estratégia de apresentação adequados durante o tempo de execução. Para a criação de regras, Horchani *et al*. propõem uma ferramenta gráfica que pode ser usada por designers e ergonomistas que não requer conhecimentos de programação.

Como trabalhos futuros, os autores propõem a exploração de outras modalidades, como feedback tátil; a checagem da consistência e da completude das regras; a aplicação de mecanismos de aprendizagem de máquina, e ampliação da ferramenta gráfica proposta de tal forma que usuários finais poderem configurar as regras em tempo de execução.

Stanciulescu *et al*. (2005) definem uma abordagem transformacional para o desenvolvimento de interfaces de usuário Web multimodais (MWUIs - *Multimodal Web User Interfaces*). A abordagem parte da elaboração de modelos de tarefa e modelos de conceitos do domínio até a produção da interface de usuário final por meio da realização de transformações nesses modelos. O processo gera grafos de transformações que são baseados em gramáticas de transformações.

Para descrever as informações pertinentes nos modelos, os autores utilizam a UsiXML (*User Interface eXtensible Markup Language*), uma linguagem baseada em XML para descrição de interfaces de usuário (*User Interface Description Language* - UIDL).

O Contexto de Uso (*Context of Use*) descreve todas as entidades que podem influenciar na geração da interface de usuário. Três aspectos são relevantes: Tipo de Usuário; Tipo de Plataforma Computacional; Tipo do Ambiente Físico.

Stanciulescu *et al*. (2005) exemplificam o uso da sua abordagem por meio da elaboração da interface de usuário para um sistema de votação, disponibilizando-o na modalidade gráfica e na modalidade falada.

Segundo os autores, as vantagens de sua abordagem estão no fato de utilizar a UsiXML para expressar todos os elementos, modelos e transformações e possibilitar ao designer a exploração de várias alternativas sem a necessidade de iniciar o processo desde o começo. Como trabalho futuro, os autores propõem estender a UsiXML para suportar outras modalidades, como interação tátil.

Talarico Neto *et al*. (2009) apresentam a Abordagem Web Multimodal (*MultiModal Web Approach* - MMWA), uma abordagem que provê um framework para o design, desenvolvimento e avaliação de projetos com interações multimodais. A abordagem segue em formato de espiral e é composta por 5 atividades que capturam e utilizam o *Design Rationale*:

1) Modelo Comportamental: consiste na elaboração de cenários, restrições e informação do ambiente nos quais as tarefas multimodais serão realizadas, obtendo informação das tarefas do domínio que serão necessárias para mapear as tarefas com as suas representações na interface multimodal;

2) Identificação de Tarefas: identifica as tarefas a serem realizadas pelo usuário na interface multimodal por meio da análise de metas, estados iniciais, atividades e procedimentos envolvidos, os problemas que podem ocorrer, o ambiente de tarefas, os usuários-alvo e sua experiência, além das entradas e saídas multimodais;

3) Representação de Tarefas: objetiva criar uma representação abstrata para as tarefas identificadas;

4) Análise de Soluções: objetiva produzir um documento de design a partir de uma análise de soluções, fornecendo as modalidades que potencialmente podem ser utilizadas para realizar cada tarefa;

5) Avaliação de Usabilidade: objetiva gerar protótipos executáveis nos quais avaliações de usabilidade podem ser realizadas. A avaliação de usabilidade é composta por um mecanismo de avaliação heurística, chamado de MMHE, e um mecanismo para geração e análise automática de log de testes com usuários remotos, chamado de MMLOG.

Segundo Lalanne *et al*. (2009) desenvolver um sistema multimodal é uma tarefa difícil, principalmente pelo fato de existirem poucas ferramentas para uso de designers e implementadores e carência de abordagens de engenharia.

Segundo Bangalore e Johnston (2009), sistemas multimodais devem considerar dados da tarefa, das preferências do usuário, do ambiente físico e social da interação, e dos dispositivos para realizar a fusão ou fissão dos dados.

Como pode ser visto, muitos frameworks são propostos na literatura para a construção de sistemas multimodais, mas a maioria deles é voltada para computadores convencionais (Knudsen e Holone, 2012). Para a Internet, o maior esforço é concentrado pela W3C, que será descrito a seguir.

2.4.3 Construção de Sistemas de Interação Multimodal para a plataforma Web

Para implementar sistemas multimodais na Web é necessário considerar a arquitetura de sistemas multimodais e a arquitetura Web. Usualmente, os sistemas Web têm arquitetura cliente-servidor, em que a máquina cliente, que possui um navegador, é responsável por mostrar as páginas Web para o usuário, receber os dados provenientes dos periféricos de interação e traduzí-los em requisições que são enviadas ao servidor. O servidor é responsável por receber e tratar essas requisições, gerando uma nova página que será retornada como resposta ao cliente.

Gruenstein, McGraw e Badr (2008) apresentam um framework para auxiliar no desenvolvimento de sistemas multimodais para a Web: o toolkit WAMI. O framework que suporta o toolkit define três componentes para o cliente (Núcleo do GUI, Controlador de GUI e Controlador de Áudio) e mais quatro componentes para o lado do servidor (Web Server, Reconhecedor de Fala, Sintetizador de Fala e *Logger*). O usuário interage com o Núcleo GUI, descrito em HTML e JavaScript, e com o Controlador de Áudio, um applet Java que recebe dados da entrada do áudio. Os dados coletados são enviados para serem tratados pelo Reconhecedor de Fala e pelo Web Server.

Comparando com a arquitetura de sistemas multimodais proposta por Dumas, Lalanne e Oviatt (2009), discutida na Seção 2.4, os componentes de Núcleo GUI, Controlador GUI, Controlador de Áudio e Reconhecedor de Fala podem ser classificados como reconhecedores e processadores de dados de entrada. O Sintetizador de Fala pode ser classificado como sintetizador de saída de dados. Os autores comentam que o toolkit WAMI focaliza os modos de fala, teclado e mouse, mas o framework pode ser expandido para incluir outros modos por meio da definição de novos componentes.

O Grupo de Trabalho de Interação Multimodal da W3C (do inglês *Multimodal Interaction Working Group,* sigla MMI-WG) está desenvolvendo desde 2002 um framework padrão para aplicações com interação multimodal. O MMI-WG propôs uma recomendação em 2012 (W3C, 2013) com três componentes para compor o Framework de Execução e mais um chamado Componentes de Modalidade que reconhecem os dados da entrada e executam uma linguagem de marcação. Os Componentes de Modalidade interagem com o componente de Gerenciador de Interação por meio de uma API (do inglês *Application Programming Interface*, cuja tradução literal é Interface de Programação de Aplicativos) de ciclo de vida baseado em evento do MMI. O componente Gerenciador de Interação é um dos componentes do Framework de Execução, responsável por enviar e receber todas as mensagens dos Componentes de Modalidade, e executar consultas e atualizações nos Componentes de Dados, caso necessário. O Componente de Dados contém o modelo de dados públicos para a aplicação multimodal. O Contexto de Entrega é um outro componente do Framework de Execução e é responsável por gravar propriedades estáticas e dinâmicas do dispositivo, condições do ambiente, e preferências do usuário; propriedades que podem ser consultadas e dinamicamente atualizadas (como o componente de gerenciamento de contexto, modelo de usuário e histórico da arquitetura proposta por Dumas, Lalanne e Oviatt (2009)). O componente de Gerenciador de Interação contém, essencialmente, a aplicação multimodal e mantém o fluxo de diálogo, estado corrente, e dados públicos (como o gerenciamento de diálogo da arquitetura proposta por Dumas, Lalanne e Oviatt (2009)).

Um aspecto importante da arquitetura multimodal da W3C é que cada componente de modalidade executa seu próprio documento de linguagem de marcação. Por exemplo, InkML (Chee *et al*., 2011), do inglês *Ink Markup Language* (em tradução literal Linguagem de Marcação de Tinta) serve como formato de dados para representar tinta proveniente da interação com caneta eletrônica ou *stylus*. Segundo Chee *et al*., a marcação possibilita a entrada e o processamento de textos manuscritos, gestos, desenhos, música e outras linguagens notacionais nas aplicações, e provê um formato comum para troca de dados da tinta entre componentes como reconhecedores de textos manuscritos ou gestos.

O Gerenciador de Interação usa a SCXML (*State Chart Extensive Markup Language*), a CCXML (*Voice Browser Call Control*, em tradução literal, Controle de Chamada à Navegação por Voz), a SMIL (*Synchronized Multimedia Integration Language*, em tradução literal, Linguagem de Integração Multimídia Sincronizada) e a VoiceXML (*Voice Extensible Markup Language*, em tradução literal, Linguagem de Marcação Extesível para Voz) para representar as mensagens enviadas entre o framework e os componentes de modalidade.

O SCXML provê um ambiente de execução genérico baseado em máquina de estados para execução que utiliza a CCXML e Tabelas de Estados Harel para prover funcionalidades de controle de chamada em aplicações que envolvem voz (especialmente no uso da VoiceXML, embora não limitados somente a essa representação de dados). Logo a SCMXL define uma máquina de estados, uma sintaxe para manipulação de eventos e um conjunto padronizado de elementos para controle de chamada.

A SMIL serve para autoria sincronizada em tempo real de mídias como áudio, vídeo e texto e é usada em aplicações multimodal para integrar as mídias, por exemplo, *streaming* de áudio e sequenciamento de imagens por tempo. VoiceXML destina-se à autoria de diálogos interativos que envolvem a fala.

As Interfaces de Contexto de Entrega (do inglês *Delivery Context Interface* ou DCI) representam propriedades dinâmicas e estáticas de dispositivos para as aplicações Web. Os registros de dados usam Xforms e DOM (*Document Object Model*) para representar o modelo de dados público.

Os Componentes de Modalidade podem usar XHTML (*eXtensible HyperText Markup Language*), SVG (*Scalable Vector Graphics*), TimedText, VoiceXML e SMIL para representar a interface de usuário e os dados.

No entanto, McCobb (2007) aponta que as linguagem XML da W3C diretamente não suportam autoria de multimodalidade, razão pela qual é preciso padronizar como combinar os arquivos XML que representam várias modalidades em um único documento quando todas as modalidades são renderizadas no cliente.

Zaguia *et al*. (2010) apresentam uma abordagem para desenvolver sistemas multimodais usando máquinas de fusão disponibilizadas por Web Services de modo que o usuário pode escolher as modalidades que melhor se adequam a sua situação ao invés de ter que usar modalidades pré-definidas.

Larson (2002) propõe três questões gerais que devem ser respondidas para averiguar se uma aplicação Web se beneficiará ao empregar um novo modo de entrada de dados: (i) o novo modo necessita agregar valor à aplicação Web; (ii) a aplicação deve aproveitar os pontos fortes do novo modo de interação e anular suas fraquezas; e (iii) os usuários necessitam ter o hardware e o software necessários para utilizar o novo modo.

No desenvolvimento de interfaces multimodais acessados por diversos dispositivos é necessário considerar a adaptação da interface de usuário no processo de fissão.

2.4.4 Desafios Relacionados a Sistemas Multimodais

Embora muito se tenha avançado em relação aos sistemas de interação multimodal desde a década de 80, muitos desafios permanecem. Bernsen e Dybkjær (2010, p. 418) apresentam trabalhos futuros para a comunidade de interação multimodal, das quais cita-se:

- Descrever modalidade, propriedades, aptidão, relacionamentos, extensão da taxonomia para as mídias de sabor e cheiro. Novas combinações de modalidade. Predição de usabilidade;
- Estabelecer linguagens de representação gerais e esquemas de código para todos os níveis de comunicação e ação humana. Disponibilizar mais coleções de dados, gratuitos e de domínio público, sobre comportamento para análise e treinamento de componente;
- Novas formas de usar métodos de usabilidade conhecidos com modalidades e combinações novas e de difícil explicação;
- Decompor de modo mais adequada e em conformidade com princípios de usabilidade em níveis superiores e nos próximos níveis. Todas as modalidades e combinações de modalidades tratadas iguais em diretivas de IHC e propostas de padrões para novas combinações de modalidade e para usabilidade geral, se possível;
- Melhorar a coleta de dados em interação com dispositivo móvel e ubíqua e testes em campo.

Considerando tais desafios, a seguir, discutimos os avanços entre 2010 e 2013.

Williamson, Crossan e Brewster (2011) descrevem que um dos desafios é desenvolver sistemas multimodais considerando questões de usabilidade e realizar testes fora de laboratório com os sistemas criados. Os autores desenvolveram um sistema de leitura de notícias que utiliza a modalidade de gestos com mão como entrada de dados e de fala para a saída de dados. O usuário controla a leitura das notícias por meio de braceletes desenvolvidos com acelerômetros e giroscópios, que possibilita que o usuário faça gestos movimentando o punho. Nesse estudo, percebeu-se que os usuários se sentem mais confortáveis em realizar os gestos nas ruas do que em transportes públicos devido ao tempo de exposição a uma pessoa desconhecida ser menor na rua do que no transporte público. Esse desconforto está relacionado com a aceitabilidade social do sistema.

Outro desafio é desenvolver sistemas multimodais de forma automatizada ou semi-automatizada. Ertl e Kaindl (2012) afirmam que, embora os gerados automáticos de interface de usuário ainda não sejam amplamente empregados, muitas pesquisas foram realizadas possibilitando uma boa compreensão do problema. Mas pouca compreensão se tem em relação à geração automática de interfaces multimodais. Os autores propõem uma abordagem que parte de um modelo de interface independente de modalidade, passando por refinamentos até produzir a interface relacionada com modalidades específicas e características da plataforma. Os autores desenvolveram algumas ferramentas que auxiliam a geração semiautomática de interfaces multimodais considerando as modalidades de fala e de ajuste de movimento (aceleração/desaceleração) de um robô, chamada de *Motion Cue*.

A interação entre humanos e robôs é outro desafio de pesquisa na qual estão sendo utilizadas técnicas de interação multimodal. Bodgdan *et al.* (2011) apresentam um estudo de aplicação da interação multimodal em robôs com interação social, no qual projetam e avaliam movimentos do corpo de um robô com suporte a comunicação, analisando a ideia de utilizar o ajuste da velocidade e orientação do robô. Para a realização de testes, escolheram um cenário onde um carrinho de compras robotizado oferece produtos por meio de áudio e GUI para o usuário enquanto ambos se movimentam em um ambiente semelhante a um supermercado, e demonstram que se o robô diminuiu a velocidade de movimento perante uma oferta, os usuários são mais propensos a reagir e levar o produto.

O uso conjunto de modalidades tem sido objetivo de pesquisas nos últimos anos. Um dos casos é o uso conjunto do reconhecimento de gestos por toque e reconhecimento de voz, chamado de reconhecimento voz-tátil (do inglês *haptic voice recognition*) (Sim *et al.*, 2012). Os autores descrevem que o "reconhecimento voz-tátil é uma interface multimodal projetada para entrada de texto eficiente e robusta nos modernos dispositivos móveis". Nessa combinação de modalidades, o computador irá reconhecer a fala do usuário, normalmente convertendo-a em texto digitado, com o apoio de informação obtida por meio do toque em teclas de um teclado virtual ou por reconhecimento de letras desenhadas por gestos. Por exemplo, Sim *et al.* (2012) descrevem que o usuário pode falar a frase e digitar a primeira letra de cada palavra no teclado virtual, possibilitando assim que o algoritmo de reconhecimento de fala use as letras recebidas/reconhecidas como informação pois oferecem indícios do início de cada palavra e também da palavra a ser reconhecida. Resultados apontados pelos autores mostram que essa técnica possui vantagens nos critérios de eficiência, acurácia e robustez em relação a técnica de reconhecimento automático de fala (do inglês *Automatic Speech Recognition*, conhecida pela sigla ASR), amplamente empregada quando se trata de reconhecimento de fala.

O uso conjunto de toque e caneta tem sido explorado por vários trabalhos (Brandl *et al.*, 2008; Hinckley *et al.*, 2010; Hamilton, Kerne e Robbins, 2012), e sistemas que empregam essas modalidades possibilitam a interação unimodal com caneta, unimodal com toque e multimodal com entrada da caneta e toque. Hinckley *et al.* descrevem que nos anos 90 os dispositivos suportavam somente o uso de toque ou de caneta, e a tecnologia não diferenciava o contato feito pelo dedo ou pela caneta. Uma nova geração de dispositivos está emergindo que são sensíveis a multitoque enquanto distingue caneta do toque, possibilitando utilizar a caneta **e** o toque ao invés de somente a caneta **ou** toque. As interações por caneta e por toque são realizadas por meio das mãos do usuário, assim estudos de combinação dessas duas modalidades envolvem: (i) as diferenças e semelhanças entre a caneta e o toque; (ii) o uso da caneta com a mesma mão que se realizará o toque (interação bimanual), ou a realização do toque com a mão que não está segurando a caneta (interação unimanual); (iii) as preferências do usuário frente as alternativas possíveis.

Hinckley *et al.* (2010) iniciam seu trabalho por meio de observações sobre o uso de papel e cadernos físicos pelas pessoas, e descrevem:

> "há um rico conjunto de comportamentos profundamente enraizados que as pessoas exibem quando se trabalha com caneta, papel, recortes, páginas e livros, tanto com *stylus* quanto por manipulação somente com caneta, que diferenciam os papéis de caneta e toque" (Hinckley *et al.*, 2010, p. 27, tradução nossa).

Segundo Hamilton, Kerne e Robbins (2012), os projetos de interação a caneta em conjunto com toque encontrados atualmente são baseados na metáfora de mesa de trabalho baseada em papel (como o trabalho de Hinckley *et al*. 2010): a mão não dominante é usada para manipular artefatos, enquanto que a mão dominante realiza interações precisas usando a caneta.

Hamilton, Kerne e Robbins (2012) descrevem que as modalidades de interação baseada em caneta, multitoque, e mouse + teclado são fundamentalmente diferentes. Logo comparações diretas de performance são difíceis de serem realizadas, e poucos estudos quantitativos foram feitos.

Estudos recentes estão utilizando sinais captados por sensores cerebrais como um modo de interação em sistemas multimodais. Nesse contexto existe uma relação entre as Interfaces Humano-Computador e as Interfaces Computador-Cérebro. Poel *et al*. (2012) descrevem que esta alternativa de uso das Interfaces Computador-Cérebro é um desafio, mas que pode enriquecer as tradicionais formas de interação. Segundo os autores, a maioria das Interfaces Computador-Cérebro não provêm uma confiável e eficiente forma para controlar entrada de dados e são difíceis de aprender e usar comparada a outros modos disponíveis. Portanto propõem a mudança do uso conceitual das Interfaces Computador-Cérebro como um ator (controle de entrada) em Interfaces Computador-Cérebro como um sensor inteligente. Os sinais obtidos por meio dos sensores podem ser utilizados para inferir o estado do usuário e adaptar a interface de um sistema.

Pesquisas estão sendo utilizadas para averiguar o potencial da multimodalidade no contexto da acessibilidade. Turunen *et al*. (2005) realizaram um estudo envolvendo três abordagens investigando o uso de um sistema multimodal com modalidade de fala desenvolvido para dispositivos móveis com o propósito de prover informação a respeito do transporte público. Knudsen e Holone (2012) apresentam um projeto com o propósito de implementar um sistema multimodal que possibilita a usuários cegos ou com visão deficiente navegar na Internet por meio de dispositivos móveis. Embora promissores e com resultados parciais positivos, os trabalhos descritos encontram-se em andamento.

Outro objeto de pesquisa que envolve a multimodalidade são as interfaces adaptativas em dispositivos móveis, na qual a modalidade de entrada é automaticamente adaptada baseada em informações contextuais como condição física do ambiente (luz e barulho) e localização (parque, escola, local público aberto, casa). David *et al*. (2011) propõem um framework para construir aplicações para dispositivos móveis com essa característica. Dumas, Solórzano e Signer (2013) propõem oito diretivas para serem aplicadas no projeto de sistemas adaptativos dependendo de contexto com interação multimodal para dispositivos móveis. Entre as diretivas propostas pode-se citar: a realização de uma análise do contexto, definição de modalidades adequadas para cada contexto, seleção de tarefas que oferecerão comportamento multimodal e definição de modalidades equivalentes para tarefas multimodais.

2.5 Adaptação da Interface de Usuário

Existem estudos dentro da área de IHC que se preocupam com a adaptação da interface de usuário conforme um conjunto de características do dispositivo utilizado, como por exemplo, o tamanho da tela ou se a tela é colorida ou preta e branca, que não estão diretamente vinculados a pesquisas de interfaces multimodais. Muitos dos trabalhos desenvolvidos nesta área de pesquisa se baseiam no desenvolvimento de algoritmos para modificação automática de páginas Web, possibilitando o acesso dessas páginas em dispositivos móveis. Exemplo temos as transformações Migração Direta, Migração Linear e Visão Geral propostas por MacKay, Watters e Duffy (2004).

Bickmore e Schilit (1997) apresentam uma proposta para adaptação da interface de páginas Web para serem exibidas apropriadamente em dispositivos de telas pequenas. A adaptação é realizada em um servidor Proxy, que ao receber uma requisição HTTP, processa a página desejada e realiza um conjunto de transformações na estrutura da página para alcançar uma melhor visualização. As transformações seguem um algoritmo baseado em heurísticas, que foram elaboradas por meio de uma análise de modificações feitas em um conjunto de páginas disponíveis na Web e que foram adaptadas para melhor visualização em dispositivos de telas pequenas.

O trabalho de Bickmore e Schilit (1997) se destaca na análise da estrutura da página Web para melhorar a visualização, mas o trabalho não considera outras mídias como áudio e vídeo, as características do usuário, do dispositivo e da rede para acessar a página Web. Bickmore e Schilit (1997) não discutem a usabilidade do sistema, limitando-se a tornar acessível a Web para dispositivos de tela pequena.

Nesse contexto, Zhang (2000) apresenta um framework para sistemas de entrega de conteúdo adaptativo em ambientes heterogêneos (ambientes que usuários possuem acesso ao sistema por meio de multi-dispositivos) cujo objetivo principal é melhorar a acessibilidade do conteúdo considerando as mudanças da rede e as condições dos dispositivos para a visualização do conteúdo. O framework é composto por módulos responsáveis por:

i) adaptar o conteúdo, seja texto, imagem, áudio ou vídeo, chamado Módulo de Adaptação;

ii) descobrir as características do equipamento do usuário, da rede e das preferências do usuário, chamado Módulo de Monitoramento de Usuário/Cliente/Rede;

iii) máquina de decisões, que analisa os dados obtidos pelo Módulo de monitoramento do Usuário/Cliente/Rede e solicita ao Módulo de Adaptação a entrega do conteúdo adaptado.

Coninx *et al*. (2003) apresentam uma proposta de framework para gerar interfaces de usuário para sistemas embarcados e dispositivos computacionais móveis que facilite o trabalho dos projetistas e dos implementadores. A abordagem é dividida em passos para:

i) especificar a tarefa e realizar anotações nas tarefas para alimentar o framework;

ii) especificar a interação da interface de usuário com a implementação das regras de negócio;

iii) gerar a interface utilizando as restrições do dispositivo, as restrições de layout e os componentes disponíveis para o dispositivo.

O objetivo de Coninx *et al*. é a elaboração de interfaces de usuário fáceis de migrar, reusáveis e consistentes, pois possibilita que uma interface de usuário seja automaticamente adaptada para novos dispositivos, oferecendo as mesmas funcionalidades, sem a necessidade de reprojeto. Coninx *et al*. comentam que mais estudos são necessários para considerar o requisito de usabilidade nas interfaces geradas.

Uma abordagem com ideia similar à de Coninx *et al*. é apresentada por Berti *et al*. (2004), que propõem o TERESA, um ambiente baseado em transformações para projetar e desenvolver interfaces para multi-dispositivos. Esse ambiente é baseado em modelagem de tarefas, contando com os seguintes passos:

i) modelagem de tarefas em alto nível do sistema multi-contexto;

ii) desenvolvimento do modelo de tarefas do sistema para as diferentes plataformas consideradas;

iii) obtenção de uma interface de usuário abstrata;

iv) geração da interface de usuário.

Pyla *et al*. (2006) discutem a consistência do sistema e da migração sem costuras da tarefa na realização de atividades usando diversos dispositivos. Ao enfatizar a consistência do sistema (seja no nível de interface de usuário, no nível de tarefa, ou simplesmente no nível de dados), Pyla *et al*. acreditam que é criado um espaço vazio na execução da tarefa, chamando esse espaço de descontinuidade da tarefa. E para minimizar o impacto dessa descontinuidade, as tarefas devem ser distribuídas de forma complementar nos dispositivos a serem utilizados. A consistência da interface de usuário somente deve ser garantida após distribuir adequadamente as tarefas.

Embora Pyla *et al*. discutam a necessidade da migração sem costuras das tarefas ao utilizar multi-dispositivos, criando o conceito de Interfaces de Usuário Contínua (*Continuous User Interfaces*), eles limitam-se a apresentar algumas funcionalidades essenciais para que seja possível a migração sem costura das tarefas (como a necessidade de alguns ou de todos os dados associados com o sistema serem transferidos de um dispositivo para outro) e não apresentam uma arquitetura para sistemas que respeitem a migração sem costuras das tarefas.

Oliveira e Rocha (2007) e Oliveira (2008) propõem prioridades de consistência no projeto de interfaces de usuário multi-dispositivos objetivando melhorar a usabilidade e a experiência do usuário quando realizar tarefas similares em diferentes dispositivos. Baseado em um modelo de atualização do modelo mental do usuário, Oliveira e Rocha propõem 3 prioridades:

i) Percepção de Tarefa, define que os mecanismos de controle para executar uma tarefa e a disposição desses controles devem ser iguais;

ii) Execução da Tarefa, define que as mesmas ações devem ser realizadas para executar uma tarefa;

iii) Personalização da Tarefa, que define possibilitar a mudança da Percepção da Tarefa e da Execução da Tarefa de acordo com as preferências do usuário.

Comentando em relação ao uso das prioridades de consistência em interfaces de sistemas para dispositivos que variam de estilo de interação (por exemplo, Máquinas ATM, Tablet PCs e telefones), Oliveira e Rocha descrevem que não há possibilidade de perceber e realizar a tarefa do mesmo modo. E que neste caso o objetivo é prover a mesma percepção e consistência no modo de execução da tarefa de uma perspectiva lógica, seja por meio de palavras digitadas, escritas ou ditas.

Em relação a navegação de páginas Web em dispositivos móveis equipados com tela sensível ao toque, Maurer *et al*. (2010, p. 742, tradução nossa) realizaram alguns estudos de usabilidade e descrevem que “mais e mais pessoas preferem usar o conteúdo original ao invés da versão móvel, especialmente para usuários de dispositivos móveis da nova geração como os celulares iPhone e Android”. Outro resultado desse trabalho aponta que os usuários preferem usar o site Web padrão ao invés de uma versão ajustada para os dispositivos móveis. Mas Schmiedl *et al*. (2009) apresentam uma opinião diferente: em suas pesquisas eles concluem que a maioria dos usuários prefere a versão ajustada. Kaikkonen (2008) acredita que as versões padrão e ajustada a dispositivos móveis de um site Web são ambos usados mas com razões diferentes.

Capítulo 3

Objetivo e Metodologia de desenvolvimento

3.1 Objetivo da Tese

No Capítulo 1 foram apresentadas a problemática e a motivação da realização deste trabalho: com o avanço da tecnologia, diversos dispositivos computacionais estão sendo construídos, muitos com poder de processamento suficiente para acessar e renderizar páginas Web e, consequentemente, acessar os ambientes de EaD. Diferentes dispositivos possibilitam diferentes experiências de interação e funcionalidades e é necessário que a interface de usuário seja projetada para aproveitar melhor o dispositivo e seus periféricos. Focalizar o desenvolvimento de interface para um dispositivo possibilita criar uma interface arrojada que aproveita as oportunidades de interação oferecidas pelo dispositivo, mas torna a tarefa dos *designers* de sistemas demasiadamente trabalhosa, necessitando criar diversas interfaces e manter um modelo conceitual consistente entre eles. Outro fator é o grande aumento do código, pois cada interface necessita ser transformada em código, tornando onerosa a manutenção das várias interfaces construídas, tanto na identificação de problemas quanto no acréscimo de novas funcionalidades.

No Capítulo 2 foram apresentados conceitos que embasam teoricamente este trabalho, envolvendo: o design de interfaces de usuário e teorias do desenvolvimento e avaliação de interfaces de usuário; o emprego de modalidades de interação e o desenvolvimento de aplicações que exploram algumas modalidades; o emprego da multimodalidade na interação e o impacto de sua adoção.

Frente à motivação de permitir que ambientes de EaD sejam acessados por uma amplitude maior de dispositivos, e aproveitar melhor as oportunidades de interação oferecidas pelos diferentes periféricos de entrada e saída, propõe-se pensar na utilização do conceito de multimodalidade na interface de usuário para que, então, seja possível que os ambientes de EaD sejam acessíveis por essa variedade de dispositivos e. também, se beneficiem das características dos periféricos de interação que compõem o dispositivo. Devido à quantidade de dispositivos computacionais existentes é difícil projetar uma interface para cada dispositivo necessitando, assim, projetar um número limitado de interfaces, ou uma única quando possível que permita a interação do usuário com boa usabilidade. Propomos a adoção da multimodalidade para possibilitar o acesso por diversos dispositivos.

Para desenvolver esta tese, este trabalho estudou o impacto das mudanças de modalidades em ambientes de EaD considerando a entrada de toque e a caneta, e estudou as possibilidades de uma interação multimodal nos ambientes de EaD. Percebeu-se que o impacto da multimodalidade pode ir além de possibilitar acesso por diversos periféricos de interação, podendo emergir novas funcionalidades nas ferramentas que apoiam práticas de ensino e aprendizagem que eram difíceis de serem realizadas nos ambientes de EaD. Ou seja, o impacto da multimodalidade pode ir além da interação entre usuário e computador, e possibilitar novas tarefas e práticas que possibilitem que determinados conteúdos sejam produzidos nos ambientes de EaD. Para implementar a interação multimodal em um ambiente de EaD é necessário uma arquitetura, proposta desta tese.

A abordagem de pesquisa adotada, além de uma revisão da literatura sobre multimodalidade, que foi apresentada no Capítulo 2, incluiu três principais grupos de atividades:

(i) Averiguar os problemas de interação que ocorrem ao acessar ambientes de EaD por modalidades que não foram consideradas durante a especificação de sua interface;
(ii) Desenvolvimento de uma ferramenta, a partir de uma ferramenta existente no ambiente, que aproveite os benefícios das modalidades de entrada a caneta e multitoque;
(iii) Proposta de uma arquitetura para ambientes de EaD com interação multimodal.

O item (i) foi feito em dois ambientes de EaD, o TelEduc e o Ae, enquanto que para o item (ii) foi a implementação de uma ferramenta no ambiente Ae, e o item (iii) propõe uma arquitetura para ambientes de EaD com interação multimodal a partir do ambiente Ae. Na próxima Seção serão detalhados cada uma dos grupos de atividades.

3.2 Condução do Trabalho

Esta tese é composta por 9 capítulos, sendo que os capítulos de 4 a 8 são compostos por artigos publicados em anais de conferências ou em revistas periódicas (*journals*). O Capítulo 9 apresenta as conclusões deste trabalho, uma reflexão a respeito dele e trabalhos futuros. Os capítulos demonstram os encaminhamentos que foram dados para desenvolver esta pesquisa preocupando-se sempre em apresentar os resultados obtidos e, via feedback da comunidade, conduzir os próximos passos.

Na sequência, são apresentadas as referências e sínteses dos artigos publicados ressaltando-se as suas contribuições:

> [Capítulo 4] DA SILVA, A.C.; DA ROCHA, H.V. (2012) Experiences in use Tablet PC to Support Student´s Activities: Five Years of an Exploratory Study. In: INTERNATIONAL CONFERENCE ON MOBILE, HYBRID, AND ON-LINE LEARNING, 4, 2012, Valência, Espanha. *Proceedings*... IARIA, 2012. p. 38-43.

Ponto inicial do desenvolvimento deste trabalho, descreve o uso do Tablet PC em atividades desempenhadas por um aluno dentro e fora da sala de aula. Dentre as atividades cita-se (i) anotações de aulas, (ii) leitura de textos destacando partes relevantes e (iii) resolução de exercícios. Três aplicativos são utilizados ao longo do estudo para apoiar as atividades: Adobe Acrobat Professional, Windows Journal e Jarnal. Os dois últimos aplicativos foram desenvolvidos especificamente para Tablet PC mas, conforme relata o artigo, contêm problemas de usabilidade que impactam na sua utilização.

Um dos problemas apontados no trabalho é a dificuldade na colaboração entre usuários com diferentes dispositivos. Troca de conteúdos, sejam anotações de aula ou resolução de exercícios, é uma atividade corriqueira entre estudantes. No estudo, um usuário utilizava o Tablet PC, sendo que os demais utilizavam computadores convencionais. O trabalho relata a dificuldade de troca de documentos entre eles, pois é necessário adotar um padrão comum aceito entre os dois dispositivos. Essa dificuldade pode limitar as ações realizadas pelos usuários ou adicionar tarefas para que a incompatibilidade de dados entre os dispositivos seja contornada. Essa complexidade a mais é indesejada em contextos onde podem haver multi-dispositivos, como é o caso dos ambientes de EaD.

Logo, uma certa transparência de tecnologia é necessária, para que a troca de dados entre os diferentes dispositivos seja realizada sem impactar na usabilidade. Mas os usuários precisam saber os periféricos disponíveis e quais as funcionalidades disponíveis para o dispositivo utilizado e, assim, poderem elaborar um modelo mental de utilização das ferramentas do ambiente de EaD.

[Capítulo 5] DA SILVA, A.C.; FREIRE, M.P.F; DA ROCHA, H.V. (2012) Identifying Cross-Platform and Cross-Modality Interaction Problems in e-Learning Environments. In: INTERNATIONAL CONFERENCE ON ADVANCES IN COMPUTER-HUMAN INTERACTIONS, 6, 2013, Nice, França. *Proceedings*... IARIA, 2013. p. 483-487.

DA SILVA, A.C.; FREIRE, M.F.P; DE ARRUDA, A.V.P.; DA ROCHA, H.V. (2013) Interaction Problems Accessing e-Learning Environments in Multi-touch Mobile Devices: a case study in TelEduc. In: IADIS MULTI CONFERENCE ON COMPUTER SCIENCE AND INFORMATION SYSTEMS - IADIS INTERNATIONAL CONFERENCE ON E-LEARNING, 2013, Praga, República Tcheca. *Proceedings*... IADIS Press, 2013. p. 199-206.

O Capítulo 5 é composto por dois artigos que tratam a identificação de problemas oriundos da mudança de modalidade. O primeiro artigo apresenta alguns problemas identificados por um especialista em interação ao analisar a interface de usuário do ambiente virtual de aprendizagem TelEduc acessando por meio de dois dispositivos com tela sensível ao toque: um celular inteligente e um Tablet PC. Ao analisar os problemas encontrados, percebeu-se que eles podem ser categorizados em (i) problemas de interação sem relação com a mudança de modalidade ou plataforma; (ii) problemas de interação relacionados com a mudança de plataforma; (iii) problemas de interação relacionados com a mudança de modalidade.

O segundo artigo possui o mesmo objetivo, entretanto utiliza testes de usuário para identificar problemas de interação ao acessar o ambiente TelEduc por meio de celulares inteligentes com tela sensível ao toque. Os problemas identificados são encaixados nas categorias definidas no artigo anterior. Para os problemas encontrados são apresentadas soluções, algumas simples, mas que possibilitam uma melhor interação com o ambiente de EaD no dispositivo utilizado.

Neste trabalho percebe-se a relação de dependência entre os periféricos de interação, o sistema operacional e os navegadores disponíveis no dispositivo. Por exemplo, os navegadores atuais e sistemas operacionais dos Tablet PCs realizam um simples mapeamento entre ações com a caneta em ações do mouse, com exceção de pequenos gestos navegacionais.

> [Capítulo 6] DA SILVA, A.C.; DA ROCHA, H.V. (2013) InkBlog: A Pen-Based Blog Tool for e-Learning Environments. *Issues in Informing Science and Information Technology*, vol. 10, p. 121-135, 2013.
>
> DA SILVA, A.C.; DA ROCHA, H.V. (2014) Improving e-Learning Environments for Pen and Multi-touch Based Interaction: A study case on blog tools and mobile devices. In: INTERNATIONAL CONFERENCE ON MOBILE, HYBRID, AND ON-LINE LEARNING, 6, 2014, Barcelona, Espanha. *Proceedings…* IARIA, 2014. p. 44-50.

O Capítulo 6, formado por dois artigos, apresenta um trabalho pioneiro no contexto da plataforma Web, ambientes de EaD e interação baseada em caneta: o InkBlog, uma ferramenta de blog que possibilita aos usuários postarem mensagens manuscritas ou desenhos elaborados por meio de uma caneta. Até o momento, as ferramentas que exploram o conceito de Pen-Based Computing eram especificamente destinadas ao dispositivo baseado em caneta e fora da plataforma Web. As funcionalidades foram extraídas na experiência obtida ao realizar o trabalho descrito no Capítulo 4, gerando uma primeira versão da ferramenta. Após estudos sobre o uso da ferramenta InkBlog em dispositivos móveis com interação a toque e a caneta, uma nova versão foi gerada, suportando também a interação multitoque.

Neste trabalho percebe-se que a dependência entre periféricos de interação, sistema operacional e navegador, pode limitar o aproveitamento dos recursos de interação. A arquitetura da Web, que embasa os navegadores e a forma de navegação nas páginas Web, impõe ao desenvolvedor restrições. É o caso da geração da tinta eletrônica nos ambientes de EaD, cuja entrada da caneta necessitou ser tratada como entrada de dados do mouse (por exemplo, eventos *mouseover* e *mousedown* da linguagem JavaScript que informam a página Web se o ponteiro está sobre um elemento ou se o botão do mouse foi pressionado).

Essa limitação tecnológica irá restringir a implementação dos ambientes de EaD com multimodalidade, pois os eventos que o navegador trata são mapeados para eventos de mouse, como será explicado logo após a síntese do Capítulo 7.

> [Capítulo 7] DA SILVA, A.C.; DA ROCHA, H.V. (2013) e-Learning Environment with Multimodal Interaction: A proposal to improve the usability, accessibility and learnability of e-learning environments. In: INTERNATIONAL CONFERENCE ON ADVANCES IN COMPUTER-HUMAN INTERACTIONS, 6, 2013, Nice, França. *Proceedings…* IARIA, 2013. p. 483-487.
>
> DA SILVA, A.C.; DA ROCHA, H.V. (2013) Architecture for e-Learning Environments with Multimodal User Interface. In: IADIS MULTI CONFERENCE ON COMPUTER SCIENCE AND INFORMATION SYSTEMS – IADIS INTERNATIONAL CONFERENCE ON INTERFACES AND HUMAN COMPUTER

INTERACTION AND GAME AND ENTERTAINMENT TECHNOLOGIES, 2013, Praga, República Tcheca. *Proceedings*... IADIS Press, 2013. p. 165-172.

Adotar multimodalidade em um software implica a adoção de componentes que tratam a multimodalidade mas, também, na criação de interfaces entre os componentes da multimodalidade e do software. No caso dos ambientes de EaD, que possuem um conjunto amplo de ferramentas, de funcionalidades dispostas nessas ferramentas, de papéis e permissões, é necessário que esses componentes se relacionem. O Capítulo 7 é composto por dois artigos, o primeiro, um artigo resumido apresentado em um workshop de teses, e o segundo, continuação do trabalho, apresenta um maior número de relacionamentos e cenários que exemplificam a necessidade do interfaceamento dos componentes.

Entretanto, multimodalidade de entrada envolve receber dados oriundos de vários periféricos de entrada, podendo esses dados serem coletados ao mesmo tempo. A implementação atual dos sistemas operacionais e navegadores torna difícil a coleta de dados a partir de diversos periféricos, por exemplo, o uso da caneta eletrônica e os dedos para entrar com os dados. Como, para essas duas modalidades, o sistema operacional “traduz” para o navegador (e, consequentemente, para a aplicação Web) como um conjunto de ações realizadas pelo mouse sobre um ponteiro (a saber, único), não é possível ao usuário utilizar essas duas modalidades ao mesmo tempo.

Essa relação entre sistema operacional e navegadores também impossibilita a criação de gestos específicos da aplicação Web, conforme aponta o Capítulo 8.

[Capítulo 8] DA SILVA, A.C.; DA ROCHA, H.V. (2013). Challenges for e-Learning Environments in m-Learning Contexts: A survey about the hardware, software and educational dimensions. In: INTERNATIONAL CONFERENCE ON MOBILE UBIQUITOUS COMPUTING, SYSTEMS, SERVICES AND TECHNOLOGIES, 7, 2013, Porto, Portugal. *Proceedings*... IARIA, 2013. p. 20-25.

Baseado nos trabalhos anteriores, este artigo discute a existência de uma influência entre as partes tecnológicas (hardware e software) e questões educacionais (atividades de ensino-aprendizagem) quando se utilizam dispositivos móveis e ambientes de EaD. A mudança da modalidade de interação, os equipamentos existentes no dispositivo, bem como o navegador e aplicativos instalados no dispositivo móvel, além de influenciar na interação entre o usuário e o ambiente, também podem influenciar nas atividades de ensino-aprendizagem.

Assim, acreditamos que, para especificar atividades de ensino-aprendizagem, é necessário compreender quais dispositivos o aluno estará utilizando para interagir. Por exemplo, celulares inteligentes possibilitam uma rápida resposta devido a sua portabilidade, mas pode impactar na elaboração de textos, ocasionando textos mais curtos. Por outro lado, pode potencializar as atividades de ensino-aprendizagem, como por exemplo, possibilitar aos alunos a criação de podcasts devido a gravação de áudio por meio dos celulares inteligentes.

Nesse contexto de diversidade de dispositivos, a tarefa do professor se torna ainda mais difícil em relação a especificação das atividades de ensino-aprendizagem. É necessário pensar nos periféricos que os alunos terão à sua disposição (podendo ser que um grupo de alunos tenha periféricos diferentes de um outro grupo). Assim, exige-se não somente a interoperabilidade de dados para que um determinado dispositivo possa visualizar o conteúdo gerado por um outro dispositivo, mas flexibilidade na atividade de ensino-aprendizagem que possibilite aos alunos desempenharem as atividades usando os dispositivos que têm.

Perceber a influência entre os periféricos e as atividades de ensino-aprendizagem, e a relação entre modalidade e curso (contexto educacional), conforme os ambientes de EaD se beneficiem de uma modalidade, possibilitará um maior sucesso em seu emprego no contexto educacional.

Capítulo 4

Uso de Tablet PC em Atividades Estudantis

Este Capítulo é composto por um artigo completo publicado na quarta edição da Conference on Mobile, Hybrid, and On-Line Learning, em 2012. Como descrito no Capítulo 3, o artigo apresenta o ponto inicial de desenvolvimento deste trabalho, descrevendo o uso do Tablet PC em atividades desempenhadas pelo autor deste trabalho dentro e fora de sala de aula.

Dentre os problemas apontados no trabalho destaca-se a dificuldade na colaboração entre usuários com diferentes dispositivos, o que pode limitar as ações realizadas ou adicionar tarefas com o objetivo de contornar a incompatibilidade de dados entre os dispositivos. Essa adição de tarefas dificulta o uso das ferramentas e a adoção do dispositivo.

eLmL 2012 : The Fourth International Conference on Mobile, Hybrid, and On-line Learning

Experiences in use Tablet PC to Support Student´s Activities:

Five Years of an Exploratory Study

André Constantino da Silva, Heloísa Vieira da Rocha
Institute of Computing, Núcleo de Informática Aplicada à Educação
UNICAMP – University of Campinas
Campinas, Brazil
{acsilva, heloisa}@ic.unicamp.br

***Abstract*— Since 2006 we are studying the use of Tablet PCs in academic environment to support learning activities in and out classrooms. This article resume a five years of Tablet PC use in many learning activities, since 2006 until 2011, describing how this device was used. In the exploratory study, the student deals with some interaction problems that are described here. One conclusion is the difficult to find software for Tablet PC that offer functions to support the broad of student´s activities, but the main contribution is the perception of the usability decrease when use software designed to desktop PC in Tablet PCs.**

Keywords-Hardware/software systems for mobile communications, computing support learning activities, usability, pen-computing.

I. Introduction

The recent technology forwards minimize part of the commons problems in the mobile computing: connectivity, length and processor capacity. This possibility brings a broad of new applications made to support many areas, inside this areas we can highlight the education. One of the mobile devices that are gained repercussion in this scenario is the Tablet PC, a device with height is similar to a notebook and have an input device similar a pen.

The paper and pen metaphor implies that tasks performed before in paper, like draw or manuscript writing, can be more natural in the Tablet PC than in the another computing devices. So, educational applications for contexts where paper and pen usually are presents can benefits from this new interaction style and from pen-computing, it offers better use experiences for teachers and students.

Backon [1] says that the keyboard allows a rapid typewriting and do texts structured by topics, but only pen can allow a creativity increment, better flexibility and a natural record of the think development process and rationale in all stages (ideal requirements for the educational environment), aging like a direct hand extension. For example in the resolution of complex problems with a large calculus solution, by using applications is possible to record all the steps and draw diagrams that explain the problem resolution, step by step.

When move the pen in the screen, the pen trace should result in electronic ink that must be treated by the application to be rendered and stored. But desktop applications, that running in the Tablet PCs, do not treat electronic ink, so it is necessary special applications, applications that treat electronic ink, to have benefices of the pen interaction style.

But more than treat electronic ink, witch more attributes differs desktop applications and Tablet PCs applications? Which design issues can be extracted from these attributes and used in the design of good usability applications for Tablet PCs?

To answer these questions, it is necessary collect and analyze the use of Tablet PC applications that support the user in the educational context, a student. So, the initial question is which available applications can be used in the student´s day a day activities? Which functions are important to support these activities? Intrigued by these questions, we started to do an exploratory study with real use the Tablet PC and some applications to support students activities. The main objective of this paper is brings some answers to these questions describing this exploratory studies until 2011 that began in the December 2006.

Section II shows the adopted methodology for the exploratory student and its context, and Section III presents the Tablet PCs models used in the study, and Section IV describes the used applications in this period. Section V presents a resume of the use and the main problems the u dealt. Section VI brings some design issues based on the related problems. The conclusions are presented in Section VII.

II. Methodology

The exploratory study goal is observing and describing how the Tablet PC can be used to support students day a day activities. So we used the following methodology:

- Identification of student activities in and out classrooms;
- Identification of Tablet PCs applications that support the identified activities;
- Use the collected applications for Tablet PC in the activities;
- Analyze the use and collect the student opinion;
- Extract interaction problems.

The study is related with the winning project [2] submitted for the 'Higher Education HP Technology for Teaching Grant Initiative Recipients' [3], an international awards where the best project receive financial support to equip a classroom with Tablet PC and do the proposed research project. The winning project was written in 2006 by a group of researches from UNICAMP interested in study

 ISBN: 978-1-61208-180-9

the use of Tablet PC in high education. This work is related with this project, that dispose a Tablet PC for integral use.

III. TABLET PCS AND USED MODELS

The Tablet PC is a computing device designed to "imitate" a notebook, allowing the user interact with a pen. Resuming, the Tablet PC has the following hardware characteristics:

(i) Pen sensitive screen;

(ii) Screen that allow different positions (Fig. 1);

(iii) Wireless network access by WLAN and bluetooth technology;

(iv) Microphones and embedded loudspeakers;

(v) Keyboard (some models the keyboard are detachable);

(vi) Batteries.

In this exploratory study, we used two HP Tablet PC models, both models have processors similar with the laptops in them build year. The first model, used between 2006 and 2008, is a HP/Compaq TC1100 (Fig. 1) [4] with a 1.2GHz processor and 1 Gb of RAM and a 10.4" screen. Initially equipped with Windows XP Tablet Edition, after it was change for Windows Vista with Portuguese manuscript recognizer. After, we adopted Linux operation system Ubuntu distribution.

The second model, a HP TouchSmart TX2-1040br (Fig. 2) [5] bought in December 2008, has a 2.2 GHz dual-core processor with 3Gb RAM and a 12" screen. The installed operation system was Windows Vista with Portuguese manuscript recognizer. This model has the design more similar to HP laptops but equipped with the hardware described to Tablet PC, and have a touchscreen.

IV. APPLICATIONS FOR TABLET PC

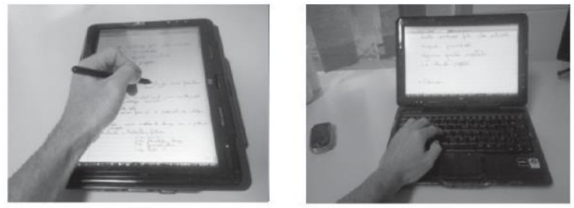
Passed by 8 year that the Tablet PCs are commercial explored, but there is few applications for them yet that can be used in the educational context. Comparing the software production for smartphones and the software production for Tablet PCs, we can see that the Tablet PC ones is timid yet.

During the study case time, tree software for Tablet PCs were used: Adobe Acrobat Professional [6], Windows Journal [7] and Jarnal [8]. All these software are for doing annotations or writing manuscript text.

The Adobe Acrobat Professional (Fig. 3) is a proprietary software that allows the creation of .pdf files and the use of some tools to typewriting text edition and doing annotations using the Pencil tool for manuscript text and the Highlighter tool for highlight parts of the text.

(a) (b)

Figure 1. Different positions to use the HP/Compaq Tablet PC model TC1100.



(a) (b)

Figure 2. Different positions to use the HP Tablet PC. model TouchSmart TX2-1040br.

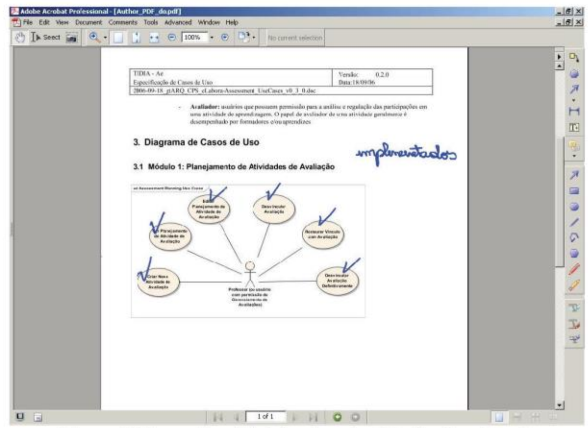


Figure 3. Example of Adobe Acrobat Professional to do annotations.

The Windows Journal (Fig. 4) is a software that comes with the Windows XP Tablet PC Edition and allow do manuscript annotations with tools to change the characteristics of the pen like fat width and color, and some functions like highlight, select and erase annotations, and commons applications functions like new file, open file, save file, copy, cut and paste text and annotations, and insert figures and text editor functions, like add or remove pages.

The Windows Journal save in a proprietary format (.jnt), and to be viewed in a desktop computer needs a special application, called Windows Journal Viewer, a software to visualize the electronic ink saved in the .jnt file. Windows Journal Viewer is distributed freely for Windows users.

The Jarnal is an application written in Java programming language with functionality similar to the Windows Journal, but it is multiplatform and free to use. Among the text editor functions and manipulate the electronic ink, Jarnal allows include files in the background so that is possible do annotations direct over the document. This function allows to use as background files like papers our presentations and highlight import parts of the text with the Highlighter tool (Fig. 5).

V. TAKING NOTES WITH TABLET PCS

The first contact with the Tablet PC occurred in a Human-Computer Interaction classes, where the students were challenged to design applications to this device considering usability attributes. One of the applications developed was the Interactive Logic Notebook [9]. But the

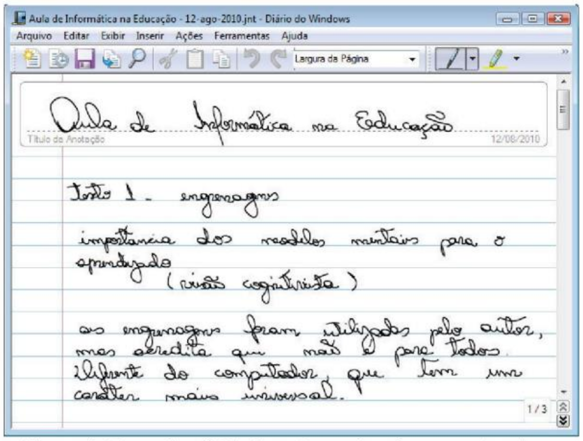


Figure 4. Example of Windows Journal to do annotations into classes.

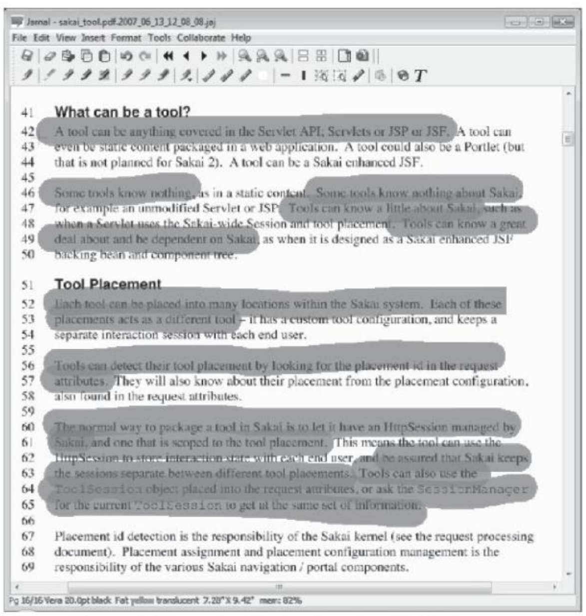


Figure 5. Example of using Jarnal to highlight text.

real use of the Tablet PC to support student activities occurred between 2006 and 2008, when we used the Tablet PC to do classes annotations and do exercises from Theory of Computing, Software Engineering, Distributed Systems and Informatics Applied in Education courses.

The first used model of Tablet PC was available for the students in December 2006, month without classes at UNICAMP. To stimulate the use and acquire experience in use Tablet PC and software, the student decided to use the device to support her in research day a day activities.

The first task the student used the Tablet PC was analyze an use case document available in .pdf format verifying which use cases about an application was implemented conform specified or identified points to update the document when the difference was result from a decision process to change the design. The student´s goal was doing marks in the text to signalize points that need to be updated or points where the implementation needs to be reviewed.

The chose software was Adobe Acrobat Professional, because the goal was do annotations in an existing .pdf file. Some interactions problems where find in this use, resulting in the adoption of another tool to make this kind of activity in classroom.

The first problem dealt by the student when she used the Adobe Acrobat Professional was the reception of input data to render the electronic ink. In the manuscript writing, the letters were rendered in a sharped form and sometimes only a trace was showed. Maybe this problem was result from the device do not match the hardware requirements that the software needs, so the Adobe Acrobat did not received all the points to draw correctly the electronic ink. This problem occurred mainly in a short time to start the writing. To resolve this problem, the student pressed the pen against the screen, waiting one or two seconds and start to write the word. But some letters become hard to recognize, because the forgotten points that compose the electronic ink. The solution adopted was written more slow.

Another interaction problem was related with the electronic ink selection function that the Pencil tool have. Pressing the pen near an existing electronic ink, the software interpret this action has a electronic ink selection. This way make hard write letters like “i”, “j”, “k”, “t”, “ç” in the manuscript way and accentuation, that require more than one trace to written the letter. When the users will write the second trace trace for one of these letters, the first trace usually is moved to other position. The user strategy was begin the second trace so far so that the second trace was recognized has a new electronic ink and not has a selection for the existing trace. But this strategy works fine with letter “t”, and does not work to write the dot for the letters “i” and “j” and write the cedilla for letter c.

Another problem deal with the student is related with the scroll bar position, disposed in the right corner. For a left-handed person, the case of this student, when interact with the scroll bar the arm embarrass the content visualization that are moving, become hard to put the content in the wished position. Another problem is the scroll bar width, that for the use with a pen must be more fat. Using the scroll bar in the default width, the student had the tendency to cross the scroll bar borders when the pen was in the second half of the screen. The pen cross the bar borders because the movement of the fist over the screen (the movement is not a straight line). When the pen cross the bar borders, the software back the visualization to the content part viewed before the scroll operation.

To scroll, the user adopted another tool that the Adobe Acrobat Professional has, the Hand tool. This tool allow press in any part of the content and scroll up or down, allowing change the document visualization. The users said the use of this tool is more efficient form to scroll the document.

To shared this use experience, the student written a manuscript text using the Jarnal application with the intent to publish the text in the Messaging Board tool of TelEduc environment, used as a communication platform among the project team. But the Internet browsers do not support electronic ink or the Jarnal format, so the student tried to publish the annotations as pictures, because her wanted to maintain the manuscript way for her readers. To convert the annotations made in the Jarnal application in pictures, it was necessary save the annotations in a .pdf format (the Jarnal does not have functions to convert annotations in pictures, but have functions to save the annotations in .pdf format), open the saved file and copy page by page using the Adobe Reader´s Instant Photography tool. After the selection and copy operations, the student paste this content in a image editor and save the file as a picture format (.jpg). So the images are uploaded to the platform as attached files and referenced by the .html file as image to be viewed by the readers. This activity was considered too hard to do and abandoned lately.

The Tablet PC pen interaction differs from the interaction styles that laptops have (keyboard, track pad) and shows a great potential to facilitate the activity of do annotations in Theory of Computing classes, because in this discipline is necessary adopt symbols that not are in the QWERTY keyboards.

In Theory of Computing classes, exercises lists was shared with students using the .pdf format. As the exercises list is a sequence of problems enunciation without space among them, this become impossible to resolve the exercises using the document as background. So the student resolve cut the enunciation exercise and past in a Windows Journal archive (Fig. 6). The student needed to adjust the pasted object dimensions to become more readability in the zoom used to resolve the exercise (page zoom). After did this steps, the student resolve the exercises. The student justify this work answering her wants to have a map between exercise and resolution for future study., for example, for an exam.

Using the problem enunciation with the resolution can facilitate the resolution, because it is possible highlight parts of the text that are important to the solution or use pictures available in the enunciation (Fig. 6). The digital content allows duplicate items that need to be more than one time in the resolution, the case of the graph draw in the Fig. 7, when the enunciation ask to the students verify how many short-ways the graph has. The student duplicated the graph picture instead of draw it.

In problem resolutions, more than the answer, is important to have the record of each step until get the answer [1]. In some resolutions, some parts needs to be duplicate to show the rationale, as the case of the Fig. 8. The exercise asks a spanning tree for the given graph. The use of the Tablet PC was specially important to use copy and paste functions, turning the answering this exercise more fast because many parts did not need to be rewritten.

A common practice among students is share classes annotations and resolutions from exercises. This practice occurs in Theory of Computing course among the student that used the Tablet PC and the others one. So her sent this annotations by e-mail in the .jnt format. But the other students needed to install the Windows Journal Viewer application to view the file content, and asked to shared the annotation in a more popular format that not require a software installation, like .pdf format. The students adopted this format for all annotations sharing.

To explain different resolutions, doubt points or indicate errors in the shared resolution, the students that did not have a Tablet PC used to typewrite text artifices when use e-mail

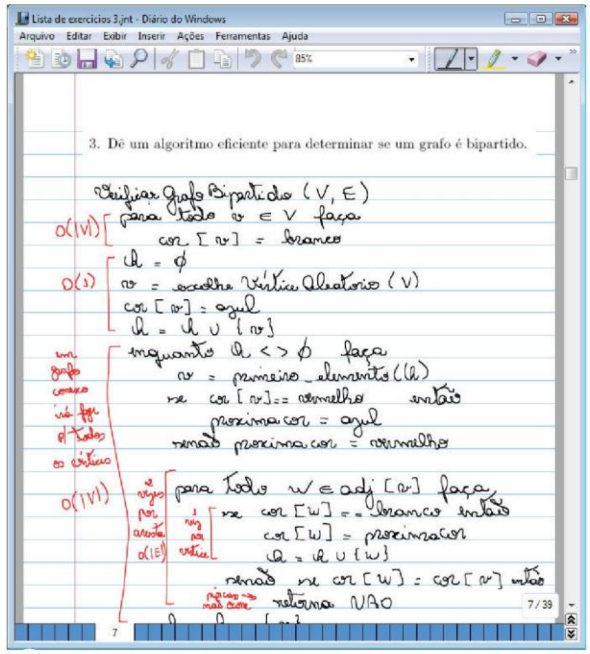


Figure 6. Example of manuscript using Windows Journal when resolving Theory of Computing exercises.

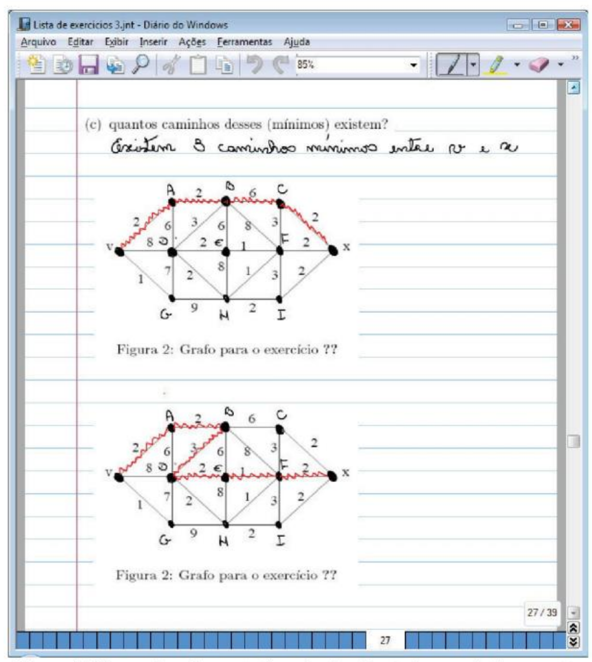


Figure 7. Example of annotations in duplicated enunciation part to resolve an exercise.

 ISBN: 978-1-61208-180-9

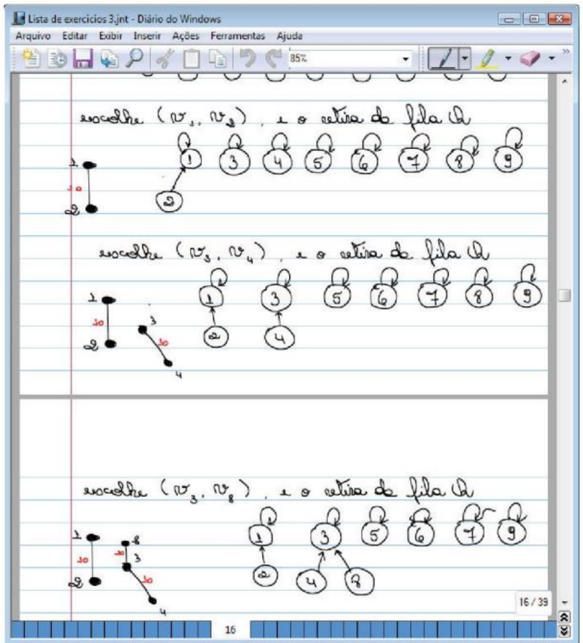


Figure 8. Example of use the copy and paste functions to duplicate annotations.

or to print the documents for do annotations with pen in meeting.

In the next course, Informatics Applied in Education (cursed in the first semester of 2008), the student used the Windows Journal (Fig. 9) to taking classes annotations, using the theme classes in the file name when save. The use of Windows Journal is justified by the better performance in receive the data input and convert into electronic ink, so the result letter is more closer than the other applications.

The Jarnal was used in the Informatics Applied in Education course to support the peer review activity. The peer review activity consist into read a text written by another student and points out how to improve the text. The available text was used as a background document, where the student did the annotations (Fig. 10).

VI. Usability Problems of the Adopted Applications

Analyzing the user interaction, it is possible identify some usability problems in the used applications. The first one described by the student is related with use the same tool to do manuscript text and select electronic ink. Considering manuscript letters, allowing select inks with the tool used to write it turns hard write letters composed by more than one trace, or have points or accentuation. So, one design issue for Tablet PC is avoid unify functions for write and for select in one tool.

To compose the traces of electronic ink, the application needs treat all input data emitted by the pen, composed by pen position on the screen and the pressure. For the letter do no be shaped and the manuscript in Tablet PC be more similar to the manuscript in paper with a pen, the application needs treat all the events generated by the pen. Techniques for pour rendering of the electronic ink while writing and a high fidelity rendering after the the writing can be used to not loose data.

One of the differences between Windows Journal and the other applications (Jarnal and Adobe Acrobat Professional) is the buttons inside the Windows Journal tool bars are softly bigger than the buttons of the other applications (Fig. 11), this make more easy to select a tool with the pen.

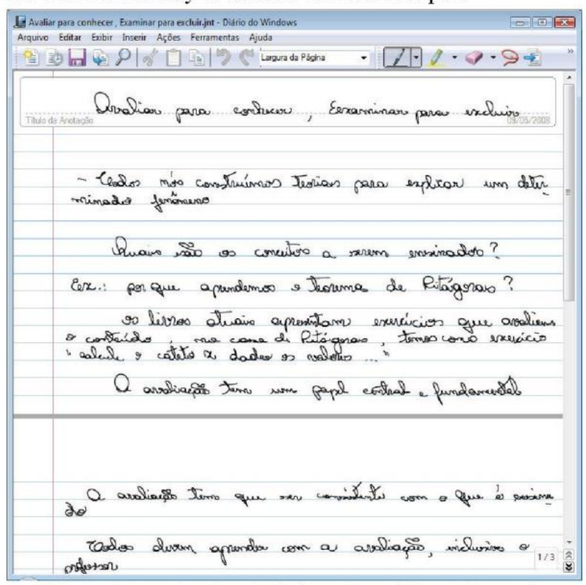


Figure 9. Example of use Windows Journal into Informatics Applied in Education classes.

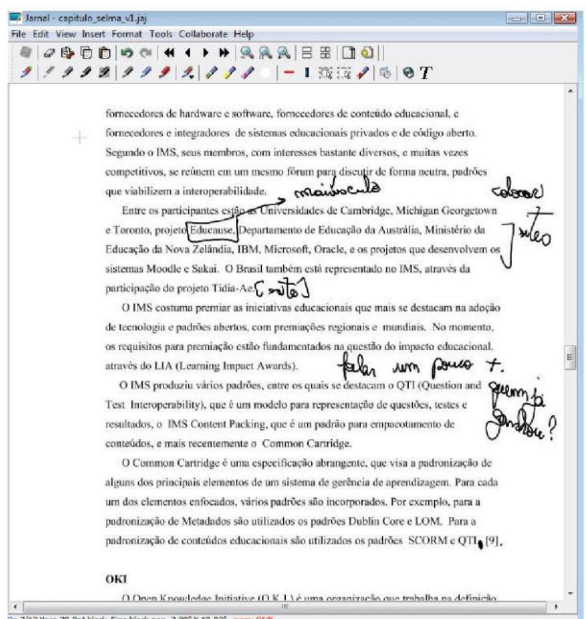
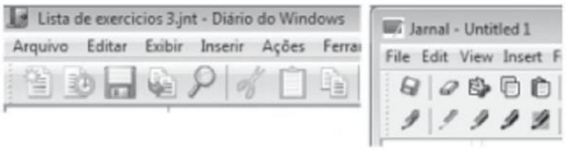


Figure 10. Example of Journal use to do annotations in a text.



(a) (b)

Figure 11. Dimensions between buttons into Windows Journal (a) and Jarnal (b) Tool bar.

Analyze the options to erase electronic ink into the Windows Journal and Jarnal applications, we perceive that the Windows Journal offers two kind of erase forms, while the Jarnal offers only one. The form offer by both applications is erase all the traces that the erase tool have contact. The Windows Journal offers too another way of erase electronic ink that erase only parts of the trace, allowing a better fine adjust of the writing.

Activities that needs collaborations among students are commons in educational context. So, the applications needs make easy the communication of many devices, by file transfer or by Internet. It is important do not restrict the communication only between Tablet PCs, but consider the possibility of many devices used like desktops, laptops and smart phones.

And for motivate the use of an application, the application needs have functions that explore the potential of digital artifacts. In this case, functions to copy, cut, paste and move the electronic ink. These functions turn motivational factors to adopt these applications.

We believe that these described issues must be considered in the design of Tablet PC applications, with the goal to explore the pen interaction and minimize usability problems.

VII. Conclusions And Future Works

The Tablet PC is a computing device designed to be similar to a notebook, and the main hardware to be used in the interaction is a pen.

For five years, we did an exploratory study where two models of Tablet PCs are used to support a graduate student activities. The main use was do classes annotations in classroom, highlight part of texts and doing exercises.

The Adobe Reader Professional, Windows Journal and Jarnal applications were used in many periods as annotations editors. Many available functions was used like color, annotation selection, copy, paste, cut of annotations and pictures.

When the pen moves over the screen, it generates input data for the applications that need to be used to generate the electronic ink. Desktop applications does not manipulate the electronic ink, resulting in loose of the Tablet PC potential because the pen is used only as a pointer device. In this case, the use of pen as a pointer device in desktop applications, there are usability problems.

We present and discuss some of them, like the dimensions of user interface components, the use of one tool to writing and selecting electronic ink and the difficult to see the content when user needs to cross the screen with her hand. The discussed issues plus the needs to explore the interaction style used by the users turns hard the design applications for multidevices, mainly when we need to consider many interaction styles.

Acknowledgment

The authors thank the HP for promote the study of Tablet PCs into classroom, and for financial support by CAPES, CNPq and FAPESP.

References

[1] J. Backon, "Student minds and pen technologies: A wonderful pedagogical marriage", in The Impact of Pen-based Technology on Education: Vignettes, Evaluations, and Future Directions, D. Berque, J. Prey, and R. Reed, Eds. USA: Purdue University Press, 2006, pp.1-11.

[2] R. J. de Azevedo and H. V. da Rocha. An annotation-based tool for collaborative learning using mobile technology. Tablet PC Technology at UNICAMP Project. Available at: http://www.ic.unicamp.br/~rodolfo/tabletpc/ , accessed November, 2011.

[3] HP Technology for Teaching 2006 grant. 2006. Available at: http://www.hp.com/hpinfo/socialinnovation/us/programs/tech_teaching/hied_global.html , accessed November, 2011.

[4] HP Compaq Tablet PC TC1100 – North America/U.S. Quick Specs. 2005. Available at: http://h18000.www1.hp.com/products/quickspecs/11755_na/11755_na.html, accessed November 11, 2011.

[5] HP TouchSmart tx2-1040br Notebook PC (NL266LA) especificações. 2009. Available at: http://h10010.www1.hp.com/wwpc/br/pt/ho/WF06b/321957-321957-3835812-3835813-3835813-3837256-3937928.html , accessed November, 2011.

[6] Adobe Acrobat Professional, 2011. Available at: http://www.adobe.com/products/acrobatpro.html , accessed November, 2011.

[7] Windows Journal. Available at: http://en.wikipedia.org/wiki/Windows_Journal , accessed November, 2011.

[8] Jarnal. Available at: http://jarnal.wikispaces.com/ , accessed November, 2011.

[9] A. C. da Silva, V. A. Neris, J. Daltio, M. Weinfurter, H. Zanetti, and H. V. da Rocha, "Caderno Interativo de Lógica – uma aplicação para TabletPC de apoio ao ensino de lógica de primeira ordem", Proc. Simpósio Brasileiro de Informática na Educação (SBIE 06), SBC, Nov. 2006. Available at: http://200.169.53.89/download/CD congressos/2006/SBIE 2006/msoftware/ms.html , accessed November, 2011.

Capítulo 5

Problemas de Interação Relacionados a Mudança de Plataforma ou de Modalidade

Este Capítulo é composto por dois artigos que tratam da identificação de problemas ocasionados pela mudança de plataforma ou mudança de modalidade. A motivação da realização dos trabalhos que geraram os artigos é que os ambientes de EaD foram projetados e avaliados considerando que os usuários irão interagir por meio de um computador convencional, equipado com teclado, mouse e monitor de tamanho médio com alta resolução. Como os dispositivos móveis possuem outras modalidades de interação, questionou-se quais são as barreiras ou problemas encontrados pelos usuários de dispositivos móveis ao acessar os ambientes de EaD. O objeto de estudo é o ambiente TelEduc acessado por celulares inteligentes (*smartphones*) e Tablet PCs, sendo que o primeiro artigo apresenta alguns problemas identificados por um especialista em interação, enquanto que no segundo foi utilizado o método de testes de usuário para identificar problemas de interação.

No primeiro artigo, publicado na sexta edição da *International Conference on Advances in Computer-Human Interactions,* os problemas encontrados são categorizados em: (i) problemas de interação sem relação com a mudança de modalidade ou plataforma; (ii) problemas de interação relacionados com a mudança de plataforma; (iii) problemas de interação relacionados com a mudança de modalidade.

No segundo artigo, publicado na *IADIS International Conference on e-Learning*, evento que compõe a *Multiconference on Computer Science and Information Systems*, os problemas identificados por meio de testes de usuários são encaixados nas categorias, reforçando as categorização criada. Os problemas coletados nos testes de usuários e apresentados no artigo são oriundos do projeto de iniciação científica do bolsista Alan Victor Pereira de Arruda, coautor do artigo, o qual co-orientei.

Pode-se concluir pela leitura dos dois artigos que, quando o usuário interage com a aplicação usando uma modalidade não considerada em seu design, problemas de interação podem ocorrer. A identificação dos problemas e a sua correção ainda não garantem que os benefícios da modalidade sejam aproveitados, mas oferecem indícios para novas funcionalidades. As soluções propostas também devem considerar as limitações dos periféricos de interação, do sistema operacional e dos navegadores disponíveis no dispositivo pois, como descrito nos artigos, restringem a implementação de soluções.

Identifying Cross-Platform and Cross-Modality Interaction Problems in e-Learning Environments

André Constantino da Silva
Institute of Computing (PG)
UNICAMP, IFSP
Campinas,Brazil, Hortolândia,Brazil
acsilva@ic.unicamp.br

Fernanda Maria Pereira Freire
NIED
UNICAMP
Campinas, Brazil
ffreire@unicamp.br

Heloísa Vieira da Rocha
Institute of Computing, NIED
UNICAMP
Campinas, Brazil
heloisa@ic.unicamp.br

***Abstract*—Web applications and sites are designed to use keyboard and mouse as input devices and a medium resolution screen as output device. Mobile devices, such as smartphones and tablets, have enough computation power to render Web pages, allowing browsing the Internet. But, their main interaction style is touching style that was not usually considered in the Web applications design. Changing the platform or interaction style can lead to interaction problems. To study these problems, we investigated the use of TelEduc, an e-Learning environment designed to Internet and to be used with keyboard and mouse, in two touchscreen devices, a smartphone and a tablet. Some problems are usability problems and do not have relation with the platform or modality, but other problems are related to the platform or modality changing.**

Keywords-Mobile devices and services; Interfaces, interactions and systems for distance education; Interface evaluation; Usability testing and evaluation;

I. INTRODUCTION

Devices, such as smartphones and tablets, are becoming increasingly popular; most of them have touch screen displays, access to the Internet and enough computing power to process Web pages. So, Web sites and Web applications, initially developed to be used with keyboard, mouse and a medium size display, are been accessed by small touch screen devices.

One kind of Web applications is e-Learning environments, as Moodle [1], SAKAI [2] and TelEduc [3], which are applications with tools to support teaching and learning activities though the Web. These tools allow users to create content, communicate with other users and manage the virtual space.

These environments take advantages of the Web to offer content with text, images, audios and videos in a hypertext document. Tools like chat, forums, portfolios, repositories are widely used, and tools that explore the audio and video resource to user communication, such as instant messenger and video-conferences, are becoming common among the environments.

Since smartphones and tablets are easy to carry, have autonomy for hours and Internet access, the e-Learning environments´ development teams are building solutions to provide access on mobile devices. Three kind of solution are emerging: specific device application, Web site specific for mobile devices, and improve the Web site for mobile and desktop access.

Browsing Web site using another device (last two solutions), the user can deal with some problems related with the platform changing (e.g., from desktop´s browser to smartphone´s browser), but some problems happen due the interaction style changing (e.g., keyboard, mouse and medium screen to a small touchscreen). We call these problems as cross-platform interaction problem and cross-modality interaction problems, respectively.

So, which cross-platform and cross-modal interaction problems users deal? How to categorize a problem? Thinking about this question in the e-Learning context, we developed this work. Though a user interface analyze by a specialist using a smartphone, a tablet and a desktop, we found some problems to browsing TelEduc using two touchscreen devices: a smartphone and a tablet. These problems were classified into categories: cross-platform problem, cross-modality problem, and platform and modality- independent problem.

The next Section describes related works. Section III presents the TelEduc Project with a brief historical view, the tools and features of TelEduc e-Learning environment. Section IV shows the material and method adopted. Section V describes some identified interaction problems, and, in the Section VI, these problems are classified into the three described categories. Section VII presents conclusion and future works.

II. RELATED WORK

The e-Learning environments´ development teams are building solutions to provide access on mobile devices. Three kind of solution are emerging: specific device application; Web site specific for mobile devices; and improve the Web site for mobile and desktop access.

Building specific device application allows designing a suitable user interface for the device and taking advantages of smartphone´s features, such as touchscreen and camera, but needs develop an application for each mobile platform, so to be developed needs specific knowledge programming team and increases the code lines number to maintain. Moodle community offers the Moodle App [4] and Moodbile

 ISBN: 978-1-61208-250-9

[5], two native mobile applications with versions for the most popular smartphone´s platforms.

Moodle, since version 2.1, offers a Web site specific to mobile devices, an example for the second type of solutions for access e-Learning environments in mobile devices. Building a specific Web site to mobile device allows designing a suitable user interface for mobile devices taking account some common characteristics, such small touchscreen, but depends of the browser to access some platform features, such GPS, and increases the code lines number to maintain too.

The latter solution considers that smartphones and tablets have enough computational power to render Web pages and to do some adaptation if it is necessary, and offer the same user interface for any device. To design this kind of user interface it is necessary to do some usability studies to found barriers or user interaction problems. Disadvantages of this solution are to depend of browsers to use the mobile features and the difficult of consider many interaction styles in the same user interface.

This solution can start from a user interface design model for desktop and be improved to consider mobile devices. So in the initial design was designed thinking the user will interact by keyboard, mouse and medium size display and, allowing users accessed these applications on mobile devices, there is an increasing of interaction styles number, such touchscreen. With the interaction hardware changing the user deals with new interaction problems. Shrestha [6] points out some problems when the users try to use mobile devices to do specific tasks into Web sites designed to desktop, so the mobile Web browsing experience needs to be improved to a more mobile friendly Web site and some mobile browser improvement (here we consider browser as one platform characteristic).

Shrestha [6] considered mobile devices equipped with joystick and a small screen. Maurer *et al.* [7] did some usability studies using touchscreen mobile devices and desktop for browsing in Web sites, shows that "more and more people prefer using original content instead of the mobile version, especially for users of new generation mobile devices like iPhone and Android phones". Another result of this work was the users prefer to use the standard Web site instead of tailored mobile versions of Web site. But Schmiedl, Seidl and Temper [8] have a different opinion; in their research they conclude most of the users still prefer tailored versions. Kaikkonen [9] shows that the standard and the tailored Web sites are both used but for slightly reasons.

Considering only the e-Learning environment context, we agree to Maurer *et al.* [7] when they argue the user prefer to use the standard version instead of mobile version of Web site, and this is one of the motivations of our work to study the third kind of solution to access e-Learning environment using mobile devices.

Here, we studied the use of touchscreen devices into e-Learning environments and distinguish the problems into categories. Shrestha [6] studied the use of joystick and a small screen to browsing into some Web sites, while Maurer *et al.* [7] considers touchscreen devices. All these researches point out some interaction problems, but do not classify them if they happen due the platform changing or due the interaction changing.

III. TELEDUC E-LEARNING ENVIRONMENT

The TelEduc is a teaching and learning environment developed by the Nucleus of Applied Informatics in Education (NIED) and the Institute of Computing (IC), State University of Campinas (UNICAMP), and adopted in several public and private institutions, like UNICAMP through Ensino Aberto project [10].

The TelEduc environment was conceited in the end of 90, born with the Cerceau´s Master dissertation (1998), with professor Heloísa Vieira da Rocha as advisor, applying constructivism theory [10][12] in situated learning [13] or in contextualized learning [14] for teacher´s continuance formation. In 2001 February, the first free version was released over GNU General Public License (GPL), an unprecedented fact in the Brazilian Educational Software scenario. Many public and private institutions adopted the TelEduc as platform, increasing the TelEduc user's community, and consequently, the development demand. This fact culminated in the release of TelEduc version 3.0 in March 2002. The version 3.0 was completely redesigned and optimized, reason for TelEduc project was awarded by ABED (Brazilian Association for Distance Education) in the "Research and Development about Distance Learning" category. In August 2011, TelEduc version 4.3 was released, with its user interface redesigned to improve user tasks and be more similar than popular Web sites.

TelEduc is a system that aggregate administration, management and communication tools designed to support teaching and learning activities. Some tools allow users to create content, other ones allow synchronous or asynchronous communication among users, and manage participants and courses. The course page of TelEduc 4.3 is structured in two parts: the left one (Fig. 1a) has a list of all tools available and in the right one (Fig. 1b) the content of the selected tool.

In the course showed at Fig. 1a, the teacher dispose the Course Dynamic, Agenda, Readings, Support Material, Activities, Chat, Mail, Discussion Forum, Frequently Asked Questions, Portfolio, Groups and other tools available. Fig. 1b shows the user interface to visualize a Support Material item, where the teacher can change the title, content, attach or remove files or links and see and write comments, and the student can see the item, download the attached files and

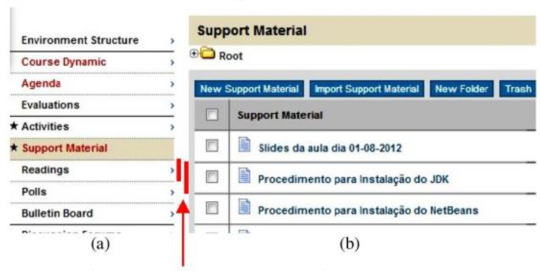


Figure 1. Height difference between menu items (a) and Support Material items (b).

visit the posted links.

Agenda is the ongoing course home page and shows the course´s program for a given period (daily, weekly, etc.). Agenda is an important tool because organize the activities that must be done in a specific period, similar teachers do in the beginning of a presence class.

The Support Material is a tool that provides an area for file storage and sharing among course participants, named Support Material Area. To store an item in this area, the user needs to be a coordinator or an instructor. When the user stores an item in support material area, she can specify the sharing type: i) not shared; ii) shared only with users who have instructor role; or iii) shared with all participants. Users with student role can access the stored items published with all participant sharing type, read their content, visit their links and download their attached files.

The Readings, Activities and Frequently Asked Questions tools have similar features and user interface with Support Material tool, but different purposes. Readings tool is used to publish relevant documents, like books, magazines, news and articles. The Activities tool is an area to publish activities to the accomplished during the course, like home work descriptions. The Frequently Asked Questions tool contains a list of the most frequently questions done by the participants during the course and their respective answers.

Tools like Discussion Forums and Mail are used to participant communication, supporting text message exchange in asynchronous mode. To synchronous communication, there is the Chat tool, its features is similar to Web chat sites.

The Portfolio is a communication tool that aims to promote the collaboration among participants through the sharing of "items" (documents, presentations, programs, links, etc.). So the Portfolio tool provides an area to item storage and sharing for each participant (user or group of users) within a course.

The Bulletin Board tool is a dedicated space where all the participants can post information considered relevant to the course content.

The Agenda, Activities, Support Material, Readings, Bulletin Board, Discussion Forums, Mail and Portfolio allow users to create text content using a text editor, the CKEditor [15]. CKEditor is a third-party WYSIWYG text editor to be used inside Web pages, bringing to the Web application common editing features found on desktop editing text applications, but CKEditor version 3.3 does not work on mobile devices. In [16], we studied some problems in use of third-party software on TelEduc to create and to visualize documents, describing that mobile compatibility needs to be considered to not prejudice mobile users.

The Agenda, Activities, Support Materials, Readings, Mail and Portfolio tools allow attaching files into them items in similar way: the user clicks in the "Attach File" button, select the file to be attached using a dialog and, after the click on the save button, the file is uploaded.

To provide content, TelEduc uses the Web infrastructure, more specifically, hypertext with images, links, audios and videos. All these media can be published as content in tools like Agenda, Support Material and Readings.

Since the e-Learning environments need to be easy to use for users with different levels of Web experience, the usability is an important nonfunctional requirement. TelEduc was designed in an iterative design-evaluation process to have good usability and the user interface does not impair teaching and learning activities, so many usability evaluations were done. The accessibility is another nonfunctional requirement desired for TelEduc, to allow impaired people to use the environment without meet barriers or obstacles.

TelEduc was designed to use a mouse and keyboard as input devices, and a medium screen size as output device. TelEduc is better visualized into 1024x200 pixels screen resolution. Visualize it into a lower screen resolution cause some user interface problems like dispose interface components in wrong position.

Since the e-Learning environments are available on the Internet, this software can be accessed by smartphones and tablets nowadays and the developers need to study how to allow all features into these devices. Access the environment in anywhere and anytime is one of the biggest attractions, but research is necessary to have a user interfaces with high usability and good user experience. To reach it, it is necessary identify cross-modality problems when TelEduc is used in touchscreen devices.

IV. MATERIALS AND METHOD

To investigate which problems happen when the users use mobile devices to browsing a Web application, we adopted the following method: a human-computer interaction (IHC) specialist, using a touchscreen device, accesses the Web application and collect interaction problems. To each collected problem, the specialist checks which ones from the three devices the problem happens and so classified into the categories: platform and modality-independent problem, cross-platform problem and cross-modality problem.

We used a Motorola Milestone smartphone [17] and a tablet PC [18]. The Motorola Milestone has a 3.7 inches multitouch display with 133MB internal storage memory expansive up to 32 GB with a memory card, 600 MHz Cortex-A8 processor and 256 MB RAM, a 5MP camera, GPS and wireless connection by Wi-Fi 802.11 b/g and bluetooth. Android 4.0.3 [19] was used as operation system (OS). Motorola Milestone has a proximity sensor, an ambient light sensor, a 3-axis accelerometer and a geomagnetism sensor to provide orientation with respect to Earth´s magnetic field. To browse in the Web application, the specialist uses the Android stock Web browser. The proximity sensor and the accelerometer can be used to interact with applications, but the stock Web browser does not use these features as an input device. Only the touchscreen was used as input device.

The Tablet PC is a computing device designed to "imitate" a notebook, allowing the user interact with a pen. Resuming, the Tablet PC has the following hardware characteristics: (i) Pen sensitive screen; (ii) Screen that allows different positions; (iii) Wireless network access by WLAN and bluetooth technology; (iv) Microphones and embedded loudspeakers; (v) Keyboard (some models the keyboard are detachable); (vi) Batteries.

In this exploratory study, we used a HP TouchSmart TX2-1040br, a 2.2 GHz dual-core processor computer with 3 GB RAM and a 12" touchscreen. The installed operation system was Windows Vista with Portuguese manuscript recognizer. This model has the design similar to HP laptops but it is equipped with the described hardware for Tablet PC. To the study case, the specialist only used touch in the interaction; she did not use the Tablet PC´s keyboard or pen. Chrome browser version 22.0.1229.94 m [20] was used to navigate through Web application.

To classify the problem, the interactions in the three devices were compared (Fig. 2). If the problem appears in all devices, the probably of the problem be a platform and modality- independent problem is high. If the problem only happens in touchscreen devices, the probably of the problem be a cross-modality problem is high. But if the problem only happens in smartphone or only in tablet, maybe it is a cross-platform problem. So, to distinguish cross-platform problem, cross-modality problem and platform and modality-independent problem we needed to use these three devices. The Tablet PC and Desktop have the same platform (Widows Vista and Chrome). Tablet PC and the smartphone are both touchscreen devices and different platform (smartphone uses Android).

V. Interaction Problems

The specialist, using both touchscreen devices, found some interaction problems. An overview of the main problems is presented in this section.

A. *Problem 1: Fat Finger Problem*

At TelEduc, the tools available to be used in a course are listed in a menu disposed in the left side, and each tool is an option in this menu (Fig. 1a). TelEduc uses vertical lists not only for the tools menu, it uses vertical list to shows items to choose, e.g., in Support Material tool, each item is an option in a vertical list (Fig. 1b). An option in items lists (36 px) is higher than an option into tools menu list (23 px). This is a small difference, but the specialist points out problem to select a tool into the menu. Due to the small option menu height the specialist had problem to select a tool, triggered

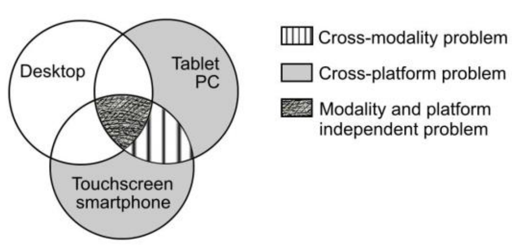


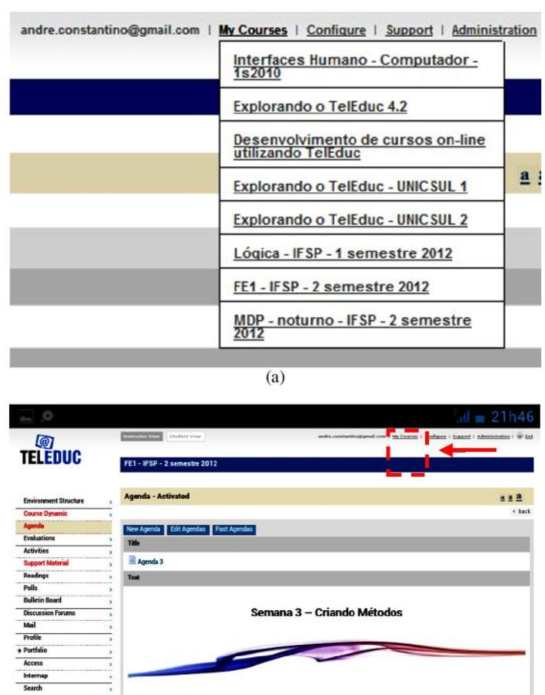
Figure 2. Categories of identified interaction problems on used devices.

the upper or down one, but this problem happen because the touched area is bigger than the clicked area when the user uses a mouse. This problem is called fat finger problem [21], when a user tries to acquire a target, the center of the contact area tends to be located a couple of millimeters off the target location—typically "below" the target [22].

The observed fat finger problem happens in both touchscreen devices, but more into smartphone probably due to page adaptation to show all content in the screen. The specialist observed the problem happens more into menu selection, because the space among menu items are not enough too large. The specialist observed the problem did not happen when selecting an item in the Portfolio or Support Material tool, probably because the distance among the items.

B. *Problem 2: Mouseover functionality*

TelEduc have My Course link to see the enrollment courses the user are and go to the Courses page. In desktop, the user can see the enrollment courses putting the mouse pointer over My Courses option, localized in screen right-top (Fig. 3a, the courses name is in Portuguese). If the user clicks in My Course link, she goes to Courses page. Touch screen devices does not show this menu since there is no *mouseover* action (Fig. 3b, the Agenda content is in Portuguese), and when the user try to see the enrollment courses, she triggers



(a)

(b)

Figure 3. Position of the My Courses menu (a) on Desktop and (b) on smartphone, the user cannot trigger the menu.

the My Course link and goes to the Courses pages.

This problem happens due the JavaScript *mouseover* function. It is common in Web applications the use of JavaScript to improve their user interfaces. But, some features of the JavaScript can cause interaction problems, like the *mouseover* function. The Android stock Web browser on smartphones makes a map between user´s action to browsing event, but there is no valid mapping to activate the *mouseover* function, since the used devices do not identify finger proximity, so it is not possible to have a feature similar to *mouseover* to do with the finger. The same problem happened on tablet.

Some of features in Table I are triggered by gesture. Using gesture into touchscreen devices the user can go forward, backward, scroll up and scroll down the page.

C. Problem 3: Gestures

Motorola Milestone and HP Tablet PC allow user interaction using gestures. To browse using these devices, the user can use one or two fingers and make gestures. Default gestures are to zooming; scroll up and scroll down a page; forward and backward pages; and select, copy and paste text. For novice users, it is not clear how to make the gesture. This is a discoverability and visibility problem as Nielsen and Norman related [23].

But, the user does not add specific gesture to use with a Web application, and the Web application cannot add specific gestures to user browser. This limitation does not allow TelEduc have gestures to create a new portfolio item, to select some Mail message or to delete a Support Material item. Since gestures have the promise to brings a powerful interaction [24], the Web application does not get all the promised potentiality, and the users gesture are limited to scroll up or scroll down a page; forward or backward pages; zooming; and select, copy and paste text.

TABLE I. MAPPING BETWEEN USER ACTIONS AND BROWSING EVENT IN THE ANDROID STOCK WEB BROWSER

Browsing Events	Using mouse	Using finger
Link activation	Left button click	Touch with one finger
Menu drop down	Right click button	Touch with one finger and hold
Scrolling text	Mouse click over scroll component interface	Touch with one finger and drag
Zooming	Not possible (needs change browser configuration)	Touch with two fingers and spread/pinch or double tap
Select text	Left mouse click over the text beginning and drag until the text end	Touch with one finger over the text and hold, release and drag the selection text component
Copy text	Select the text, click on right mouse button and choose copy	Select the text, click the option button and choose copy
Paste text	Right mouse button and choose paste	Touch with one finger and hold, choose paste option

D. Problem 4: Device features

Android platform specification defines four physical buttons: Back, Menu, Home and Search. Android platform allows developers customize the reaction of these standard buttons, like use search to find in application data or show the application menu instead of the default menu. This is one of the differences between Web applications and Android applications. Web applications do not have this possibility, and the action buttons are defined by the Web browser. So the search button, instead of searching into Web application data, opens the URL field (search a page).

Smartphones have a lot of features and the Web application cannot use. The specialist tried to post a photo and a video in her Portfolio, but the browser does not upload them and does not show a message error.

E. Problem 5: Third-party text editor

TelEduc uses CKEditor to allow users write rich text instead of simple plain text (Fig. 4a). But CKEditor does not work in Android devices (Fig. 4b), but works in the tablet.

F. Problem 6: menu activation

TelEduc menu item is only activated when the user click in the menu item text, a little different from the computers menu interaction. This difference is easy to understand when browsing in the menu, since the mouse pointer does not change when the pointer is over the menu item. In the case of smartphone and tablet, the user can try many times touching the space in front of the text to understand there she must tap the text.

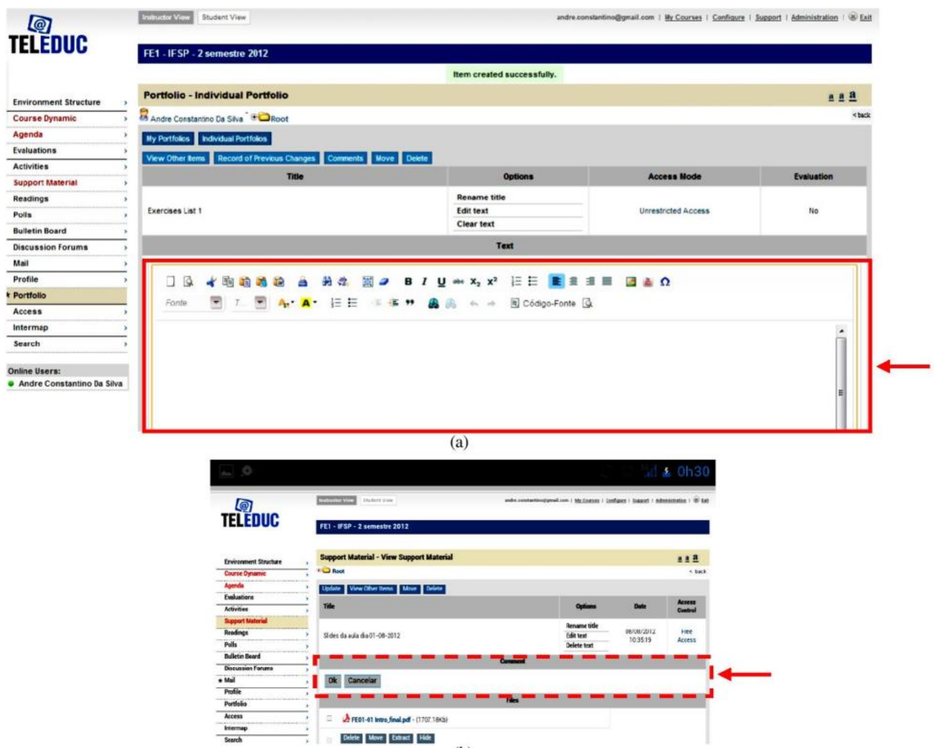
Table II summarizes the identified problems and the platform where they happen. The next section discusses and classifies these problems.

VI. CROSS-PLATFORM AND CROSS-MODALITY PROBLEMS

The fat finger problem (Problem 1), *mouseover* functionality (Problem 2), gesture integration (Problem 3) and menu activation (Problem 6) happened in both smartphone and tablet. Gesture integration can be classified

TABLE II. IDENTIFIED PROBLEMS AND THEIR CLASSIFICATION

	Problem	Touchscreen smartphone	Tablet PC	Desktop	Classification
1	Fat finger problem	yes	yes	no	cross-modality
2	*Mouseover* functionality	yes	yes	no	cross-modality
3	Gesture integration	yes	yes	no	cross-platform
4	Device features gap	yes	no	no	cross-platform
5	Third-party text editor	yes	no	no	cross-platform
6	Menu activation	yes	yes	yes	platform and modality-independent



(a) (b)

Figure 4. Position of the rich text editor in TelEduc a) the editor is displayed on Desktop and tablet PC and b) it is not displayed on Android-based devices.

as platform problem, though it happens on both devices, this problem is not directly related with touch; this is a lack of functionality. Fat finger problem and the use of *mouseover* functionality were considered as cross-modality problems. They happened when the user interface, designed to be used with an interaction style, is accessed by different interaction styles. If the application was designed to be used with touchscreens, the designer would choose the suitable interface components size and the spacing among them, decreasing the occurrence of the fat finger problem. In the case of *mouseover* functionality, the designer would choose a better way to show the menu, e.g., when the My Course link is activated, the menu is showed instead of go to Courses page.

Menu activation (Problem 6) was considered as platform and modality-independent problem, since the problem happened in desktop computers too, but when the user is using a smartphone the problem turns more severe and be easier to find.

Device features gap (Problem 4) only happens in smartphone due the difference between the Android platform and Desktop Operating Systems. This problem was classified as cross-platform problem.

The problem with no display the CKEditor (Problem 5) was classified as cross-platform problem. CKEditor works fine in the tablet, though the specialist has difficulty using a text editor.

VII. Conclusion And Future Work

Almost Web sites and applications are developed thinking to be used with mouse and keyboard as input device, and medium size screen as output devices. But smart phones with touchscreen have enough computational power to access Internet, so these devices allow users browse into Web applications by touching. When a smartphone is used to access Web applications we can have modality and/or platform changing.

Changing the interaction style or platform brings interaction problems, impacting on usability. Which problems, if we consider the software have good usability in desktop computers, appears when we used a non-specified interaction style to browse a Web application? And which problems are not related with the platform changing, just only with the modality changing? We call cross-modality interaction problems. We developed this work trying to identify cross-platform and cross-modality interaction problems using TelEduc, an e-Learning environment developed to be used with mouse and keyboard as input devices, in touchscreen devices.

To identify problems, a HCI specialist analyzed the interaction and observed problems, classified them into platform and modality-independent problems, cross-platform problems and cross-modality problems. Through this work we prove software designed to be used with some interaction styles may have problems to be used with other interaction style. In this paper, we presented 6 problems to distinguish cross-platform and cross-modality problems, and show interaction problems that are more severe when the interaction style is changed.

It is important highlight we used devices with limitation, e.g., the algorithm who decides which user interface component the user touched. The algorithm accuracy may prejudice the interaction problem identification. Disregarding these limitations, it is clear the need to a better integration between platform and Web applications to increase the user experience, gestures need be more explored when user browsing.

Another important result is the perception of the changing interaction styles allows highlight existing usability problems.

As future works, we planned to study solution for these problems, the relationship between the problem and the code and study cross-modality problems for other input devices, like pen.

ACKNOWLEDGMENT

The authors thank the CAPES and CNPq for financial support and for FAPESP through TIDIA-Ae Project (n. 05/60572-1), which provided the equipment.

REFERENCES

[1] Moodle Trust. Moodle.org: open-source community-based tools for learning. Available at <http://moodle.org>. Accessed in January 2013.

[2] SAKAI Environment. Sakai Project | collaboration and learning - for educators by educators. Available at <http://sakaiproject.org>. Accessed in January 2013.

[3] TelEduc Environment. TelEduc Ensino à Distância. Available at <http://www.teleduc.org.br>. Accessed in October 2012.

[4] Moodle Trust. Moodle App. Available at: <http://docs.moodle .org/dev/Mobile_app>. Accessed in January 2013.

[5] M. J. Casany, M. Alier, E. Mayol, J. Piguillem, N. Galanis, F. J. García-Peñalvo, and M. Á. Conde. "Extending Moodle Services to Mobile Devices: The Moodbile Project". Proc. of the 6th International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM 2012), IARIA, 2012, pp. 24-28.

[6] S. Shrestha. "Mobile Web Browsing: Usability Study". Proc. of the 4th International Conference on Mobility Technology, Applications and Systems (Mobility 2012), ACM Press, 2007, pp. 187-194, doi: 10.1145/1378063.1378094.

[7] M.E. Maurer, D. Hausen, A. De Luca, and H. Hussman. "Mobile or Desktop Website? Website Usage on Multitouch Devices". Proc. of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries (NordiCHI 2010), ACM Press, 2007, pp. 739-742, doi: 10.1145/1868914.1869018.

[8] G. Schmiedl, M. Seidl, and K. Temper. "Mobile Phone Web Browsing –A Study on Usage and Usability of the Mobile Web". Proc. of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI´09), ACM Press, 2009, doi: 10.1145/1613858.1613942.

[9] A. Kaikkonen. "Full or tailored mobile web – where and how do people browse on their mobile?". Proc. of the International Conference on Mobile Technology, Applications and Systems (Mobility 2008), ACM Press, 2008, pp. 111-111, doi: 10.1145/1506270.1506307.

[10] M. A. Franco, L. M. Cordeiro, and R. A. F. del Castillo. "O ambiente virtual de aprendizagem e sua incorporação na Unicamp", Educação e Pesquisa, vol. 29, issue 2, Jun-Jul 2003, pp. 341-353.

[11] S. Papert, "Constructionism: a new opportunity for elementary science education". Proposal to The National Science Foundation. Massachusetts: Cambridge, 1986.

[12] J. A. Valente, "Computadores e conhecimento: repensando a educação". Campinas, Brazil: Gráfica da UNICAMP., 1993.

[13] J. Lave and E. Wenger, "Situated Learning – Legitimate peripheral participation". Cambrigde: Cambrigde University Press, 1991.

[14] J. A. Valente, "Formação de Professores: Diferentes Abordagens Pedagógicas". In O Computador na Sociedade do Conhecimento, J. A. Valente, Ed. Campinas, Brazil: UNICAMP/NIED, 1999, pp. 131-156.

[15] CKSource - Frederico Knabben. CKEditor - WYSIWYG Text and HTML Editor for the Web. Available at <http://www.ckeditor.com>. Accessed in January 2013.

[16] A. C. da Silva, H. V. da Rocha. "Improving e-Learning Environment With GPL Software - The visualization content case for computers and mobile devices". Proc. of the 6th International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM 2012), IARIA, 2012, pp. 222-227.

[17] Motorola Mobility Inc., 'Milestone Android – Phone Without Compromise - Tech Specs - Motorola Mobility India', Available at: <http://www.motorola.com/Consumers/IN-EN/Consumer-Products-and-Services/MobilePhones/ci.Motorola-MI LESTONE-IN-EN.alt>. Accessed in January 2013.

[18] HP TouchSmart tx2-1040br Notebook PC (NL266LA) specification. 2009. Available at: <http://h10025.www1.hp.com/ewfrf/wc/product?cc=us&lc=en&dlc=en&tmp_geoLoc=true&product=3937930>. Accessed in January 2013.

[19] Google. 'Android'. Available at: <http://www.android.com/> Accessed in January 2013.

[20] Google Inc. Chrome Browser. Available at: <http://www.google.com/chrome>. Accessed in January 2013.

[21] D. Vogel, and P. Baudisch. "Shift: A Technique for Operating Pen-Based Interfaces Using Touch". Proc. of the 25th International Conference on Human Factors in Computing Systems (CHI'07), ACM Press, 2007, pp. 657–666, doi: 10.1145/1240624.1240727.

[22] C. Holz and P. Baudisch. "The Generalized Perceived Input Point Model and How to Double Touch Accuracy by Extracting Fingerprints". Proc. of the 28th International Conference on Human Factors in Computing Systems (CHI '10), ACM Press, Apr. 2010, pp. 581–590, doi: 10.1145/1753326.1753413.

[23] D. A. Norman and J. Nielsen. "Gestural Interfaces: A Step Backwards In Usability", Interactions, vol. 17, issue 5, Sep.-Oct. 2010, pp. 46-49, doi: 10.1145/1836216.18362287.

[24] D. A. Norman. "Natural User Interfaces Are Not Natural", Interactions, vol. 17, issue 3, May-Jun. 2010, pp. 6-10, doi: 10.1145/1744161.1744163.

INTERACTION PROBLEMS ACCESSING E-LEARNING ENVIRONMENTS IN MULTI-TOUCH MOBILE DEVICES: A CASE STUDY IN TELEDUC

André Constantino da Silva[1], Fernanda Maria Pereira Freire[2], Alan Victor Pereira de Arruda[3] and Heloísa Vieira da Rocha[4]

[1]*Institute of Computing (PPG)/UNICAMP, IFSP, Campinas, Brazil / Hortolândia, Brazil*

[2]*NIED/UNICAMP, Campinas, Brazil*

[3]*ProFIS/UNICAMP, Campinas, Brazil*

[4]*Institute of Computing/UNICAMP, Campinas, Brazil*

ABSTRACT

e-Learning environments offer content, such text, audio, video, animations, using the Web infrastructure and they are designed to users interacting with keyboard, mouse and a medium-sized screen. Mobile devices, such as smartphones and tablets, have enough computation power to render Web pages, allowing browsing the Internet and access e-Learning environments. But, their main interaction style is touching, bringing some problems, cosmetics or severe ones. To identify these problems and correct them to allow mobile users access the TelEduc, an e-Learning environment, we planned a user test case with four volunteers using smartphones for access it. The collected data was analyzed and some problems were identified, giving some insights about the problems and barriers multi-touch mobile users dealt and can be easily corrected to give them a better interaction experience. To have an e-Learning environment with good usability in mobile devices is a first step to have a virtual learning environment for sharing data between desktop and mobile devices and understand the behavior of mobile users in this context.

KEYWORDS

Mobile devices and services; Interfaces, interactions and systems for distance education; Interface evaluation; Usability testing and evaluation.

1. INTRODUCTION

e-Learning environments, as Moodle (2013), SAKAI (2013) and TelEduc (2013), are Web applications with tools to support teaching and learning activities though the Web, taking advantages of the Web structure to offer content with text, images, audios and videos in a hypertext document. The environment´s tools allow users to create content, communicate with other users and manage the virtual space, so tools like chat, forums, portfolios, repositories are widely used, and tools that explore the audio and video resource to communication, such as instant messenger and video-conferences, are becoming common among the environments.

Mobile devices, such as smartphones and tablets, are becoming increasingly popular; most of them have touch screen displays, are easy to carry, have autonomy for hours, Internet access and enough computing power to process Web pages. So, Web sites and Web applications, initially developed to be used with keyboard, mouse and a medium-sized display, are been accessed by small touch screen devices. The e-learning environments fit in this context. Three kind of solution to allow mobile users access e-Learning environment are emerging: specific device application, Web site specific for mobile devices, and improvement of the environment for mobile and desktop access. Each solution had vantages and disadvantages. But, in the case of virtual teaching and learning environments, we need ask: is it fundamental have a specific environment for only mobile users? Since Web technology has a power to integrate different technologies, we believe in a virtual teaching and learning environment that allow any device accesses.

ISBN: 978-972-8939-88-5 © 2013 IADIS

Maurer *et al.* (2007) did some usability studies using touchscreen mobile devices and desktop for browsing in Web sites, shows that "more and more people prefer using original content instead of the mobile version, especially for users of new generation mobile devices like iPhone and Android phones". Another result of this work was the users prefer to use the standard Web site instead of tailored mobile versions of Web site. But Schmiedl *et al.* (2009) have a different opinion; in their research they conclude most of the users still prefer tailored versions. Kaikkonen (2008) shows that the standard and the tailored Web sites are both used but for slightly reasons.

Since there is not a clearly opinion in the researches about the users preference in using tailored mobile version or the standard interface, it is important to find and destroy barriers that mobile users meet when uses standard Web interface. In e-Learning context, which are the barriers or the problems the user dealt? To identify them in TelEduc environment we did user tests. Through this work we identified important design issues to allow mobile users access e-Learning environments and Web applications. Section 2 presents related work, Section 3 presents the TelEduc environment. The materials and method are described in Section 4. The identified problems are presented in Section 5, and Section 6 discusses some solutions. Section 7 presents the conclusion and future work.

2. RELATED WORK

The e-Learning environments´ development teams are building solutions to provide access on mobile devices. Three kind of solution are emerging: specific device application; Web site specific for mobile devices; and improve the Web site for mobile and desktop access.

Building specific device application allows designing a suitable user interface for the device and taking advantages of smartphone´s features, such as touchscreen and camera, but needs develop an application for each mobile platform, so to be developed needs specific knowledge programming team and increases the code lines number to maintain. Moodle community offers the Moodle App (2013) and Moodbile (2012), two native mobile applications with versions for the most popular smartphone´s platforms.

Moodle, since version 2.1, offers a Web site specific to mobile devices, an example for the second type of solutions for access e-Learning environments in mobile devices. Building a specific Web site to mobile device allows designing a suitable user interface for mobile devices taking account some common characteristics, such small touchscreen, but depends of the browser to access some platform features, such GPS, and increases the code lines number to maintain too.

The latter solution considers that smartphones and tablets have enough computational power to render Web pages and to do some adaptation if it is necessary, and offer the same user interface for any device. To design this kind of user interface it is necessary to do some usability studies to found barriers or user interaction problems. Disadvantages of this solution are to depend of browsers to use the mobile features and the difficult of consider many interaction styles in the same user interface.

This solution can start from a user interface design model for desktop and be improved to consider mobile devices. So in the initial design was designed thinking the user will interact by keyboard, mouse and medium size display and, allowing users accessed these applications on mobile devices, there is an increasing of interaction styles number, such touchscreen. With the interaction hardware changing the user deals with new interaction problems. Shrestha (2012) did user tests with volunteers using mobile devices equipped with joystick and a small screen and points out some problems when the users try to use mobile devices to do specific tasks into Web sites designed to desktop, so the mobile Web browsing experience needs to be improved to a more mobile friendly Web site and some mobile browser improvement (here we consider browser as one platform characteristic).

Considering only the e-Learning environment context, we agree to Maurer *et al.* (2010) when they argue the user prefer to use the standard version instead of mobile version of Web site, and this is one of the motivations of our work to study the third kind of solution: to access e-Learning environment using mobile devices. Shrestha (2012) studied the use of joystick and a small screen to browsing into some Web sites, while Maurer *et al.* (2010) considers touchscreen devices. All these researches point out some interaction problems, but the context was not e-Learning environments. In this work, we identified, through users test, problems when users use touchscreen mobile devices to access the TelEduc e-Learning environment.

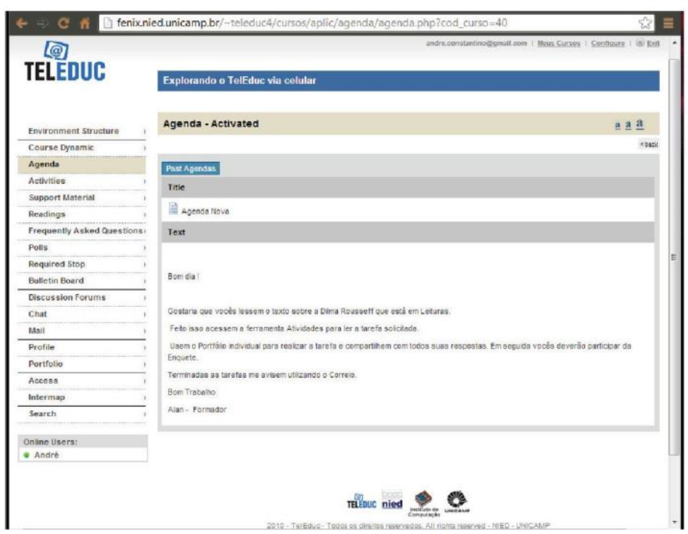
3. THE TELEDUC E-LEARNING ENVIRONMENT

The TelEduc is a teaching and learning environment developed by the Nucleus of Applied Informatics in Education (NIED) and the Institute of Computing (IC), State University of Campinas (UNICAMP), and adopted in several public and private institutions, like UNICAMP through Ensino Aberto project (Franco *et al.*, 2003).

The TelEduc environment was conceited in the end of 90, born with the Cerceau´s Master dissertation (1998), with professor Heloísa Vieira da Rocha as advisor, applying constructivism theory (Papert, 1986)(Valente, 1993) in situated learning (Lave and Wenger, 1991) or in contextualized learning (Valente, 1999) for teacher´s continuance formation. In 2001 February, the first free version was released over GNU General Public License (GPL), an unprecedented fact in the Brazilian Educational Software scenario. Many public and private institutions adopted the TelEduc as platform, increasing the TelEduc user's community, and consequently, the development demand. This fact culminated in the release of TelEduc version 3.0 in March 2002. The version 3.0 was completely redesigned and optimized, reason for TelEduc project was awarded by ABED (Brazilian Association for Distance Education) in the "Research and Development about Distance Learning" category. In August 2011, TelEduc version 4.3 was released, with its user interface redesigned to improve user tasks and be more similar than popular Web sites.

TelEduc is a system that aggregate administration, management and communication tools designed to support teaching and learning activities. Some tools allow users to create content, other ones allow synchronous or asynchronous communication among users, and manage participants and courses. The course page of TelEduc 4.3 is structured in two parts: the left one (Fig. 1a) has a list of all tools available and in the right one (Fig. 1b) the content of the selected tool.

In the course showed at Fig. 1a, a print screen of the course made for our user tests (the content is in Portuguese), the teacher dispose the Course Dynamic, Agenda, Activities, Support Material, Readings, Frequently Asked Questions, Polls, Required Stop, Bulletin Board, Discussion Forums, Chat, Mail, Profile, Portfolio, Access, Intermap and Search tools. Agenda is the ongoing course home page and shows the course´s program for a given period (daily, weekly, etc.). Agenda is an important tool because organize the activities that must be done in a specific period, similar teachers do in the beginning of a presence class. Fig. 1b shows the Agenda tool visualized by a student, where the student can read the agenda content and access the previous agendas.



(a) (b)

Figure 1. A desktop print screen of TelEduc course page (a) tool menu and (b) tool content.

ISBN: 978-972-8939-88-5 © 2013 IADIS

The Readings, Activities, Frequently Asked Questions and Support Material tools have similar features and user interface, but different purposes. Readings tool is used to publish relevant documents, like books, magazines, news and articles. The Activities tool is an area to publish activities to the accomplished during the course, like home work descriptions. The Frequently Asked Questions tool contains a list of the most frequently questions done by the participants during the course and their respective answers. Support Material offers a virtual space for teachers publish useful information about the course and the proposed activities.

Tools like Discussion Forums and Mail are used to participant communication, supporting text message exchange in asynchronous mode. To synchronous communication, there is the Chat tool, its features is similar to Web chat sites.

The Portfolio is a communication tool that aims to promote the collaboration among participants through the sharing of "items" (documents, presentations, programs, links, etc.). So the Portfolio tool provides an area to item storage and sharing for each participant (user or group of users) within a course. The Bulletin Board tool is a dedicated space where all the participants can post information considered relevant to the course content.

To provide content, TelEduc uses the Web infrastructure, more specifically, hypertext with images, links, audios and videos. All these media can be published as content in tools like Agenda, Support Material and Readings. To create text content, the user interacts with a rich text editor, the CKEditor (2013). CKEditor is a third-party WYSIWYG text editor to be used inside Web pages, bringing to the Web application common editing features found on desktop editing text applications, but CKEditor version 3.3 does not work on mobile devices. In previous work (Da Silva, & da Rocha, 2012), we studied some problems in use of third-party software on TelEduc to create and to visualize documents, describing that mobile compatibility needs to be considered to not prejudice mobile users.

The Agenda, Activities, Support Materials, Readings, Mail and Portfolio tools allow attaching files into them items in similar way: the user clicks in the "Attach File" button, select the file to be attached using a dialog and, after the click on the save button, the file is uploaded.

Since the e-Learning environments need to be easy to use for users with different levels of Web experience, the usability is an important nonfunctional requirement. TelEduc was designed in an iterative design-evaluation process to have good usability and the user interface does not impair teaching and learning activities, so many usability evaluations were done. The accessibility is another nonfunctional requirement desired for TelEduc, to allow impaired people to use the environment without meet barriers or obstacles.

TelEduc was designed to use a mouse and keyboard as input devices, and a medium screen size as output device. TelEduc is better visualized into 1024x200 pixels screen resolution. Visualize it into a lower screen resolution cause some user interface problems like dispose interface components in wrong position. Beside the problems happen due the resolution changing, which problems and barriers the mobile users met when interact with TelEduc? To answer this question, we developed a user test case described in the next section.

4. MATERIALS AND METHOD

To identify interaction problems when an e-Learning environment is accessed by a mobile device we planned a study case composes with four users tests. The main activities to do the study case were: (i) tools to be investigated and tasks definition; (ii) hardware to be used; (iii) questionnaires planning; (iv) preparation of the course.

The following sub-sections explain briefly each activity. We started choosing the tools to be evaluated and defined the tasks for user test sessions using the smartphones available, an Android-based and an iPhone.

Based on our experience in using e-Learning environments, we chose the tools: Agenda, Activities, Readings, Polls, Mail and Portfolio. For each tool was defined one task, considering the most important feature of the tool. All tools were used considering the student´s view. The defined tasks were: 1. Login in TelEduc; 2. Read the Agenda; 3. Read the Reading and do it; 4. Read the Activity; 5. Post her activity in her portfolio; 6. Answer a poll; 7. Send e-mail to the teacher; 8. Logout of TelEduc.

To do the task 5, the volunteers need to search into the Web for an article. The purposeful activity is to allow us identify interaction problems when the users need to navigate in two web pages, visualizing one page and going to another. In the tasks planning we found some interaction problems that impair the user to conclude the tasks, so we shaped the defined activities such a way the user does not meet the identified problems with the goal to find new interaction problems. The problems identified on task planning are discussed in Section 5.1.

We used a Motorola Milestone smartphone (Motorola Mobility, 2013) and an Apple iPhone 3GS (Apple, 2013). The Motorola Milestone has a 3.7 inches multi-touch display with 133MB internal storage memory expansive up to 32 GB with a memory card, 600 MHz Cortex-A8 processor and 256 MB RAM, a 5MP camera, GPS and wireless connection by Wi-Fi 802.11 b/g and bluetooth. Android 4.0.3 (Google, 2013) was used as operation system (OS). Motorola Milestone has a proximity sensor, an ambient light sensor, a 3-axis accelerometer and a geomagnetism sensor to provide orientation with respect to Earth´s magnetic field. To browse in the Web application, the volunteers use the Android stock Web browser.

The Apple iPhone 3GS (Apple, 2013a) has a 3.5 inches multi-touch display with 32 Gb internal storage memory, 600 MHz Cortex-A8 processor and 256 MB eRAM, a 3MP camera, GPS and wireless connection by Wi-Fi 802.11 b/g and bluetooth. The installed operation system is iOS version 5 (Apple, 2013b). iPhone has too a proximity sensor, an ambient light sensor, a 3-axis accelerometer and a geomagnetism sensor. To browse in the Web application, the volunteers use the Safari browser. In both devices, the proximity sensor and the accelerometer can be used to interact with applications, but the browsers do not use these features as input. Only the touchscreen was used as input device.

To collect information about the volunteers and them opinion we elaborated two questionnaires. The first one was to collect data about user age, genre, internet usage and mobile skills. The second one was to get personal satisfaction in do the tasks. Due we did an experiment involving people, we prepared a consent form to be signed by the volunteers informing her about the study purposes, the anonymity of the volunteer, and allowing recording the interaction to further analysis.

After the task planning, we prepared a course in the TelEduc installed in our server with the agenda content, the activity description, an article published on Readings tool, and the poll to be answered. This preparation is to simulate a course in the environment with chained activities. Since the user tests laboratory and the server share the same local network, we did not worry about Internet lacks or fail.

We invited four undergraduate students to participate, one man and three women, whom filled the profile questionnaire before the participation. All volunteers have 20 years old, three of them have smartphone and is mobile users for more than six month, the last one does not have a smartphone and her consider not be a mobile user.

5. DATA ANALYSIS

In this section we discuss about the found problems when mobile users access the TelEduc Environment. We found one problem during the tasks definition, and discuss it here since impose a barrier and do not allow mobile users conclude their goal. Analyzing the user tests data, we found 5 relevant problems for mobile users and other usability problems that is device independently. We focused here in the mobile problems and in the end of the section we present solutions.

5.1 Problem found in the Task Planning

TelEduc uses CKEditor to allow users write rich text instead of simple plain text, but CKEditor does not work in iPhone and in Android-based devices, bringing a barrier to user write rich text. For example, it is not possible put content in a Portfolio item because the content field does not appear (*Problem 1*). Looking in the environment for others barriers caused by CKEditor, we found that it is possible to write message in Mail tool and post Bulletin Board posts, but the content field appears to write simple text since CKEditor is not displayed. In the other environment tools the barrier was found.

Focused on discovery new interaction problems we planned the tasks such a way the users do not match this problem. The next sub-section presents the problems discovery through the user tests.

ISBN: 978-972-8939-88-5 © 2013 IADIS

5.2 Problems found in the User Tests

The first identified problem in user tests (*Problem 2*) is related with the web site does not follow the smartphone interface standard. The first time the user 2 tried to go to another tool, she touched in the space between the wanted menu item the symbol '>' (Fig. 2a shows the point tapped by the user in the interface), since no action was triggered, she tried a lot of times again in the same place. Asked about the unexpected system response, the user touched in the tool label and the system shows the tool content. Maybe the user thought the TelEduc interface follows the smartphones standard, which is possible to trigger an action touching in the space nearby the menu item label (Fig. 2b). In desktops it is softened due to change the mouse cursor from an arrow to a pointing hand, giving a visual clue that it does not have in touchscreen mobile.

(a) (b)

Figure 2. Print screens marked with the place tapped by the user in (a) TelEduc tool menu and no action was triggered and (b) in iPhone menu settings, where an action was triggered.

Another problem observed during the user tests was related with the selection function (*Problem 3*). Three users needed help to select the desirable URL to perform the copy and paste functions. Maybe due the users are familiar with desktops, they were expected for a menu or a toolbar with options. Asked about how they interact with the device, they realized about the touch interface and they tried to tap the desirable text, and, in a second try, tapped the text and hold the finger over the text, and the interface responses with the options. The problem is about the visibility and discoverability of the actions. We do not realize that we will meet this problem, since the all volunteers are the generation who use mobile devices since childhood, but they have a mental model more related with the desktop computers than mobile devices.

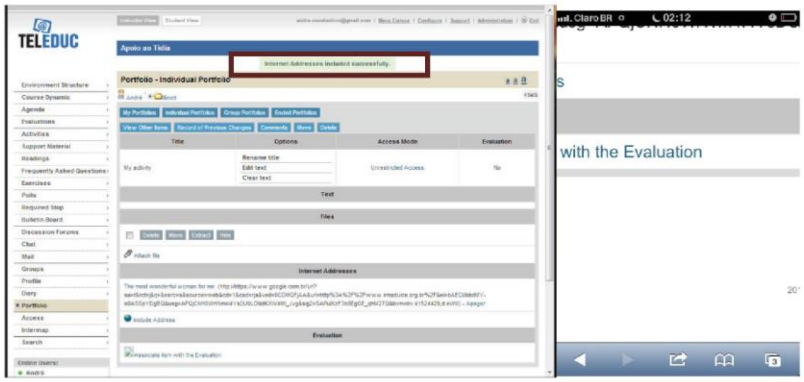
After copy the URL to do the defined activity in task 4, two users had another problem: how to return to TelEduc, since another tab was opened (*Problem 4*). The button to trigger the action to show opened tabs is at bottom on iPhone and is in at top on Android-based smartphones. One of the users prefers typing the TelEduc address again instead of searching for how to go to an opened tab. The other user asked for help.

Most of TelEduc feedback appears on the page top (Fig. 3a), and to draw the user´s attention the feedback blinks. In mobile devices, with the user is visualizing another part, the feedback is not visible (*Problem 5*). One example happens when the user is posting a URL in the Portfolio item (doing the task 5), and after to tap the "Ok" button, the system response with the feedback "Internet Addresses included successfully." in the page top, but the user does not see since she was interaction with the elements in the bottom (Fig. 3b). It is important to highlight that the TelEduc form to show feedbacks differs from the smartphones form to do it. In smartphones feedbacks were showed as a popup dialog.

The last identified problem is about the time a feedback message is displayed (*Problem 6*). When the user sends a message, the TelEduc shows a feedback page to inform that the message was successful sent for 2000 milliseconds. Since the smartphones take more time to render the page and shows all the page content, the waiting time expires; the browser closes the feedback page almost immediately, the user can read the feedback message, giving the impression that an error happed.

Analyzing the user opinion questionnaire and considering all identified problems, the users agree with TelEduc have good usability. Three of the four volunteers described feel difficult to interact, and the last one related does not feel any difficult to interact.

5.3 Solutions



(a) (b)

Figure 3. Print screens of TelEduc feedback on (a) desktop and (b) mobile devices after the user click in the "Ok" link to post the link in the Portfolio item to do the task 5.

Problem 1 is about the mobile user cannot write the content in most of TelEduc due the CKEditor is not displayed on mobile devices. A good solution is to allow mobile user only write plain text, similar happen to write messages in TelEduc Mail Tool. Analyzing the Portfolio tool code, we discovery that the content field, a textarea tag, was not showed due a style property (display:none) that CKEditor changes when it loads. Since mobile devices do not load the CKEditor, this property is not changed. Changing the code to get the content from the value property when CKeditor is not load, and get the content from CKEditor when it loaded, resolve the problem, destroying this barrier.

Problem 2 (user tapping in the space between menu item label and the symbol '>' to trigger an action) is easily to resolve. The tool menu is implemented with table and anchor tags, just changing the order putting the li tag inside the <a> tag, becoming all the cell area a link.

Problem 3 (trigger a text selection) and Problem 4 (return to a previous tab) are related with the smartphones operation system. Norman and Nielsen (2010) discuss discoverability and visibility problems in touchscreen devices, and they argue there is a lack of established guidelines for gestural control.

A solution for the Problem 5 (TelEduc feedback is fixed at the top) is show feedback messages using JavaScript dialog when the device have JavaScript support. Since browsers for mobile devices have JavaScript engine inside, the browser have the responsibility to show the message in the visual standard used in the mobile device. This solution may be applied to the Problem 6 (feedback time), giving the user the feeling of control over the application.

6. CONCLUSION AND FUTURE WORK

Mobile devices, such as smartphones and tablets, are becoming increasingly popular; most of them have touch screen displays, Internet access and enough computing power to process Web pages, allowing mobile users access Web applications, such e-Learning environments. The e-Learning environments were initially designed to be used with keyboard, mouse and a medium size display, and now have to consider their access in mobile devices. We believe that Internet is a meeting point for all devices, so the Web applications need to be access by any device. Which kind of barriers and problems do mobile users meet when browsing in an e-Learning environment? We did a user test case with four volunteers accessing the TelEduc environment through iPhone and Android-based smartphones, and discovery 6 problems that impact in the usability. Do usability studies with users is a hard work, but it is necessary due the lack of techniques to do this kind of evaluation. The discovered problem is related with the consistence between device interfaces and TelEduc interface, discoverability and visibility of the features and code structure.

e-Learning environments have a lot of contributions to m-Learning, offering a platform to teachers and students do their teaching and learning activities, but some technologic gaps need to be filled. With this work, we started with the first step finding and eliminating barriers and interaction problems that become hard mobile users access e-Learning features. Another step is taking advantages of the mobile device features in

ISBN: 978-972-8939-88-5 © 2013 IADIS

pro learning activities such create content like photos and videos. More research is need, since the interaction style changed and different features are available, e.g., the mobile users write text with fewer words? Analyzing the mail message sent from our volunteers to teachers to inform the activity was done, we noted just a sentence was write. Is it happened due to the profile of our younger volunteers or due the mobile device restrictions? How this influences in the activities in the environments? Mobile users prefer record a video or write a big text? These questions needs more research to be explained, but needs an e-Learning environment without barrier to be accessed on mobile devices.

ACKNOWLEDGMENT

The authors thank the CAPES and CNPq for financial support and for FAPESP through TIDIA-Ae Project (n. 05/60572-1), which provided the equipment.

REFERENCES

Apple, 2013a. *iPhone 3GS – Technical Specification*s. Available at: <http://support.apple.com/kb/SP565> [Accessed in January 2013].

Apple, 2013b. iOS 5. Available at: < http://developer.apple.com/library/ios/#releasenotes/General/WhatsNewIniPhoneOS /Articles/iOS5.html> [Accessed in January 2013].

Casany, M. J. et al, 2012. Extending Moodle Services to Mobile Devices: The Moodbile Project. *Proceedings of the 6th International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies*, Barcelona, Spain, pp. 24-28.

CKSource - Frederico Knabben, 2013. *CKEditor - WYSIWYG Text and HTML Editor for the Web*. Available at <http://www.ckeditor.com>. [Accessed in January 2013].

Da Silva, A. C. and da Rocha, H. V., 2012. Improving e-Learning Environment With GPL Software - The visualization content case for computers and mobile devices. *Proceedings of the 6th International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies*, Barcelona, Spain, pp. 222-227.

Franco, M. A. et al, 2003. O ambiente virtual de aprendizagem e sua incorporação na Unicamp. *In Educação e Pesquisa*, Vol. 29, No. 2, pp. 341-353.

Google, 2013. *Android*. Available at: <http://www.android.com/> [Accessed in January 2013].

Kaikkonen, A., 2008. Full or tailored mobile web – where and how do people browse on their mobile?. *Proceedings of the International Conference on Mobile Technology, Applications and Systems*, I-Lan, Taiwan, pp. 111-111.

Lave, J. and Wenger, E., 1991. *Situated Learning – Legitimate peripheral participation*. University Press, Cambrigde, USA.

Maurer, M.E. et al, 2007. Mobile or Desktop Website? Website Usage on Multitouch Devices. *Proceedings of the 6th Nordic Conference on Human-Computer Interaction*, Reykjavik, Iceland, pp. 739-742.

Moodle Trust, 2013a. *Moodle.org: open-source community-based tools for learning*. Available at <http://moodle.org>. [Accessed in January 2013].

Moodle Trust, 2013b. *Moodle App*. Available at: <http://docs.moodle .org/dev/Mobile_app>.[Accessed in January 2013].

Motorola Mobility Inc., 2013. *Milestone Android – Phone Without Compromise - Tech Specs - Motorola Mobility India*, Available at: <http://www.motorola.com/Consumers/IN-E N/Consumer-Products-and-Services/MobilePhones/ci.Motoro la-MI LESTONE-IN-EN.alt>. [Accessed in January 2013].

Normal, D A. and Nielsen, J., 2010. Gestural Interfaces: A Step Backward In Usability. *In Interactions*, Vol. 17, No. 5, pp. 46-49.

Papert, S., 1986. *Constructionism: a new opportunity for elementary science education*. Proposal to The National Science Foundation, Massachusetts, USA.

SAKAI Environment, 2013. *Sakai Project | collaboration and learning - for educators by educators*. Available at <http://sakaiproject.org>. [Accessed in January 2013].

Schmiedl, G. et al, 2009. Mobile Phone Web Browsing –A Study on Usage and Usability of the Mobile Web. *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services*, Bonn, Germany.

Shrestha, S., 2007. Mobile Web Browsing: Usability Study. *Proceedings of the 4th International Conference on Mobility Technology, Applications and Systems*, Singapore, pp. 187-194.

TelEduc Environment, 2013. *TelEduc Ensino à Distância*. Available at <http://www.teleduc.org.br>. [Accessed in October 2012].

Valente, J. A., 1993. *Computadores e conhecimento: repensando a educação*. Gráfica da UNICAMP, Campinas, Brazil.

Valente, J. A., 1999. Formação de Professores: Diferentes Abordagens Pedagógicas. *In O Computador na Sociedade do Conhecimento*, Editor Valente, J. A. UNICAMP/NIED, Campinas, Brazil, pp. 131-156.

Capítulo 6

Interação Baseada em Caneta e Baseada em Multitoque em Ambientes de EaD

Este Capítulo, conforme descrito na Seção 3.2, é formado por dois artigos apresentando um trabalho pioneiro no contexto da plataforma Web, ambientes de EaD e interação baseada em caneta: o InkBlog, uma ferramenta de blog que possibilita os usuários postarem mensagens manuscritas ou desenhos elaborados por meio de uma caneta. Até o momento, as ferramentas que exploram o conceito de Pen-Based Computing eram especificamente destinadas ao dispositivo baseado em caneta e fora da plataforma Web.

Para o levantamento das funcionalidades implementadas considerou-se a experiência obtida ao realizar o trabalho descrito no Capítulo 4, gerando uma primeira versão da ferramenta. O resultado da implementação do InkBlog foi publicado como artigo na *Technology Enhanced Learning Environments* que compõe a *Informing Science + IT Education Conferences*. O artigo foi premiado como um dos melhores artigos da conferência e foi publicado no *Journal Issues in Informing Science and Information Technology*.

Após estudos do uso da ferramenta InkBlog em dispositivos móveis com interação a toque e a caneta, uma nova versão da ferramenta foi gerada, suportando também a interação multitoque. Os resultados foram submetidos e aceitos para publicação na sexta edição da *International Conference on Mobile, Hybrid, and On-Line Learning*.

Para receber os dados provenientes da caneta e do toque na elaboração de desenhos e rascunhos foi criado um componente em linguagem JavaScript, o InkEditor. Esse componente pode ser integrado a outras ferramentas do ambiente, ou outra ferramenta Web, possibilitando que novas ferramentas sejam criadas.

Issues in Informing Science and Information Technology Volume 10, 2013

InkBlog: A Pen-Based Blog Tool for e-Learning Environments

André Constantino da Silva
Instituto Federal de São Paulo, Hortolândia, São Paulo, Brazil
CPG-IC, State University of Campinas, Campinas, São Paulo, Brazil

andre.constantino@ifsp.edu.br

Heloísa Vieira da Rocha
IC – NIED, State University of Campinas, Campinas, São Paulo, Brazil

heloisa@ic.unicamp.br

Abstract

Weblogs are communication and collaborative tools disposed in Web that aims to publish texts and sharing opinions, typically displaying the post in reverse chronological order, and allowing visitors to leave comments. Since devices equipped with pen are becoming common, such as tablets, and have enough computational power to render Web pages, users can access blog tools through pen-based devices, writing new posts and read and comment published posts. But tablets have a different input style and the user need to type the post instead of handwriting it, decreasing the usability and make boring the writing task in these devices. Pen-based computing refers to a computer user-interface using a pen, rather than devices such as a keyboard or a mouse, and is a promise research area to enhance Education, but it is need to properly consider the pen. We purpose the InkBlog, a blog tool that receive input data from pen so that the user can handwriting her posts. In this work we present the InkBlog architecture and the main component, the InkEditor for Web pages. Using InkBlog the user can write equations or sketches faster than using a mouse or a keyboard, getting the best of the pen-based devices.

Keywords: pen-based interaction, e-Learning environments, user interface design, usability.

Introduction

Blog software is a communication and collaborative tool disposed in Web that aims to promote the sharing of messages among participants through an area named blog. Users can publish texts, images, audio, videos and links and sharing opinions in posts typically displayed in reverse chronological order (the most recent post appears first) and allowing visitors to leave comments. In this way, blogging can be seen as a form of social networking.

In education, blogs can be used as instructional resources, referred to as edublogs. Edublogs archive and support student and teacher learning by facilitating reflection, questioning by self and others, collaboration (Ray, 2006) and by providing contexts for engaging in higher-order thinking. Ray (2006) cites four ways to incorporate blogs into the classroom, including: (i) using them to communicate information to students and parents; (ii) to provide instructional resources and useful links; (iii) to allow students the opportunity to collaborate with one another on various projects without being in the classroom itself; and (iv) to showcase student work and projects, like poetry and photographs of project work.

In e-Learning environments, each course participant (student or teacher) owns a blog, where she can post or exclude messages. Each participant may access other course participants' blogs, comment posted messages by the owner of the blog. In Weblogs, to write a post, the users interact with in a text editor that allows users without HTML (HyperText Markup Language) skills to write rich text.

Mobile devices, such as smartphones and tablets, are becoming increasingly popular; most of them have touch screen displays, Internet access and enough computing power to process Web pages. So it is possible to access Weblogs, read the messages, post new messages and write comments. But it is important consider that these pages are developed to be used with keyboard, mouse and a medium size display, and the changing of interaction styles brings interaction problems (da Silva, Freire, & da Rocha, 2013) and does not take advantages of the interaction style features. For example, to post a message in a Weblog using a pen-based device the user needs to type each letter pressing the respective key in a virtual keyboard. This way of writing the text take a lot of time, make boring the writing task and do not take the mainly pen purpose: handwriting and do sketches easily.

We believe that the e-learning environment needs to be improved to be more easy to use in some contexts, e.g., areas which need sketches or drawing such Mathematics. To make it easier to handwrite or sketched posts in pen-based devices, we add features to manipulate electronic ink into a blog tool and called InkBlog. This work describes the InkBlog tool and presents some uses in e-Learning. The next Section presents a literature review about electronic ink technology, the e-Learning environments focusing in Weblog. The following Section presents the InkBlog, describing how the technologies are employed to allow users handwriting posts and comments. The last Section presents the conclusion and future works.

Literature Review

The recent technology forwards brings a broad of new applications made to support many areas, such Education. One of the mobile devices that are gained repercussion in this scenario is the Tablet PC (Tablet Personal Computer), a device with height as similar to a notebook and had an input device similar a pen. The paper and pen metaphor implies that tasks performed before in paper, like draw or manuscript writing, can be more natural in the Tablet PC than in the another computing devices. Resuming, the Tablet PC has the following hardware characteristics: (i) Pen sensitive screen; (ii) Screen that allow different positions; (iii) Wireless network access by WLAN (Wireless Local Area Network) and bluetooth technology; (iv) Microphones and embedded loudspeakers; (v) Keyboard (some models the keyboard are detachable); (vi) Batteries.

Pen-based Computing or Pen Computing refers to a computer user-interface using a pen, rather than devices such as a keyboard or a mouse. User interfaces for pen computing can be implemented in several ways, like using the pen as a pointing input device, but takes better advantages when the application is developed considering direct manipulation, handwriting or gesture recognition. Educational applications for contexts where paper and pen usually are presents can benefits from this new interaction style and from pen-based computing, offering better experiences of use for teachers and students.

Berque, Bonebright and Whitesell (2004) describes the use of pen-based electronic classrooms covering topics from the undergraduate computer science curriculum such as two-dimensional arrays, pointer-based linked lists, binary search trees, directed and undirected graphs, digital logic diagrams, and finite state automata. The authors say that these concepts are very difficult to communicate quickly and extemporaneously using a keyboard and they are also difficult to describe orally, studying the use of Tablet PC to cover these topics. Through a questionnaire, they conclude that pen-based computers are well-received in the Computer Science classroom.

Florea and Radu (2007) present a study using the pen-based technology to improve the quality and efficiency of the instructional process in a computer science course using concept maps, showing how the pen-based technology can support several learning styles. They developed the Pen Annotator, a tool to do graphical annotation on documents, and @Graph, a tool to develop and manage concept maps. Classroom Presenter (Anderson *et al*., 2007) is a Tablet PC-based classroom interaction system that supports the sharing of digital ink on slides between instructors and students to increase the instructor's flexibility while lecturing.

Benlloch, Buendía and Cano (2009), using the Classroom Presenter, discuss an approach with six steps based on a teaching method that incorporates active learning techniques and supports in-class teacher/student and student/student collaborations. They implemented in a first-year Computer Engineering Course, using a wireless classroom equipped with 25 Tablet PCs for students and one additional for the instructor, which is connected to the data projector. Benlloch, Buendía and Cano (2009) propose sketching in the early stages of project-based approaches using Tablet PCs and polls to gather formative information and in turn providing immediate feedback, activities supported by Classroom Presenter.

Pargas et al. (2007) presents a Tablet PC application, called OrganicPad, which enable an instructor to engage her students in class by sending them Chemistry problems to solve. The students develop answers to the problems guided by tips and hints provided by the software. Labahn *et al.* (2008) propose MathBrush, a tool that allows users to draw math input using a pen-input device on a tablet computer, recognizes the math expression, and then supports mathematical transformation and problem solving. The authors argue that entering mathematics on a computer is problematic, it is more natural write the formulae than inputting the latex form, maple form or mathematic form.

McCormick (2007) points out that, based on a study with primary (3-11 years) classrooms in the United Kingdom, a high degree of interaction and collaboration and a greater ownership over their work when use interactive whiteboards and other pen-based learning technologies on pedagogical practice. Backon (2006) says that the keyboard allows a rapid typewriting and do texts structured by topics, but only pen can allow a creativity increment, better flexibility and a natural record of the think development process and rationale in all stages (ideal requirements for the educational environment), aging like a direct hand extension. For example in the resolution of complex problems with a large calculus solution, by using applications is possible to record all the steps and draw diagrams that explain the problem resolution, step by step.

When the user moves the pen in the screen, the pen trace should result in electronic ink that must be treated by the application to be rendered and stored. But desktop applications, that running in the Tablet PCs, do not treat electronic ink, so it is necessary special applications to treat the electronic ink to have benefices of the pen interaction style. This is the case of Classroom Presenter and the other described pen-based software. In our previous work (da Silva & da Rocha, 2012) we studied three applications (Professional Adobe Acrobat, Windows Journal and Jarnal) for do classes annotations or writing manuscript in student activities using a Tablet PC, describing some identified usability problems. Desktop applications do not manipulate the electronic ink, resulting in loose of the Tablet PC potential because the pen is used only as a pointer device. In this case, the use of pen as a pointer device in desktop applications brings some usability problems, such the difficult to see the content when the user needs to cross the screen with her hand. There is similar lack of technologies too in web browsers, and since pen based devices has computational power enough to render Web pages, there is a need to fill these gap.

The World Wide Web has changed since its invention from a static to a highly dynamic media in the recent years, so the term "Web 2.0" was coined in 1999 to describe web sites that use technology beyond the static pages and its uses for collaborative, user-centric content production and interactive content access (O´Reilly, 2005). Safran, Helic and Gütl (2007) describe that in literature the marks of Web 2.0 includes (1) social phenomena such as the Web for participation, (2) technology for significant change in web usage, and (3) design guidelines for loosely coupled services. The Web 2.0 allows users to interact and collaborate with each other in social networking sites, weblogs, podcasts, wikis, video sharing sites and other sort of tools.

One kind of Web applications is e-Learning environments, as Moodle (http://moodle.org), SAKAI (http://sakaiproject.org) and Ae (http://tidia-ae.iv.org.br/), applications with tools to support teaching and learning activities though the Web and have some Web 2.0 characteristics. Tools in these environments allow users to create content, communicate with other users and manage the virtual space. Tools like chat, forums, portfolios, repositories are widely used, and tools that explore the audio and video resource to user communication, such as instant messenger and videoconferences, are becoming common among the environments. These environments take advantages of the Web to offer content with text, images, audios and videos in a hypertext document.

The mainly Web technology, the HTML, evolved too and the improvement defined to the last version, the HTML5, are related with support multimedia, keep it easily readable by humans and consistently understood by computers and devices (Berjon *et al.*, 2012). HTML5 adds the new <video>, <audio> and <canvas> tag elements, as well as the integration of Scalable Vector Graphics (SVG, a vector image format for two-dimensional graphics based on eXtended Markup Language - XML) content and MathML (Mathematical Markup Language is a XML based-format to describing mathematical notations) to integrate mathematical formulae into Web pages. These features are designed to make it easy to include and handle multimedia and graphical content on the web without having proprietary plugins and APIs (Application Programming Interface) installed.

The <canvas> tag allows for dynamic, scriptable rendering of 2D shapes and bitmap images, allowing a drawable region defined in HTML code with height and width attributes. JavaScript code may access the area through a full set of drawing functions similar to those of other common 2D APIs, thus allowing for dynamically generated graphics.

W3C, the World Wide Web Consortium who standardizes HTML, defines XML formats for non-primitive data to allow exchange of a wide variety of data on the Web and elsewhere, and one example is InkML (Chee *et al.*, 2011). The InkML (Ink Markup Language) provides a common format for the exchange of ink data between components such as handwriting and gesture recognizers, signature verifiers, sketches, music and other notational languages in applications. The InkML serves as the data format for representing ink entered with an electronic pen or stylus. It is possible to find some uses of InkML, such in Microsoft Word 2010 support electronic ink in text review and the InkML JavaScript Library (http://inkml.codeplex.com/), that offers some functions to allows InkML digital ink to be referenced within Web pages and rendered directly into the HTML5 <canvas> tag.

Considering the technology breakthrough that HTML5 purposes, most of the web sites uses HTML5 to impress the user in the content exhibition and few take care about the user input interaction styles, developing the web pages considering the user is interacting with keyboard and mouse in desktop computers. But this scenario is changing with the smartphone and tablet popularization: the web designers need to think about the other interaction styles, such as touchscreen and pen-sensitive devices.

Another characteristic of Web 2.0 is that they would be used by a variety of devices and platforms. Devices have Internet access and have enough computing power to render Web pages and using good practices to design Web pages is possible to render in small devices. But some devices have interaction hardware different from desktops; some have touchscreen and other pen-sensitive screen, such Tablet PCs.

Berque, Bonebright and Whitesell (2004), Florea and Radu (2007), Anderson *et al.* (2007), McCormick (2007), Benlloch, Buendía and Cano (2009) points out benefits in using Tablet PC in classrooms. These works and da Silva and da Rocha (2011) studied the use of Tablet PC in teaching or learning activities in or out classrooms in presence courses. But how to get the pen computing advantages in Web tools, mainly in e-Learning environments? Users with pen-based devices, such as tablet, can access Weblogs and easily read the posts, but writing blog post is not too easy: the user needs to type the text; even she is using a pen. To make it easier to handwrite posts and allow users post sketches in pen-based devices, we add features based on HTML5 technologies to manipulate electronic ink into a blog tool and called InkBlog. The remainder of this section presents the Ae environment and its Weblog tool, the software we used as base to create the InkBlog.

Ae Environment

The TIDIA-Ae Project (TIDIA-Ae is the acronym for "*Tecnologia da Informação para o Desenvolvimento da Internet Avançada – Aprendizado Eletrônico*", in English "Information Technology for Development of Advanced Internet – Electronic Learning") was initiated by FAPESP (the State of São Paulo Research Foundation) with the main goal of developing an e-Learning environment that can explore the potential of Advanced Internet and can provide support to different educational context needs. To provide a flexible system the Ae project defined a layered component-based architecture for Ae environment. The advantages of using component-based software architecture, and which justify its adoption, are related to the following characteristics that are obtained:

(i) **Abstraction**: explicit separation between the specification and the implementation of functionalities; and

(ii) **Uniform composition**: in order to assemble functionalities implemented by distinct components, there must be communication between these components' provided and required interfaces, thereby providing uniform composition. In this way, it is possible to reuse components (Reusability) developed by other educational institutions, bringing increased productivity and quality of the produced software.

Layered component-based software architecture was defined for the TIDIA-Ae Project (Beder, Casagrande, & Rubira, 2004) with the following layers:

- Presentation layer: provides the application user interface;
- System layer: provides an interface for the application functionality;
- E-Learning layer: provides the component interfaces that implement the application's business rules, which can be used by various applications;
- Infrastructure layer: implements a set of infrastructure services such as, for example, data persistence;
- Common services layer: has the public services that can be utilized and accessed by all other architecture layers, except the presentation layer.

Following this architecture the Ae project developed a set of tool that was integrated with Sakai environment nucleus to compose the Ae environment. One Ae´s tool is Weblog, an edublog tool. The next session explain its purpose and its architecture.

The Weblog Tool

The Weblog tool makes possible to the user share with other participants texts, video or audio messages. The posted messages may be commented by the readers of the Weblog. Each participant (student or teacher) has a blog are for each course (Figure 1). The participant can post messages in her blog, clicking in the link "Compose a new post" in the user blog page. A page within a form to be filled will be displayed (Figure 2). After fill up the form, the user click in the "Confirm" button and the system will redirect the user to the user blog page displaying the new message at top.

The user can read messages posted by other participants accessing the participant´s blog, and comment the message posted by other participants in their blog area clicking in the link "Post a comment". The system will display a form to be filled in a new window (Figure 3), and to submit the new comment the user needs to click in the "Confirm" button. The new comment will be displayed when the user will click in the "Show comments" link next to the commented message.

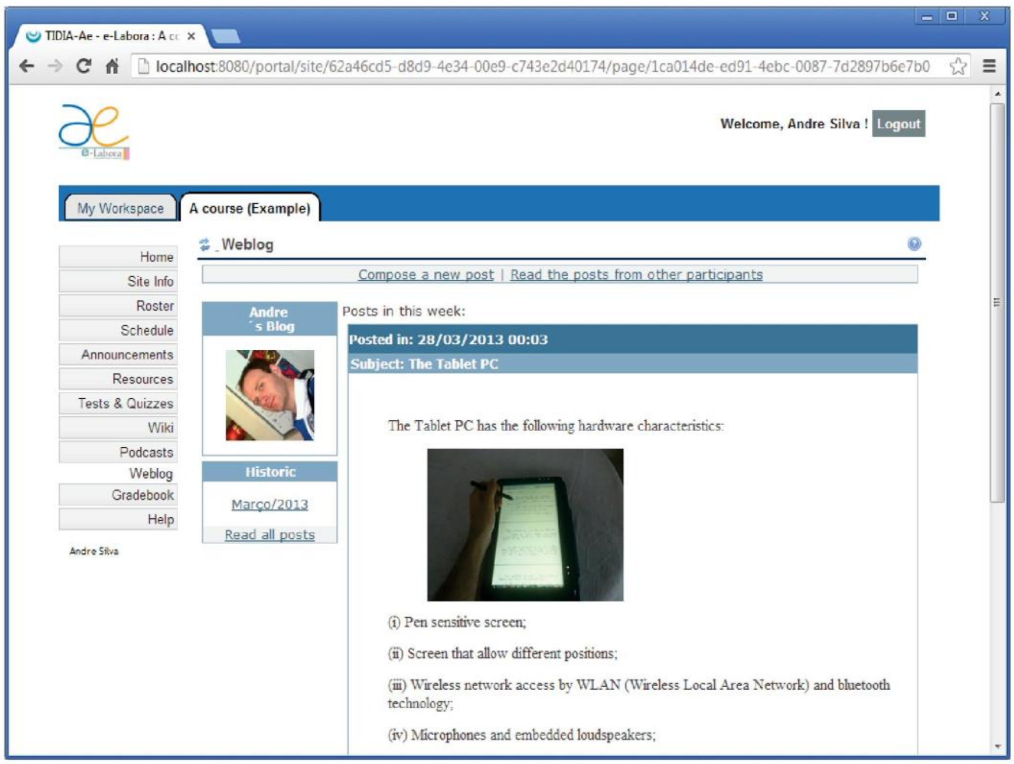


Figure 1: A participant weblog area in Ae´s Weblog tool rendered by Google Chrome.

Weblog tool was developed following the Ae architecture and developing platform (Beder *et al.*, 2007), described in the above Section, and its presentation layer was codified using Struts and Tiles (da Silva *et al.*, 2006). Struts is a framework to extends the Java language to develop software based on a Model–View–Controller (MVC) architecture and Tiles is a template mechanism that allows the presentation layer to be composed from independent header, footer, and content components making more easy develop Web pages avoiding code replication.

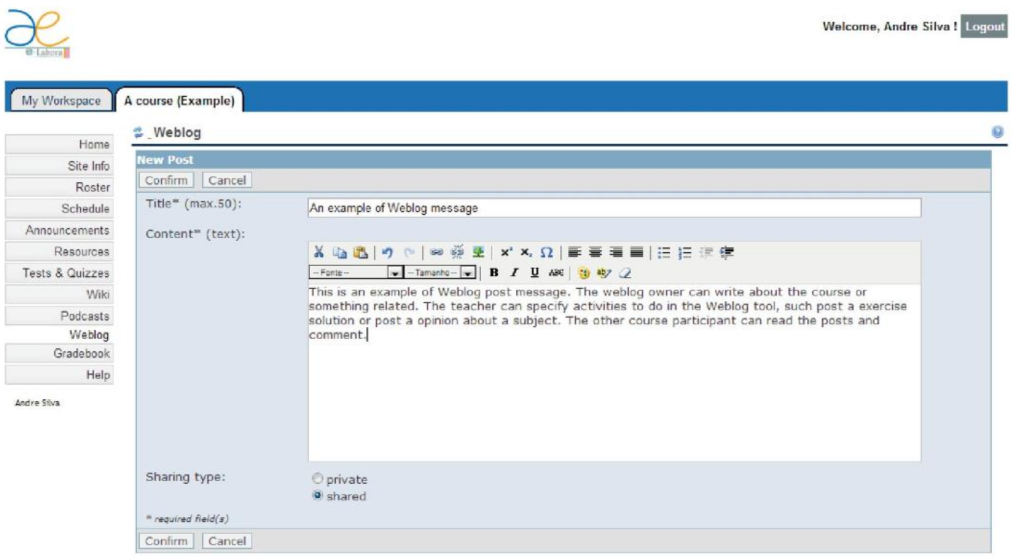


Figure 2: User posting a new message in her weblog.

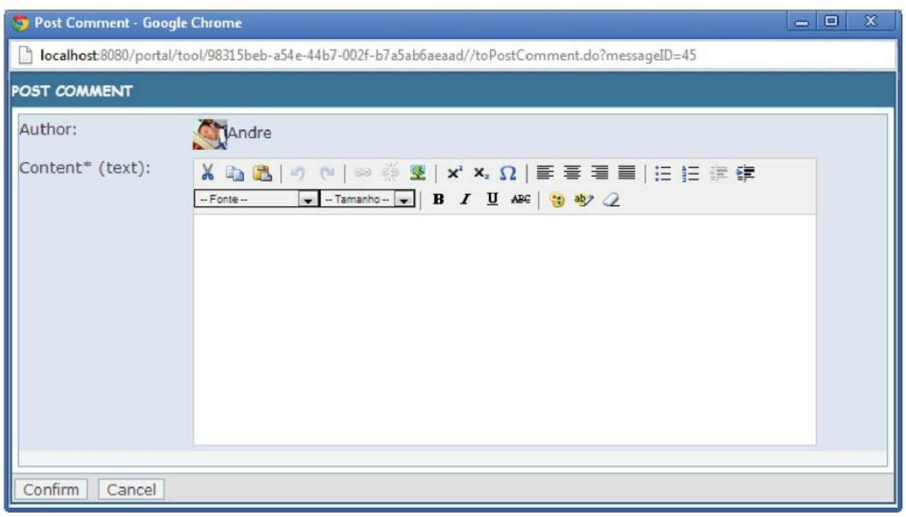


Figure 3: A participant writing a comment for a posted message in other participant weblog area.

The Weblog Tool Architecture

This section presents the results from the phase of mapping the software components of the Weblog tool onto the architecture defined by the TIDIA-Ae project. Figure 4 presents the components responsible for the implementation of the functionalities of this tool distributed in the layers defined by the Ae architecture.

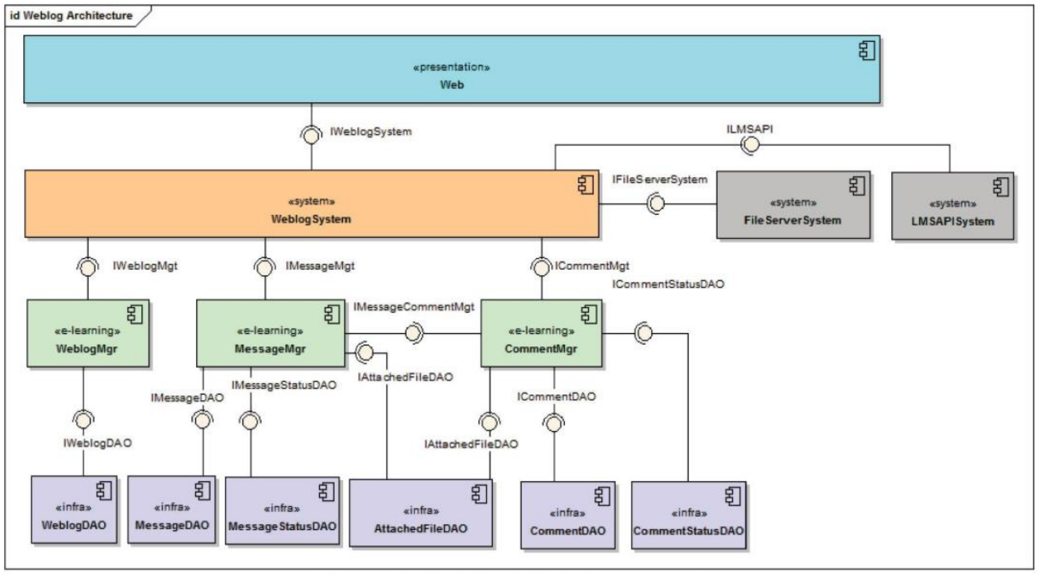


Figure 4: Weblog Tool Components.

We used J2EE (Java Enterprise Edition, an API and runtime environment for developing and running enterprise software in Java language) as the development platform in the implementation of the Ae tools. In relation to this technology, the following decisions were taken:

- Tomcat (http://tomcat.apache.org) as the J2EE application server;
- Java Beans were used in the system and e-learning layers according to the decisions of the TIDIA-Ae Architecture Group;
- Hibernate (http://www.hibernate.org) was also adopted as the main persistence framework.

The InkBlog Tool

The InkBlog was created to make it easier to handwrite posts and comments using pen. We increased the Weblog tool with components to generate and manipulate the electronic ink in the user interface, representing the electronic ink in InkML format. Figure 5 shows the InkBlog architecture, based on Weblog architecture. The client device needs to have a compatible HTML5 browser to run the InkEditor, a user interface component that will render the electronic ink and

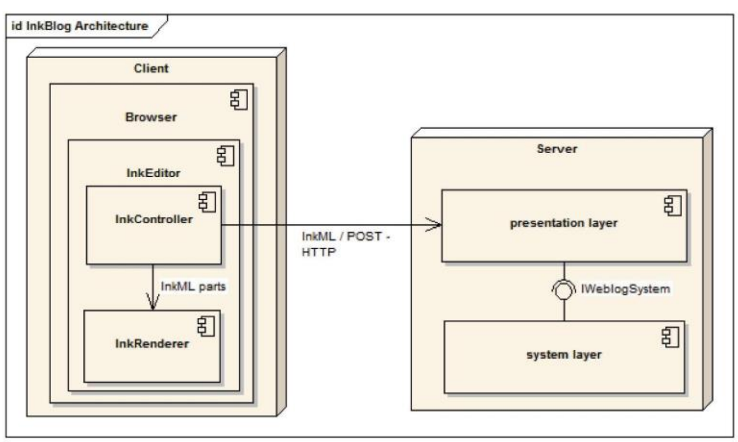
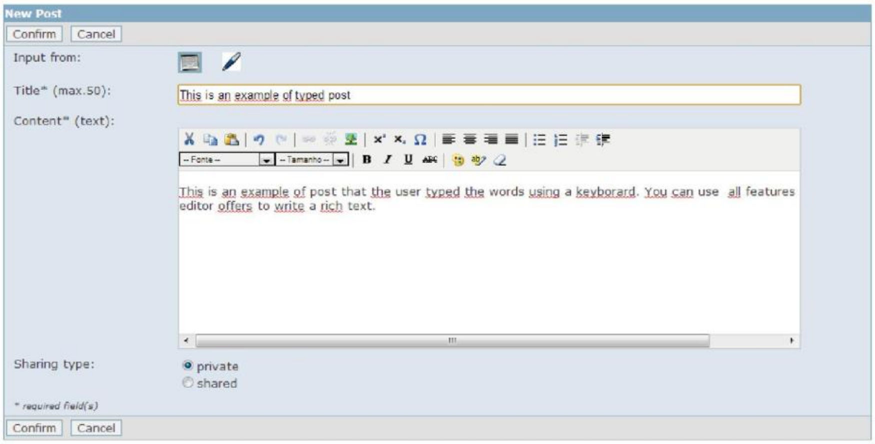
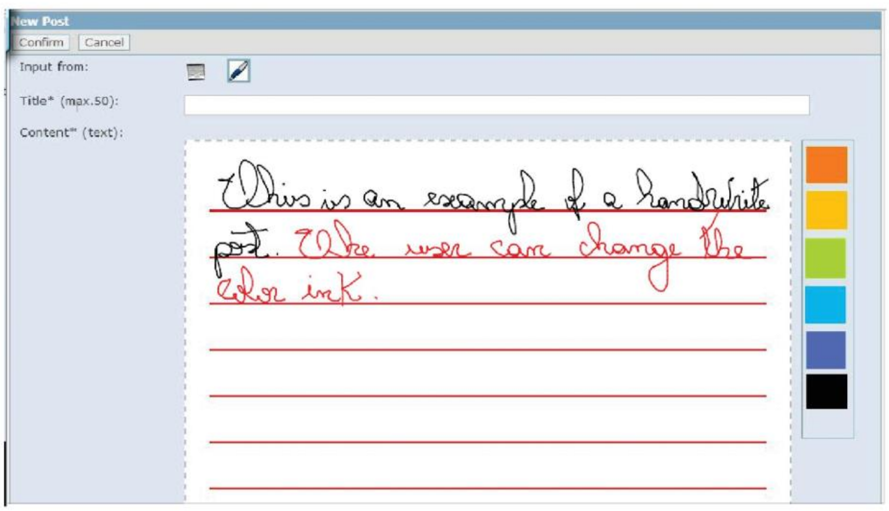


Figure 5: InkBlog Tool Components used in a handwritten or sketched post.

will receive the input data generated by the pen. The InkEditor uses InkML to represent the handwriting data and the Canvas HTML attribute to draw the traces in the screen.



(a)



(b)

Figure 6: New post composite page to (a) type the message using keyboard and (b) handwrite the message using a pen.

To post a new message, the user can choose, selecting the icon on "input from:" field, to type the text using a keyboard (Figure 6a) or to handwrite a post, selecting the Pen option in the New Post page (Figure 6b). When the user chooses post a handwriting post, the browser will hidden the text editor and show the InkEditor, where the user will use a pen to handwriting. When the user touches the InkEditor within the pen and draw a trace, the InkController, an InkEditor component, it will listen the user actions, getting the dots that compose the trace. Each dot is recorded and a line connecting the preceding point to the new point is drawn in the canvas until the user releases the pen. After the pen release, the InkController will generate the InkML´s trace node for the new trace. The user can draw how many traces she wants, all them will be stored and will compose the InkML file. When finalize to handwrite the post, the user will click in the Confirm button, and the generated InkML file will be sent to the server to be stored.

The pen input data is received by the InkController, who transform each point of the trace into coordinate points following InkML format. The user can choose the trace´s color and the width selecting the buttons options in the right side. When the user points out and presses the pen into a color or width button, the next traces will have the brush attribute set to look likes the selected options.

The InkRenderer, another InkEditor component, draws the traces of the handwrite posts in the user screen, following the posts display order (Figure 7). InkRenderer´s code, the electronic ink data in InkML format and the HTML page are sent to the client over the HTTP protocol (Figure 8). After all the data and code arrived in the client, the InkRenderer reads the InkML data resident in the canvas´ data-inkml attribute, and draw the electronic ink for each trace, taking account the ink formatting. The InkRenderer was developed using the InkML JavaScript Library.

Some changes were need in the server to distinguish the textual content from the typewriting content and show the correct editor in the post view. The changes are done in the presentation component; the other components were not changed.

It is possible too handwrite comments and post them: the process is similar to the described process above.

Figure 7: The user´s posts page.

To test the InkBlog we used a HP TouchSmart TX2-1040br, a 2.2 GHz dual-core processor computer with 3 GB RAM and a 12” touchscreen with Windows Vista operating system and with Windows 8 operating system. This model has the design similar to HP laptops but it is equipped with the described hardware for Tablet PC, for our purposes, a pen sensitive screen. Google Chrome browser version 22.0.1229.94m and Mozilla Firefox version 15.0.1 were used to navigate through the InkBlog and demonstrated good performance in both browsers.

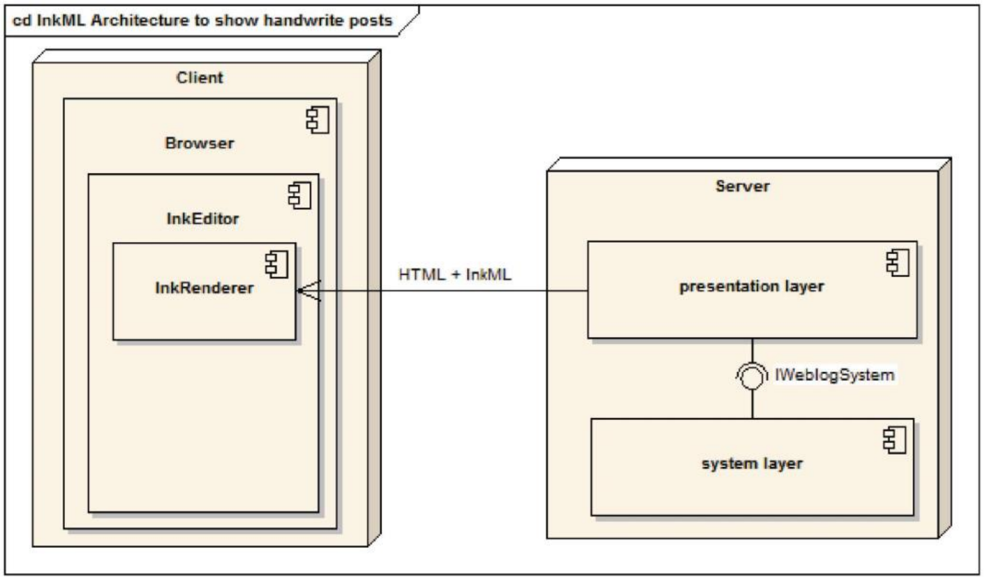


Figure 8: InkBlog Tool Components used to render a handwritten or sketched post.

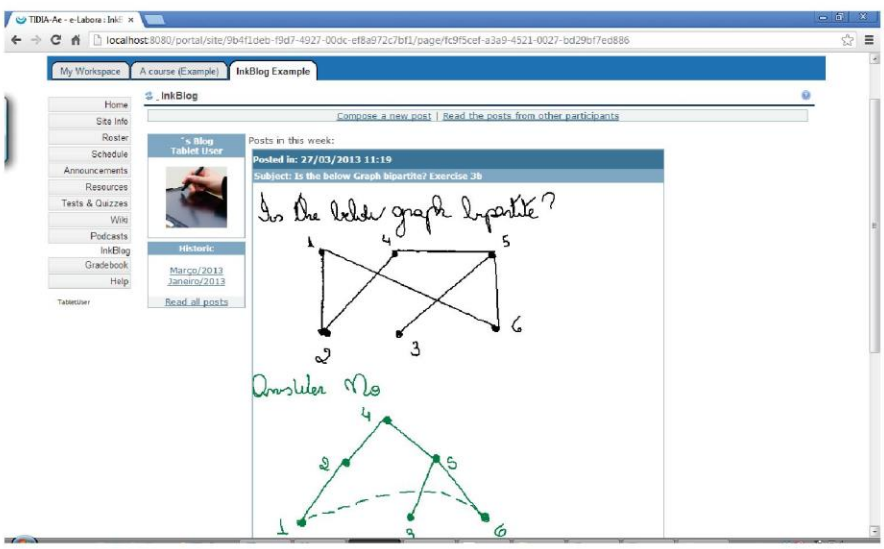


Figure 9: Using InkBlog to share a solution for an exercise about Graph Theory rendered by Chrome Browser.

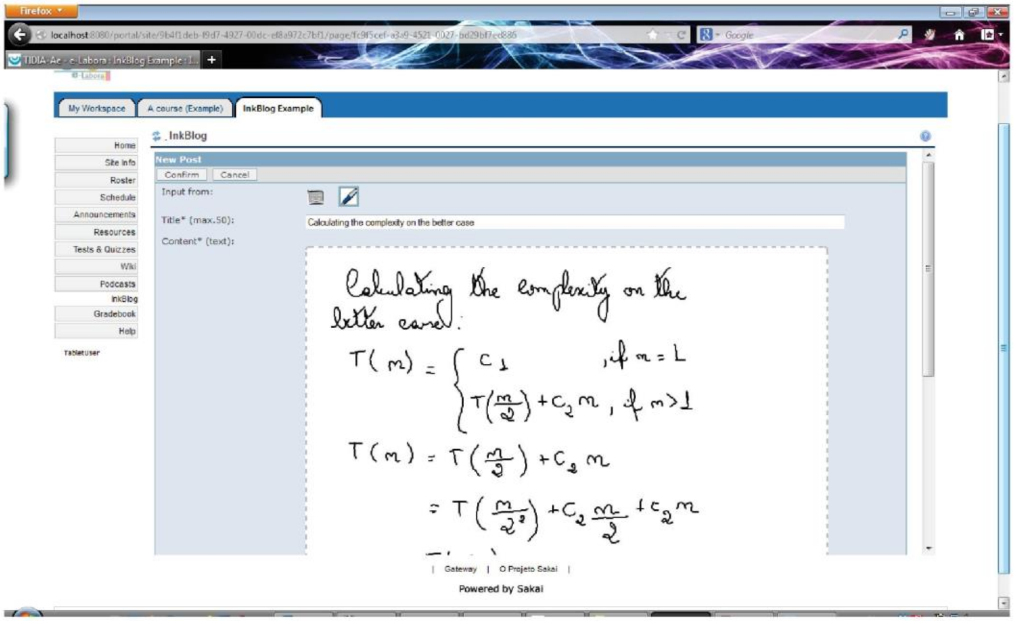
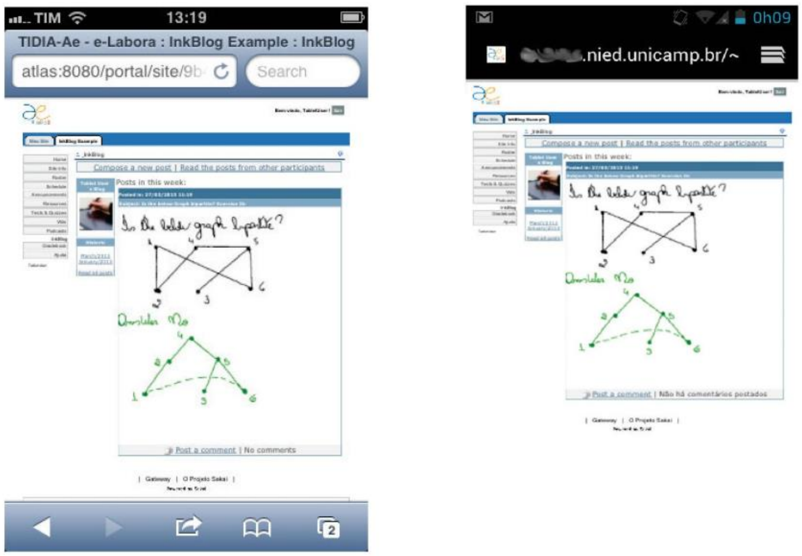


Figure 10: Using InkBlog to handwrite a reasoning to prove the better case of an algorithm in Computer Theory course rendered by Mozilla Firefox.

Using the InkBlog is possible a better support to disciplines like Graph Theory and Computer Theory. Without the InkBlog, the user needs to use paper and pencil to resolve an exercise and use specific hardware, such scanner, to digitalize the solution. Or the user needs to use a special application to draw a graph. In both solutions, the user posts the picture as an attached file in weblog post. Using the InkBlog the user can sketch the graph direct on the weblog tool through direct manipulation (Figure 9). Figure 10 shows the use InkBlog to handwrite a reasoning to prove the best case for an algorithm.

We tested the InkBlog in iPhone and Android devices (Figure 11a and 11b) using the stock browser in each device. Both devices display correct the posts, but due the platform does not distinguish between touch and stylus press, it is not possible to handwrite a post in these smartphones; both devices recognize the input as page scrolling.



(a) (b)

Figure 11: InkBlog rendered by stock browser on (a) iPhone and (b) Android based-device.

Conclusion and Future Works

Blog software is a communication and collaborative tool disposed in Web that aims to promote the sharing of messages among participants, and some of them are used to educational purposes, called EduBlogs. Blog can be used in many ways in classes (Ray, 2006) and it is one of the available tools in e-Learning environments, such as Weblog in Ae environment. Weblog tools use a text editor to allow users typing rich texts without HTML skill, but when the user interact with a pen she needs to typing each letter, so the usability, most specific, the efficiency decreases and making boring the writing task. Another problem is the difficult to draw sketches using the mouse. These usability problems were deal by many researches before in other contexts, e.g., when the user needs to write mathematic formulae or draw a diagram. The studied researches got the solution on pen-based computing. These researches developed solution to be installed on the computer and studied in the classroom, concluding that pen-based computing may enhance the Education in classroom. How to improve the e-Learning environment to get advantages of the Pen-based computing and enhanced the Education in distance learning? One solution is, when the user access the environment using a pen-based device, allow the user does handwriting posts and draws sketches to do her activities. We improve the Ae´s Weblog with components to generate and manipulate electronic ink, calling this new tool as InkBlog.

In the InkBlog development, we chose well-defined and promising technologies, such as W3C InkML and HTML5, allowing to codify a light electronic ink editor for Web pages, the InkEditor. Some modifications are done in the Weblog for distinguish the correct editor to be displayed in the new post composite page and display handwriting posts and comments. We think allowing the Web application to manipulate the electronic ink we are powering the application with the best that the pen can offer. Using the InkBlog, users can handwriting and sketches posts and comments, powering the blog usages in areas like mathematics, architecture, fashion and music. But we need to validate this insight with some studies.

For future works, we will increase the InkEditor functionalities, such as ink selection, copy and paste features. Another future work is implements gesture to trigger functionalities by the integration with the environment with a gesture recognizer, and study the gesture´s draw for each function. And the integration with a handwriting recognizer to recognize the word and allow a better integration with another environment tools, such search tool. Another feature to be implemented is use a post or a document as background, allowing teachers or other student to do annotation over a post.

Acknowledgment

The authors thank the CAPES and CNPq for financial support and for FAPESP through TIDIA-Ae Project (n. 05/60572-1), which provided the equipment.

References

Anderson, R., Anderson, R., Davis, P., Linnell, N., Prince, C., Razmov, V., et al. (2007). Classroom presenter: Enhancing interactive education with digital ink. *Computer*, *40*(9), 56-61. doi:10.1109/MC.2007.307

Backon, J. (2006). Student minds and pen technologies: A wonderful pedagogical marriage. In D. A. Berque, J. C. Prey, & R. H. Reed (Eds), *The impact of pen-based technology on education: Vignettes, evaluations, and future directions* (pp.1-11). West Lafayette, IN: Purdue University Press.

Beder, D. M., Casagrande, L., & Rubira, C. (2004). *Proposta de Arquitetura de Software para o Projeto TIDIA-Ae* [TIDIA-Ae Project Software Architecture Proposal], Internal Document, TIDIA-Ae Project, São Paulo, Brazil. Retrieved January 25, 2012.

Beder, D. M., da Silva, A. C., Otsuka, J. L., da Silva, C. G., & da Rocha, H. V. (2007, July). A case study of the development of e-learning systems following a component-based layered architecture. *Proceedings of the 7th IEEE International Conference on Advanced Learning Technologies,* Niigata, Japan, 21-25. doi: 10.1109/ICALT.2007.4

Benlloch, J., Buendía, F., & Cano, J. (2009, July). Tablet PC-based learning approach on a first-year computer engineering course. *Proceedings of the 9th IEEE International Conference on Advanced Learning Technologies,* Riga, Latvia, 86-87. doi: 10.1109/ICALT.2009.155

Berjon, R., Leithead, T., Navara, E. D., O'Connor, E., & Pfeiffer, S. (2012, December). *HTML5 - A vocabulary and associated APIs for HTML and XHTML W3C candidate recommendation*. Retrieved March 25, 2013 from http://www.w3.org/TR/html5/

Berque, D., Bonebright, T., & Whitesell, M (2004, March). Using pen-based computers across the computer science curriculum. *Proceedings of the 35th SIGCSE Technical Symposium on Computer Science Education*, Norfolk, USA, 61-65. doi: 10.1145/1028174.971324.

Chee, Y., Franke, K., Froumentin, M., Madhvanath, S., Magaña, J., Pakosz, G., et al. (2011, September). I*nk markup language (InkML) W3C recommendation*. Retrieved March 25, 2013 from http://www.w3.org/TR/InkML/ .

da Silva, A. C., Buchdid, S. B., dos Santos, D. T., & da Rocha, H. V. (2006, November). A process to implement the user interface improved by struts and tiles. *Proceedings of the III Workshop TIDIA FAPESP*, São Paulo, Brazil. Retrieved March 25, 2013 from http://www2.icmc.usp.br/~francisco/works/IIITidia workshop.pdf .

da Silva, A. C., & da Rocha, H. V. (2012, February) Experiences in use Tablet PC to support student's activities: Five years of an exploratory study. *Proceedings of 4th International Conference on Mobile, Hybrid, and On-line Learning*, Valencia, Spain, 38-43. Retrieved March 25, 2013 from http://www.thinkmind.org/index.php?view=article&articleid=elml_2012_2_30_50076

da Silva, A. C., Freire, F. M. P., & da Rocha, H. V. (2013, February) Identifying cross-platform and cross-modality interaction problems in e-learning environments. *Proceedings of 6th International Conference* *on Advances in Computer- Human Interactions*, Nice, France, 243-249. Retrieved March 25, 2013 from http://www.thinkmind.org/index.php?view=article&articleid=achi_2013_9_50_20417

Florea, A. M., & Radu, S. (2007, May). Enhancing pen-based experiences with the use of concept maps. *Proceedings of 1st International Workshop on Pen-Based Learning Technologies*, Catania, Italy, 1-6. doi: 10.1109/PLT.2007.10

Labahn, G., Lank, E., MacLean, S., Marzouk, M. S., & Tausky. D. (2008, September). MathBrush: A system for doing math on pen-based devices. *Proceedings of the 8th IAPR International Workshop on Document Analysis Systems*, Nara, Japan, 599-606. doi: 10.1109/DAS.2008.21

McCormick, D. (2007, May). Interactive whiteboards in the classroom: A primary teacher's perspective. *Proceedings of 1st International Workshop on Pen-Based Learning Technologies*, Catania, Italy, 1-4. doi: 10.1109/PLT.2007.12

O'Reilly, T. (2005). *What Is Web 2.0 - Design patterns and business models for the next generation of software*. Retrieved March 25, 2013 from http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html

Pargas, R., Cooper, M., Williams, C., & Bryfczynski, S. (2007, May). OrganicPad: A tablet PC based interactivity tool for organic chemistry. *Proceedings of 1st International Workshop on Pen-Based Learning Technologies*, Catania, Italy, 1-4. doi: 10.1109/PLT.2007.4

Ray, J. (2006). *Welcome to the blogosphere: The educational use of blogs*. Kappa delta pi record, p. 175-177.

Safran, C., Helic, D., & Gütl, C. (2007, September). E-Learning practices and Web 2.0. *Proceedings of International Conference on Interactive Collaborative Learning*, Villach, Austria. Retrieved March 25, 2013 from http://www.iicm.tugraz.at/iicm_papers/ICL2007_csafran_final.pdf . doi: 10.1.1.161.8553

Biographies

MSc. **André Constantino da Silva** is Professor on Institute Federal of São Paulo (IFSP), Campus Hortolândia, and he is a post graduate student in Institute of Computing (IC) of State University of Campinas (UNICAMP). His topics of interest include e-Learning, m-Learning, mobility, multimodalities, educational software and educational games development.

Dr. Heloísa Vieira da Rocha is Professor in Institute of Computing (IC) and researcher at the Nucleus of Informatics Applied on Education (NIED) of State University of Campinas (UNICAMP). She lectured in the areas of artificial intelligence, human-computer interaction and informatics applied on education at both undergraduate and postgraduate levels. Her main topic of interest is the junction of the Information Technologies in Education and Human-Computer Interfaces areas.

Improving e-Learning Environments for Pen and Multi-touch Based Interaction

A study case on blog tools and mobile devices

André Constantino da Silva
Institute of Computing (PG)
UNICAMP, IFSP
Campinas, Brazil, Hortolândia, Brazil
acsilva@ic.unicamp.br

Heloísa Vieira da Rocha
Institute of Computing, NIED
UNICAMP
Campinas, Brazil
heloisa@ic.unicamp.br.br

***Abstract*— e-Learning environments are applications that use the Web infra-structure to support teaching and learning activities; they are designed to have good usability using a desktop computer with keyboard, mouse and high resolution medium-size display. Devices equipped with pen and touch sensitive screen have enough computational power to render Web pages and allow users to navigate through the e-learning environments. But, pen-based or touch sensitive devices have a different input style; decreasing the usability of e-learning environments due the interaction modality change. To work on mobile contexts, e-learning environments must be improved to consider the interaction through pen and touch. In a previous work, we presented the InkBlog, a blog tool that receives input data from pen allowing users to handwrite posts. In this work, we present an extension of InkBlog to receive input from multi-touch screens. We described the changes over the InkBlog architecture and implemented components to treat data generated by pen and touch.**

Keywords-Human-Computer Interaction; Electronic Learning Environment; Mobile Devices; Interaction Styles.

I. INTRODUCTION

e-Learning environments, such as Moodle [1], SAKAI [2], TelEduc [3], Ae [4], are applications that use the Web infra-structure to support teaching and learning activities. The e-Learning environments are designed to support a variety of users and learning contexts, but they are designed to conventional computers, usually equipped with keyboard and mouse as input and a medium screen and speakers as output; a limited set of interaction styles for nowadays devices. These modalities, and the technology that support them, shape the teaching and learning activities done in the e-Learning environments; they focus on reading and writing skills.

One example of tools available at e-learning environments is Weblog, a communication and collaborative tool that aims to promote the sharing of messages among participants through an area named blog. Users can publish texts, images, audio, videos and links, sharing their opinions, in posts typically displayed in reverse chronological order (the most recent post appears first) and allowing visitors to leave comments. In this way, blogging can be seen as a form of social networking.

Mobile devices, such as smartphones and tablets, are becoming increasingly popular; most of them have touch screen displays, Internet access and enough computing power to process Web pages. So, it is possible to access blog tools, read the messages, post new messages and write comments through mobile devices. But, it is important to consider that these tools (and so their Web pages) are developed to be accessed by desktop computers equipped with keyboard, mouse and a medium size display; in your previous work we described that when a user interface designed for a set of interaction styles is accessed by a different set of interaction styles the users face interaction problems [5]. Another problem is that it is not possible to take advantage of the interaction style features; for example, in a desktop computer, users use the keyboard to typing the post text. In a pen-based computer without handwrite recognition, users need to type each letter pressing the pen in the respective key in a virtual keyboard. This way of writing texts takes a lot of time, make boring the writing task and do not take the mainly pen purpose: handwriting and do sketches easily. In the case of touch sensitive screen, the user can touch in the virtual keyboard to write the post, but it is not possible to do sketches.

So, we believe that blog tools and e-learning environments need to be improved to be easier to use in some devices and some contexts, e.g., areas which need sketches or drawing such Mathematics. In our previous work [6], we developed the InkBlog tool to easily write handwrite or sketched posts in pen-based devices by adding features to manipulate electronic ink into a blog tool from Ae, an e-Learning environment. But, the InkBlog have some limitations: display handwriting posts correctly in smartphones, but it is not possible to sketch post. In this work, we add some features to improve the InkBlog to allow users to write messages in many devices: desktop computers, smartphones and tablets.

Section II presents a literature review about electronic ink technology and e-Learning environments, focusing on Weblog tool. Section III presents the InkBlog, describing how the technologies are employed to allow users handwriting posts and comments using pen-based devices. In Section IV, we present the modifications done in InkBlog to allow users to interact using multiple fingers. The last section presents the conclusion and future work.

II. E-LEARNING ENVIRONMENTS AND INTERACTION STYLES

The World Wide Web has changed since its invention from a static to a highly dynamic media in the recent years; so, the term “Web 2.0” was coined in 1999 to describe the

 ISBN: 978-1-61208-328-5

web sites that use technology beyond the static pages and its uses for collaborative, user-centric content production and interactive content access [7]. Safran, Helic and Gütl [8] describe that in literature the marks of Web 2.0 include: (i) social phenomena, such as the Web for participation, (ii) technology for significant change in web usage, and (iii) design guidelines for loosely coupled services. The Web 2.0 allows users to interact and collaborate with each other in social networking sites, weblogs, podcasts, wikis, video sharing sites and other sort of tools.

One kind of Web applications that have some Web 2.0 features is e-Learning environments, as Moodle [1], SAKAI [2] and Ae [3]. They are applications with tools to support teaching and learning activities though the Web. Tools in these environments allow users to create content, communicate with other users and manage the virtual space. Tools like chat, forums, portfolios, repositories are widely used, and tools that explore the audio and video resource to user communication, such as instant messenger and video-conferences, are becoming common among the environments.

HyperText Markup Language (HTML) is used for any web application to describe the page interface and its content. Usually, in web applications that users post text, there is a rich text editor to allow users write formatted text without HTML skills. In desktop computers, the users use the keyboard to typewrite the letters, and use the mouse to select format functionalities (some of them have shortcuts to be triggered by the keyboard). Since the rich text editors have a direct manipulation interface similar as text editor applications, it is easy to be used in desktop computers equipped with mouse and keyboard.

The HTML have some improvement defined in the last version, the HTML5, related with support multimedia, keep it easily readable by humans and consistently understood by computers and devices [9]. HTML5 adds the new <video>, <audio> and <canvas> tag elements, as well as the integration of Scalable Vector Graphics (SVG, a vector image format for two-dimensional graphics based on eXtended Markup Language - XML) content and Mathematical Markup Language (MathML, a XML based-format to describing mathematical notations) to integrate mathematical formulae into Web pages. These features are designed to easily include and handle multimedia and graphical content on the web without having proprietary plugins and Application Programming Interface (APIs) installed.

The <canvas> tag allows for dynamic, scriptable rendering of 2D shapes and bitmap images; it is a drawable region defined in HTML code with height and width attributes. JavaScript code may access the area through a full set of drawing functions similar to those of other common 2D APIs, thus allowing for dynamically generated graphics.

Another evolution in HTML is standardizing how the browser must handle events from touch and pointer inputs [10]. The W3C Group specified that "The Touch Events specification defines a set of low-level events that represent one or more points of contact with a touch-sensitive surface, and changes of those points with respect to the surface and any DOM (Document Object Model) elements displayed upon it (e.g., for touch screens) or associated with it (e.g., for drawing tablets without displays)". This specification was done thinking in devices equipped with stylus such as tablet and define event types for: (i) when a user touch the surface (touchstart), (ii) when a user remove a touch point from the surface (touchende), (iii) when a user moves a touch point along the surface (touchmove), (iv) to indicate a touch point has been disrupted (touchcancel). Having different event types for input data generated by each modality gives flexibility for the developers define the actions to be trigger for each modality.

W3C defines XML formats for non-primitive data to allow exchange of a wide variety of data on the Web and elsewhere; one example is Ink Markup Language (InkML) [11]. The InkML provides a common format for exchange ink data between components such as handwriting and gesture recognizers, signature verifiers, sketches, music and other notational languages in applications. The InkML serves as data format for represent ink gathered by an electronic pen or stylus. It is possible to find some uses of InkML, such in Microsoft Word 2010 support electronic ink in text review and the InkML JavaScript Library [12], that offers some functions to allows InkML digital ink to be referenced within Web pages and rendered directly into the HTML5 <canvas> tag.

Considering the technology breakthrough that HTML5 proposes, most of the web sites use HTML5 to impress users through content exhibition; few developers take care about the user input interaction styles, so they develop web pages for users with keyboard and mouse in desktop computers and not appropriate for touch devices. But, this scenario is changing with the smartphone and tablet popularization: the web designers need to think about the other interaction styles, such as touchscreen and pen-sensitive devices.

In pen-based devices when the user moves the pen in the screen, the pen trace should result in electronic ink that must be treated by the application to be rendered and stored. But, desktop applications, that running in the Tablet PCs, do not treat electronic ink, so it is necessary to incorporate special applications to treat the electronic ink to have benefices of the pen interaction style. This is the case of Classroom Presenter [13] and others pen-based software. In our previous work [14] we studied three applications (Professional Adobe Acrobat, Windows Journal and Jarnal) to do classes annotations or writing manuscript texts in student activities using a Tablet PC, describing some identified usability problems. We conclude that desktop applications that do not manipulate the electronic ink loose the Tablet PC potential because the pen is used just as a pointer device. This is valid for web applications.

In our previous work, trying to develop web applications that do a better use of pen interaction, we developed the InkBlog for Ae environment, described in the next Section. We chose a blog tool because edublogs, blogs with educational purpose, archiving and supporting student and teacher learning by facilitating reflection, questioning by self and others, collaboration [15] and by providing contexts for engaging in higher-order thinking. Ray [15] cites four ways to incorporate blogs into the classroom, including: (i) using them to communicate information to students and parents, (ii) to provide instructional resources and useful links, (iii) to allow students the opportunity to collaborate with one another on various projects without being in the classroom itself, and (iv) to showcase student work and projects, like poetry and photographs of project work. Another motivation is that users can share draws and images, and the stylus allows users to draw in the device.

III. THE AE ENVIRONMENT AND INKBLOG TOOL

The TIDIA-Ae Project (TIDIA-Ae is the acronym for "Tecnologia da Informação para o Desenvolvimento da Internet Avançada – Aprendizado Eletrônico", in English "Information Technology for Development of Advanced Internet – Electronic Learning") was initiated by FAPESP (the State of São Paulo Research Foundation) with the main goal of developing an e-Learning environment that can explore the potential of Advanced Internet and can provide support to different educational context needs: the Ae environment. Several tools are developed for Ae environment such as Portfolio, Mail, Discussion Forum, Chat and Weblog.

In e-Learning environments, each course participant (student or teacher) owns a blog, where she can post or exclude messages sharing texts, video or audio messages with other participants. Each participant may access other course participants' blogs, comment posted messages by the owner of the blog.

To publish a message in the Ae´s Weblog, the participant clicks in the link "Compose a new post" in her blog page. A new page within a form to be filled will be displayed. After filling up the form, the user clicks on the "Confirm" button and the system will save the new post and redirect the user to the user´s blog page displaying the new message at the top.

The Weblog uses a rich text editor to allow users to easily write formatted text without having knowledge of HTML. For desktop computers the rich text editor works well, the user type the text in the keyboard and use the mouse or shortcuts to trigger format features. But, in the case of pen-based tablets the usability decreases: the user needs to typewrite each letter pressing the pen over the respective key in a virtual keyboard. It is not possible to handwriting or sketching, actions done with a pen and paper.

Ae´s Weblog tool was designed in a design-evaluation process. The first step to create the Weblog´s user interface was to develop a high-fidelity prototype considering the user will interact in a desktop computer equipped with mouse, keyboard and a high-resolution medium size display. This user interface was evaluated and some modifications were done before codifing phase. A usability test was done with the first implement version, and some modifications (e.g., removing some functionallities and changing some labels) were recommended. The final version is presented in Figure 1.

In our previous work, we presented the InkBlog [6] (Fig. 1). It was created to make easier to handwrite posts and comments using a stylus in pen-based devices. The approach to develop the InkBlog was to increase the Weblog tool with components to generate and manipulate the electronic ink in the user interface, representing the electronic ink in InkML format; before a usability test was done to identify problems when user interacts with pen. Changes in the Weblog´s architecture (Fig. 2) and user interface (Fig. 3) were done to support input data from stylus. In the architecture we added a component to receive data from pen, the InkController component, and a component to renderer this data as electronic ink, the InkRenderer component. Both

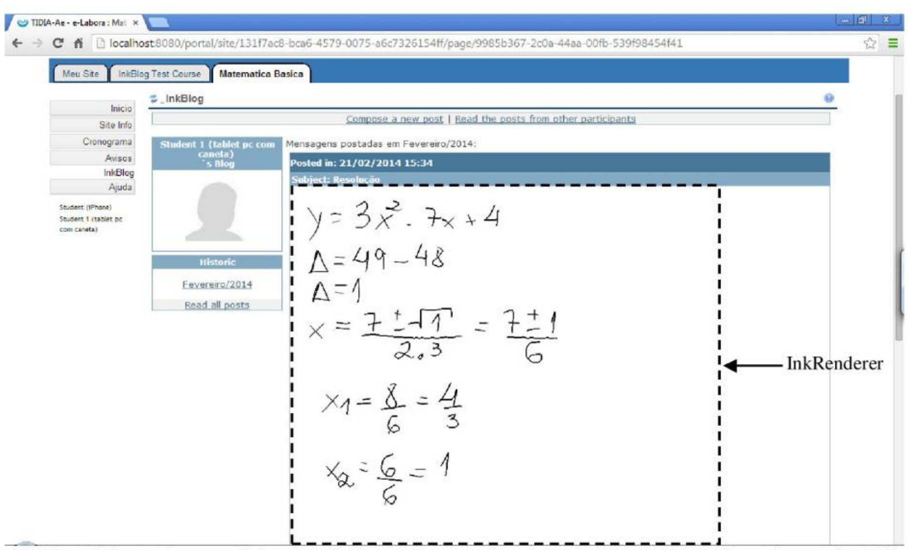


Figure 1. Using InkBlog to share a resolution about finding root for quadratic equation rendered by Google Chrome browser in a Tablet PC.

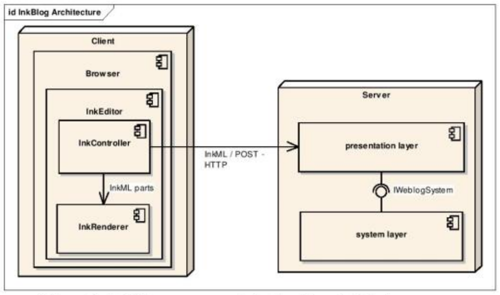


Figure 2. InkBlog components to treat input data from pen.

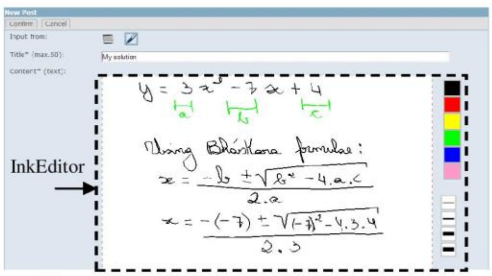


Figure 3. Using InkBlog to handwrite a post using a stylus.

components, InkController and InkRenderer, compose the InkEditor: a handwritten text editor for web pages that render the electronic ink and receive the input data generated by stylus.

The pen input data is received by the InkController, which transforms each point of the trace into coordinate points following InkML format. The user can choose the trace´s color and the width selecting the buttons options in the right side (Fig. 3). When the user points out and presses the pen into a color or width button, the next traces will have the brush attributes set to look like the selected options.

The InkRenderer, the other InkEditor component, draws the traces of a handwritten post in the user screen (Fig. 1). The InkRenderer´s code, the electronic ink data in InkML format and the HTML page are sent to the client over the HTTP protocol (Fig. 2) to display the page with posts. After all the data and code arrived in the client, the InkRenderer reads the InkML data resident inner the tag canvas, and draws the electronic ink for each trace, taking into account the ink formatting. The InkRenderer was developed using the InkML JavaScript Library.

To post a new message, the user can choose the input hardware (keyboard or pen) selecting the icon on "input from:" field, to type the text using a keyboard or to handwrite a post with a stylus (Fig. 3). When the user chooses the pen, she will write a handwriting post, the browser will hidden the text editor and show the InkEditor, where the user will use the stylus to handwrite. When the user touches the InkEditor within the pen and draws a trace, the InkController will listen to the user actions, getting the dots that compose the trace. Each dot is recorded and a line connecting the preceding point to the new point is drawn until the user releases the pen. After the pen is released, the InkController will generate the InkML´s trace node for the new trace. The user can draw as many traces she wants, all them will be stored and will compose the InkML data. When finished handwriting the post, the user will click in the Confirm button and the generated InkML data will be sent to the server to be stored.

Some changes were needed in the Web application to distinguish textual content from typewriting content and to show the correct editor in the post view. The changes are done in the presentation layer; the other layers have not been changed.

The client device needs to have a compatible HTML5 browser to run the InkEditor. The InkEditor uses InkML to represent the handwriting data and the Canvas HTML attribute to draw the traces in the screen.

It is also possible to handwrite comments and post them: the process is similar to the described process above.

IV. Improving Blog Tool For Pen and Touch Interaction

To support touch interaction, we developed a new version of the InkBlog, extending it with features to manipulate data from touchscreens. To include touch interaction, we adopted the same evolutionary approach as when we included pen interaction: usability test with user interaction with pen (two devices were considered, a smartphone and a tablet) and modifications of the user interface, following usability tests.

Multi-touch in web applications is more common on games; Johnson [16] presents a tutorial to include features of multi-touch in web applications. Analyzing the InkBlog code, we modified the InkController code to support touch devices including the handling of touch events defined by W3C. The first modification was handling the touchstart, touchmove and touchend event types. Since we wanted the users draw with their fingers in touch sensitive devices, these events types call functions to start a line, to compose the line, and to stop to draw a line, respectively. This modification allows users to interact with one finger per time.

To allow multi-touch, it was necessary to use a <inkml:trace> to store the data from each finger in a array. The browser send to the function that will handle the user interaction an event object with the *changedTouches* attribute, a collection with data from one or more modified touch points. To identify finger´s move it is possible to use the event´s *identifier* attribute; this value was used as index in the array of <inkml:trace> to put the data in the correct line.

To avoid the browser to scroll the page when the user moves the fingers on the screen, it was called the event´s functions *preventManipulation()* and *preventDefault()*.

To test the new version of InkBlog we used a Tablet PC, a tablet and two smartphones. The Tablet PC is a HP TouchSmart TX2-1040br, 2.2 GHz dual-core processor with 3 GB RAM and a 12" multi-touch screen with Windows Vista operating system and with Windows 8 operating system. This model has the design similar to HP laptops but it is equipped with the described hardware for Tablet PC, for our purposes, a pen sensitive and touch sensitive screen. Google Chrome browser version 31.0.1650.57m was used to navigate through the e-learning environment. So, the used Tablet PC has pen and multi-touch modalities.

The tablet used in the session tests is a Brazilian Positivo Ypy AB10EC, 1.0 GHz ARM Cortex A9 processor with 1 GB RAM and a 10" multi-touch screen 800 x480 pixels of resolution with Android version 4.1.1 operating system. In the iPhone, the used browser was the Google Chrome 31.0.1650.59.

The smartphones used in the session tests are an iPhone 3GS and a Motorola Milestone. The iPhone 3GS has a 833 MHz Samsung S5L8920 ARM Cortex-A8 processor with 256 MB RAM and a 3.5" multi-touch screen 480 x320 pixels of resolution and iOS version 6.1.3 as operating system. In the iPhone, the used browser was the Safari. The Motorola Milestone has a 600 MHz Cortex-A8 processor with 256 MB RAM and a 3.7" multi-touch display 854 x 480 pixels of resolution and Android version 4.0.3 as operation system. To browse in the Web application, the Android stock Web browser was used.

In the post message task, the new version of InkBlog demonstrated to have good performance in all devices. Fig. 4 shows the use of the Positivo Tablet to draw in the InkBlog using ten fingers at the same time (each line was done by one finger, draw here with different colors but the InkBlog just allow user to draw with one). The HP Tablet PC with Windows 8 recognized 4 fingers, instead of 5 fingers described in the operation system user manual. The Motorola Milestone and iPhone (Fig. 5) recognized 5 fingers.

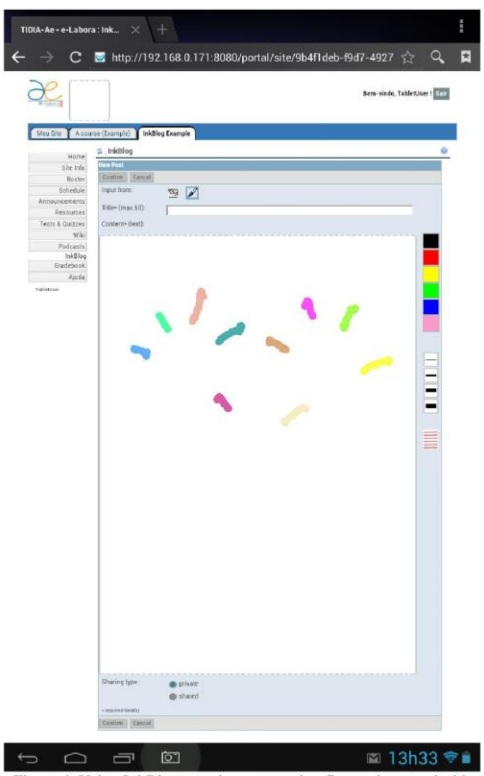


Figure 4. Using InkBlog to write a post using fingers in an android-based tablet in a landscape position.

All devices displayed the posts correctly (Fig. 6), but when the number of drawing posts increases, more time is necessary to render. The render process depends on the hardware, so different devices take different time to complete the task.

In all devices, smartphones or tablets, the touch input data was correctly identified by the web application. In Windows 8 using Google Chrome, the input data from the pen is recognized as mouse input.

To manipulate the iPhone and the Android-based smartphone we used fingers and a capacitive stylus. Since these devices cannot distinguish touches by fingers from touches by stylus, it is not possible to handling the input data considering the data origin. In this case, we prefer to allow users drawing on the surface instead of using the data to scroll the screen. We decided considering the user goal: to post a message.

Another handling is possible: to combine the power of input modalities, e.g., in Tablet PCs, devices that have capacity to distinguish the origin of the input data, it is possible to use the data from the pen to generate the electronic ink and the data from touch to scroll the screen or to trigger another gesture such as selection and zoom.

We did one study case based on user test (following Human-Computer Interaction methodology) planned to apply a profile questionnaire, the performance of four tasks and, in the end, apply a opinion questionnaire. In the opinion questionnaire we adopted the Likert scale for multiple choice

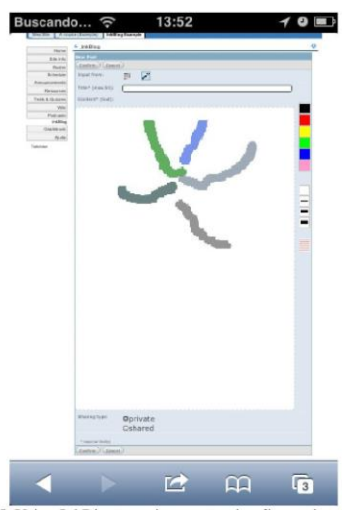


Figure 5. Using InkBlog to write a post using fingers in an iPhone smartphone in a landscape position.

questions. We invited three volunteers: one woman and two men. One of them have post-graduation (33 years old) and the other two are undergraduate student in Computer Science (22 and 27 years old). All of volunteers consider themselves experts and use touchscreen devices every day for more than one year. None of the volunteers used pen-sensitive devices. All volunteers access Internet, one of them for more than ten hours per day and two of them uses between five to ten hours per day.

About e-learning environments, all volunteers have used: one of them does not know the name of the used environment and the other two had used TelEduc and Moodle. Two of them reported that they used for more than one year, and the another volunteer reported that she used between one to five months. Two users reported that they are experts about e-learning environment and another one reported have intermediate knowledge.

The volunteers were questioned about how to help a colleague that asked for help by e-mail to resolve an exercise that needs to apply the Bháskara formulae. One volunteer used special character to represent the power and the square in the formulae. Another one said that she probably sent a link with the formulae and some resolved exercises. The last volunteers answered that probably she will search on the Internet to remember the Bháskara formulae.

All volunteers performed four tasks: i) login (with a given username and password) in the Ae environment; ii) post a message in the InkBlog with the resolution of an exercise that need to find the roots of a quadratic equation; iii) aim a colleague to find errors in her resolution; iv) logout of the environment. The volunteers are asked to perform the four tasks in three devices: i) the Tablet PC using the pen; ii) the iPhone using fingers; and iii) the Positivo tablet using fingers.

By observation, it was possible to identify several problems. The volunteers had some difficult to do small traces (e.g., - and =); this traces are done quickly, and most of time, the interval between the start and the end is not long enough to generate the data; sometimes, the trace is recognized by the browser as gesture (e.g., back or forward). Dots are considered by all volunteers difficult to be done; since they are small, the user needs to press and move slowly the pen (or the finger) in circular move to do a dot. Maybe, these are the reason for the volunteers answered in the questionnaire, when questioned "It was easy write a message using a pen", one user marked the option "strongly agree", another one marked the option "agree", and the last one marked "disagree". One volunteer said that it was difficult to write as legibly as when she writes on the paper.

In the Tablet PC using the pen, the volunteers tried to scroll the page touching the pen in a white area (similar as they did when interact in smartphones). But, this feature is not implemented in the environment and the Chrome browser either. Maybe, this is the reason a volunteers mark "agree" when asked about if her had difficult to scroll the page. The volunteers reported that had some problems to use the virtual keyboard due the small keys.

About the interaction using the iPhone, the volunteers just used one finger to write in the InkEditor. The volunteers used two fingers only to do gesture for zooming. The

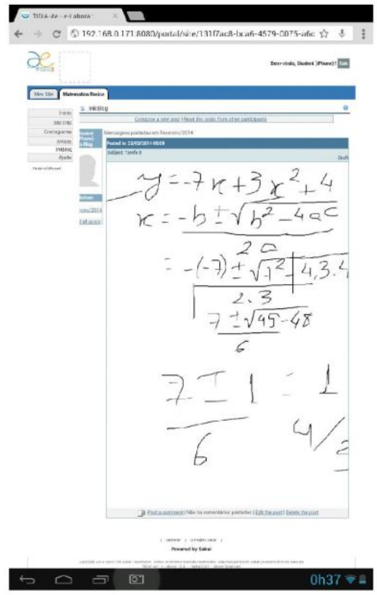


Figure 6. A resolution made by one volunteer with her finger about finding root for a quadratic equation rendered by Google Chrome browser in the Positivo tablet.

volunteers described some difficulty to read the screen, but they did not have difficult to trigger links and icons. One volunteer reported difficulty to scroll the page. About writing the resolution using the finger (question "it was easy write my message through the finger"), a volunteer marked "strongly agree", another volunteer marked "agree" the last one marked "strongly disagree". It is important to highlight that the exercise is easier to be done using a pen, due the pen characteristics. In the open question, one volunteer said "the iPhone screen is too small, and my fingers are large turn hard to do precise strokes and that fit in the available space" (Fig. 6). One volunteer reported the impossibility to do the gesture for zooming in the InkEditor component (InkEditor recognize as electronic ink).

About the interaction using the Positivo tablet, the volunteers reported that they did not have difficult to trigger links and icons, but one volunteer reported had some difficult to read the screen. No volunteer had difficulty to scroll the page (since a touch in anywhere in the screen can generate a scroll). About writing the message within a finger in the Positivo tablet (question "It was easy write my message using a finger"), one user reported "strongly agree", another one reported "disagree", and the last one "strongly disagree". One volunteer reported (and it was observed too) that, after each trace done, the visual focus returned to the field text, and the browser scroll up the page, causing some risks in unwanted places in the InkEditor due the user started the next trace.

V. Final Considerations and Future Work

e-Learning environments are applications that use the Web infra-structure to support teaching and learning activities. To post text, there is a rich text editor to allow users writing formatted text without HTML (HyperText Markup Language) skills. This solution has good usability on desktop computers, but when the user interacts with a pen or touching, she needs to type each letter using a virtual keyboard, so the usability, most specifically, the efficiency, decreases and makes the writing task boring. Another problem is the difficulty to draw sketches using the mouse.

In our previous work we improved the Ae´s Weblog with components to generate and manipulate electronic ink, allowing users handwrite posts and draw sketches in pen-based devices, calling this new tool InkBlog (available to test at [17]). For the InkBlog development, we chose well-defined and promising technologies, such as W3C InkML and HTML5, allowing to codify a light electronic ink editor for Web pages, the InkEditor. Some modifications are done in the Weblog for distinguishing the correct editor to be displayed in the new post composite page and display handwriting posts and comments. However, InkBlog has some limitations: it is not possible, in touchscreen device, to use fingers to draw.

We believe that with InkBlog it is possible to have a better support to disciplines like Graph Theory and Computer Theory. Without the InkBlog, the user needs to use paper and pencil to resolve an exercise and use specific hardware, such a scanner, to digitalize the solution. Or the user needs to use a special application to draw a graph. In both solutions, the user posts the picture as an attached file in a weblog post. Using the InkBlog the user can sketch the graph directly on the weblog tool through direct manipulation.

In this work, we modified InkBlog to handling touch events, allowing users with multi-touchscreen devices post messages with draw done by fingers. The new version of InkBlog was tested in 4 devices (HP Tablet PC, Positivo tablet, Motorola Milestone smartphone and iPhone) and in 3 operation system (Windows Vista and 8, iOS and Android). Usability test was done with three volunteers. Some usability problems were found and they will be corrected in the next version. For future works, we will: 1) do a performance evaluation and evaluated in a class; 2) improve other tools by adding the developed components; 3) support gestures to trigger functionalities by the integration with the environment with a gesture recognizer, and study the gesture´s drawing for each function; 4) integrate with a handwriting recognizer to recognize the word and allow a better integration with another environment tools, such as search tools; 5) using a post or a document as background, allowing teachers or other student to do annotations about a post.

Acknowledgment

The authors thank the CAPES and CNPq for grants and for FAPESP through TIDIA-Ae Project (n. 05/60572-1), which provided the equipment.

References

[1] Moodle Trust, “Moodle.org: open-source community-based tools for learning,” Available at <http://moodle.org>. [retrieved: Jul. 2013]

[2] SAKAI Environment, “Sakai Project | collaboration and learning - for educators by educators,” available at <http://sakaiproject.org>. [retrieved: Jul. 2013]

[3] TelEduc Environment, “TelEduc Ensino à Distância,” available at <http://www.teleduc.org.br>. [retrieved: Jul. 2013]

[4] Ae Project. “Ae - Aprendizado Eletrônico Environment,” available at <http://tidia-ae.iv.org.br/>. [retrieved: Nov. 2013]

[5] A. C. da Silva, F. M. P. Freire, and H. V. da Rocha “Identifying Cross-Platform and Cross-Modality Interaction Problems in e-Learning Environments,” Proc. of 6th International Conference on Advances in Computer-Human Interactions (ACHI 2013), February 2013, pp. 243-249.

[6] A.C. da Silva and H.V. da Rocha, “InkBlog: A Pen-Based Blog Tool for e-Learning Environments,” in Bridging Disciplinary Boundaries: Issues in Informing Science and Information Technology, vol. 10, May 2013, pp. 121-135. ISBN: 1932886699.

[7] O'Reilly, T. “What Is Web 2.0 - Design Patterns and Business Models for the Next Generation of Software,” available at <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>. [retrieved: Nov. 2013]

[8] C. Safran, D. Helic, and C. Gütl., “E-Learning practices and Web 2.0,” Proc. of International Conference on Interactive Collaborative Learning (ICL 2007), Sep. 2007, 8p.

[9] R. Berjon, T. Leithead, E. D. Navara, E. O´Connor, and S. Pfeiffer, “HTML5 - A vocabulary and associated APIs for HTML and XHTML W3C Candidate Recommendation, “ available at <http://www.w3.org/TR/html5/>. [retrieved: Nov. 2013]

[10] D. Schepers, S. Moon, M. Brubeck, and A. Barstow, “Touch Events” available at < http://www.w3.org/TR/touch-events/>. [retrieved: Nov. 2013]

[11] Y. Chee, K. Franke, M. Froumentin, S. Madhvanath, J. Magaña, G. Pakosz, G. Russell, M. Selvaraj, G. Seni, C. Tremblay, and L. Yaeger “Ink Markup Language (InkML) W3C Recommendation,” available at <http://www.w3.org/TR/InkML/> [retrieved: Nov. 2013].

[12] T. Underhill. “InkML JavaScript Library,” available at <http://inkml.codeplex.com/>. [retrieved: Nov. 2013]

[13] R. Anderson, R. Anderson, P. Davis, N. Linnell, C. Prince, V. Razmov, and F. Videon, “Classroom Presenter: Enhancing Interactive Education with Digital Ink”. Computer, vol. 40, September 2007, pp. 56-61. doi:10.1109/MC.2007.307.

[14] A.C. da Silva, and H.V. da Rocha, “Experiences in use Tablet PC to Support Student´s Activities: Five Years of an Exploratory Study,” Proc. of 4th International Conference on Mobile, Hybrid, and On-line Learning (eL&mL 2012), pp. 38-43.

[15] J. Ray, “Welcome to the blogosphere: the educational use of blog”, Kappa delta pi record, v. 42, n. 4, 2006, pp. 175-177.

[16] T. Johson, “Handling Multi-touch and Mouse Input in All Browsers – IEBlog – Site Home – MSDN Blogs,” available at < http://blogs.msdn.com/b/ie/archive/2011/10/19/handling-multi-touch-and-mouse-input-in-all-browsers.aspx> [retrieved: Nov. 2013].

[17] Ae Environment with InkBlog. “Ae - Aprendizado Eletrônico Environment,” available at <http://tidia-ae.nied.unicamp.br/portal>. [retrieved: Mar. 2013]

Capítulo 7

Arquitetura para Ambientes de EaD com Interação Multimodal

Baseado na experiência com as modalidades obtida na realização deste trabalho e na literatura sobre multimodalidade, propomos uma arquitetura para construir um ambiente de EaD com interação multimodal. Adotar multimodalidade em um software implica a adoção de componentes que tratam a multimodalidade mas, também, a criação de interfaces entre os componentes da multimodalidade e do software. No caso dos ambientes de EaD, que possuem um conjunto amplo de ferramentas, de funcionalidades dispostas nessas ferramentas, de papéis e permissões, é necessário que esses componentes se relacionem.

Este Capítulo é composto por dois artigos, o primeiro, um artigo resumido apresentado em um workshop de teses que ocorreu na sexta edição da *International Conference on Advances in Computer-Human-Interactions,* na qual apresentamos a primeira versão da arquitetura para ambientes de EaD com interação multimodal, baseada na arquitetura de Dumas, Lalanne e Oviatt (2009), apresentada na Seção 2.4

O segundo artigo, gerado a partir da continuação do trabalho anterior e publicado na *IADIS International Conference on Interfaces and Human-Computer Interaction and Game and Entertainment Techonologies*, apresenta um maior número de relacionamentos entre os componentes e cenários que exemplificam a necessidade do interfaceamento dos componentes. Assim, este Capítulo mostra a construção e a evolução de uma proposta de arquitetura para os ambientes de EaD suportarem a multimodalidade, usando como base o ambiente Ae, selecionado por ter documentação de sua arquitetura.

Ressalta-se que a multimodalidade nos ambientes de EaD não depende somente da implementação do ambiente. Os navegadores e sistemas operacionais desempenham papel importante na coleta e envio de dados para os componentes do ambiente de EaD, e a tecnologia atual ainda está em evolução, trazendo desafios para a área, conforme será descrito no Capítulo 8.

ACHI 2013 : The Sixth International Conference on Advances in Computer-Human Interactions

e-Learning Environment with Multimodal Interaction

A proposal to improve the usability, accessibility and learnability of e-learning environments

André Constantino da Silva
Institute of Computing (PG)
UNICAMP, IFSP
Campinas, Brazil, Hortolândia, Brazil
acsilva@ic.unicamp.br

Heloísa Vieira da Rocha
Institute of Computing, NIED
UNICAMP
Campinas, Brazil
heloisa@ic.unicamp.br

***Abstract*—The Human-Computer Interaction is challenging the use of many modalities to interact with an application. The e-Learning environments interfaces are been exposed to this diversity of modalities, but they are designed to be used with a limited set. The impact is that users have interaction problems caused by the cross modality. The e-Learning environments need to evolve allowing users to interact with a more broadly interaction styles. One solution is adopt Multimodal Interaction concepts, improving the usability and accessibility of the e-Learning environment and make possible to embrace better learning contexts, property that we define as learnability.**

Keywords-Human-Computer Interaction; Interaction Styles; Multimodal Interaction; Electronic Learning Environment.

I. Introduction

The Human-Computer Interaction is challenging the replacement of the mouse and actual interfaces for interfaces that works with natural interaction, non-tactile interfaces, speech recognition, facial and movement recognition and gestures [1][2]. So, there are many ways to interact with digital artifacts and applications, like keyboards, mouse, small, medium and big displays, voice, touch and gestures. Many hardware components are available supporting different interaction styles.

The Multimodal Interaction is a solution to possibility the use of an application in this diversity of interaction styles, allowing users interact with computers by many input modalities (e.g., speech, gesture, eye tracking) and output channels (e.g., text, graphics, audio) [3]. Multimodal Interaction is proposed to turn the human-computer interaction more natural, i.e., more close to the human-human interaction. The main benefices are the increase of application's usability, accessibility, flexibility and convenience [3]. But, building a multimodal interaction system is not a trivial task yet, because the literature does not have sufficient information about how to design this kind of system and there is a lack of technologies to support them.

e-Learning environments like Moodle [4], SAKAI [5], TelEduc [6], Ae [7] are applications that use the Web infra-structure to support teaching and learning activities. The e-Learning environments are designed to support a variety of users and learning contexts, but they are designed to support a limited interaction styles, usually keyboard and mouse as input and a medium screen as output.

To attend this demand, the e-Learning environment needs to have good usability, accessibility, mobility and learnability. Considering all these dimensions is not a trivial task. Does the multimodality can improve these requirements on e-Learning environments?

Section II presents a literature review about e-Learning environments, multimodality and multimodal interfaces. Section III presents the research problem that we want to deal, and Section IV our hypothesis and methods. Section V some preliminary results and expected contributions.

II. Literature Review

The actual versions of e-Learning environments take advantages of the Web to offer content with text, images, audios and videos in a hypertext document. Tools like chat, forums, portfolios, repositories are widely used, and tools those explore the audio and video resource to user communication, such as instant messenger and video-conferences, are becoming common among the environments.

Due the diversity of users whom may use the e-Learning environments, these systems need to have good usability so that the user interface does not prejudice the teaching and learning activities. Accessibility is another important requirement to allow disabled people to use the environment. So usability and accessibility are two desired requirements to the e-Learning environments.

Devices, such as smartphones and tablets, are becoming increasingly popular; most of them have touch screen displays, access to the Internet and enough computing power to process Web pages. So, Web sites and Web applications, initially developed to be used with keyboard, mouse and a medium size display, are been accessed by small touchscreen devices. This is another aspect of accessibility, so the environments´ development teams are building solutions to provide access on mobile devices. Three kind of solution are emerging: specific device application; web site specific for mobile devices; and improve the web site for mobile and desktop access [8].

Two motivations allow the participants interacting anytime and anywhere with the content and each other; but, due to the device restrictions, there are needs to obtain better design solutions. The actual user interface design techniques take account just a limit set of input and output hardware, limited to the context, such as techniques to design user interface for desktop or for mobile platforms. But, there is a

 ISBN: 978-1-61208-250-9

lot of input or output hardware in these devices and these techniques are asked to consider all of them. Some input and output devices are: touchscreen, microphones, pen sensitive screen, touchpad, TrackPoint, accelerometers, joysticks, loudspeakers, small screen, large screen, printers, etc.

Another e-Learning environment characteristic is to be used in many of learning context, e.g., teacher training, undergraduate courses, and team training in all areas of knowledge. We call these as learnability. But, the actual hardware increases the difficulty to use the environment to produce content for any area and support student activities.

To attend this demand, the e-Learning environment needs to be usable and accessible for many users in many social, physical, technological and learning contexts. So, e-Learning environment needs to be evaluated in the usability, accessibility, mobility and learnability dimensions, a not trivial task.

Since the e-Learning environments were building to Web, they have a common architecture: the client-server. Client is responsible to render the user interface through a browser. It is in the client side that the user interacts with the system using input and output hardware. The server is responsible to process client´s requests and data persistence. The server knows few about the input and output devices in client side.

Ae is an e-Learning environment developed by a consortium of Brazilian universities using J2EE technology and component-based development process. Layered component-based software architecture was defined for Ae environment [9] with the following layers:

Presentation layer: provides the application user interface;

System layer: provides an interface for the application functionality, that is, it is a façade for the application business rules;

e-Learning layer: provides the component interfaces that implement the application's business rules, which can be used by various applications and which use services and functionalities from the infrastructure layer to implement the business rules;

Infrastructure layer: implements a set of infrastructure services such as, for example, data persistence;

Common services layer: has the public services that can be utilized and accessed by all other architecture layers, except the presentation layer.

Multimodal interaction is a research proposal to turn the interaction between humans and machines more natural, i.e., more close to the interactions between two humans, and have the benefits to increase the usability, flexibility and convenience [3].

Modality is the used term to define a mode what the user data input or a system output is expressed. The communication mode refers to the communication model used by two different entities to interact [10]. Nigay and Coutaz [11] define modality as an interaction method that an agent can use to reach a goal, and it can be described in general terms such "speech" or in specific terms such "using microphones".

For monomodal systems, designers are not limited to choose only one modality. But, in multimodal systems, they can choose many modalities, that used together, increasing the system flexibility and gives other benefices. Interfaces with this characteristic are called as multimodal interfaces and the system are called multimodal interaction systems. Mayes [12] defines multimodal interaction systems as a system with the capacity to communicate with the user by many different communication modes, using more than one modality, automatically gives or extracts mean.

According to Oviatt [10] "Multimodal interfaces process two or more combined user input modes (such as speech, pen, touch, manual gesture, gaze, and head and body movements) in a coordinated manner with multimedia system output".

Lalanne *et al.* [3] describe multimodal interaction systems, or multimodal systems, allow users to interact with computers though many data input modalities (e.g., speech, gesture, eye gaze) and output channels (e.g., text, graphics, sound, avatars, voice synthesis).

Bangalore and Johnston [13] say the critical advantage of multimodal interfaces is that they allow user input and system output to be expressed in the mode or modes to which they are best suited, given the task at hand, user preferences, and the physical and social environment of the interaction. Allowing users interact with many modes, it is possible to improve the accessibility because a multimodal interface can be used for a disabled person using the interaction mode that she can handle.

Multimodal content is common on multimedia systems. The research problem that we want proposes a solution is to use multimodality on user interaction and get benefits of multimodal content too.

Dumas, Lalanne and Oviatt [14] present a generic architecture for multimodal systems, turning more easy to understand the mainly components of the multimodal systems: a fusion engine, a fission module, a dialog manager and a context manager, which all together form what is called the "integration committee". The authors define "input modalities are received though various recognizers, which output their results to the fusion engine in charge of giving a common interpretation of the inputs. A fusion machine gives an interpretation for the data and it communicates it to the dialog manager, in charge of identifying the dialog state, the transition to perform, the action to communicate to an application, and/or the message to return through the fission machine. The fission machine returns a message to the user through the most adequate modality or combination of modalities, depending on the user profile and context of use. For this reason, the context manager, in charge of tracking the location, context and user profile, closely communicates any changes in the environment to the three other components, so that they can adapt their interpretations". Multimodal systems need to take account of all input done by the user to identify and process the solicited action.

Developing multimodal interaction systems is a complex task [14]; but, to turn more easy the development there are some frameworks, such as the ICARE framework [15], FAME [16] and special approaches [17][18].

Bouchet, Nigay and Ganille [15] define formally the CARE properties to characterize and assess aspects of multimodal interaction: the Complementarity, Assignment, Redundancy, and Equivalence that may occur between the interaction techniques available in a multimodal user interface. To aim build multimodal system, the authors propose the ICARE framework.

Larson [19] shows three general questions to response with a web application will be improved with a new mode of input: the new mode needs to add value to the web application, the application leverages the strengths of the new mode and avoids its weaknesses, and the users need to have the required hardware and software.

To implement multimodal system for web it is necessary consider both architecture: for multimodal systems and for web systems. Gruenstein, McGraw and Badr [20] present a framework to develop multimodal interfaces for web, the WAMI Toolkit. The framework defines tree client-side components (Core GUI, GUI Controller and Audio Controller) and more four server-side components (Web Server, Speech Recognizer, Speech Synthesizer, and Logger). The user interact with the Core GUI, described at HTML and JavaScript, and the Audio Controller, a Java Applet to receive the audio input. The collected data is sent to server to be treated by the Speech Recognizer and the Web Server components.

Zaguia *et al.* [21] present an approach to develop multimodal systems using fusion machines dispose on web services in such a way the user can choose the modalities that she sees fit to her situation instead of already pre-defined modalities from the beginning.

But we need to not only build a multimodal system; we are worry about the environment usability so that the interface does not prejudice the teaching and learning activities. Nielsen [22] defines usability as a combination of five elements: easy to learning, efficient, easy to remember, low probability of users do mistakes and user satisfaction.

III. The Research Problem

Since the number of devices accessing the Web grows, the e-Learning environments are exposed to a variety of interaction styles, including ones that are not considered in the design time. Just supporting these interaction styles causes cross modality interaction problems [8], limitation in the multimodality use, and do not take advantages from the interaction style in use. So it is necessary to develop a solution without these limitations.

Thinking about the learning domain, we ask "how do the users interact with a multimodal e-Learning environment? Is this a solution to allow an application be accessed by different devices with a diversity of users, physical and social contexts avoiding cross modality problems and get the better use of the interaction styles?"

Since there are no one multimodal interaction e-Learning environment to aim us to response these questions, we need built one. So we want to know how to develop a multimodal interaction e-Learning environment? Which are the cross modality problems related with this context? How can we identify?

It is the main problem, but several others derive: Is it possible to improve an application to be used by many interaction styles and get advantages? Which kind of modifications is necessary to get the best use of an interaction style? How to guaranties that the usability will not be affected? After the modifications, the application has a better accessibility? Allowing many interaction styles, do we have a new kind of application? How to distinguish the applications that have these characteristics from the other that does not have?

Due the multimodality allows users interact with many modes, maybe the user will use this mode not only to interact with the application, but to create content. How this impact in the environment architecture?

The special interest in the learning domain is the necessity to improve the e-Learning environment to better attend the teaching and learning activities in the variety of learning context, reaching out the maximum of learnability.

IV. The Research Hypothesis and Methods

We argue that it is possible improve the usability, accessibility and learnability of an e-Learning environment adding multimodality concepts in the environment´s user interface.

We planned to use empirical method to validate our hypothesis, building a multimodal interaction e-Learning environment prototype and doing user study, collecting interaction data and user opinions by observation and questionnaires. Not one interaction style will be studied per time; we want study various interaction styles being used to complete some tasks in the environment. The collected data will be used to verify the e-Learning environment in three dimensions: usability, accessibility and learnability.

Due to the quantity of interaction styles and the complexity to build a multimodal system, we need to define an incremental process to build the prototype, taking one or a limited number of interaction styles per time. But, it is important to reduce the development efforts for the next iteration, when another interaction style will be selected. Here, we will apply some software engineering techniques, like software components. We proposed these steps for this process:

1. Select an interaction style;
2. Do user studies to collect interaction problems using the selected interaction style and how the manner to realize the tasks change;
3. Redesign the software user interface to defining better solutions for identified problems;
4. Analyze the software to identify the components responsible to user interface and components that manipulate user data;
5. Find or implement recognizers and synthesizers for the selected interaction style;
6. Change the software architecture to have the main components of a generic multimodal system;
7. Do tests to collect errors and fix them.

We believe if we can do these steps two or three times; so, there is possible to repeat it until all interaction styles is considered.

 ISBN: 978-1-61208-250-9

To prove our solution, we propose to implement the Multimodal Interaction concepts in the Ae e-Learning environment [7], and use this new kind of environment to research advantages and disadvantages of multimodal interfaces in learning. We planned study the interaction using pen, touch and gesture in two tools of the Ae e-Learning environment: the Weblog and the Whiteboard. Both tools are used to construct the content and are selected based on our premises that they are good choices to study the multimodality.

V. Preliminary Results and Expected Contributions

One of the preliminary results were some problems in the TelEduc environment that happened when users interact with touchscreen devices [8], i.e., some problems happened due the platform changing (when user access the environment using a smartphone) and some problems happened due the modality changing. The problems identified in TelEduc due the modality changing will happen on Ae environment.

Other contribution is related with how the multimodal concept changes the architecture of an application and how to find solutions. The Ae architecture needs to be changed to adopt the multimodal concepts. Considering the Web-Accessible Multimodal Interfaces architecture [20] and the architecture of multimodal systems [14], we redesign the Ae architecture (Fig. 1). Due to the fact that the browser has the responsibility to show the GUI components in the client side, some components are added to treat the data input from the input devices. The user interaction data is sent to the server, who have the responsibility to process this data and gives an interpretation for the received data. After the input data interpretation process, the correspondent action is sent to the system component. So, the fusion and fission machines will be on the server. The fusion machine will be called when the server receive the data from the client-side components, and the fission machine will be called when the system response the request, after the data processing. Since, we are using a component-based architecture, the presentation layer and the multimodal components can run in a proper server, increasing the scalability and performance.

The main expected contribution is to know more about the relationship between usability, accessibility and multimodality. For the educational context, we want to know if the support more interaction styles there is an impact in the learning contexts, so define the learnability.

The prototype is another contribution, because there is no one multimodal interaction e-Learning environment, we will call it |Ae|.

Acknowledgment

The authors thank the CAPES and CNPq for financial support and for FAPESP through TIDIA-Ae Project (n. 05/60572-1), which provided the equipment.

References

[1] L. Kugler. “Goodbye, Computer Mouse”. Communications of ACM, vol. 51, n. 9, September 2008, pp. 56.

[2] B. Hayes. “Gartner´s Seven IT Grand Challenges”. Communications of ACM, vol 51, n. 7, July 2008, pp. 10.

[3] D. Lalanne, L. Nigay, P. Palanque, P. Robinson, J. Vanderdonckt, and J. Ladry. “Fusion Engine for Multimodal Input: A Survey”. Proc. of the 11th International Conference on Multimodal Interfaces (ICMI-MLMI´09), ACM Press, Nov. 2009, pp. 153-160, doi: 10.1145/1647314.1647343.

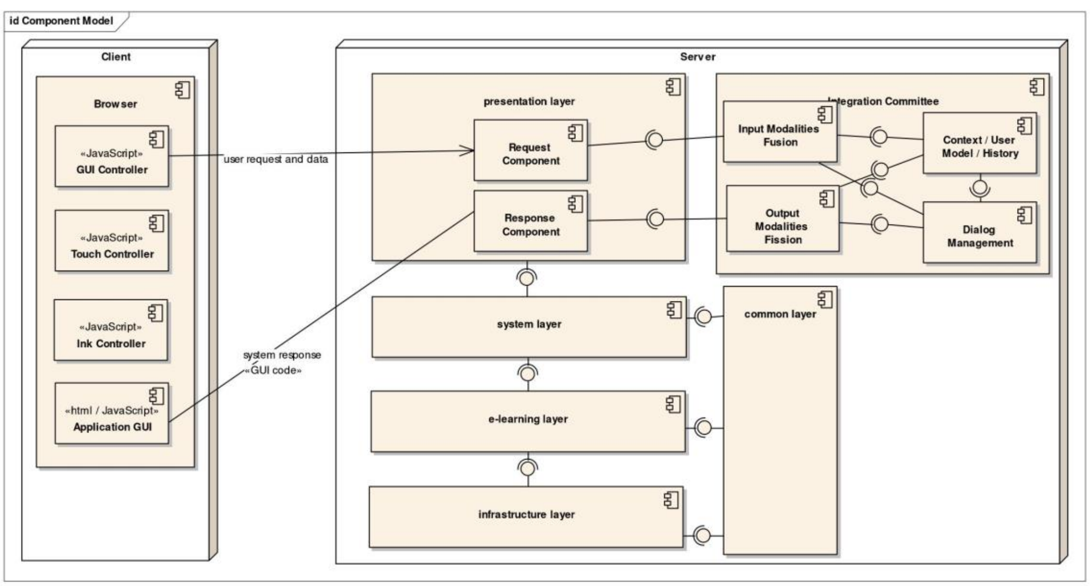


Figure 1. Ae architecture with multimodal components.

[4] Moodle Trust. "Moodle.org: open-source community-based tools for learning". Available at <http://moodle.org>. [retrieved: Jan. 2013]

[5] SAKAI Environment. "Sakai Project | collaboration and learning - for educators by educators". Available at <http://sakaiproject.org>. [retrieved: Jan. 2013]

[6] TelEduc Environment. "TelEduc Ensino à Distância". Available at <http://www.teleduc.org.br>. [retrieved: Jan. 2013]

[7] Ae Project. "Ae - Aprendizado Eletrônico Environment". Available at <http://tidia-ae.iv.org.br/>. [retrieved: Jan. 2013]

[8] A. C. da Silva, F. M. P. Freire, and H. V. da Rocha. "Identifying Cross-Platform and Cross-Modality Interaction Problems in e-Learning Environments". Proc. of the 6th International Conference on Advances in Computer-Human Interactions (ACHI 2013), IARIA, 2013. / in press /

[9] D. M. Beder, A. C. da Silva, J. L. Otsuka, C. G. Silva, and H. V. da Rocha. "A Case Study of the Development of e-Learning Systems Following a Component-based Layered Architecture". Proc. of the 7th IEEE International Conference on Advanced Learning Technologies (ICALT 2007), IEEE Press, Jul. 2007, pp. 21-25, doi: 10.1109/ICALT.2007.4.

[10] S. L. Oviatt. "Advances in Robust Multimodal Interface Design", in IEEE Computer Graphics and Applications, vol. 23, no. 5, Sep. 2003, pp. 62-68, doi: 10.1109/MCG.2003.1231179.

[11] L. Nigay and J. Coutaz. "A Generic Platform for Addressing the Multimodal Challenge". Proc. of 13th Conference on Human Factors in Computing Systems (SIGCHI 1995), ACM Press / Addison-Wesley Publishing Co., May 1995, pp. 98-105, doi: 10.1145/223904.223917.

[12] T. Mayes. "The 'M' Word: Multimedia interfaces and their role in interactive learning systems", in Multimedia Interface Design in Education, A. D. N. Edwards and S. Holland, Eds. Berlin: Springer-Verlag, 1992, pp. 1-22, doi: 10.1007/978-3-642-58126-7_1.

[13] S. Bangalore and M. Johnston. "Robust Understanding in Multimodal Interaction". Computational Linguistic, Cambridge, Massachusetts, vol. 35, n. 3, Sep. 2009, pp. 345-397.

[14] B. Dumas, D. Lalanne, and S. Oviatt. "Multimodal Interfaces: A Survey of Principles, Models and Frameworks", in Human-Machine Interaction, D. Lalanne and J. Kohlas, Eds. Berlin: Springer Berlin / Heidelberg, 2009, pp. 3-26, doi: 10.1007/978-3-642-00437-7_1.

[15] J. Bouchet, L. Nigay, and T. Ganille. "ICARE: Software Components for Rapidly Developing Multimodal Interfaces". Proc. of 6th International Conference on Multimodal Interfaces (ICMI 2004), ACM Press, Oct. 2004, pp. 251-258, doi: 10.1145/1027933.1027975.

[16] C. Duarte and L. Carriço. "A Conceptual Framework for Developing Adaptive Multimodal Applications". Proc. of 11th International Conference on Intelligent User Interfaces (IUI 2006), ACM Press, Jan. 2006, pp. 132-139, doi: 10.1145/1111449.1111481.

[17] A. Stanciulescu, Q. Limbourg, J. Banderdonckt, B. Michotte, and F. Montero. "A Transformational Approach for Multimodal Web User Interfaces based on UsiXML". Proc. of 7th International Conference on Multimodal Interfaces (ICMI 2005), ACM Press, Oct. 2005, pp. 259-266, doi: 10.1145/1088463.1088508.

[18] A. T. Neto, T. J. Bittar, R. P. M. Fortes, and K. Felizardo. "Developing and Evaluating Web Multimodal Interfaces – A Case Study with Usability Principles". Proc. of 24th ACM Symposium on Applied Computing (SAC 2009), ACM Press, Mar. 2009, pp. 116-120, doi: 10.1145/1529282.1529306.

[19] J. A. Larson. "Should You Build a Multimodal Interface for Your Web Site?". Available at < http://www.informit.com/articles/article.aspx?p=29024>. [retrieved: Jan. 2013]

[20] A. Gruenstein, I. McGraw, and I. Badr. "The WAMI Toolkit for Developing, Deploying, and Evaluating Web-Accessible Multimodal Interfaces". Proc. of 10th International Conference on Multimodal Interfaces (ICMI 2008), ACM Press, Oct. 2008, pp. 141-148, doi: 10.1145/1452392.1452420.

[21] A. Zaguia, M. D. Hina, C. Tadj, and A. Ramdane-Cherif. "Using Multimodal Fusion in Accessing Web Services", in Journal of Emerging Trends in Computing and Information Sciences, vol. 1, n. 2, Oct. 2010, pp. 121-137.

[22] J. Nielsen. Usability Engineering. EUA: Morgan Kaufmann, 1993. 362 p.

IADIS International Conferences Interfaces and Human Computer Interaction and
Game and Entertainment Technologies 2013

ARCHITECTURE FOR E-LEARNING ENVIRONMENTS WITH MULTIMODAL USER INTERFACE

André Constantino da Silva[1] and Heloísa Vieira da Rocha[2]
[1]*Institute of Computing (PPG)/UNICAMP, IFSP, Campinas, Brazil, Hortolândia, Brazil*
[2]*Institute of Computing / UNICAMP, Campinas, Brazil*

ABSTRACT

Multimodal interaction is a proposal to turn the interaction between humans and machines more natural, increasing the usability, flexibility and convenience of the application. Improve an application with multimodal features impacts on its architecture and, to describe the main components to treat the multimodality, some architecture models are proposed in the literature, including for Web multimodal systems. e-Learning environments are Web-based systems and need a good usability, flexibility and convenience, requirements that can be improved with implementation of multimodal features on them. Since they have their own peculiarities, we need a more specific multimodal architecture model described in such a way to reuse the components built for multimodal systems and connect them with the e-learning environment components. This work proposes an architecture for e-learning environments and show their impact on the Ae e-learning system, an e-learning environment developed using a component-based development process.

KEYWORDS

Interaction Styles; Multimodal Interaction; Learning Systems Platforms and Architectures.

1. INTRODUCTION

Devices, such as smartphones and tablets, are becoming increasingly popular; most of them have touch screen displays, access to the Internet and enough computing power to process Web pages. The number of input hardware on desktop computers is increasing, with computers equipped with touchscreen devices, microphones and cameras. But Web sites and Web applications are still designed to be used with keyboard, mouse and a high resolution medium sized monitor. One kind of Web application is e-learning environments, systems used to support teaching and learning activities that must have good usability and accessibility to support diversity of users and contexts. But the actual version of e-learning environments does not receive input data from many modes to control the application; they only consider trigger actions by keyboard and mouse.

Multimodal systems are present in the Human-Computer Interaction (HCI) literature, and to turn easier to understand and build multimodal systems some works present general architecture model for these systems. Multimodality can aim increase the usability, accessibility, convenience and flexibility of an application (Dumas, B. et al., 2009), four desirable requirements for e-learning environments. But to build a multimodal e-learning environment it is not a trivial task, it is necessary to deeply understand the e-learning environments, their use and their technology, and define an architecture that considers components of multimodal interaction, of native e-learning environments on the Web platform restrictions. To define these components and their communications we propose an architecture for multimodal e-learning systems.

Section 2 presents a literature review about multimodal systems and architecture for this kind of systems. Section 3 presents our proposal to an architecture model for multimodal e-learning systems and some use cases. Section 4 presents conclusion and future work.

ISBN: 978-972-8939-90-8 © 2013 IADIS

2. LITERATURE REVIEW

The actual versions of e-learning environments, such as Moodle (2013), TelEduc (2013) and SAKAI (2013), take advantages of the Web to offer content with text, images, audios and videos in a hypertext document. Chat, forums, portfolios, repositories are tools found on e-learning environments, and tools those explore the audio and video resource to user communication, such as instant messenger and video-conferences, are becoming common among the environments.

The e-learning environments are built to Web so they have a common architecture model: the client-server. Client is responsible to render the user interface through a browser. It is in the client side that the user interacts with the application using input and output hardware. The server is responsible to process client´s requests and data persistence. In this architecture, an action done by the user in the client side (browser) will trigger a request in the server side, so it is deterministic since the programmer defines, for each page, the mapping between the user action and the server response. In the client side, the user will interact with the browser, an application to render Web pages and send requests for the server. To implement a web page it is important consider the browsers limitations and their configuration.

The server knows few about the input and output devices in client side. All the communication done between server and client is using HTTP protocol and HTML to specify the user interface. Besides the client-server architecture model, a Web application can adopt another architecture model to define and structure the client or server components. For example, the Ae environment (2013), an e-learning environment developed by a consortium of Brazilian universities using J2EE technology, has a layered component-based software architecture model (Beder et al., 2007) with the following layers:

- **Presentation layer**: provides the application user interface;
- **System layer**: provides an interface for the application functionality, that is, it is a façade for the application business rules;
- **e-Learning layer**: provides the component interfaces that implement the application's business rules, which can be used by various applications and which use services and functionalities from the infrastructure layer to implement the business rules;
- **Infrastructure layer**: implements a set of infrastructure services such as, for example, data persistence;
- **Common services layer**: has the public services that can be utilized and accessed by all other architecture layers, except the presentation layer.

Each developed tool following the Ae architecture model has one or more components for each layer. Usually, the presentation layer is treated like a component for each tool, and there is one component for the system layer, but it is not a rule in the Ae architecture.

The actual version of Ae environment is based on SAKAI nucleus (2013), so the course and user management, user authentication and session management are from the SAKAI environment, they have functionalities to manage data about courses, users, and where an online user is browsing. To define what a user can do in a page (e.g. a student called Marie posting a blog message in a course about HCI), it is necessary analyze data about who is the user (Marie), which role she have (a student), which course she is browsing (a course about HCI) and which tools she has access in these context (blog tool). Fig. 1 presents the Ae architecture model highlighting these components in the system layer.

The Ae e-learning environment have flexibility to allow the administrator install only a set of tools, so tools that do not follow the institution rules or overload the server can be removed. Flexibility allows developers and community produce tools similar as plugins; they can be installed in any server and be separately distributed. The flexibility to choose tools is given for teacher too, who can choose which installed tools are available in the course. This flexibility is not trivial to treat, since the environment development team needs to specify how the environment will discover new tools and their functionalities.

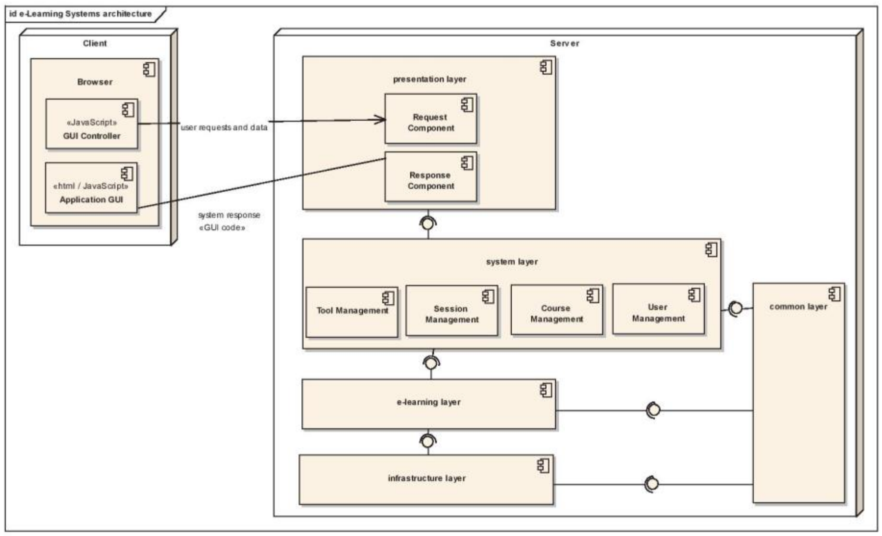


Figure 1. The Ae e-Learning environment architecture.

The environments nowadays allow users to publish content in many media; participants (teachers or students) use video, text with images, hypertext or other multimedia to the content. So the multimodality can be found on content. But the interaction continue the same, deterministic and each mode per time, focused on mouse and keyboard to input and output in a high resolution medium size screen and sometimes the use of alert sounds. But even to produce content, the e-learning environments do not have an easily approach. For example, to produce a video and publish on these e-learning environments, it is necessary to use an application installed on the client to record the video and, when finish to record, go to the environment and upload the file. Other approach is to use videos available on dedicated sites, such YouTube, but it is necessary to browse in the site to find the media, copy their URL and paste it on the environment´s tool, a difficult task for people whom know few about Internet and technology. Beside these problems, the environments have a good usability considering the technology restrictions, but it is need to improve their usability in a new context: many hardware devices to interact.

For example, computers equipped with pen are become common nowadays, e.g., Tablets PCs, a device with height as similar to a notebook and had an input device similar a pen. The paper and pen metaphor implies that tasks performed before in paper, like draw or manuscript writing, can be more natural in the Tablet PC than in another computing devices. Pen-based Computing or Pen Computing refers to a computer user-interface using a pen, rather than devices such as a keyboard or a mouse. User interfaces for pen computing can be implemented in several ways, like using the pen as a pointing input device; in this case the user can interact with the e-learning systems as similar and deterministic as using a mouse, but with more limitations: the hand over the screen can decrease the usability. But the user takes better advantages over the pen based interaction when the application is developed considering direct manipulation, handwriting or gesture recognition.

When the user moves the pen in the screen, the pen trace should result in electronic ink that must be treated by the application to be rendered and stored. But desktop applications, that running in the Tablet PCs, do not treat electronic ink; it is necessary special applications to treat the electronic ink to have benefices of the pen-based interaction style. The actual environments do not support electronic ink either, so in our previous work (da Silva and da Rocha, 2013), we proposed a pen-based blog tool for Ae e-learning environment that allow users to handwrite posts or sketches and do comments using pen (Fig. 2). To implement the InkBlog, we specified the InkEditor, an electronic ink editor, composed by two components: **InkController** and **InkRenderer**. The pen input data is received by the InkController, who transform each point of the trace into coordinate points following InkML format (Chee *et al.*, 2011). The user can choose the trace´s color and the width selecting the options in the right side (Fig. 2).

ISBN: 978-972-8939-90-8 © 2013 IADIS

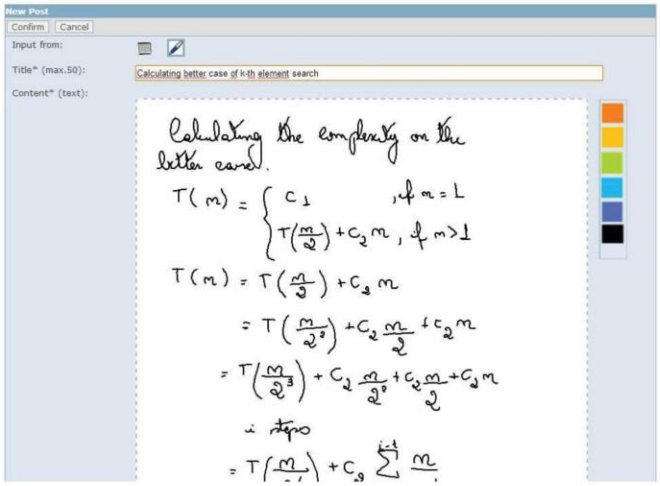


Figure 2. Using InkBlog in the Ae Environment to handwrite a reasoning to prove the better case of an algorithm in a Computer Theory course.

When the user points to and presses the pen into a color or width button, the next traces will have the brush attribute set to look likes the selected options. The InkRenderer, the another InkEditor´s component, draws the traces of the handwrite posts in the user screen.

The InkBlog turn easier to produce content using the pen by direct manipulation, but the interaction does not have all the advantages of pen based computing: the pen is used just as a mouse to trigger actions (e.g., trigger a button action or to navigate to another page using a link), no gesture support and it is not possible to use two or more modalities together; the e-learning environments needs specific modules to treat the handwriting and gesture recognition, and the results can be used not only for trigger an action, but for support multimodality too.

Some models of Tablet PCs are equipped with touchscreen too, so user can interact with the keyboard, mouse, track pad, pen or using her fingers. Since the touch is become common to interact with digital application, mainly on mobile devices, e-learning environments need to be improved to manipulate data from this input device. Even using the InkEditor and InkBlog, it is not possible to have input from pen and from fingers in the same time, the environments cannot process more than one modality per time.

Multimodal interaction is a research proposal to turn the interaction between humans and machines more natural, i.e., more close to the interactions between two humans, and have the benefits to increase the usability, flexibility and convenience (Lalanne *et al.,* 2009). Mayes (1992) defines multimodal interaction systems as a system with the capacity to communicate with the user by different communication modes, using more than one modality, automatically gives or extracts mean. According to Oviatt (2003) "Multimodal interfaces process two or more combined user input modes (such as speech, pen, touch, manual gesture, gaze, and head and body movements) in a coordinated manner with multimedia system output". Lalanne *et al.* (2009) describe multimodal interaction systems, or multimodal systems, allow users to interact with computers though many data input modalities (e.g., speech, gesture, eye gaze) and output channels (e.g., text, graphics, sound, avatars, voice synthesis).

Multimodal systems need to process all input done by the user to identify and process the desired action and generate the output using the appropriate modes. Dumas *et al.* (2009) present a generic architecture for multimodal systems composed by the components: i) **input recognizers & processors**, ii) **output synthesizers**, iii) **fusion** engine, iv) **fission** module, v) **dialog management** and vi) **context, user model and history manager**. The last four components (components iii, iv, v, vi) compose the Integration Committee.

The authors define how these components communicate with each other "*input modalities are received though various recognizers, which output their results to the fusion engine in charge of giving a common interpretation of the inputs. A fusion machine gives an interpretation for the data and it communicates it to the dialog manager, in charge of identifying the dialog state, the transition to perform, the action to* *communicate to an application, and/or the message to return through the fission machine. The fission machine returns a message to the user through the most adequate modality or combination of modalities, depending on the user profile and context of use. For this reason, the context manager, in charge of tracking the location, context and user profile, closely communicates any changes in the environment to the three other components, so that they can adapt their interpretations*" (Dumas *et al.*, 2009).

To implement multimodal system for web it is necessary consider both multimodal architecture and Web architecture. Gruenstein *et al.* (2008) present a framework to aim to develop multimodal interfaces for web, the WAMI Toolkit. The framework defines tree client-side components (Core GUI, GUI Controller and Audio Controller) and more four server-side components (Web Server, Speech Recognizer, Speech Synthesizer, and Logger). The user interacts with the Core GUI, described at HTML and JavaScript, and the Audio Controller, a Java Applet to receive the audio input. The collected data is sent to server to be treated by the Speech Recognizer and the Web Server components. The components Core GUI, GUI Controller, Audio Controller and Speech Recognizer can be classified as the **input recognizers & processors** of the Dumas *et al.*´s architecture**.** The Speech Synthesizer can be classified as **output synthesizers** of the Dumas *et al.*´s architecture. The WAMI toolkit is focused on speech plus keyboard and mouse modes, but the framework can be expanded to include other modes through definition of new components.

The W3C Multimodal Interaction Working Group (MMI-WG) is developing a standard framework for such applications since 2002 and it purposes a recommendation in 2012 (W3C, 2013) with three components to compose the Runtime Framework and more one called Modality Components which recognizes the input data and running a markup language. The Modality Components interact with the Interaction Manager component through the event-based MMI life-cycle API. The Interaction Manager is one of the Runtime Framework components, who is responsible to sends and receives all messages from Modality Components, and queries and updates the Data Component as needed. The Data Component contains the public data model for the multimodal application. The Delivery Context is another Runtime Framework component, it is responsible to store static and dynamic device properties, environmental conditions, and user preferences; properties which can be queried and dynamically updated (as the Dumas *et al.*´s **context, user model and history manager** component). The Interaction Manager component essentially contains the multimodal application and it maintains the dialog flow, current state, and public data (as the Dumas *et al.*´s **dialog management**).

An important aspect of the W3C Multimodal Architecture is that each modality component runs its own markup language document, but McCobb (2013) points out that W3C XML languages directly not support multimodal authoring and there is a lack for standard to combine the XML that represent various modalities into a single document when all modalities are rendered on the client.

Due the difficulties to build a multimodal system and the technology to build Web pages, the actual e-learning systems does not have multimodal interface. One step forward is to have an architecture model. To propose an architecture for multimodal e-learning we will use as basis the Ae environment due the characteristic of have well-defined interfaces for each environment´s component, turning easier to build an e-learning system with multimodal interface, but the functionalities to manage users, courses and roles are in any e-learning environment.

3. AN ARCHITECTURE FOR MULTIMODAL E-LEARNING ENVIRONMENT

Considering the W3C Multimodal Architecture (W3C, 2013), Web-Accessible Multimodal Interfaces architecture (Gruenstein *et al.*, 2008) and the architecture of multimodal systems (Dumas *et al.*, 2009), we redesign the Ae architecture model (Fig. 1) and propose an architecture for multimodal e-learning environments (Fig. 3).

Due to the fact that the browser has the responsibility to show the GUI components in the client side, the **input recognizers & processors** and **output synthesizers** components needs to be distributed in the client and the server. **Input recognizers** components are added to treat the data input from the client side, these components receive the data input stream and describe it in the proper XML format. In Fig. 3, there are components to treat data arising from pen and touch. More components need to be added to consider other interaction modes.

ISBN: 978-972-8939-90-8 © 2013 IADIS

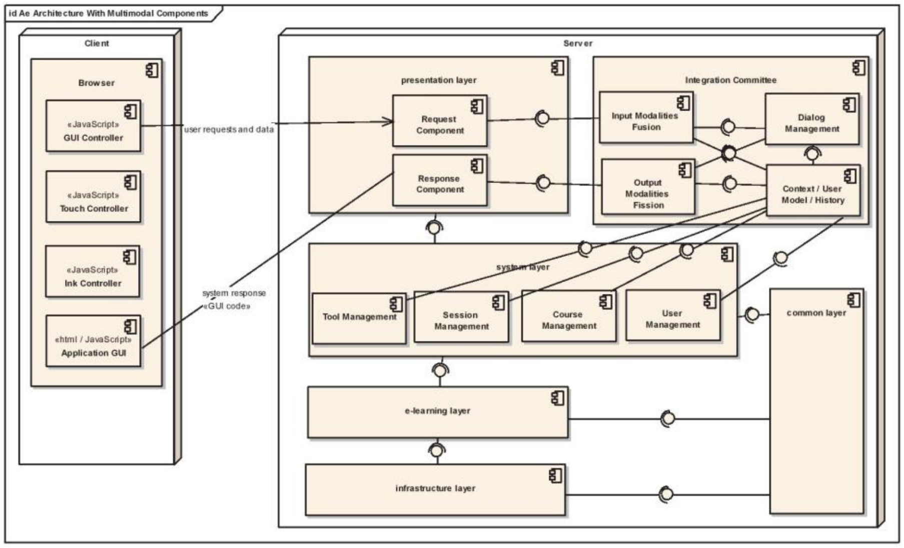


Figure 3. The proposed architecture for multimodal e-Learning environments.

The user interaction data is sent to the server by modality components, and the server has the responsibility to process this data and gives an interpretation for it by the integration committee. So, the fusion and fission machines will be on the server. The fusion machine will be called when the server receive the data from the client-side components, and the fission machine will be called when the system response the request, after the data processing. Since we are using a component-based architecture, the presentation layer and the multimodal components can run in a proper server, increasing the scalability and performance. After the input data interpretation process, the correspondent action is invoked on the Ae´s system component.

e-Learning environments are a complex system, composed by a lot of tools and features. The user usually deals with a set of the functionalities related with the tool she is using, and some features she can use. For example, the student is reading a material published in a content tool but want to write a message to the teacher. The student can invoke the new message screen without to change to Mail tool. But user cannot delete an item in the tool without specify clearly which item is, she cannot say just "delete it", unless she is seeing the item. So the Dialog Management component needs to be integrated with to Tool Management and Session Management to identify the meaning to given for a user request and system response. The data about functionalities needs to be updated each time the user browsing to another tool or another course, since some functionalities change when the user browse to another tool or course. But some functionalities, e.g., change the course, logoff, needs to be available every time. This is a peculiarity of e-learning systems; in the literature the multimodal system have the same features during all the user interaction.

Another characteristic with impact on tasks that can be performed by users is related with the available tools in the server. It is necessary a mechanism to tools inform for the dialog management about the functionalities the user can trigger. One solution is writing code to describe the tool´s functionalities and process it on the tool´s deployment. The tool developers need specify and write the code about the tool´s functionalities in the developing time.

The Context, User Model and History component needs to interact with Tool Management, Session Management, Course Management and User Management components of the Ae environment that will provide useful data for update the context model in the integration committee and necessary data about the tasks users can perform.

A multimodal e-learning environment needs to treat the input data for trigger action, but since there are components to receive data from a different channel, this data can be used as content too.

For example, if the user is interacting with a pen, the generated electronic ink can be used to trigger action such item selection or can be used to produce content such a sketch in a blog tool. The user can trigger a delete action over a Portfolio item or can record an audio content, like a podcast post.

Different from existing examples of multimodal systems, where the data is recognized to typed text and stored in the database, the e-learning systems need to store data from the modalities, not only store typed text. So multimodal e-learning environments join functionalities from multimodal systems and multimedia systems to the user interaction and produce content. But some tools access data from other tools, e.g., the search tool. The search tool needs to read the content stored in other tools and return for the user each occurrence. Search tool in e-learning environments uses string compare techniques, so search tool needs access recognizers when searching in non-typed content. Some studies must be done to verify the necessity of search mechanism must have non-typed content as input.

3.1 Use Cases

Consider the following case: Marie is an undergraduate student and she had a Tablet PC that she uses to browsing into Internet and to handwrite class annotations using a proprietary application. In this semester, Marie is doing the Computer Theory course and the teacher adopted an e-learning environment to support the teaching and learning activities. In the adopted environment, the teacher usually publishes slides, articles, suggested books to read, activities and exercises. In the middle of the course, Marie read the content and started to do exercises. To resolve the exercises is necessary to write equations, a hard task to be done using keyboard and mouse, but Marie uses a proprietary tool to resolve exercises using the pen in the Tablet PC. But she had some doubts in the exercises resolution and due to the environment restrictions, she asks for help sending a mail message and attached the file in a format that she thinks the teacher had software to open it (e.g., .pdf format). She also posted her resolutions on the blog tool as pictures (.jpg format) to share her reasoning with the other students.

John is the Marie´s teacher on the Computer Theory course. He read the Marie´s message asking for help, and to answer her, he printed the attached file writing over the document. He prefers print her resolutions due to write new equations and correct some equations written by Marie. After to do the annotations, John scans the sheets and responses the Marie´s mail attaching the file with scanned sheets. Marie read the John´s message and now knows how to correctly resolve the exercises.

In this context, the use of InkBlog will brings some efficient to the process, reducing time and resource. Marie could write her solutions direct on the InkBlog message, and the teacher writes him suggestions and explanations as a comment in the blog message. But it is necessary a more detailed vision to understand the improvement given by a multimodal interaction. For example, Marie wants a new blog message, so she uses the browser to access the environment, do the log in, goes to the Computer Theory course page, select the blog tool and click in the “New post message” button. If the environment has support to recognize gestures, she can draw a symbol, e.g. “B” letter, over the environment page with the pen to trigger the new blog message action. But the reaction of the environment will depends the page where the Marie draw: if it is over a course page, the new post will be publish on this course, but if the user draw over her space page, the environment needs to ask which course the new post will be published. This example shows the necessity to consider the context (in this case, the page where the user is) to give the meaning for the user´s request and the integration between the multimodal components and the environment components on the Fig. 3.

Another example is delete an item published in a tool. The user can draw the letter “D” or draw an “X” to trigger the action to delete an item, but, to execute the operation, the environment needs to know which item the action is to be done. The draw can be done over the page that shows one item, or can be done over the page that shows many items or a page that this gesture does not have a meaning. In the case of a page with many items, the environment needs to ask for the user to mark the items to be deleted.

When Marie was writing their responses using InkBlog, she used the pen to write her answer and to change the color or line thickness. If the environment had components to receive and treat touch input, she can use her fingers to trigger the functions about colors or line thickness, or can use the pen to do a selection and after move the selected sketches with the fingers. With components to receive audio input, actions for change the color or thickness line can be triggered by voice and the selection and move can be done using a pen or a finger.

ISBN: 978-972-8939-90-8 © 2013 IADIS

4. CONCLUSION AND FUTURE WORK

e-Learning environments are system to support teaching and learning activities through the Web, allowing course participants communicate with each other, build and publish content and manage the virtual space. With the increasing of computational power and interaction styles in some devices the environments are accessed by users using interaction styles that were not considered in the design time. A solution is adopting a multimodal interface to increase the usability, accessibility and convenience of e-learning environments in this new context. The first impact is on the e-learning environment architecture, which needs components to deal with many channels and do the data fusion and fission.

We proposed an architecture for multimodal e-learning environments based on generic architectures for multimodal systems and specificities of e-learning. Due the Web platform, it was necessary to have components in the client to receive the input data and render the result. Another need was integrate components of multimodal architecture with components of the e-learning environment, e.g., the context, user model and history management of the multimodal architecture with the Tool Management, Session Management, Course Management and User Management components of the e-learning environment.

As future work we want to implement the proposed architecture in an e-learning environment considering the touch and pen channels and do some measurements.

ACKNOWLEDGMENT

The authors thank the CAPES and CNPq for financial support and for FAPESP through TIDIA-Ae Project (n. 05/60572-1), which provided the equipment.

REFERENCES

Ae Project, 2013. *Ae - Aprendizado Eletrônico Environment*. Available at <http://tidia-ae.iv.org.br/>. [retrieved: Jan. 2013]

Beder, D. M. et al., 2007. A Case Study of the Development of e-Learning Systems Following a Component-based Layered Architecture. *Proceedings of the 7th IEEE International Conference on Advanced Learning Technologies (ICALT 2007)*, Niigata, Japan, pp. 21-25, doi: 10.1109/ICALT.2007.4.

Chee, Y. et al., 2011. *Ink Markup Language (InkML) W3C Recommendation*. Available at: <http://www.w3.org/TR/InkML/>. [retrieved: Jun. 2013]

Da Silva, A.C. and da Rocha, H.V. 2013. InkBlog: A Pen-Based Blog Tool for e-Learning Environments. *Issues in Informing Science and Information Technology*, Vol. 10, pp. 121-135.

Dumas, B. et al., 2009. Multimodal Interfaces: A Survey of Principles, Models and Frameworks. *In Human-Machine Interaction*, Lalanne, D. and Kohlas, J., Eds. Springer Berlin / Heidelberg, Berlin, German, pp. 3-26, doi: 10.1007/978-3-642-00437-7_1.

Gruenstein, A. et al., 2008. The WAMI Toolkit for Developing, Deploying, and Evaluating Web-Accessible Multimodal Interfaces. *Proceedings of 10th International Conference on Multimodal Interfaces (ICMI 2008)*, Chania, Greece, pp. 141-148, doi: 10.1145/1452392.1452420.

Lalanne, D. et al., 2009. Fusion Engine for Multimodal Input: A Survey. *Proceedings of the 11th International Conference on Multimodal Interfaces (ICMI-MLMI'09)*, Cambridge, USA, pp. 153-160, doi: 10.1145/1647314.1647343.

Mayes, T., 1992. The 'M' Word: Multimedia interfaces and their role in interactive learning systems. *In Multimedia Interface Design in Education*, Edwards, A. D. N., Holland, S. Eds. Springer-Verlag, Berlin, German, pp. 1-22.

McCobb, G., 2013. *The W3C Multimodal Architecture, Part 2: the XML specification stack*. Available at <http://www.ibm.com/developerwo rks/web/library/wa-multimodarch2/index.html>. [retrieved: Jan. 2013]

Moodle Community, 2013. *Moodle.org: Open-source community-based tools for learning*. Available at <http://www.moodle.org>. [retrieved: Jan. 2013].

Oviatt, S. L, 2003. Advances in Robust Multimodal Interface Design. *In IEEE Computer Graphics and Applications*, Vol. 23, No. 5, pp. 62-68, doi: 10.1109/MCG.2003.1231179.

SAKAI Community, 2013. *Sakai Project | collaboration and learning - for educators by educators*. Available at <http://sakaiproject.org>. [retrieved: Jun. 2013].

TelEduc Project, 2013. *TelEduc. Ensino a Distância*. Available at <http://www.teleduc.org.br>. [retrieved: Jun. 2013].

W3C, 2013. *Multimodal Architecture and Interfaces.* Available at < http://www.w3.org/TR/2012/REC-mmi-arch-20121025/ >. [retrieved: Jan. 2013]

Capítulo 8

Desafios da Multimodalidade em Ambientes de EaD

Este Capítulo, composto por um artigo publicado na *International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies,* se baseia na experiência obtida na realização deste trabalho e discute a existência de uma influência entre as partes tecnológicas (hardware e software) e questões educacionais (atividades de ensino-aprendizagem) quando se utilizam dispositivos móveis e ambientes de EaD. A mudança da modalidade de interação, os equipamentos existentes no dispositivo, bem como o navegador e aplicativos instalados no dispositivo móvel, além de influenciarem na interação entre o usuário e o ambiente, também podem influenciar nas atividades de ensino-aprendizagem.

Assim, para especificar atividades de ensino-aprendizagem no contexto de multi-dispositivos, é necessário compreender quais dispositivos o aluno estará utilizando para interagir. Por exemplo, celulares inteligentes possibilitam uma rápida resposta devido a sua portabilidade, mas pode impactar na elaboração de textos, ocasionando textos mais curtos. Por outro lado, podem potencializar as atividades de ensino-aprendizagem, como por exemplo, possibilitar aos alunos a criação de podcasts por meio da gravação de áudio usando celulares inteligentes.

Nesse contexto de diversidade de dispositivos, a tarefa do professor se torna ainda mais difícil em relação à especificação das atividades de ensino-aprendizagem. É necessário pensar nos periféricos que os alunos terão à sua disposição (podendo ser que um grupo de alunos tenha periféricos diferentes de um outro grupo). Assim, exige-se não somente a interoperabilidade de dados para que um determinado dispositivo possa visualizar o conteúdo gerado por um outro dispositivo, mas que haja flexibilidade na atividade de ensino-aprendizagem de modo a possibilitar que os alunos desempenhem as atividades usando os dispositivos de que dispõem.

Percebendo a influência entre os periféricos e as atividades de ensino-aprendizagem, e a relação entre modalidade e curso (contexto educacional) na produção de um conteúdo, conforme os ambientes de EaD se beneficiem de uma modalidade, possibilitará um maior sucesso em seu emprego no contexto educacional.

Challenges for e-Learning Environments in m-Learning Contexts

A survey about the hardware, software, and educational dimensions

André Constantino da Silva
Institute of Computing (PG)
UNICAMP, IFSP
Campinas, Brazil, Hortolândia, Brazil
acsilva@ic.unicamp.br

Heloísa Vieira da Rocha
Institute of Computing, NIED
UNICAMP
Campinas, Brazil
heloisa@ic.unicamp.br

***Abstract*— e-Learning environments are applications that use the Web infra-structure to support teaching and learning activities. Their user interfaces were designed to have good usability using a desktop computer with keyboard and mouse as input devices and a high resolution medium-size display and speak loud as output devices. Devices, such as tablets and smartphones, have computational power enough to render Web pages, allowing mobile users navigate through the e-Learning environments and closing the e-Learning environments to the m-Learning contexts. But to have an effective use in the mobile context, the e-Learning environments need to be analyzed in three dimensions: hardware, software, and educational. This paper presents a survey about works in merging e-Learning environments and m-Learning and discusses about challenges such hardware and software integration, integration between the browser and the environment and changes on the teaching and learning activities.**

Keywords-Human-Computer Interaction; Electronic Learning Environment; Mobile Devices; Interaction Styles.

I. INTRODUCTION

e-Learning environments, such as Moodle [1], SAKAI [2], TelEduc [3], Ae [4], are applications that use the Web infra-structure to support teaching and learning activities. The e-Learning environments are designed to support a variety of users and learning contexts, but they are designed to conventional computers, usually equipped with keyboard and mouse as input and a medium screen and speakers as output; a limited interaction style for nowadays devices. These modalities and the technology shape the teaching and learning activities done in the e-Learning environments; they focus on reading and writing skills. Despites these technology impact, the e-Learning environments need to have good usability, accessibility, performance, security, availability and other software attributes.

Devices, such as smartphones and tablets, are becoming increasingly popular; most of them have touch screen displays, access to the Internet and enough computing power to process Web pages. So, Web sites and Web applications, initially developed to be used with keyboard, mouse and a medium size display, are being accessed by small touchscreen devices. This can be considered as another aspect of accessibility, so the environments´ development teams are building solutions to provide access on mobile devices, and the use of mobile devices with educational purposes is called m-Learning. Two motivations for m-Learning are to allow participants interact anytime and anywhere with the content and with each other.

The Internet was built to connect many technologies, so allowing mobile users to access e-Learning environments make the boundary between e-Learning and m-Learning not so clear. This brings some challenges for e-Learning environments developers and m-Learning practitioners. This paper presents a survey about the works to merge these fields. Section II presents a literature review about e-Learning, m-Learning and e-Learning environments. Section III presents the three dimensions to analyze the use of e-Learning environments to support mobile users. Section IV presents challenges to have an effective use of the e-Learning environments in the m-Learning context. Section V presents final considerations.

II. E-LEARNING AND M-LEARNING

e-Learning refers to the use of Information and Communication Technologies (ICT) in education. So, e-Learning is any distributed learning experience through the Internet, Intranet, Extranet, CD or DVD-ROM, because the e-Learning main concern is related with the teaching form and not with technology [5]. The e-Learning term refers broadly the terms Web-based learning, Internet-based learning, online learning, distributed learning, and computer-based learning. e-Learning is suited to distance learning, but it can also be used in conjunction with face-to-face teaching, calling blended learning.

Online systems that support e-Learning through the Web are called e-Learning environments or Virtual Learning Environments (VLE) or Learning Management Systems (LMS). Moodle [1], SAKAI [2], TelEduc [3] and Ae [4] are examples of e-Learning environments.

The first e-Learning environments are designed to integrate the content with communication tools arranged on the Internet (such as chat, mail and forum). They evolved increasing the number of tools for content publishing, participant´s communication and course administration, taking advantages of the Web to offer content with text, images, audios and videos in a hypertext document. Tools like chat, forums, portfolios, repositories are widely used, and tools those explore the audio and video resource to user communication, such as instant messenger and video-conferences, are becoming common among the environments.

The environments´ development teams are building solutions to provide access on mobile devices, and the use of mobile devices with educational purposes is called m-Learning.

m-Learning is any kind of learning that happens when the learner is not at a fixed, predetermined location, or learning that happens when the learner takes advantage of the learning opportunities offered by mobile technologies [6]. m-Learning Technology includes handheld computers, MP3 players, notebooks, mobile phones and tablets; devices that have owner autonomy and it is easy to portable. Two motivations for m-Learning is allow the participants interacting anytime and anywhere with the content and with each other.

The recent technology forwards brings a broad of new applications made to support many areas, such Education. One of the mobile devices that are gained repercussion in this scenario is the Tablet PC (Tablet Personal Computer), a device with height as similar to a notebook and had an input device similar a pen. The paper and pen metaphor implies that tasks performed before in paper, like draw or manuscript writing, can be more natural in the Tablet PC than in the another computing devices. Resuming, the Tablet PC has the following hardware characteristics: (i) Pen sensitive screen; (ii) Screen that allows different positions; (iii) Wireless network access by Wireless Local Area Network (WLAN) and bluetooth technology; (iv) Microphones and embedded loudspeakers; (v) Keyboard (some models the keyboard are detachable); (vi) Batteries. Usually, the screen size ranges from 9 inches to 12 inches.

Other devices used in m-Learning are smartphones, which have touchscreen with computation power enough to render Web pages. Usually, the screen size ranges from 3 inches to 5 inches. Resuming, the smartphones have the following hardware characteristics: (i) touch sensitive screen; (ii) Screen that allow two different positions (landscape and portrait); (iii) Wireless network access by WLAN (Wireless Local Area Network), by 3G and bluetooth technology; (iv) Microphones and embedded loudspeakers; (v) Virtual keyboard (some models have a physical keyboard); (vi) Batteries.

Gay *et al.* [7] suggest that the introduction of wireless computing resources in learning environments can potentially affect the development, maintenance, and transformation of learning communities.

It is possible to use e-Learning environments for support m-Learning and three kind of solution are emerging: i) specific device application; ii) web site specific for mobile devices; and iii) improve the web site for mobile and desktop access [8]. Each solution has its vantages and disadvantages due the hardware and software restrictions of the mobile device.

Building specific device application allows designing a suitable user interface for the device and taking advantages of smartphone´s features, such as touchscreen and camera, but needs develop an application for each mobile platform. So, the applications to be developed needs specific knowledge programming skills and increases the code lines number to maintain. Moodle community offers the Moodle App [9] and Moodbile [10], two native mobile applications with versions for the most popular smartphone´s platforms.

Moodle, since version 2.1, offers a Web site specific to mobile devices, an example for the second type of solutions for access e-Learning environments in mobile devices. Building a specific Web site to mobile device allows designing a suitable user interface for mobile devices taking account some common characteristics, such small touchscreen, but depends of the browser to access some platform features, such GPS, and increases the code lines number to maintain too.

The latter solution considers that smartphones and tablets have enough computational power to render Web pages and to do some adaptation if it is necessary, and offer the same user interface for any device. But the interaction styles may vary, so to design this kind of user interface it is necessary to do some usability studies to found barriers or user interaction problems for each interaction style. Disadvantages of this solution are to depend of browsers to use the mobile features and the difficult of consider many interaction styles in the same user interface. Da Silva, Freire and da Rocha [8] point out some problem that happen when a user interface designed to be used with specific interaction hardware is used with other interaction hardware.

Another e-Learning environment characteristic is to be used in many of learning contexts, e.g., teacher training, undergraduate courses, and team training in all areas of knowledge. We call this property as learnability. But, the actual hardware of conventional computers increases the difficulty to use the environment to produce content for any area and support student activities, e.g., to post a mathematic exercise that needs write formulas to resolve, the user need to use a specific software and post the file since the e-Learning environments do not support directly this kind of content.

The e-Learning environments need to be usable and accessible for many users in many social, physical, technological and learning contexts. So, e-Learning environments need to be evaluated in the usability, accessibility, mobility and learnability dimensions.

III. Dimensions for Allowing e-Learning and m-Learning Merging

Khan [11] defines eight dimensions for a good e-Learning environment: Resource Support, Ethical, Institutional, Pedagogical, Technological, Interface Design, Evaluation and Management. Since in our work we want to use e-Learning environments in m-Learning contexts, the Technological, Interface Design, and Pedagogical dimensions need to be discussed due the impact caused by the device changing. We will join the Technological and Interface Design dimensions, since there is an intrinsic relation between them, and our focus is on mobile users. The Pedagogical dimension is related about the teaching and learning activities developed by the teachers and done by the students.

For technological dimension, we propose the e-learning environments needs to be evaluated about their usability, accessibility, mobility, and other software requirements, such performance, scalability and availability. Due the diversity of users whom may use the e-Learning environments, these systems need to have good usability so that the user interface does not prejudice the teaching and learning activities, so the user interface cannot be a barrier between the student and the content or the users and their goals or injure the course activities.

Nielsen [12] defines usability as a combination of five elements: easy to learning, efficient, easy to remember, low probability of users do mistakes and user satisfaction. Nielsen proposes a method for evaluating the user interface usability, and other methods can be found on the Human-Computer Interaction literature [13] [14].

Accessibility, another important requirement, is about to allow disabled people use the environment, and can be understood as to be accessed by anyone, whatever their hardware or software. So to have a high level of accessibility, a Web page needs to be accessed by computers or any other computational device, such the smartphones and Tablet PCs, which can have assistive technology installed. Some methods can be found to evaluate the application accessibility, mainly for Web pages and applications, e.g., evaluating the interface conformance with the W3C guidelines [15].

In the educational dimension, the teacher needs to dispose content, to plain activities, to ask questions of the students about the activities and the content, to evaluate activities and other tasks. Pedagogically, the teacher needs to select the tools to be used to dispose the content and to run the activities. Usually, due the e-Learning environment tools and the hardware used to interact, a desktop computer, the activities are discussions by Forums, writing individual or collaborative texts. The texts may have pictures, but it is similar as works done in presence education, where the work can be delivered in a printed format. This kind of work does not take advantages of multimedia and the hardware available in the mobile devices. We discuss these problems in the next sections as challenges for e-Learning environments in the mobile contexts.

IV. CHALLENGES FOR E-LEARNING ENVIRONMENTS IN THE M-LEARNING

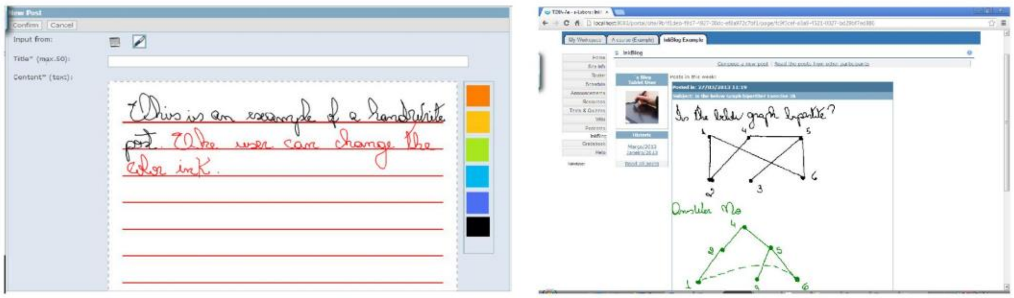
About the technology issues, it is need to analyze the hardware and the software and, since there are a variety of computing devices, aspects of human-machine interaction (such as ergonomics) and the device characteristics need to be considered. The most visible problem in browsing e-Learning environments using mobile device is the user interface adaptation. One of the adaptation issues is the content readability. For tablets it is not a big problem, since it is possible render the web page without injure de readability; the size of the tablet´s screen is not so small compared to the desktop display. But, in the case of smartphones, the screen size is not large enough to ensure readability, so it is necessary to have a page adaptation or techniques to visualize the entire page and use zooming to see the page details.

The user interface adaption is not a trivial task, a challenge not only for e-Learning environments, but for the Human-Computer Interaction, is to build system with adaptive user interface. Bickmore and Schilit [16] present a heuristic-based approach for Web page adaptations to be rendered in mobile devices with small screens, but their work does not consider audio and video adaptation. Zhang [17] purposes a framework to do content adaptation for systems accessed by a sort of devices (multidevices), increasing accessibility and doing a distribution optimization over the network. Oliveira and da Rocha [18] purpose priorities for consistence in the adaptation to maintain the same conceptual model for mutidevices, whom done a study case over an e-learning environment. Pyla *et al.* [19] discuss about the task migration between devices: the user starts the task in one device, go to another one to perform some sub-task, and change the device until the task be done.

But, Web applications should consider the Web architecture model, i.e., the client-server, to have a better adaptation. Client is responsible to render the user interface through a browser. It is in the client side that the user interacts with the system using input and output hardware. The server is responsible to process client´s requests and data persistence, but it knows few about the input and output devices in client side. Since the server is responsible to produce the user interface code to be displayed on the client, the server needs to know about the user device, the user location and user preferences to do a better adaptation. The generated interface needs to have good usability.

For example, smartphones are good to read and write small texts, such as post-it notes. The tablet with touchscreen allows users have mobility and a good readability, but it is not so efficient to write texts. The tablets equipped with pen sensitive screen should be used to write formulas or do sketches, instead of only as a pointing device. Da Silva and da Rocha [20] propose the InkBlog tool, a blog tool that receives input data from stylus in a pen sensitive tablet so that the user can handwriting her posts (Fig. 1a).

The InkBlog is a tool for the Ae e-Learning environment, and take advantages of the Pen-based computing to allow users interact with the e-Learning environment with a pen. The authors comment that using the InkBlog is possible a better support to disciplines such Graph Theory and Computer Theory. Without the InkBlog, the user needs to use paper and pencil to resolve an exercise and use specific hardware, such scanner, to digitalize the solution. Or the user needs to use a special application to draw a graph. In both solutions, the user posts the picture as an attached file in weblog post. Using the InkBlog the user can sketch the graph direct on the weblog tool through direct manipulation (Fig. 1b). InkBlog was tested in iPhone (Fig. 2) and Android devices using the stock browser in each device. Both devices display correct the posts, but due the platform does not distinguish between touch and stylus press, it is not possible to handwrite a post in these smartphones; both devices recognize the input as page scrolling. This is another challenge, the web pages depends on the browser, because the browser is responsible to the capture of the user input and the page rendering. So, the hardware and the browser needs be more integrated to enhance the Web application.



(a) (b)

Figure 1. Using InkBlog to (a) handwriting a post and (b) to share a solution for an exercise about Graph Theory. Both pages rendered by Chrome Browser.

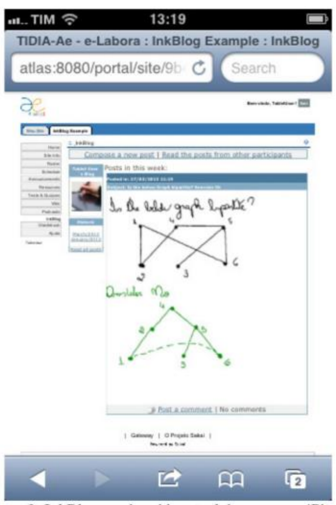


Figure 2. InkBlog rendered by stock browser on iPhone.

In the case of use e-Learning environments on mobile devices, is important to study the interaction problems that happen, since the actual environments was designed to be used in a desktop computer, and mobile phones have other interaction devices. da Silva, Freire, Arruda and da Rocha [21] present some problems when the users uses an mobile phone to interact with the Teleduc environment using an Android-based mobile device and a iPhone. Da Silva, Freire and da Rocha [8] present some problems when user uses an android-based smartphone and a Tablet PCs. They present a taxonomy about the interaction problems: cross-platform problem, cross-modality problem and platform and modality-independent problem. These problems are barriers or difficulties to the user navigate thought the environment. So one challenge is identify these problems and correct.

The environments need to have good usability and accessibility, but the content created inside the environments needs to. Since the environments are used by a diversity of people, most of them do not have knowledge about Web accessibility, they can create content with low accessibility in mobile devices. Fig. 3 shows an agenda on a course in TelEduc created by the teacher to describe about the topics and activities to be done in a week. Some problems when the agenda is rendered in iPhone (b) can be viewed with compared with the desktop computer (a). So, another challenge is to develop authoring tools that easily create accessible contents and to develop features to allow the mobile users to visualize the published content [22].

Allowing access by mobile devices is not the only challenge to facing it. The mobile devices have specific hardware that can be used to produce content and to use to interact with the environment. The camera can be used to take photos or do videos and the microphone can be used to create audio files to be published by the teacher or by the students as content in the e-Learning environment. To allow this it is necessary to have a better integration between the mobile hardware, the browser and the environment to easily allow users create content using camera, microphone or any other input device. These devices can too be used as communication if integrated with communication software. The actual versions of e-Learning environments only dispose an action button that trigger a dialog box where the user can choose the photo or video he want upload. Depending on the solution adopted to integrate the environment in the mobile device, it is possible to turn the publishing task easier. In the case of a specific device application it is possible to customize the mobile Operation System and include sharing option in the photo and video gallery, similar as Facebook app does (Fig. 4). So, the mobile user has options to publish the media in her virtual space on the e-learning environment, like a Portfolio item on TelEduc environment, or a resource item on SAKAI and Ae.

The TelEduc e-Learning environment has a notification tool that send e-mails for the course participants describing what happen since the last logging. This feature could be integrated with the mobile device advisor management system. So one more challenge is identify e-Learning features that can be integrated with mobile devices features.

The e-Learning environment has data about the course participants, whom can be added on the mobile device

 ISBN: 978-1-61208-289-9

(a) (b)

Figure 3. A content created by a user in TelEduc with accessibility problems displayed on desktop computers (a) and on a mobile device (b).

Contacts tool to turn easily sent messages between the participants. Another one feature is synchronizes the activities deadline or important dates to the mobile device´s Calendar tool.

Allowing student to create multimedia content in activities, the teacher needs to select the best media or a set of them to planning the activities. The student´s works not need be more only texts, allowing them to produce videos, audios, photos or a multimedia work. If the teachers allow user to use the mobile phone, it is important elaborate activities that explore the mobile hardware, such take photos, make videos or record an audio. The student can write some text too, but it is important notice (and this is must be clear on the activity) that usually it is not expected a small texts for student´s activities; since writing on smartphones can be a difficult task, the teacher must specify how long the text must be. This perception impacts on the pedagogical use of the e-learning environment.

The actual user interface design techniques take account just a limit set of input and output hardware, limited to the context, such as techniques to design user interface for desktop or for mobile platforms. But, there is a lot of input or output hardware in these devices and these techniques are asked to consider all of them. Some input and output devices are: touchscreen, microphones, pen sensitive screen, touchpad, TrackPoint, accelerometers, joysticks, loudspeakers, small screen, large screen, printers, etc. One solution to deal with this variety of devices is use multimodal techniques on the e-Learning environment´s user interface. Multimodal interaction is a research proposal to turn the interaction between humans and machines more natural, i.e., more close to the interactions between two humans, and have the benefits to increase the usability, flexibility and convenience [23]. According to Oviatt [24] "multimodal interfaces process two or more combined user input modes (such as speech, pen, touch, manual gesture, gaze, and head and body movements) in a coordinated manner with multimedia system output". But, developing multimodal interaction systems is a complex task [25]; da Silva and da Rocha propose the |Ae| environment, an e-Learning environment with multimodal user interface, discuss how multimodal architecture changes the e-Learning environment architecture. Despite these works, the challenge is to have an

Figure 4. The integration between iPhone´s Photo Gallery and Facebook App.

e-Learning environment to deal with the variety of input and output modes in the device, using it in a proper way.

V. Final Considerations

e-Learning is the use of the Web structure to support teaching and learning activities. There is some Web-based software, the e-Learning environments, which support these activities in courses. Due the technology development, it is possible to use mobile devices to teaching and learning activities, calling m-Learning. The mobile devices can access e-Learning platform, but to an effective course accessed by mobile devices it is necessary that the e-Learning does not have barriers or difficulties for the mobile devices, and the content and activities must be shaped to consider the mobile devices and their hardware, such camera and microphone.

We believe that e-Learning and m-Learning will merge in the case of have an environment to support the teaching and learning activities and brings benefices on learning. But to plain an effective course, it is important to consider the hardware and software that the students will used. So, the technology needs to be considered in the instructional design process, together with the educational issues. Future works are detail each part of these dimensions.

 ISBN: 978-1-61208-289-9

ACKNOWLEDGMENT

The authors thank the CAPES and CNPq for financial support and for FAPESP through TIDIA-Ae Project (n. 05/60572-1), which provided the equipment.

REFERENCES

[1] Moodle Trust, "Moodle.org: open-source community-based tools for learning," Available at <http://moodle.org>. [retrieved: Jul. 2013]

[2] SAKAI Environment, "Sakai Project | collaboration and learning - for educators by educators," available at <http://sakaiproject.org>. [retrieved: Jul. 2013]

[3] TelEduc Environment, "TelEduc Ensino à Distância," available at <http://www.teleduc.org.br>. [retrieved: Jul. 2013]

[4] Ae Project. "Ae - Aprendizado Eletrônico Environment," available at <http://tidia-ae.iv.org.br/>. [retrieved: Jul. 2013]

[5] J. R. Lima and Z. Capitão, e-Learning e e-Conteúdos, Portugal: Inova, 2003.

[6] M. Sung, J. Gips, N. Eagle, A. Madan, R. Caneel, et al., "Mobile-IT Education (MIT. EDU): m-learning applications for classroom settings," in Journal of Computer Assisted Learning, vol. 21, May 2005, pp. 229–237.

[7] G. Gay, M. Stefanone, M. Grace-Martin, and H. Hembrooke, "The effects of wireless computing in collaborative learning environments," in International Journal of Human–Computer Interaction, vol. 13, n. 2, 2001, pp. 257–276, doi: 10.1207/S15327590IJHC1302_10.

[8] A. C. da Silva, F. M. P. Freire, and H. V. da Rocha, "Identifying Cross-Platform and Cross-Modality Interaction Problems in e-Learning Environments," Proc. of the 6th International Conference on Advances in Computer-Human Interactions (ACHI 2013), IARIA, Feb. 2013, pp. 243-249.

[9] Moodle Trust, "Moodle App," available at: <http://docs.moodle .org/dev/Mobile_app>. [retrieved: Jul. 2013].

[10] M. J. Casany, M. Alier, E. Mayol, J. Piguillem, N. Galanis, et al., "Extending Moodle Services to Mobile Devices: The Moodbile Project," Proc. of the 6th International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM 2012), IARIA, Sep. 2012, pp. 24-28.

[11] B. H. Khan, "A framework for Web-based learning," in Web-based training, B. H. Khan, Ed. Englewood Cliffs, NJ: Educational Technology Publications, 2001, pp. 75-98.

[12] J. Nielsen, Usability Engineering, EUA: Morgan Kaufmann, 1993, 362 p.

[13] M. G. Helander, T. K. Laundauer, and P. V. Prabhu (Editors), Handbook of Human-Computer Interaction, Amsterdam, Elsevier Science B. V., 1997.

[14] J. Rubin, Handbook of Usability Testing: How to Plan, Design and Conduct Effective Tests, New York: John Wiley & Sons, 1994, 330 p.

[15] W3C, "Web Content Accessibility Guidelines (WCAG) Overview," available at: <http://www.w3.org/TR/WCAG/>. [retrieved: Jul. 2013]

[16] T. W. Bickmore and B. N. Schilit, "DIGESTOR: Device-independent Access to the World Wide Web," Proc. of 6th International World Wide Web Conference, Elsevier Science Publishers Ltd., Sep. 1997, pp. 1075-1082, doi: 10.1016/S0169-7552(97)00026-3.

[17] H. Zhang, "Adaptive Content Delivery: A New Research Area in Media Computing," Keynote Speech on Proc. of the 2000 International Workshop on Multimedia Data Storage, Retrieval, Integration and Applications, Jan. 2000. Available at: <http://research.microsoft.com/apps/pubs/default.aspx?id= 68791 >. [retrieved: Jul. 2013]

[18] R. Oliveira and H. V. da Rocha, "Consistency Priorities for Multi-device Design," Proc. of 11th IFIP TC 13 International Conference on Human-Computer Interaction (INTERACT), Springer-Verlag, Sep. 2007. pp. 426-429.

[19] P. Pyla, M. Tungare, and M. Perez-Quiñones, "Multiple user Interfaces: Why consistency is not everything, and seamless task migration is key," Proc. of CHI 06 Workshop on the Many Faces of Consistency in Cross-platform Design, vol. 24, CEUR-WS, Apr. 2006. pp. 55-60.

[20] A. C. da Silva and H. V. da Rocha, "InkBlog: A Pen-Based Blog Tool for e-Learning Environments," in Bridging Disciplinary Boundaries: Issues in Informing Science and Information Technology, vol. 10, May 2013, pp. 121-135. ISBN: 1932886699.

[21] A. C. da Silva, F. M. P. Freire, A. V. P. de Arruda, and H. V. da Rocha, "Interaction Problems Accessing e-Learning Environments in Multi-touch Mobile Devices: a case study in TelEduc," Proc. of the 7th International IADIS Conference on e-Learning, IADIS Press, Jul. 2013, pp. 199-206.

[22] A. C. da Silva and H. V. da Rocha, "Improving TelEduc Environment With GPL Software: The Visualization Content Case for Computers and Mobile Devices," Proc. of the 6th International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM 2012), IARIA, Sep. 2012, pp. 222-227.

[23] D. Lalanne, L. Nigay, P. Palanque, P. Robinson, J. Vanderdonckt, et al., "Fusion Engine for Multimodal Input: A Survey," Proc. of the 11th International Conference on Multimodal Interfaces (ICMI-MLMI´09), ACM Press, Nov. 2009, pp. 153-160, doi: 10.1145/1647314.1647343.

[24] S. L. Oviatt, "Advances in Robust Multimodal Interface Design," in IEEE Computer Graphics and Applications, vol. 23, n. 5, Sep. 2003, pp. 62-68, doi: 10.1109/MCG.2003.1231179.

[25] B. Dumas, D. Lalanne, and S. Oviatt, "Multimodal Interfaces: A Survey of Principles, Models and Frameworks," in Human-Machine Interaction, D. Lalanne and J. Kohlas, Eds. Berlin: Springer Berlin/Heidelberg, 2009, pp. 3-26, doi: 10.1007/978-3-642-00437-7_1.

Capítulo 9

Conclusão

A IHC é “a disciplina preocupada com o design, avaliação e implementação de sistemas computacionais interativos para uso humano e com o estudo dos principais fenômenos ao redor deles” (ACM SIGCHI, 1992, p. 6; Rocha e Baranauskas, 2003). Segundo a IHC, projetar uma interface, ou seja, fazer um design de interface, é uma atividade que envolve: (i) conhecer a tecnologia e suas restrições; (ii) conhecer os usuários, seus conhecimentos, suas preferências e limitações; (iii) a área da aplicação e as tarefas que serão realizadas por meio do dispositivo computacional; (iv) o contexto de uso, além de ferramentas adequadas para apoiar o processo de desenvolvimento da interface de usuário.

Assim é necessário compreender algo que também evolui: o hardware, as áreas de aplicação e, principalmente, os anseios das pessoas em relação ao que podem realizar com um dispositivo computacional. Se antes projetava-se interfaces para um número conhecido de usuários, hoje projeta-se para milhões de pessoas. Se antes projetava-se interfaces para apoiarem o trabalho, hoje a computação é aplicada em vários nichos diferentes: educação, saúde, governo, entretenimento (Shneiderman, 2006). Atualmente, um grande desafio é possibilitar que vários dispositivos acessem a mesma aplicação. Logo é necessário ir além de possibilitar o acesso a ferramentas de TIC usando um dispositivo qualquer. É necessário considerar que o número de dispositivos está aumentando ao longo dos anos e possibilitando que usuários interajam com as aplicações de forma diferente da habitual. Conforme descrito no Capítulo 1, Preece *et al.* (1994) já apontavam esses problemas: como dar conta da rápida evolução tecnológica? Como garantir que os design ofereçam uma boa IHC ao mesmo tempo que exploram o potencial e as funcionalidade da nova tecnologia?

É nesse contexto, cujo objetivo é adaptação a evolução tecnológica, que este trabalho se insere. A variedade de dispositivos computacionais, composta por uma relação de periféricos que vão além do mouse, teclado físico e monitor de tamanho médio, e incluem tela sensível a um ou vários toques, tela sensível a caneta, microfone, câmera e sensores de movimento, aumentaram significativamente os modos como os usuários podem interagir com as aplicações, trazendo desafios para a Computação, inclusive para a área de IHC. Esses periféricos possibilitam ao usuário interagir com diferentes modalidades: voz, gestos, escrita manuscrita, toque, entre outras. Cada modalidade oferece uma experiência diferente ao usuário e, quando comparadas, apresentam diferentes pontos fortes e pontos fracos, e também, diferentes relações com os sistemas perceptual, cognitivo e emocional do ser humano (Bernsen, 2008).

O número de dispositivos que podem ser criados pode ser muito grande, tornando-se quase impossível desenvolver uma aplicação que funcione todos os dispositivos. Como alternativa, propomos o suporte a diversas modalidades como solução ao *design* de interface para uma aplicação que será utilizada em uma variedade de dispositivos.

Aplicações que possibilitem ao usuário interagir por diversas modalidades, ou seja, multimodalidade, é o campo de pesquisa da Interação Multimodal. O trabalho seminal neste campo é atribuído a Bolt (1980), que desenvolveu uma aplicação para mover objetos em uma tela cujos dados de entrada eram oriundos das modalidades de voz e gesto. Diversos avanços nessa área de pesquisa foram conquistados desde o trabalho de Bolt, podendo citar: estudo da combinação de duas modalidades; construção de protótipos de sistemas multimodais; definição de arquitetura para ambientes multimodais; processos de engenharia e *frameworks* para construção de sistemas multimodais; levantamento de impactos da interação multimodal; variáveis que interferem na interação multimodal, como tempo e local, conforme descrito no Capítulo 2.

As interfaces multimodais diferem das interfaces monomodais por terem um grau de incerteza da ação que o usuário deseja e são, portanto, denominadas probabilísticas (Dumas, Lalanne e Oviatt, 2009). O grau de incerteza pode ocorrer, por exemplo, na atribuição de significado dos dados oriundos de uma modalidade no processo de reconhecimento, pela falta de informação de uma modalidade, ou até mesmo pela contradição dos dados de modalidades diferentes. Embora haja problemas de interação ocasionados pelo grau de incerteza, pesquisadores apontam como benefício da multimodalidade a facilidade de uso, aumento na eficiência, na flexibilidade e na conveniência e preferência dos usuários em utilizar aplicações multimodais ao invés de aplicações monomodais. Esses benefícios são os componentes da usabilidade, conforme definido por Nielsen (1993). O emprego da multimodalidade também impacta no aumento da acessibilidade, pois o usuário pode utilizar a modalidade de interação mais apropriada considerando as suas necessidades específicas.

Frente ao desafio de desenvolver aplicações para diversos dispositivos aumentando a usabilidade e a acessibilidade da aplicação e frente aos promissores resultados das pesquisas de Interação Multimodal, este trabalho propôs-se a estudar a aplicação da multimodalidade em ambientes de EaD. Para realizar esta investigação, a abordagem de investigação usada envolvia incrementar um determinado ambiente de EaD desenvolvido para computadores convencionais (portanto, modalidade de entrada composta por teclado e mouse, sendo a modalidade de saída uma tela de tamanho médio e alta resolução) com modalidades não consideradas em seu projeto (a modalidade de entrada de dados a caneta e a toque), adotando-se os seguintes passos:

1. Investigação individual das modalidades de caneta e de toque:

1a. Estudo da modalidade de entrada por caneta, considerando também a interação por gestos gerados pelo movimento da caneta;

1b. Estudo da modalidade de entrada por toque, considerando também a interação por gestos gerados pelo movimento do dedo;

2. Desenvolvimento de soluções para os problemas encontrados no passo 1:

2a. Correção dos problemas de interação identificados;

2b. Desenvolvimento de um editor para a entrada de textos manuscritos e sua integração ao ambiente;

2c. Desenvolvimento de um componente de interface para a entrada de gestos por caneta ou por toque e sua integração ao ambiente;

3. A proposta de uma arquitetura para ambientes de EaD com interação multimodal, considerando os componentes desenvolvidos.

Para a investigação dos problemas de interação utilizou-se o ambiente TelEduc, enquanto que para a elaboração da arquitetura optou-se pelo ambiente Ae, por este possuir uma arquitetura documentada.

A Seção a seguir apresenta uma síntese das contribuições, que estão diretamente relacionadas com a abordagem empregada. As lições aprendidas são descritas na Seção 9.2, e os trabalhos futuros na Seção 9.3. A Seção 9.4 apresenta as considerações finais.

9.1 Síntese das Contribuições

Adotar várias modalidades em um ambiente de EaD impacta na sua arquitetura de software. Logo, considera-se como a principal contribuição deste trabalho a apresentação de uma arquitetura de ambientes de EaD com interface multimodal. A partir de estudos sobre arquiteturas multimodais e arquiteturas de ambientes de EaD, identificou-se relações entre os componentes do ambiente de EaD e dos componentes que tratam da multimodalidade. Os componentes que tratam da multimodalidade necessitam de dados que lhes informem das funcionalidades existentes no sistema e permissões de quem pode executá-las, dos usuários, seus históricos e preferências. Dados que podem ser encontrados em componentes dos ambientes de EaD, como gerenciadores de cursos, de usuários e de ferramentas. Portanto, esses componentes precisam ter um interfaceamento que possibilite que eles troquem dados.

Outra contribuição deste trabalho foi a categorização dos problemas de interação que ocorrem quando uma determinada interface é acessada por um dispositivo não considerado durante o seu projeto. Percebeu-se que quando se altera o dispositivo, pode ocorrer uma mudança de plataforma (por exemplo, sistema operacional ou navegador) e/ou uma mudança de modalidade (por exemplo, do teclado mais mouse para o toque). Três categorias foram levantadas:

(i) Problemas de interação sem relação com a mudança de modalidade ou plataforma;
(ii) Problemas de interação relacionados com a mudança de plataforma;
(iii) Problemas de interação relacionados com a mudança de modalidade.

Acessar uma aplicação por modalidades que não foram consideradas durante o projeto da interface ocasionará, na maioria das vezes, em problemas de interação. Esses problemas são indesejados pois diminuem a usabilidade e da aplicação e portanto devem ser identificados e solucionados. Neste trabalho adotou-se dois métodos de avaliação de usabilidade para identificação dos problemas: inspeção da interface por especialista e testes com usuário. Os métodos mostraram-se eficazes na identificação de problemas.

Uma contribuição que consideramos pioneira no contexto da plataforma Web, ambientes de EaD e interação baseada em caneta é o InkBlog, uma ferramenta de blog que possibilita os usuários postarem mensagens manuscritas ou desenhos elaborados por meio de uma caneta. Até o momento, as ferramentas que exploram o conceito de *Pen-Based Computing* eram especificamente destinadas ao dispositivo baseado em caneta e fora da plataforma Web. A ferramenta foi estendida para também receber dados da modalidade toque. Os componentes InkEditor e InkRenderer, utilizados para tratar os dados das diferentes modalidades e exibir a tinta eletrônica, podem ser utilizados na geração de novas ferramentas.

Uma última contribuição foi uma discussão dos limites da tecnologia atual frente a multimodalidade e apresentação de desafios para a área. A implementação atual dos sistemas operacionais e navegadores torna difícil a coleta de dados a partir de diversos periféricos, e, consequentemente, modalidades. Os estudos realizados neste trabalho foram limitados ao par sistema operacional e navegador pois, pelo fato do sistema operacional mapear as ações das modalidades escolhidas para movimentos do ponteiro, não foi possível obter dados resultantes das duas modalidades ao mesmo tempo. A evolução dos sistemas operacionais e navegadores, de modo a repassarem os dados oriundos dos periféricos concomitantemente é necessária para se ter a multimodalidade nas aplicações Web e, assim, possibilitar uma melhor experiência de navegação em dispositivos com várias modalidades.

Os ambientes de EaD são projetados para que sejam utilizados em qualquer curso (contexto de ensino). Embora seja possível, a sua adoção é mais difícil nos cursos em que a modalidade de escrita digitada (usando o periférico teclado) não se encaixa. Adotando outras modalidades, possibilita que as atividades de ensino e aprendizagem sejam melhor realizadas nos ambientes de EaD, pois os usuários podem utilizar a modalidade mais conveniente, dentre as suas restrições, para elaborar o conteúdo e interagir com o ambiente.

As pesquisas da área de Interação Multimodal se preocupam em empregar a multimodalidade na interação entre usuário e computador, ou seja, usar as modalidades de entrada para acionar funcionalidades e dados necessários para a execução das funcionalidades, e as respostas sejam emitidas nas modalidades de saída. Neste trabalho, os dados oriundos de uma modalidade podem ser também empregados como conteúdo da aplicação. Por exemplo, os dados oriundos da caneta podem ser utilizados para acionar funcionalidades ou adicionar informações para a execução da funcionalidade (exemplo, usar a caneta para acionar um link ou uma funcionalidade por meio de gestos) ou usar a caneta para produzir conteúdo (exemplo, postar um desenho na ferramenta InkBlog).

9.2 Reflexões Finais e Lições Aprendidas

Este trabalho explorou as modalidades de caneta e toque ao acessar um ambiente de EaD, propondo uma arquitetura para a elaboração de um ambiente de EaD com interface multimodal.

A arquitetura mostra os componentes possibilitando que seja tratada a multimodalidade. O recebimento de dados dos vários periféricos e o tratamento desses dados para dar significado a eles e, assim, identificar a ação desejada pelo usuário, é de responsabilidade dos componentes da interface multimodal. Mas somente ter componentes que tratam da multimodalidade não é suficiente para manter a usabilidade de uma aplicação. Por exemplo, considere o problema citado no Capítulo 5, no qual o usuário teve dificuldades em acionar uma ferramenta no menu de ferramentas do ambiente TelEduc (Figura 9.1). O local tocado pelo usuário não aciona o link que abre a ferramenta; o toque deveria ser realizado sobre o nome da ferramenta. O problema é que não existe ação para ser acionada no espaço tocado pelo usuário. E a simples adição de um componente que trate do toque não resolveria o problema encontrado pelo usuário. Desta forma, a usabilidade da aplicação na modalidade toque estaria prejudicada. Portanto, a interface necessita ser avaliada usando a nova modalidade para identificar problemas. Uma outra modalidade poderia minimizar o impacto deste problema, como por exemplo, a voz. O usuário poderia dizer "Abrir ferramenta Material de Apoio" após tentar, seguidas vezes sem sucesso, acionar a ferramenta por toque.

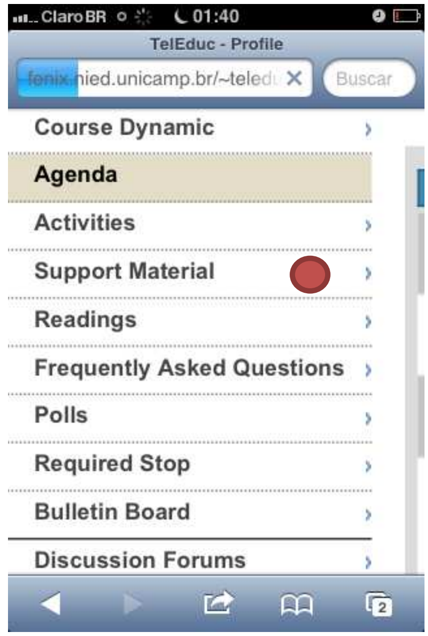


Figura 9.1 – Tela do TelEduc ao ser acesso por um dispositivo móvel com marcação de local tocado pelo usuário ao tentar acionar a ferramenta Material de Apoio.

Uma das limitações deste trabalho está no fato de, embora não tivesse a intenção de fazê-lo, identificar exaustivamente os problemas de interação decorrentes da mudança de modalidade. A pesquisa realizada foi feita para demonstrar que uma interface projetada para uma modalidade pode ter sua usabilidade prejudicada quando se alteram as modalidades. E a simples mudança de modalidade não necessariamente poderá aproveitar os recursos oferecidos pela nova modalidade.

Outra limitação foi a não aplicação da arquitetura em outro ambiente de EaD que não seja o ambiente Ae. Embora suponhamos uma semelhança de componentes arquiteturais entre os ambientes de EaD (componentes como gerenciadores de curso, gerenciadores de usuário, gerenciadores de papéis e permissões), não foram realizados testes para implantar a multimodalidade em outro ambiente.

Por meio deste trabalho percebeu-se a relação de dependência entre os periféricos de interação, o sistema operacional e os navegadores disponíveis no dispositivo. Por exemplo, os navegadores atuais e sistemas operacionais dos Tablet PCs realizam um simples mapeamento entre ações com a caneta em ações do mouse, com exceção de pequenos gestos navegacionais.

Ressalta-se que cada modalidade oferece uma experiência diferente ao usuário, e, portanto, novas funcionalidades surgem quando se adota uma nova modalidade em um software. Percebeu-se que o impacto da multimodalidade pode ir além de possibilitar acesso por diversos periféricos de interação, conforme estudado pela área de Interação Multimodal, podendo emergir novas funcionalidades nas ferramentas que apoiam conteúdos educacionais que sejam difíceis de serem produzidos nos ambientes de EaD. Ou seja, o impacto da multimodalidade pode ir além da interação entre usuário e computador, e possibilitar novas tarefas e práticas que possibilitem que determinados conteúdos sejam melhor trabalhados pelos ambientes de EaD.

Acreditamos que existe uma relação entre modalidade de interação e as atividades de ensino aprendizagem usadas em um ambiente de EaD. Embora muito do conhecimento humano esteja disponível em linguagem verbal escrita, as outras modalidades também possuem sua utilidade. Com a popularização de modalidades que antes não existiam no computador convencional, haverá uma mudança nas atividades de ensino aprendizagem nos contextos que aplicam ambientes de EaD? Shneiderman (2006) comenta sobre a habilidade da escrita:

> “O papel tem uma função ainda mais poderosa do que armazenar conhecimentos. Adquire um poder enorme quando é uma folha em branco, convidando o aluno a criar. Mas a percepção transformadora de que os alunos devem ser mais que meros copiadores levou séculos para vir à tona, e sua promoção deve ter sido difícil. Atualmente, os professores consideram que os alunos devem escrever tão bem quanto ler.” (Shneiderman, 2006, p. 137)

Concordamos com o autor sobre a importância dos alunos dominarem a língua escrita, em particular nos ambientes de EaD, nos quais as ferramentas ainda priorizam a escrita digitada por teclado. Mas acreditamos que, não muito distante, com a influência dos materiais multimídia e a popularização dos diferentes periféricos de interação e com o acesso à Internet por meio de redes de banda larga, os alunos terão que mostrar destreza na elaboração de materiais multimídias e nas várias modalidades para se expressarem. Deverão saber construir esboços, mostrar o desenvolvimento de suas ideias até chegarem ao produto final, seja uma apresentação ou seja um trabalho escrito. Focalizar-se-á no processo de desenvolvimento da ideia e do aprendizado, e não somente ao produto final. Um ambiente de EaD multimodal permitirá realizar melhor este processo? Acreditamos que sim.

9.3 Trabalhos Futuros

Como trabalhos futuros podemos citar:

- Investigação de outras modalidades no ambiente de EaD multimodal: neste trabalho explorou-se a modalidade a caneta e a modalidade de interação a toque. Outras modalidades, como voz, rastreamento de olho, gestos com o corpo, podem ser investigadas, identificando os problemas de interação que ocorrem ao acessar por meio da nova modalidade, e evoluindo os componentes da arquitetura para coletar e processar os dados da nova modalidade. As modalidades de saída diferentes das gráficas podem ser exploradas, como, por exemplo, sintetização de voz;

- Investigação de outros dispositivos para acessar o ambiente de EaD multimodal: neste trabalho utilizou-se o Tablet PC, mas outros dispositivos compostos por diferentes periféricos de interação (logo, outras modalidades) estão se tornando populares como lousas e paredes interativas. Embora empreguem interfaces gráficas, a dimensão da interação é de outra grandeza, o que pode impactar na interação;

- Impacto de novas modalidades nas ações pedagógicas: no Capítulo 8 descrevemos que existe uma relação entre as atividades de ensino e aprendizagem empregadas nos ambientes de EaD com as modalidades existentes no dispositivo empregado, e que a educação poderá se beneficiar do emprego de variadas modalidades. Essa pesquisa deve ser mais aprofundada por pesquisadores da educação considerando casos de uso reais para se identificar os impactos da nova modalidade no contexto educacional e a adoção da modalidade por parte dos seus usuários.

- Realização de testes de viabilidade da arquitetura proposta e a sua adoção em outros ambientes de EaD: a arquitetura proposta foi parcialmente implementada no ambiente de EaD Ae para apenas confirmar sua viabilidade. Pesquisas são necessárias para melhor especificar as interfaces entre os componentes, possibilitando ser implementada em outros ambientes de EaD;

- Realização de estudos de uso fora de laboratório: a maioria dos estudos realizados neste trabalho foram realizados em laboratório. Estudos realizados fora de laboratório podem apresentar diferentes respostas por conta das várias características que influenciam na escolha da(s) modalidade(s) por parte do usuário;

- Algoritmos para ajustar a interface multimodal: no Capítulo 2 apresentamos algumas pesquisas que focalizam o ajuste de interfaces multimodais conforme características do contexto e preferências do usuário. Nesse sentido, pesquisas podem ser realizadas para estudar as características do contexto (físicas ou sociais) que podem impactar na interação do usuário com o sistema de EaD multimodal, bem como as preferências do usuário ao interagir com um determinado dispositivo e em qual contexto educacional. Apresentamos que uma modalidade pode ser preferida pelo usuário em determinados contextos educacionais ou atividades de ensino aprendizagem, logo é necessário identificar e explorar as características que influenciam na escolha da modalidade;

- Algoritmos e representações dinâmicas para identificação de ações possíveis de serem realizadas: Os ambientes de EaD possuem dinamicidade nas ações disponíveis aos usuários (por exemplo, para apagar uma mensagem de Weblog é necessário estar em um curso e acessando a tela com as mensagens postadas em seu weblog, caso a ferramenta Weblog esteja disponível para o curso). Acreditamos que existem funcionalidades que comuns a várias ferramentas (por exemplo, colar, copiar, recortar, fazer logout, trocar de curso), funcionalidades relacionadas ao curso que podem ser acionadas a qualquer momento (por exemplo, escrever uma nova mensagem de correio), funcionalidades específicas para uma tela (por exemplo, selecionar itens de Portfólio para serem movidos ou apagados);

- Ferramentas de computação baseada em caneta na Web: utilizando o InkEditor é possível construir novas ferramentas para a Web que aproveitam os recursos da computação baseada em caneta, por exemplo, um editor de palheta de música, um anotador de conteúdo digital (por exemplo, sobre um vídeo ou slides de apresentação);

- Modelo mental: o usuário ao interagir com uma plataforma constrói um modelo metal dela, que será utilizado em interações futuras com a plataforma. O usuário pode empregar também modelos mentais de outras plataformas, como é o caso dos usuários novatos no uso dos celulares inteligentes com sistema operacional Android que são experientes nos computadores convencionais. Esses usuários, ao não quererem utilizar mais uma aplicação, procuram por um elemento que irá acionar o fechamento do programa. Entretanto, o Android possui um outro modelo: o sistema operacional gerencia os aplicativos em uso e, em caso de um longo período sem utilização ou necessitando de memória, o sistema operacional fecha o aplicativo sem qualquer ação do usuário. Surge, assim, a necessidade de pesquisar sobre os modelos mentais dos usuários ao interagir com uma interface multimodal;

- Interação multimodal e design universal: alguns pesquisadores estão discutindo a relação entre multimodalidade e design universal, conforme apresentado no Capítulo 2. Um ambiente de EaD multimodal possibilita fomentar esta discussão, pois, como acreditamos, um ambiente de EaD deve ter um alto nível de usabilidade e acessibilidade para atingir um grande número de usuários.

9.4 Considerações Finais

Este trabalho se iniciou com a necessidade de utilizar o periférico caneta disposto nos primeiros modelos de Tablet PC para acessar os ambientes de EaD. Ao longo de seu delineamento, afirmou-se a complexidade em desenvolver software na época atual: o surgimento de diversos periféricos de interação. O toque começou a ser usado em massa, principalmente devido à popularização dos celulares inteligentes, e empregado em novos modelos de Tablet PC. Assim, nos deparamos com a variedade de modalidades.

A variedade de dispositivos, periféricos de interação que eles possuem, a diversidade de usuário, o alto número de usuários que utilizam aplicações de determinados nichos, a boa usabilidade e acessibilidade, e, no caso dos ambientes de EaD, os vários contextos educacionais existentes, mostraram-se um problema complexo como espaço de pesquisa.

Propôs-se aplicar nos ambientes de EaD o conceito de multimodalidade e investigar os impactos e, ao longo da pesquisa, percebeu-se a importância de certas modalidades em determinados contextos educacionais, o que nos levou a perceber a facilidade de um conteúdo para um contexto educacional ser produzido dentro do ambiente de EaD com multimodalidade em comparação aos ambientes de EaD monomodais.

A multimodalidade foi então utilizada não somente para a interação entre o homem e interface de usuário, conforme propõe a área de Interface Multimodal mas, também para produzir conteúdo. Essa percepção possibilitou a criação de uma ferramenta dentro do ambiente de EaD, o InkBlog, cujo conceito de computação baseada em caneta pode ser aplicado a outras ferramentas, criando ferramentas que melhor aproveitem essa modalidade ou modalidades de interação.

Por fim, ressalta-se a importância da diversidade de dispositivos, cada um empregado conforme necessidade do usuário e a tarefa a ser realizada. Shneiderman (2006) já apontava essa necessidade:

> “Inspirados pelo gosto de Leonardo [da Vinci] por cadernos de bolso e pastas de esboços e por seus afrescos, os usuários e os projetistas podem imaginar a necessidade de uma linha abrangente de computadores, desde dispositivos pequenos, porém práticos e elegantes, até pomposos computadores de mesa e impressionantes modelos do tamanho de uma parede. Mantendo o espírito de Leonardo, cada novo modelo de equipamento de informática seria deliciosamente divertido e necessariamente útil.” (Shneiderman, 2006, p. 19)

Além de desenvolver uma ferramenta que aplica conceitos de Computação Baseada em Caneta para ambientes de EaD, acreditamos que demos um passo tecnológico importante ao possibilitar o uso de uma diversidade de dispositivos nos ambientes de EaD, empregando a Interação Multimodal nesse ambiente por meio de uma arquitetura específica. Esse é um dos objetivos da IHC segundo Hewett *et al.* (1992 *apud* Barbosa e da Silva, 2010, p. 10): arquitetura de sistemas computacionais e da interface com usuários. Outro objetivo da IHC tratado neste trabalho foi um processo de desenvolvimento preocupado com o uso, originando um espaço de pesquisa para explorar a multimodalidade nas atividades de ensino e aprendizagem nos demais objetivos inter-relacionados da IHC: o uso de sistemas interativos situado em contexto e características humanas.

Referências Bibliográficas

ACM SIGCHI. *Conference on Human Factors in Computing Systems*, 10., 1992, Monterey, Califórnia: EUA. *Proceedings of...* BAUERSFELD, P.; BENNETT, J.; LYNCH, G. (Eds.). Califórnia, EUA: ACM Press. Mai. 1992.

ALTY, J.; MCCARTNEY, C. Design of a Multi-Media Presentation System For A Process Control Environment, In: EUROGRAPHICS MULTIMEDIA WORKSHOP, Session 8: Systems, Stockholm, Suécia, 1991.

BACKON, J. Student minds and pen technologies: A wonderful pedagogical marriage. In: BERQUE, D.A.; PREY, J. C.; REED, R. H. (Eds.), *The Impact of Pen-based Technology on Education:* Vignettes, Evaluations, and Future Directions. West Lafayette, IN, EUA: Purdue University Press, 2006. p. 1-11.

BANGALORE, S.; JOHNSTON, M. Robust Understanding in Multimodal Interaction. *Computational Linguistic*, Cambridge, Massachusetts, EUA, v. 35, n. 3, p. 345-397. Set. 2009.

BARBOSA, S.D.J.B; DA SILVA, B.S. Interação Humano-Computador. Rio de Janeiro: Campus/Elsevier Editora Ltda, 2010. 408 p.

BERNSEN, N.O. Multimodality Theory. In: TZOVARAS, D. (Ed.), *Multimodal User Interfaces*: From signal to interaction. Berlim, Alemanha: Springer-Verlag Berlin Heidelberg, 2008. p. 5-28.

BERNSEN, N.O; DYKJÆR, L. Multimodal Usability. Berlim, Alemanha: Springer-Verlag Berlin Heidelberg, 2010. 431 p.

BERTI, S.; CORREANI, F.; MORI, G.; PATERNÒ, F.; SANTORO, C. TERESA: a transformation-based environment for designing and developing multi-device interfaces. In: CHI 2004 EXTENDED ABSTRACTS ON HUMAN FACTORS IN COMPUTING SYSTEMS, 22., 2004, Vienna, Austria. *Proceedings of...* Nova Iorque, EUA: ACM, Abr. 2004. p. 793-794.

BICKMORE, T. W.; SCHILIT, B. N. DIGESTOR: Device-independent Access to the World Wide Web. In: INTERNATIONAL WORLD WIDE WEB CONFERENCE, 6., 1997, California, USA. *Proceedings of...* Essex, Reino Unido: Elsevier Science Publishers Ltd., 1997. p. 1075-1082.

BODGDAN, C.; ERTL, D.; HÜTTENRAUCH, H.; GÖLLER, M.; GREEN, A.; SEVERINSON-EKLUNDH, K.; FALB, J.; KAINDL, H. Evaluation of robot body movements supporting communication. In: DAUTENHAHN, K.; SAUNDERS, J. (Eds.), *New Frontiers in Human–Robot Interaction*, Amsterdã, Holanda: John Benjamins Publishing Company, 2011. p. 185–210.

BOLT, R. A. Put-that-there: voice and gesture at the graphics interface. In: INTERNATIONAL CONFERENCE ON COMPUTER GRAPHICS AND INTERACTIVE TECHNIQUES, 7., 1980, Seattle, EUA. *Proceedings of...* Nova Iorque, EUA: ACM, 1980. p. 262-270.

BOUCHET, J.; NIGAY, L.; GANILLE, T. ICARE: Software Components for Rapidly Developing Multimodal Interfaces. In: INTERNATIONAL CONFERENCE ON MULTIMODAL INTERFACES, 6., 2004, Pensilvânia, EUA. *Proceedings of*... Nova Iorque, EUA: ACM, 2004. p. 251-258.

BRANDL, P.; FORLINES, C.; WIGDOR, D.; HALLER, M.; SHEN, C. Combining and measuring the benefits of bimanual pen and direct-touch interaction on horizontal interfaces. In: WORKING CONFERENCE ON ADVANCED VISUAL INTERFACES, 9., 2008, Napoli, Itália. *Proceedings of*... Nova Iorque: ACM, 2008. p. 154-161.

CACEFFO, R.E.; DA ROCHA, H.V.; DE AZEVEDO, R. J. Ferramenta de apoio para o aprendizado ativo usando dispositivos com caneta eletrônica. *Revista Brasileira de Informática na Educação*, v. 19, n. 2, Brasil, p. 25-41, 2011. DOI: 10.5753/RBIE.2011.19.02.25.

CHEE, Y.; FRANKE, K.; FROUMENTIN, M.; MADHVANATH, S.; MAGAÑA, J.; PAKOSZ, G.; RUSSELL, G.; SELVARAJ, M.; SENI, G.; TREMBLAY, C.; YAEGER, L. Ink Markup Language (InkML) W3C Recommendation. Setembro, 2011. Disponível em: http://www.w3.org/TR/InkML/ . Acesso em: 25 de março de 2013.

CHITTARO, L. Distinctive Aspects of Mobile Interaction and Their Implications for the Design of Multimodal Interfaces. *Journal on Multimodal User Interfaces*, Heidelberg, Alemanha, v. 3, n. 3, p. 157-165. Abr. 2010.

COHEN, P. R.; JOHNSTON, M.; MCGEE, D.; OVIATT, S.; PITTMAN, J.; SMITH, I.; CHEN, L.; CLOW, J. QuickSet: Multimodal Interaction for Distributed Applications. In: ACM INTERNATIONAL CONFERENCE ON MULTIMEDIA, 5., 1997, Washington, EUA. *Proceedings of*... Nova Iorque, EUA: ACM, 1997. p. 31-40.

CONINX, K.; LUYTEN, K.; VANDERVELPEN, C.; BERGH, J.V.D.; CREEMERS, B. Dygimes: dynamically generating interfaces for mobile computing devices and embedded systems. In: HUMAN-COMPUTER INTERACTION WITH MOBILE DEVICES AND SERVICES, 5., 2003, Udine, Itália. *Proceedings of*... Alemanha: Springer Berlin / Heidelberg, 2003. p. 256-270.

DAVID, L.; ENDLER, M.; BARBOSA, S.D.J.; FILHO, J.V. Middleware Support for Context-Aware Mobile Applications with Adaptive Multimodal User Interfaces. In: INTERNATIONAL CONFERENCE ON UBI-MEDIA COMPUTING, 4., 2011, São Paulo, Brasil. *Proceedings of*... Los Alamitos, CA, EUA: IEEE Computer Society, 2003. p. 106-111. DOI: 10.1109/U-MEDIA.2011.50.

DUARTE, C.; CARRIÇO, L. A Conceptual Framework for Developing Adaptive Multimodal Applications. In: INTERNATIONAL CONFERENCE ON INTELLIGENT USER INTERFACES, 11., 2006, Sidney, Austrália. *Proceedings of*... Nova Iorque, EUA: ACM, 2006. p. 132-139

DUMAS, B.; INGOLD, R.; LALANNE, D. Benchmarking Fusion Engines of Multimodal Interactive Systems. In: INTERNATIONAL CONFERENCE ON MULTIMODAL INTERFACES, 11., 2009, Massachusetts, EUA. *Proceedings of*... Nova Iorque, EUA: ACM, 2009. p. 169-176.

DUMAS, B.; LALANNE, D.; OVIATT, S. Multimodal Interfaces: A Survey of Principles, Models and Frameworks. In: LALANNE, D.; KOHLAS, J. (Org.). *Human-Machine Interaction*. Alemanha: Springer Berlin / Heidelberg, 2009. p. 3-26.

DUMAS, B.; SOLÓRZANO, M.; SIGNER, B. Design Guidelines for Adaptive Multimodal Mobile Input Solutions. In: INTERNATIONAL CONFERENCE ON HUMAN-COMPUTER INTERACTION WITH MOBILE DEVICES AND SERVICES (MobileHCI 2013), 15., 2013, Munique, Alemanha. *Proceedings of...* Nova Iorque, EUA: ACM, 2013. p. 285-294.

ERTL, D.; KAINDL, H. Semi-automatic Generation of Multimodal user Interfaces for Dialogue-based Interactive Systems. In: INTERNATIONAL CONFERENCE ON MULTIMODAL INTERACTION (ICMI 2012), 14., 2012, Santa Monica, Califórnia, EUA. *Proceedings of...* Nova Iorque, EUA: ACM, 2012. p. 441-444.

FLIPPO, F.; KREBS, A.; MARSIC, I. A Framework for Rapid Development of Multimodal Interfaces. In: INTERNATIONAL CONFERENCE ON MULTIMODAL INTERFACES, 5., 2003, Columbia, Canadá. *Proceedings of...* Nova Iorque, EUA: ACM, 2003. p. 109-116.

FOSTER, M. E. State of the Art Review: Multimodal Fission. Disponível em: http://groups.inf.ed.ac.uk/comic/documents/deliverables/Del6-1.pdf. 2002. Acesso em: março 2014.

GRUENSTEIN, A; MCGRAW, I.; BADR, I. The WAMI Toolkit for Developing, Deploying, and Evaluating Web-Accessible Multimodal Interfaces. In: INTERNATIONAL CONFERENCE ON MULTIMODAL INTERFACES (ICMI 2008), 10., 2008, Chania, Grécia. *Proceedings of...* Nova Iorque, EUA: ACM, 2008. p. 141-148, doi: 10.1145/1452392.1452420.

HAMILTON, W.; KERNE, A.; ROBBINS, T. High-Performance Pen + Touch Modality Interactions: A real-time strategy game eSports context. In: ANNUAL ACM SYMPOSIUM ON USER INTERFACE SOFTWARE AND TECHNOLOGY (UIST 2012), 25., 2012, Cambridge, MA, EUA. *Proceedings of...* Nova Iorque, EUA: ACM, 2012. p. 309-318, doi: 10.1145/1866029.1866036.

HAYES, B. Gartner´s Seven IT Grand Challenges. *Communications of the ACM,* Nova Iorque, EUA: ACM, v. 51, no. 7, p. 10. Jul. 2008.

HINCKLEY, K.; YATANI, K.; PAHUD, M.; CODDINGTON, N.; RODENHOUSE, J.; WILSON, A.; BENKO, H.; BUXTON, B. Pen + Touch = New Tools. In: ANNUAL ACM SYMPOSIUM ON USER INTERFACE SOFTWARE AND TECHNOLOGY (UIST 2010), 23., 2010, Nova Iorque, EUA. *Proceedings of...* Nova Iorque, EUA: ACM, 2010. p. 27-36, doi: 10.1145/1866029.1866036.

HIX, D.; HARTSON, H. R. *Developing User Interfaces: Ensuring Usability Through Product & Process*. EUA: Wiley & sons, 1993. 416 p.

HORCHANI, M.; CARON, B.; NIGAY, L.; PANAGET, F. Natural Multimodal Dialogue Systems: A Configurable Dialogue and Presentation Strategies Component. In:

INTERNATIONAL CONFERENCE ON MULTIMODAL INTERFACES, 9., 2007, Aichi, Japão. *Proceedings of...* Nova Iorque, EUA: ACM, 2007. p. 291-298.

KAIKKONEN, A. Full or tailored mobile web – where and how do people browse on their mobile?. In.: INTERNATIONAL CONFERENCE ON MOBILE TECHNOLOGY, APPLICATIONS AND SYSTEMS, 6., 2008, I-Lan, Taiwan. *Proceedings of...* Nova Iorque, EUA: ACM, 2008, p. 111. DOI: 10.1145/1506270.1506307.

KARPOV, A.; CARBINI, S.; RONZHIN, A.; VIALLET, J.E. Two SIMILAR Different Speech and Gestures Multimodal Interfaces. In: TZOVARAS, D. (Ed.), *Multimodal User Interfaces*: From signal to interaction. Berlim, Alemanha: Springer-Verlag Berlin Heidelberg, 2008. p. 155-184.

KNUDSEN, L.E.; HOLONE, H. A Multimodal Approach to Accessible Web Content on Smartphones. In.: INTERNATIONAL CONFERENCE ON COMPUTERS HELPING PEOPLE WITH SPECIAL NEEDS, volume II, 13., 2012, Linz, Austria. *Proceedings of...* Berlim, Alemanha: Springer-Verlag Berlin, Heidelberg, 2012, p. 1-8. DOI: 10.1007/978-3-642-31534-3_1.

KOONS, D. B.; SPARREL, C. J.; THORISSON, K. R. Integrating Simultaneous Input from Speech, Gaze, and Hand Gesture. *Intelligent Multimedia Interfaces*, Califórnia, EUA: American Association for Artificial Intelligence, 1993. p. 257-276.

KRAHNSTOEVER, N.; KETTEBEKOV, S.; YEASIN, M.; SHARMA, R. A Real-Time Framework for Natural Multimodal Interaction with Large Screen Displays. In: INTERNATIONAL CONFERENCE ON MULTIMODAL INTERFACES, 4., 2002, Pensilvânia, EUA. *Proceedings of...* Washington, EUA: IEEE Computer Society, 2002. p. 349-354.

KUGLER, L. Goodbye, Computer Mouse. *Communications of the ACM*, Nova Iorque, EUA: ACM, v. 51, n. 9, p. 56. Set. 2008.

KURTENBACH, G. Pen-Based Computing. *XRDS*, Nova Iorque, EUA: ACM, v. 16, n. 4, p. 14-20. verão 2010. DOI: 10.1145/1764848.1764854.

LALANNE, D.; NIGAY, L.; PALANQUE, P.; ROBINSON, P.; VANDERDONCKT, J.; LADRY, J. Fusion Engine for Multimodal Input: A Survey. In: INTERNATIONAL CONFERENCE ON MULTIMODAL INTERFACES, 11., 2009, Massachusetts, EUA. *Proceedings of...* Nova Iorque, EUA: ACM, 2009. p. 153-160.

LARSON, J. A. Should You Build a Multimodal Interface for Your Web Site?. Disponível em: http://www.informit.com/articles/article.aspx?p=29024. 2002. Acessado em: janeiro 2013.

LIMA, J.R.; CAPITÃO, Z. *e-Learning e e-Conteúdos:* aplicações das teorias tradicionais e modernas de ensino e aprendizagem à organização e estruturação de e-cursos. Famalicão, Portugal: Centro Atlântico, 2003. 287 p.

MACKAY, B.; WATTERS, C.; DUFFY, J. Web page transformation when switching devices. In: MOBILE HUMAN-COMPUTER INTERACTION, 6., 2004, Glasgow, Reino Unido. *Proceedings of...* Alemanha: Springer Berlin / Heidelberg, 2004. p. 228-239.

MAYES, T. The ‘M’ Word: Multimedia interfaces and their role in interactive learning systems. Edwards, A. D. N.; Holland, S. (Eds.), *Multimedia Interface Design in Education*, p. 1-22. Berlim, Alemanha: Springer-Verlag. 1992.

MAURER, M.E.; HAUSEN, D.; DE LUCA, A.; HUSSMANN, H. Mobile or Desktop Website?: Website usage on multitouch devices. In: NORDIC CONFERENCE ON HUMAN-COMPUTER INTERACTION, 6., 2010, Reykjavik, Islândia. *Proceedings of...* Nova Iorque, EUA: ACM, 2010. p. 739-742. DOI: 10.1145/1868914.1869018.

MCCOBB, G. The W3C Multimodal Architecture, Part 2: the XML specification stack. Disponível em: http://www.ibm.com/developerworks/web/library/wa-multimodarch2/index.html. 2007. Acessado em: janeiro 2014.

MENDONÇA, H.; LAWSON, J. L.; VYBORNOVA, O.; MACQ, B.; BANDERDONCKT, J. A Fusion Framework for Multimodal Interactive Applications. In: INTERNATIONAL CONFERENCE ON MULTIMODAL INTERFACES, 11., Nov. 2009, Cambridge, Massachusetts, EUA. *Proceedings of...* Nova Iorque, EUA: ACM, 2009. p. 161-168.

MORENO, R.; MAYER, R. Interactive Multimodal Learning Environments – Special Issue on Interactive Learning Environments: contemporary issues and trends. *Journal of Educational Psychology Review*, v. 19, n. 3, p. 309-326. Set. 2007.

NEAL, J. G.; SHAPIRO, S. C. Intelligent Multimedia Interface Technology. In: J. Sullivan; S. Tyler (Eds.) Intelligent User Interfaces. Nova Iorque, EUA: ACM Press, 1991. p. 11-43. 416 p.

NIELSEN, J. *Usability Engineering*. EUA: Morgan Kaufmann, 1993. 362 p.

NIGAY, L.; COUTAZ, J. A Generic Platform for Addressing the Multimodal Challenge. In: SIGCHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, 13., 1995, Colorado, EUA. *Proceedings of...* Nova Iorque, EUA: ACM Press / Addison-Wesley Publishing Co., 1995. p. 98-105.

NORMAN, D. A.; DRAPER, S. W. *User Centered System Design: New Perspectives on Human-Computer Interaction*. Nova Jersey, EUA: Lawrence Erlbaum Associate Publishers, 1986.

NORMAN, D. A.; NIELSEN, J. Gestural Interfaces: A step backward in usability. *Interactions*, Nova Iorque, EUA: ACM, v. 17, n. 5, p. 46-49, setembro 2010.

OLIVEIRA, R. Design Multi-Dispositivo em Contextos de Uso Alternado e Migração de Tarefas. 2008. Tese (Doutorado em Ciência da Computação) – Instituto de Computação, Universidade Estadual de Campinas, Campinas, Brasil. 2008.

OLIVEIRA, R.; ROCHA, H. V. Consistency Priorities for Multi-device Design. In: HUMAN-COMPUTER INTERACTION - INTERACT, 11., 2007, Rio de Janeiro, Brasil. *Proceedings of...* Alemanha: Springer Berlin, 2007. p. 617-623.

OVIATT, S. L. Ten Myths of Multimodal Interaction. *Communications of the ACM,* Nova Iorque, EUA: ACM, v. 42, no. 11, p. 74-81. Nov. 1999.

PATERNÒ, F. *Model-based Design and Evaluation of Interactive Application*. Spring Verlag. 1999.

PETEJAN, E. D. Automatic Lipreading to Enhance Speech Recognition. 1984. Tese (Doutorado) – University of Illinois at Urbana-Champaign, Champaign, EUA. 1984.

POEL, M.; NIJBOER, F.; BROEK, E.L. VAN DEN; FAIRCLOUGH, S.; NIJHOLT, A. Brain Computer Interfaces as Intelligent Sensors for Enhancing Human-Computer Interaction. In: INTERNATIONAL CONFERENCE ON MULTIMODAL INTERACTION (ICMI 2012), 14., 2012, Santa Monica, CA, EUA. *Proceedings of...* Nova Iorque, EUA: ACM, 2012. p. 379-382.

PORTILLO, P. M.; GARCÍA, G. P.; CARREDANO, G. A. Multimodal Fusion: A New Hybrid Strategy for Dialogue Systems. In: INTERNATIONAL CONFERENCE ON MULTIMODAL INTERFACES, 8., 2006, Banff, Canadá. *Proceedings of...* Nova Iorque, EUA: ACM, 2006. p. 357-363.

PREECE, J.; ROGERS, Y.; SHARP, H.; BENYON, D.; HOLLAND, S.; CAREY, T. *Human-Computer Interaction: Methods for User-Centred Design*. Addison-Wesley. 1994.

PYLA, P.; TUNGARE, M.; PÉREZ-QUIÑONES, M. Multiple user Interfaces: Why consistency is not everything, and seamless task migration is key. In: CHI 06 WORKSHOP ON THE MANY FACES OF CONSISTENCY IN CROSS-PLATFORM DESIGN, 24., 2006, Quebec, Canadá. *Proceedings of...* CEUR-WS, 2006. p. 55-60.

REEVES, L. M.; LAI, J.; LARSON, J. A.; OVIATT, S.; BALAJI, T. S.; BUISINE, S. P.; COLLINGS, P.; COHEN, P.; KRAAL, B.; MARTIN, J. C.; MCTEAR, M.; RAMAN, T.; STANNEY, K. M.; SU, H.; WANG, Q. Y. Guidelines for Multimodal User Interface Design. *Communications of the ACM,* Nova Iorque, EUA: ACM, v. 47, no. 1, p. 57-59. Jan. 2004.

ROCHA, H. V.; BARANAUSKAS, M. C. C. *Design e Avaliação de Interfaces Humano-Computador*. São Paulo, Brasil: NIED/UNICAMP. 2003. 244 p. Disponível em: http://pan.nied.unicamp.br/publicacoes/livros.php . Acessado em: janeiro 2014.

SCHMIEDL, G.; SEIDL, M.; TEMPER, K. Mobile Phone Web Browsing –A Study on Usage and Usability of the Mobile Web. In: INTERNATIONAL CONFERENCE ON HUMAN-COMPUTER INTERACTION WITH MOBILE DEVICES AND SERVICES, 11., 2009, Bonn, Alemanha. *Proceedings of...* Nova Iorque, EUA: ACM, 2009. 2 p. DOI: 10.1145/1613858.1613942.

SHARP, H.; ROGERS, Y.; PREECE, J. *Interaction Design: beyond human-computer interaction*. 2. edição. Chichester, EUA: Wiley & sons. 2007. 800 p.

SHNEIDERMAN, B. *O Laptop de Leonardo*: Shneiderman, 2006. 228 p. Tradução de Vera Whaterly, Rio de Janeiro: Nova Fronteira. 2006.

SHNEIDERMAN, B.; PLAISANT, C. *Designing the User Interface: Strategies for Effective Human-Computer Interaction.*. 4. edição, 672 p. Reading, MA, EUA: Addison-Wesley. 2004.

SIM, K.C; ZHAO, S.; YU, K.; LIAO, H. ICMI´12 Grand Challenge: haptic voice recognition. In: INTERNATIONAL CONFERENCE ON MULTIMODAL INTERACTION (ICMI 2012), 14., 2012, Santa Monica, CA, EUA. *Proceedings of...* Nova Iorque, EUA: ACM, 2012. p. 363-370.

STANCIULESCU, A.; LIMBOURG, Q.; VANDERDONCKT, J.; MICHOTTE, B.; MONTERO, F. A Transformational Approach for Multimodal Web User Interfaces based on UsiXML. In: INTERNATIONAL CONFERENCE ON MULTIMODAL INTERFACES, 7., 2005, Trento, Itália. *Proceedings of...* Nova Iorque, EUA: ACM, 2005. p. 259-266.

SUNG, M.; GIPS, J.; EAGLE, N.; MADAN, A.; CANEEL, R.; DEVAUL, R.; BONSEN, J.; PENTLAND, A. Mobile-IT Education (MIT. EDU): m-learning applications for classroom settings. *Journal of Computer Assisted Learning*, v. 21, maio 2005, p. 229–237.

TALARICO NETO A.; BITTAR, T. J.; FORTES, R. P. M.; FELIZARDO, K. Developing and Evaluating Web Multimodal Interfaces – A Case Study with Usability Principles. In: SYMPOSIUM ON APPLIED COMPUTING, 24., 2009, Honolulu, Hawaii. *Proceedings of...* Nova Iorque, EUA: ACM, 2009. p. 116-120.

TEICHE, A.; RAI, A. K.; YANC, C.; MOORE, C.; SOLMS, D.; ÇETIN, G.; RIGGIO, J.; RAMSEYER, N.; D´INTINO, P.; MULLER, L.; KHOSHABEH, R.; BEDI, R.; BINTAHIR, M.T.; HANSEN, T.; ROTH, T.; SANDLER, S. *Multitouch Technologies*. NUI Group. 2009. 90 p. Disponível em: http://nuigroup.com/log/nuigroup_book_1/ .

TURUNEN, M.; HAKULINEN, J.; SALONEN, E.; KAINULAINEN, A.; HELIN, L. Spoken and multimodal bus timetable systems: design, development and evaluation. In: INTERNATIONAL CONFERENCE SPEECH AND COMPUTER (SPECOM 2005), 10., 2005, Patras, Grécia. *Proceedings of...* Patras, Grécia: University of Patras, 2005. p. 389-392.

W3C. Web Content Accessibility Guidelines (WCAG) Overview. 2005. Disponível em: http://www.w3.org/TR/WCAG/. Acessado em: julho 2013.

W3C. Multimodal Architecture and Interfaces. 2013. Disponível em: http://www.w3.org/TR/2012/REC-mmi-arch-20121025/ . Acessado em: janeiro 2013.

WILLIAMSON, J.R.; CROSSAN, A.; BREWSTER, S. Multimodal Mobile Interactions: usability studies in real world settings. In: INTERNATIONAL CONFERENCE ON MULTIMODAL INTERFACES (ICMI 2011), 13., 2011, Alicante, Espanha. *Proceedings of...* Nova Iorque, EUA: ACM, 2011. p. 361-368. DOI: 10.1145/2070481.2070551.

ZAGUIA, A.; HINA, M. D.; TADJ, C.; RAMDANE-CHERIF, A. Using Multimodal Fusion in Accessing Web Services. *Journal of Emerging Trends in Computing and Information Sciences*, v. 1, no. 2, outubro 2010, p. 121-137.

ZHANG, HONGJIANG. Adaptive Content Delivery: A New Research Area in Media Computing. In.: KEYNOTE SPEECH ON PROCEEDINGS OF THE 2000 INTERNATIONAL WORKSHOP ON MULTIMEDIA DATA STORAGE, RETRIEVAL, INTEGRATION AND APPLICATIONS, 2000, Hong Kong. *Proceedings of...* 2000. Disponível em: http://research.microsoft.com/research/pubs/view.aspx?type=Publication&id=905 .