

Avaliação Expedita de Livros Falados Digitais

Joana Pereira

Dep. Informática,

Faculdade de Ciências da Univ. de Lisboa
Campo Grande, Edifício C6, 1749-016 Lisboa
joanaper@netcabo.pt

Luís Carriço

LaSIGE & Dep. Informática,

Faculdade de Ciências da Univ. de Lisboa
Campo Grande, Edifício C6, 1749-016 Lisboa
lmc@di.fc.ul.pt

Sumário

Este artigo versa a aplicação de arquitecturas cognitivas para a avaliação expedita de Livros Falados Digitais. O trabalho, ainda em curso, passa pela identificação das características dessas arquitecturas e dos seus problemas na especificação das variantes de utilizadores e tarefas. Nessa base, pretende-se estendê-las no intuito de permitir uma fácil modelação de utilizadores com diferentes tipos de deficiências e de tarefas normalmente realizadas na utilização desses livros. Pretende-se ainda, através de uma validação comparativa com testes de usabilidade com utilizadores reais, compreender os equilíbrios das capacidades de percepção desse tipo de utilizadores e definir estereótipos de modelos que possam ser usados em avaliações simuladas posteriores.

Palavras-chave

Usabilidade, Cognição, Acessibilidade, Arquitecturas Cognitivas, Livros Falados Digitais

1. INTRODUÇÃO

A avaliação de interfaces é, hoje em dia, considerada um requisito fundamental no desenvolvimento de aplicações computacionais [Nielsen03, Brinck02, Preece02]. Essa avaliação, contribuindo definitivamente para melhorar o desempenho dos utilizadores na realização de tarefas, através da detecção de falhas de usabilidade e posterior correcção do desenho das interfaces, é determinante na aceitação e conseqüente sucesso dessas aplicações. No entanto, por diversas razões subjacentes, normalmente encapsuladas sob o epítáfio do argumento “custo”, a avaliação de usabilidade é relegada para segundo plano no desenvolvimento de aplicações. A sua integração nos projectos é por isso fortemente circunscrita - embora não seja difícil provar que os ganhos a prazo são superiores aos dispêndios [Nielsen03]. Assim sendo, uma escolha criteriosa das metodologias, métodos e técnicas dos processos de avaliação, que se adequem aos projectos e custos disponíveis, torna-se extremamente pertinente.

Neste sentido, alguns trabalhos têm tentado apontar as várias dimensões dos processos de avaliação de usabilidade [Coutaz94, Ivory01]. Para além da identificação de fases, competências e técnicas a usar, emergem implicitamente dessas taxinómias três grupos de actores fundamentais: os utilizadores, as tecnologias de suporte e os investigadores. Cada um deles, em particular o primeiro, tem um peso determinante no custo e tempo do processo de avaliação, mas também na validade dos resultados obtidos. É nesse contexto, pois, que surgem aproximações que tentam capitalizar nos aspectos teóricos, propondo técnicas e ferramentas que permitam evitar o re-

curso generalizado a utilizadores. São particularmente interessantes as que tentam automatizar (ainda que parcialmente) o processo de avaliação [Ivory01], não só pela ausência daqueles actores, mas também pela promessa de exaustão nos testes efectuados.

Surge então, o paradigma das arquitecturas de modelação cognitiva. Simulando o processo cognitivo dos utilizadores e também os seus comportamentos durante a interacção com o sistema, estas arquitecturas cognitivas pretendem constituir uma base de sustentação fiável à criação de indicadores e métricas relativos às características do desempenho humano. Neste contexto, distinguem-se três arquitecturas principais: ACT-R/PM [Byrne01], Soar [Lewis93] e EPIC [Kieras97a].

No caso particular do trabalho apresentado neste artigo, pretende-se adoptar esta perspectiva à avaliação de Livros Falados Digitais (LFD) [Duarte04]. Estes, por se orientarem essencialmente para um nicho particular de utilizadores, com deficiências visuais, e porque o acesso a esse nicho é dificultado por diversos factores, tornam-se um caso de estudo por excelência para a aplicação de processos expeditos de avaliação. Por outro lado, a falta de directivas e heurísticas claras e definitivas, sobre os aspectos de usabilidade de aplicações multimodais, como é o caso dos LFD, e para a comunidade alvo a que se destinam, enfatiza a avaliação como primordial ao processo de desenvolvimento.

O problema que se apresenta desde logo a estas arquitecturas é: como simplificar a especificação dos modelos de utilizador e tarefas, de modo a que a avaliação possa

pragmaticamente ser feita. No que respeita aos LFD, o caso complica-se pela necessidade de poder variar com facilidade as características perceptuais nos modelos de utilizador ou até mesmo como encontrar equilíbrios e modelar os desequilíbrios entre as modalidades da percepção. Estas arquitecturas não dispõem de um modo expedito de configuração dessas características. Para obviar esta limitação, pretende-se desenvolver uma forma de efectuar a correcta configuração da arquitectura EPIC, com base em diagnósticos simples das deficiências dos utilizadores. O trabalho é complementado por um estudo e ambiciona ainda, à definição de estereótipos representativos dos utilizadores em causa e das próprias aplicações, validados por testes de usabilidade com utilizadores reais que estão a ser levados a cabo no âmbito do projecto IPSOM [Duarte04].

Neste artigo descreve-se o trabalho realizado até ao momento, focando-se as características das arquitecturas cognitivas existentes, nomeadamente do EPIC, que potencialmente serão o óbice e o objecto de melhoramento, evolução e ajuste ao tipo de modelação que se pretende. Na secção seguinte apresentam-se genericamente essas arquitecturas. Depois apontam-se os requisitos que emergem do trabalho com os LFD e que têm impacto directo no processo de avaliação que se pretende automático. Finalmente, discutem-se as dificuldades e problemas que o EPIC, em particular, apresenta ao tipo de modelação que se pretende, traçando, para concluir, direcções de trabalho futuro.

2. AVALIAÇÃO EXPEDITA DE USABILIDADE

A avaliação de usabilidade é habitualmente concretizada através da realização de testes sobre a interface a avaliar [Ivory01] que consistem em promover interacção de utilizadores reais com ela. Esta interacção é baseada na execução de tarefas previamente definidas por um grupo de avaliadores, de forma a explorar um máximo de funcionalidades oferecidas pela interface. Este processo de avaliação, uma vez que envolve recursos humanos desempenhando funções quer de avaliador quer de utilizador, revela-se normalmente bastante moroso e dispendioso, e exhibe uma natureza subjectiva, já que tanto a interacção como os juízos avaliativos dependem de factores ambientais, culturais, etc.

As arquitecturas de modelação cognitiva propõem-se a substituir os utilizadores por representações do processo de interacção, usando modelos de utilizador e/ou de desenho de interface e/ou de tarefas. Nesse sentido, a redução dos custos e tempo de avaliação inerentes aos utilizadores humanos é uma consequência imediata. Por outro lado, estas arquitecturas permitirão prever de uma forma sistemática a interacção de utilizadores com diferentes características (desde que modeláveis), alargando significativamente o espectro de informação a considerar na avaliação dos sistemas. Outra vantagem, consiste na possibilidade de realizar avaliações em fases precoces do desenvolvimento de sistemas, permitindo detectar e corrigir problemas de usabilidade do desenho, minimizando os desvios a nível dos custos, prazos, etc. Finalmente, as

próprias arquitecturas ao simularem a interacção do utilizador, reportam os resultados sob a forma de um conjunto de medidas de desempenho [Ivory01].

Estas arquitecturas estabelecem uma plataforma de conjuntos computacionais de funções, métodos e estruturas de dados reutilizáveis, que representam com precisão algumas capacidades e limitações do raciocínio humano. O seu objectivo principal subsiste em caracterizar os aspectos invariantes do raciocínio humano e estabelecer uma base estável para a construção de modelos de cognição. Normalmente dispõem de componentes computacionais individualizados para caracterizar a percepção, a cognição, a memória e as acções motoras. A comunicação entre a unidade central cognitiva e os módulos perceptuais e motores pode assumir duas formas dependendo do sentido em que circula a informação: as acções geradas podem determinar um comando sobre um módulo perceptual e/ou motor e vice versa, um módulo pode alterar a memória declarativa com dados provenientes das informações que recolhe do exterior. O conhecimento sobre a forma de execução das tarefas a modelar, [Kieras97b], é representado através de regras de produção IF-THEN, que quando interpretadas geram comportamentos cognitivos. Internamente, estas arquitecturas, consideram que os modelos cognitivos são construídos para executarem tarefas e não instâncias (casos particulares) de tarefas. Deste modo, para qualquer combinação estímulo-contexto que define uma instância, o modelo cognitivo responde correctamente.

Entre as arquitecturas cognitivas mais utilizadas encontram-se: EPIC (Executive Process-Interactive Control) [Kieras97a], Soar [Lewis93], ACT-R/PM (Atomic Components of Thought with Rational Analysis and Perceptual/Motor) [Byrne01].

2.1 Soar

Relativamente à arquitectura Soar, distinguem-se dois componentes principais [Lewis93]: os processos cognitivos e os módulos perceptuais e motores. No que concerne aos processos cognitivos, salientam-se três estruturas: (1) a memória subdividindo-se na memória de reconhecimento, com carácter perene, e na memória de trabalho declarativa; (2) as primitivas de processamento que asseguram as alterações dos conteúdos na memória de trabalho declarativa e as operações de correspondência e particionamento da memória de reconhecimento; e por fim, (3) a estrutura de controlo que estabelece a combinação entre o processo de reconhecimento de correspondências e o ciclo de decisão. Os *Módulos Perceptuais e motores* da arquitectura Soar efectuem a actualização dos factos presentes na memória de trabalho e a sua ligação com a central cognitiva.

2.2 ACT-R/PM

A arquitectura ACT-R/PM, estabelece quatro módulos motores perceptuais, coordenados pela unidade cognitiva central. O *Sistema de Produção* [Byrne01] desta arquitectura compreende três tipos de memória: uma (1) memória declarativa que contém factos, uma (2) memória de

produção ou procedimentar, onde se encontram os mecanismos de suporte à cognição, isto é, as regras de produção, e ainda uma (3) pilha de objectivos a concretizar. Distinguem-se ainda quatro tipos de *Módulos Motores Perceptuais*: o módulo de visão, o módulo motor e os módulos auditivo e de discurso que não apresentam um nível de desenvolvimento tão avançado como os anteriores [Byrne01].

2.3 EPIC

A arquitectura EPIC [Kieras97a] [Kieras97b] é constituída por três tipos de processadores que suportam diferentes funcionalidades. O *Processador Cognitivo* é o responsável pela actividade principal de transformação da informação e para isso, socorre-se de um interpretador de regras de produção e de três estruturas de armazenamento que correspondem às memórias perceptuais, de longo prazo e de produção. As funções de recepção e codificação de informação e estímulos provenientes do ambiente são concretizadas pelos *Processadores Perceptuais*. Neste campo, identificam-se três processadores especializados em três áreas sensoriais: a visão, a audição e o tacto. Uma vez recebido e codificado um estímulo, é transmitida ao processador cognitivo, a informação a ele referente para posterior computação. Finalmente, os *Processadores Motores* executam as respostas e actividades motoras previamente determinadas pelo processador cognitivo e contemplam três modalidades: visual, vocal e manual.

3. AVALIAÇÃO DOS LFD

Os Livros Falados Digitais têm o intuito de fornecer acesso a livros a pessoas com deficiências visuais. No caso de sistemas destinados à interacção de pessoas com limitações visuais, não basta aplicar os padrões de acessibilidade geral. É necessário que a arquitectura suporte navegação adequada, e que permita ao utilizador memorizar facilmente uma estrutura do sistema. Para facilitar a produção de novos LFD, a NISO em colaboração com o consórcio DAISY desenvolveu uma norma [Ansi/Niso02] que especifica as características principais dos LFD bem como a sua arquitectura.

A aplicação de testes de usabilidade a utilizadores com dificuldades de visão requer um conjunto adaptações e características de rigor que os dificultam sobremaneira [Gerber02]. Por se tratar de uma amostra de utilizadores com limitações visuais, é necessário garantir a proporcionalidade dos diferentes tipos de limitações observados em toda a população, que já por si só é diminuta ou de acesso difícil. Este facto é ainda mais limitativo na medida em que não podem ser recrutados utilizadores que já tenham feito algum teste, pois a sua experiência anterior implica a não representatividade da amostra. De forma a minimizar a perturbação sentida durante a avaliação, os testes devem ser realizados em locais familiares aos próprios utilizadores, sendo uma técnica habitual o recurso a protocolos verbais (e.g. "pensar em voz alta"). Finalmente, as posturas defensivas, já de si vulgares em utilizadores comuns, e os tratamentos excessivamente cuidadosos, por parte de quem conduz os estudos, são, normalmente,

extremados, tornando a fase de adaptação mais longa e os resultados mais enviesados.

Por outro lado, os hábitos associados às comunidades em questão e o próprio balanceamento entre órgãos sensoriais decorrente da dificuldade de visão, tornam falsos alguns pressupostos tomados para utilizadores não deficientes. Exemplo disso são algumas observações feitas em testes com invisuais sobre LFD [Duarte04], em que a voz gravada para a locução do livro, por uma locutora profissional, em condições óptimas, foi considerada "estridente". Na verdade, para além das razões técnicas, um dos juízos que levou a fazer e usar essa gravação controlada, em detrimento das cassetes áudio, foi a dificuldade e a sensação desagradável ("estridente") que os indivíduos sem deficiência visual tiveram ao ouvir estas últimas.

4. ARQUITECTURAS COGNITIVAS E OS LFD

Apesar das vantagens enumeradas para o processo de simulação baseado em arquitecturas cognitivas, algumas desvantagens devem ser assinaladas. A primeira, inerentemente ligada à própria aproximação, diz respeito à necessidade da discriminação das actividades envolvidas nas tarefas, a nível muito baixo. O desenvolvimento desses modelos de tarefas revela-se portanto um processo moroso que requer perícia por parte do modelador. Todavia, a adaptação a um tipo particular de aplicações, como os LFD, em que as formas de interacção são usualmente demarcadas, permitirá sem dúvida simplificar o processo, mas também reutilizar em grande parte a especificação desses modelos.

Infelizmente, no entanto, uma segunda desvantagem, relativa ao estado actual dessas arquitecturas, manifesta-se inconvenientemente: a modelação requer normalmente conhecimentos exactos acerca da estrutura e modo de processamento interno das arquitecturas. A modelação é por vezes feita ao nível do código, com repercussões nem sempre determinísticas. O nível de conhecimento exigido ao modelador apresenta-se pois em campos diametralmente opostos. É nesse sentido desejável nestas arquitecturas a integração de ferramentas de especificação de mais alto nível que permitam ao modelador concentrar-se a um nível de abstracção adequado, directamente relacionado com o processo de interacção.

Finalmente, e no caso dos LFD ou em geral de aplicações orientadas para pessoas com deficiências cognitivas (em sentido lato), estas arquitecturas exibem dificuldades e mesmo limitações, no que respeita à modelação de utilizadores, em particular daqueles com características particulares como os da comunidade alvo dos LFD. Para além disso, e mais uma vez, os mecanismos de configuração relativos aos módulos perceptuais, encontram-se embebidos na declaração e implementação dos métodos dos processadores sensoriais, dotando todo o processo de adaptação da arquitectura de complexidade acrescida.

A EPIC, por exemplo, permite especificar os tempos de reconhecimento e transmissão de estímulos dos três processadores sensoriais por valores numéricos (milissegundos). Estes podem, por exemplo, ser alterados de acordo

com as capacidades dos utilizadores em causa. No entanto, e especialmente no módulo visual, os parâmetros relativos à qualidade da informação percebida podem tomar apenas dois valores: *exacta e indisponível*. Na modelação de utilizadores com deficiências visuais que não representem a perda total da visão, o facto de a qualidade da informação visual assumir unicamente dois valores resulta na impossibilidade da concretização do processo de configuração, e consequentemente, na correcta simulação dos processos cognitivos.

Inerente ao modo de adaptação da arquitectura ao utilizador a simular, encontra-se o mecanismo de compensação das limitações. No caso de deficientes visuais, isto reflecte-se principalmente ao nível da velocidade de reconhecimento e processamento de estímulos auditivos e do prazo de decaimento dos dados presentes em memória. Por exemplo, de acordo com as observações referidas atrás, deverá poder-se configurar os intervalos de timbre, volume, etc, considerados adequados ou aprazíveis e mesmo velocidades de apreensão. A própria norma para a construção de LFD [Ansi/Niso02] estabelece como requisito a leitura em velocidade acelerada, também como forma de acomodar os níveis de processamento auditivo, resultantes da especialização deste sentido. Outros estudos [Gerber02] mostram que muitas vezes a perda da faculdade visual é compensada pela retenção em memória dos factos durante mais tempo.

5. TRABALHO EM CURSO E FUTURO

Na sequência das limitações da arquitectura EPIC acima apresentadas, está a elaborar-se um estudo acerca das alterações dos processos cognitivos de pessoas com capacidades visuais reduzidas. Com base na investigação e no apuramento de resultados, proceder-se-á à introdução dos modificadores perceptuais e cognitivos na arquitectura, promovendo-se também uma forma de especificação de mais alto nível. Finalmente submeter-se-á um sistema de LFD a uma avaliação de usabilidade com interacção simulada comparando-a com os resultados obtidos com utilizadores.

6. REFERÊNCIAS

- [Ansi/Niso02] ANSI/NISO, Specifications for the Digital Talking Book, (2002) <http://www.niso.org/standards/resources/z39-86-2002.html>
- [Busse99] Busse, D. K. and Johnson, C. W. Investigating Pilots' Cognitive Processes in Aviation Accidents, Proceedings of the International Conference on Human Interfaces in Control Rooms, Cockpits, and Command Centres, (June 1999).
- [Brink02] Tom Brink. Darren Gergle, Scott Wood. Usability for the Web. *Morgan Kaufmann*, 20012.
- [Byrne01] M. D. Byrne. ACT-R/PM and menu selection: applying a cognitive architecture to HCI. *International Journal of Human Computer Studies*, 55, 1 (2001), 41-84.
- [Coutaz94] J. Coutaz and S. Balbo. Evaluation des interfaces utilisateur: taxonomie et recommandations. In Proceedings of the IHM'94, Human Computer Interaction Conference, France, (December 1994).
- [Duarte04] Duarte, C. and Carriço, L. Usability Evaluation of Digital Talking Books. Artigo submetido para a 1ª Conferência Nacional em Interação Pessoa-Máquina, (2004).
- [Gerber02] Elaine Gerber, Conducting Usability Testing With Computer Users Who Are Blind or Visually Impaired, Proceedings of the 17th Annual International Conference of California State University Northridge (CSUN) "Technology and Persons with Disabilities", (March 2002).
- [Ivory01] Ivory, M. Y., and Hearst, M. A. The state of the art in automating usability evaluation of user interfaces. *ACM Computing Surveys* 33, 4 (December 2001), 470-516.
- [Kieras97a] Kieras, D. and Meyer, D. An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human-Computer Interaction*, 12 (1997), 391-438.
- [Kieras97b] Kieras, D., Wood, S. D., and Meyer, D. Predictive engineering models based on the EPIC architecture for a multimodal highperformance human-computer interaction task. *ACM Transactions on Computer-Human Interaction*, 4 (1997), 230-275.
- [Kieras00] Kieras, D. and Meyer, D. The role of cognitive task analysis in the application of predictive models of human performance. Schraagen, Chipman and Shalin (Eds.) *Cognitive Task Analysis*, Lawrence Erlbaum Associates (2000), 237-260.
- [Lewis93] R. L. Lewis, An architecturally-based theory of human sentence comprehension. Tech. Rep. CS-93-226, Computer Science Department, Carnegie Mellon University, Pittsburgh, PA, (December 1993).
- [Nielsen03] Jakob Nielsen, USABILITY 101, (2003) <http://www.useit.com/alertbox/20030825.html>
- [Preece02] J. Preece, Y. Rogers and H. Sharp. Interaction Design: beyond human-computer interaction. John Wiley & Sons, Inc. 2002.