

LA VALUTAZIONE AUTOMATICA DELL'USABILITA' DI SITI WEB

L.Paganelli, F.Paternò, M.Pizzolati

CNUCE-C.N.R.

Via G.Moruzzi 1, Pisa

L'usabilità dei siti web si sta affermando come un elemento determinante per il loro successo. Tuttavia l'usabilità è un concetto complesso a molte dimensioni che interagiscono tra di loro e spesso gli sviluppatori di siti web ne hanno una conoscenza superficiale o hanno poco tempo da dedicare alla sua valutazione. Ecco che strumenti automatici che aiutino il progettista assumono una forte importanza. In questo articolo introduciamo un ambiente progettato e sviluppato nel gruppo di interazione uomo-macchina del CNUCE che consente di ottenere una valutazione di come gli utenti interagiscono con un sito, effettuata con il supporto di modelli di compiti, seguita da un'analisi automatica delle pagine che sono trovate problematiche secondo criteri di usabilità.

MOTIVAZIONI

L'usabilità è un concetto a molte dimensioni, dove il peso di ciascun aspetto dipende anche dal dominio applicativo considerato. Se consideriamo la definizione data dallo standard ISO 9241 (la misura in cui un prodotto può essere usato da utenti per raggiungere obiettivi specifici con efficacia, efficienza e soddisfazione in un certo contesto di uso) ci rendiamo facilmente conto che mentre il concetto è chiaramente espresso, allo stesso tempo lascia completamente aperta la scelta degli aspetti concreti da considerare per valutare in termini più precisi l'usabilità.

Molti tipi di metodi sono stati proposti per la valutazione dell'usabilità: metodi basati su ispezioni, dove degli esperti analizzano un prototipo od un sistema usando specifici criteri (esempi sono il cognitive walkthrough o la heuristic evaluation di Nielsen); altri metodi sono basati sull'uso di modelli per, ad esempio, predire in quanto tempo l'utente sarà in grado di raggiungere un obiettivo (come succede con valutazioni di tipo GOMS/KLM), in altri casi ancora ci si basa sulla esperienza diretta degli utenti, che può venire analizzata con il supporto di varie apparecchiature.

Il successo del Web è basato sulla sua capacità di comunicare facilmente informazione ad utenti che sono sparsi ovunque sul globo. Vi sono vari

strumenti che consentono di realizzare facilmente siti Web, tuttavia comunicare tale informazione in modo che sia facile da accedere, interpretare ed usare per utenti potenziali richiede ancora un certo sforzo che molto spesso non viene intrapreso come si può constatare navigando, anche per poco tempo, sul Web.

Da queste problematiche è nato l'interesse per strumenti automatici capaci di supportare una valutazione di usabilità di siti Web. Lo scopo di questi strumenti è di fornire un insieme di informazioni che rendono più facile la valutazione dell'usabilità richiedendo uno sforzo molto minore al progettista/valutatore. In questo modo si riduce il costo della valutazione dell'usabilità, aumenta la consistenza nell'identificazione delle parti problematiche, si riduce il bisogno di esperti e si aumentano le tipologie delle caratteristiche del sito valutate.

ALCUNI ESEMPI DI STRUMENTI AUTOMATICI

Per esemplificare le possibilità degli strumenti automatici consideriamo tre strumenti. Il nostro commento sui tre strumenti automatici di valutazione di siti web si divide in due dimensioni: la prima riguarda le problematiche sottostanti questi strumenti automatici - con particolare riferimento al concetto di usabilità e valutazione di usabilità - mentre la seconda si riferisce alle analisi realizzate. Questi tre

strumenti sono prodotti da istituzioni eterogenee: Doctor HTML è prodotto da una società (Imagiware Inc.); Web Sat è sviluppato dal NIST (National Institute of Standards and Technology), un istituto pubblico che fa capo al Dipartimento del Commercio statunitense; Bobby, infine, è stato creato da CAST (Centre for Applied Special Technology) un'organizzazione no-profit.

Per quanto riguarda la prima dimensione cui abbiamo accennato, rileviamo varie differenze.

Doctor HTML fa riferimento a problemi che possono avere effetto sul layout e sui tempi di caricamento e si dà come obiettivo quello di fornire un rapporto chiaro e facile da usare per perfezionare una pagina web.

Web Sat è l'unico dei tre strumenti che fa esplicito riferimento alla *valutazione di usabilità* di un sito Web; i controlli effettuati sono sviluppati a partire da diverse linee guida riguardanti il web, linee guida riguardanti interfacce utenti e studi sull'usabilità dei siti web. L'obiettivo che si pone è quello di essere un meccanismo di avviso di potenziali problemi di usabilità.

Bobby fa invece riferimento all'*accessibilità*. Il suo obiettivo è essere un aiuto per identificare e riparare significative barriere all'accesso per individui con invalidità. Questo strumento è stato implementato basandosi sulle linee guida per l'accessibilità del contenuto web del World Wide Web Consortium.

Per quanto riguarda l'analisi che gli strumenti sono in grado di svolgere automaticamente, Doctor HTML fornisce, sotto forma di tabelle, il risultato dei seguenti controlli: ortografia del testo, sintassi HTML delle immagini, comandi INPUT dei form, diverse proprietà delle immagini, attributi HTML delle tabelle, analisi sintattica dell'intero codice HTML, link, tempo di caricamento della pagina, dimensione del testo e dei comandi HTML, segnalando l'errore con il colore rosso e corredandolo del numero di linea di codice HTML corrispondente.

L'analisi di WebSat si struttura su sei categorie di linee guida individuate: accessibilità, uso dei form, rendimento, mantenibilità, navigazione, leggibilità.

Per ogni categoria viene data una spiegazione dettagliata, la descrizione delle verifiche effettuate sul codice HTML e i riferimenti bibliografici dai quali sono state ricavate le linee guida. Lo strumento segnala se e quante volte una linea guida non è rispettata.

L'analisi di Bobby è strutturata sulle raccomandazioni W3C e quindi, come queste, basa la sua analisi su tre livelli di priorità che potremmo definire: da soddisfare assolutamente, preferibilmente da soddisfare, a cui fare attenzione. Per ogni priorità, lo strumento riporta gli errori individuati e i controlli da eseguire manualmente. Riporta, infine, gli errori di compatibilità del browser e i tempi di download. Da l'approvazione solo alle pagine o ai siti web che non riportano nessun errore individuato del primo livello di priorità.

LA NOSTRA PROPOSTA

Come abbiamo visto, gli strumenti automatici disponibili, anche quelli più avanzati, non considerano molti aspetti che hanno un impatto sull'usabilità. Allo scopo di superare tale limitazione, nel gruppo di interazione uomo-macchina del CNUCE-C.N.R. (<http://giove.cnuce.cnr.it>) è in corso la progettazione ed implementazione di due strumenti automatici per la valutazione di siti Web. I due strumenti sono complementari uno all'altro: uno si basa sull'analisi di pagine HTML e si propone di verificare se una serie di aspetti (come il layout, la scelta delle tecniche di interazione, l'uso della multimedialità) soddisfa criteri di usabilità, mentre l'altro strumento (WebRemUSINE) consente l'analisi automatica dei log contenenti le interazioni degli utenti.

Per quanto riguarda l'analizzatore di pagine Web abbiamo stilato una lista di *guidelines* suddivisa nei cinque aspetti di layout, tecniche di interazione, navigazione, multimedia e documentazione; individuato, poi, la sintassi di codice HTML corrispondente a ciascuna linea guida. La verifica del rispetto delle *guidelines* viene effettuata mediante analisi delle pagine del sito Web attraverso i tag del codice HTML. Questo strumento vuole essere flessibile, permettendo all'utente/progettista/valutatore che vaglia l'usabilità di una pagina o di un sito Web di operare alcune scelte sui controlli da effettuare.

D'altra parte, WebRemUSINE (Paternò e Paganelli, 1991) è uno strumento per la valutazione dell'usabilità remota (quando la valutazione si effettua con il valutatore e l'utente distanti in tempo e/o spazio). Questo strumento si è basato su una nostra precedente esperienza: RemUSINE (Lecerof e Paternò, 1998) (Paternò e Ballardini, 2001), uno strumento basato sull'uso di modelli di task e informazioni automaticamente generate indicanti le interazioni effettuate dall'utente (file di log) allo scopo di supportare la valutazione dell'usabilità di applicazioni grafiche.

WEBREMUSINE

La prima questione che abbiamo vagliato nella progettazione di questo strumento è quali tipi di file di log avremmo dovuto considerare. Molti approcci all'analisi e valutazione di siti Web considerano i logs che si possono trovare nei server. Noi abbiamo scartato questa soluzione perchè soffre di varie limitazioni: alcune interazioni su una pagina Web non sono rilevate dai server, in alcuni casi la presenza di proxy server o l'uso di indirizzamento IP dinamico rendono difficile individuare l'utente, inoltre la memoria cache dei browser rende difficile rilevare tutti gli accessi alle pagine fatte dall'utente dal lato server. Quindi è stato implementato in JavaScript uno strumento di logging, da eseguire al lato browser, che è in grado di memorizzare tutte le interazioni effettuate dall'utente (come interazioni con il mouse, input da tastiera, selezione di link, ...) ed anche il comportamento del browser, come l'inizio e la fine del caricamento di una pagina.

Figura 1. Esempio di applicazione Web con task relativi.

Altro punto fondamentale è capire le intenzioni dell'utente durante la visita del sito Web. Nello strumento precedentemente sviluppato per valutare applicazioni grafiche si era partiti dall'analisi degli errori dell'utente. In particolare, si considerava quando l'utente cercava di eseguire delle azioni che non erano abilitate a generare reazioni del sistema in quel momento. Questo significava che molto probabilmente l'utente voleva eseguire il task associato all'azione non abilitata ma non aveva capito che doveva fare qualcosa prima per poter eseguire quell'azione. Questo tipo di analisi è meno significativa in applicazioni Web: un link è sempre abilitato. Quindi diventa più difficile catturare automaticamente le intenzioni dell'utente. Per catturare anche questo tipo di informazione durante il test viene visualizzata costantemente anche la lista dei task ad alto livello che è possibile realizzare durante l'accesso al sito Web (vedi Figura 1), richiedendo all'utente di indicare il task che vuole realizzare. Anche questa informazione verrà registrata nel log ed utilizzata nella fase di analisi automatica. Il modello dei task viene specificato con la notazione ConcurTaskTrees (Paternò, 1999) usando uno strumento automatico disponibile pubblicamente a <http://giove.cnuce.cnr.it/ctte.html>

Per poter essere applicato WebRemUsine richiede una fase di preparazione dove il valutatore crea una corrispondenza tra le azioni contenute nei log ed i task di base (quelli elementari). Si crea quindi una corrispondenza che consentirà di analizzare il comportamento dell'utente utilizzando la struttura del modello di task. In questo modo si può capire come vengono eseguiti i task ad alto livello supportati dal sistema e se vi sono dei problemi di usabilità.

In particolare, WebRemUSINE confronta i log con il modello dei task per poi fornire ai progettisti diversi risultati che sono relativi sia ai task che l'utente intende svolgere che alle pagine Web del sito considerato e alle loro interrelazioni (per ogni task, la lista di pagine associate e viceversa).



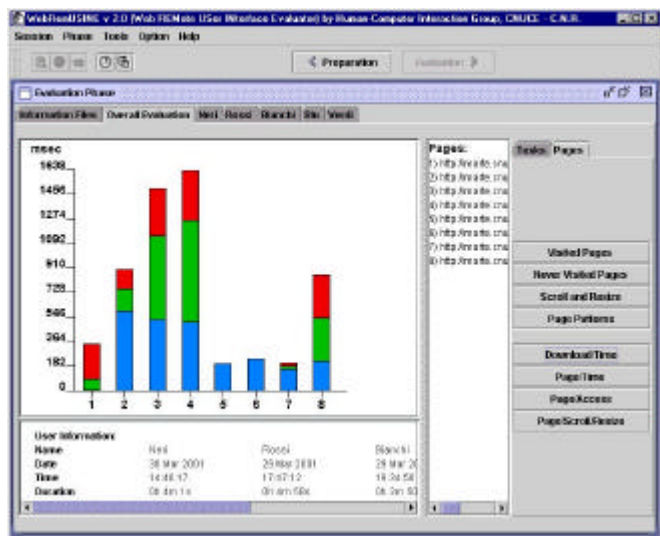


Figura 2. Esempi di risultati forniti da WebRemUSine.

Questi risultati possono essere utilizzati dal progettista per analizzare l'usabilità del sito Web da entrambi i punti di vista, per esempio confrontando il tempo impiegato dall'utente per realizzare un task su di una pagina con il tempo impiegato per scaricarla.

Attraverso l'uso di WebRemUSINE è possibile evidenziare diversi tipi di possibili problemi di usabilità del sito. Ad esempio, esaminando i task con un tempo molto elevato o i task eseguiti non correttamente è possibile identificare le pagine del sito che creano problemi all'utente.

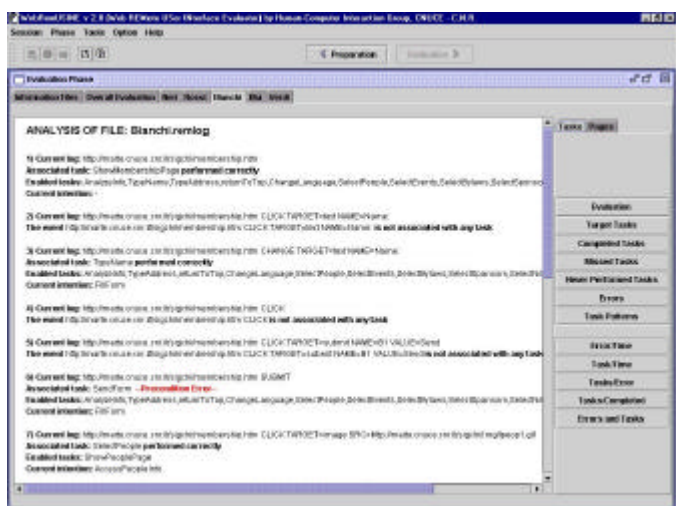


Figura 3. Esempio di analisi di una sessione di utente.

Lo strumento identifica le sequenze di task eseguite e pagine accedute ed è in grado di

identificare eventuali patterns, capire se l'utente ha seguito la sequenza più efficiente per i suoi obiettivi e contare le azioni inutili che vengono effettuate. In aggiunta lo strumento è in grado di indicare i task completati, quelli iniziati ma non completati, e quelli mai provati. L'ultima informazione è data anche per le pagine Web: pagine mai accedute possono indicare o pagine con contenuto poco importante o pagine difficili da trovare. Più in generale varie tipologie di informazioni vengono date anche relativamente alle pagine del sito (vedi Figura 2).

Una sessione dell'utente può venire successivamente analizzata alla luce degli obiettivi dell'utente e dei task che vengono dinamicamente abilitati e disabilitati a seconda delle interazioni effettuate (vdi Figura 3).

E' possibile analizzare il comportamento degli utenti per vedere se la loro performance migliora durante la sessione (così mostrando un'interfaccia facile da imparare). Inoltre si analizzano azioni degli utenti (quali movimenti delle barre di scorrimento e cambiamento delle dimensioni delle finestre) per analizzare se le dimensioni delle pagine sono ottimali o troppa informazioni è fornita. Lo strumento è capace di analizzare sia sessioni singole che di gruppi di utenti, consentendo così al valutatore di capire se un problema si verifica spesso o è limitato a certi utenti in particolari circostanze.

In aggiunta, si può usare WebRemUSine in modo complementare all'altro strumento, in corso di sviluppo che è stato introdotto all'inizio dell'articolo, il cui scopo è fare un'analisi automatica del codice HTML di un sito. In particolare, quest'altro strumento si può utilizzare per quelle pagine che tramite WebRemUSine sono state identificate come particolarmente problematiche per gli utenti.

CONCLUSIONI

Abbiamo discusso le motivazioni e le problematiche relative alla valutazione automatica di usabilità. Quindi, abbiamo introdotto gli strumenti che sono stati sviluppati presso il gruppo di interazione uomo-macchina del CNUCE-C.N.R.. La nostra intenzione è di ottenere alla fine un unico ambiente di valutazione di siti Web che utilizza diverse tecniche per scopi complementari. Comunque, gli strumenti sviluppati, ed in particolare WebRemUSine, hanno

già un buon livello di sviluppo tali da poter essere usati in contesti applicativi reali. A questo scopo il nostro gruppo è disponibile per collaborazioni di ricerca.

RIFERIMENTI BIBLIOGRAFICI

Bobby <http://www.cast.org/bobby/>

DoctorHTML

<http://www2.imagiware.com/RxHTML>

A.Lecerof, F.Paternò, Automatic Support for Usability Evaluation, *IEEE Transactions on Software Engineering*, Vol.24, N.10, October 1998, IEEE Press.

F.Paternò, G.Ballardin, RemUSINE: a Bridge between Empirical and Model-based Evaluation when Evaluators and Users are Distant, *Interacting with Computers*, Vol.13, N.2, pp. 151-167, 2000.

F.Paternò, Model-based Design and Evaluation of Interactive Applications, *Springer Verlag*, November 1999

F.Paternò, L.Paganelli, Remote Evaluation of Web Sites Based on Task Models and Browser Monitoring, Proceedings CHI'2001, Seattle.

WebSat <http://zing.ncsl.nist.gov/webmet/sat/websat-process.html>

RINGRAZIAMENTI

Ringraziamo Domenico Natale (Sogei) per le utili discussioni sugli argomenti trattati nell'articolo.