

## 使いやすさ評価システム”UI テスタ”の提案

旭 敏之、井関 治

(日本電気(株) 関西C&C研究所)

ユーザに対して、一定レベル以上の使いやすさを備えた製品を安定して供給するためには、評価テストの実施が不可欠である。ユーザインタフェースは本質的に多面性を有しているため、相補的性質を持った各種評価方式を状況に応じて使い分ける必要がある。本稿では、現行の各種評価方式を概観し、実用的な評価体系が必要なことを示す。この要求を踏まえ、ユーザ/装置間対話の大局的把握を可能にし、かつ自動評価環境の実現を目標にした、“UI テスタ”を新たに提案する。本システムの基本概念およびそのベースとなる対話履歴表示機能について紹介する。

### "UI Tester": A Tool for Computer Aided Usability Testing

Toshiyuki Asahi, Osamu Iseki

Kansai C&C Laboratory, NEC Corporation

Usability testing is one of the most important technique to assure users of product usability. It is required to select testing methods appropriately to meet each situation because a user-interface can be viewed from so many aspects. This paper shows usability testing methods now in use, and claims necessity of developing new testing methods which have supplemental characteristics with them. For satisfying this requirement, "UI Tester" which is the system to analyze user-machine dialogue and to develop the automatic testing environment is newly proposed. The fundamental concept of "UI Tester" and its dialogue visualization function is also described.

## 1. はじめに

ユーザによる評価テスト（ユーザビリティ・テスト）は、製品開発において「使いやすさ」を作りこんでいくための必須技術となっている。ユーザ・インタフェースに要求される性質は各製品の使用目的や環境、ユーザの嗜好などに強く依存しているため、基準や指針をトップダウンに与えることが難しい<sup>1)</sup>。したがって、現実の製品においてユーザに一定品質以上のUIを安定して提供するためには、[設計]→[試作]→[テスト]を繰り返す反復的設計手法が必要となる。設計と評価の試行錯誤を繰り返していくことで、UIを望ましいレベルにまで漸近的に brush up することが、現実的な方策として求められている。

反復的設計手法において評価のステップは、製品に内在する問題点を明確にし、改善に必要な基礎データを提供するためにある。したがってここで得られる評価結果が最終的な製品の品質を大きく左右する。個々の評価手法の効果や限界を明確に把握し、最適な運用心がける必要がある。

本稿前半では、多面的なアプローチが要求される評価テストの背景と、すでに製品開発の現場で応用実績のある各種評価手法の特性を比較検討する。後半では、これら既存手法と評価範囲において相補的な関係にあり、さらに評価の自動化を目的とした“UI テスタ”を提案する。

## 2. ユーザ・インタフェースの品質保証

### 2.1 評価の多面性

ユーザインタフェースの大きな特徴としては、その多面性を挙げることができる<sup>2)</sup>。「使いやすさ」は、本質的には人間が主観的に判定する（感じる）ものであり、これに影響を与える要因は各製品あるいは各対話の局面ごとに無数に存在する。当然ながら、その見方は個々のユーザや設計者によっても多岐に変化する。例えば、一般にメニューによる対話方式は初心者理解を助けると言えるが、熟練者にとってはかえって煩雑である場合もある。さらに、実際の製品を対象にしている以上、開発期間やコストなどの問題も無視することができない。

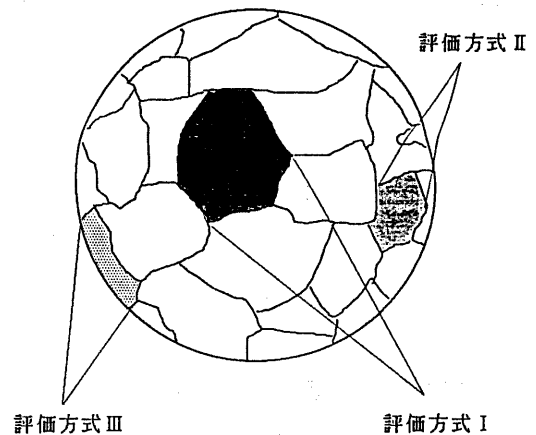


図1 評価の多面的アプローチ

実施時期	企画, 仕様, 設計, 試作, 製造, 検査, 市場投入後
評価の特性	定量的/定性的, 発見的/分析的, 個人差の影響, 製品改善効果
テスト実施方式	工数, 専門家の必要性, 必要なテスト環境/設備
評価の対象	画面, 入出力装置, 対話手順, マニュアル, メッセージ
使い易さの側面	分かり易さ, 効率, 学習の容易さ, おもしろさ, 生理的/心理的

表1 評価方式選択のための視点

表2 テスト方式の比較

	パフォーマンス評価	主観的評価	操作性評価
手法	・タスク遂行時間 ・エラーレート/復帰時間 ・学習時間	・インタビュー ・アンケート	・プロトコル解析 ・モニタリング
長所	実施が容易 定量的比較が容易 自動的データ収集可能	幅広い評価が可能 多数ユーザの傾向を把握するのに適している	幅広い評価が可能 問題点の発見が容易 →製品改善効果大 記憶に依存しない
短所	評価する側面が一面的 改善点の指摘が困難 絶対評価不可能	ユーザの記憶に依存 統計処理の過程でデータが埋もれる可能性 新規な問題点の発見は困難	データ解析が困難 解析者による結果のばらつきがある テストユーザを増やすことが困難

したがって評価の場面において、限られた手法だけで全ての要件をカバーすることは不可能であり、ケースバイケースで最適な手法を選択していく必要がある。製品開発過程で生じる様々な要求に対応できるように、様々な相補的特質を持った評価方式を開発していくことが重要となる。表1は、特質を考慮する際のいくつかの視点を示したものである。

## 2.2 従来評価手法

製品の使いやすさ評価テストにおいて、従来用いられてきた代表的な方式とその特徴を比較した(表2)。ただし、ここではあくまで「ユーザを用いたテスト」に絞ったため、モデリングやチェックシートによる机上評価方式などは含めていない。

## 2.3 技術課題

表2において、パフォーマンス評価や主観的評価は定量的分析が可能であり、母集団を大きくとることで評価結果の客観性を高めることができる。その反面、あらかじめ評価項目を設定することが原則であるため、製品を総合的に見た問題点の発見的評価が難しい。操

作性評価方式は、UIの幅広い範囲で問題点を発見することが可能である。特に即時的な発話内容を分析するプロトコル解析手法は、製品改善に適したきめ細やかな問題点抽出ができるため、積極的に応用技術が開発されてきた<sup>3)</sup>。

プロトコル解析手法の長所としては、

- ・問題点の洗い出しに有効  
→製品に対するフィードバック効果大
- ・致命的な欠点からマイナーな欠点まで数多く指摘できる
- ・マニュアルを含んだ評価が可能
- ・少ないユーザで多くのデータが収集できる
- ・評価対象を限定しない

などを挙げることができる。このため、筆者らの他にも多くの実用例が見られる<sup>4) 5)</sup>が、逆に次のような欠点も指摘されている<sup>6)</sup>。

- ・テスト、データ解析に時間がかかる
- ・評価者によって結果の解釈がばらつく
- ・ユーザ個人差の影響が大きい
- ・対話の流れを把握することが難しい

我々は、様々な側面から評価の要求に対応するため、実用的かつ網羅的な評価体系の構築を最終的な研究目標としている。

現在その一環として、特に上記プロトコル解析の問題点を考慮し、かつ製品開発工程のできるだけ上流（できれば設計段階）で実施できるテスト手法を構築するため、計算機支援環境の開発を検討している。以下にその概要について説明する。

### 3. UI テスタ

“UI テスタ”は、従来評価者のスキルや煩雑な作業に依存する割合が高かった評価テストにおいて、計算機支援環境を実現し、段階的に自動評価システムを開発していくためのベースマシンである。製品開発における実用性の観点から、

- ・テスト／データ解析工数削減
- ・評価結果の品質均質化／客観化
- ・専門家不要のテスト環境実現
- ・テスト参加ユーザの増加  
→ 個人差の影響低減

等の効果を狙う。さらに従来方式では難しかったユーザとシステムの対話そのものを対象とした評価手法開発を試みる。

### 3.1 概念

被計測系にプローブを挿入し、ある物理量を計測するようなテストとのアナログを考える。ユーザが持つ操作イメージと装置側のUI構造の一致度を判定することによって、使いやすさの計測を試みる。

- ・被計測系→ユーザと機器（実機あるいは仮想的機器）を含む対話環境
- ・計測系 →UI テスタと評価者
- ・計測量 →対話シーケンス
- ・出力 →使い易さ（ユーザの操作イメージとUI構造との一致度）
- ・計測形式→無侵襲計測

当然ながら、物理的実体のないユーザの操作イメージを直接的に計測することは不可能である。ただし、一般にユーザは、いくつかのレベルにおいてゴールを設定し、これを達成するというサブタスクを積み重ねて、最終的に目的の作業を達成している（図3<sup>7)</sup>）。したがって対話の履歴は、ある時点でユーザのイメージしたゴールやサブタスクを反映しており、誤りや冗長操作も含めることにより、操作イメージの一表現形態であると考えることができる。

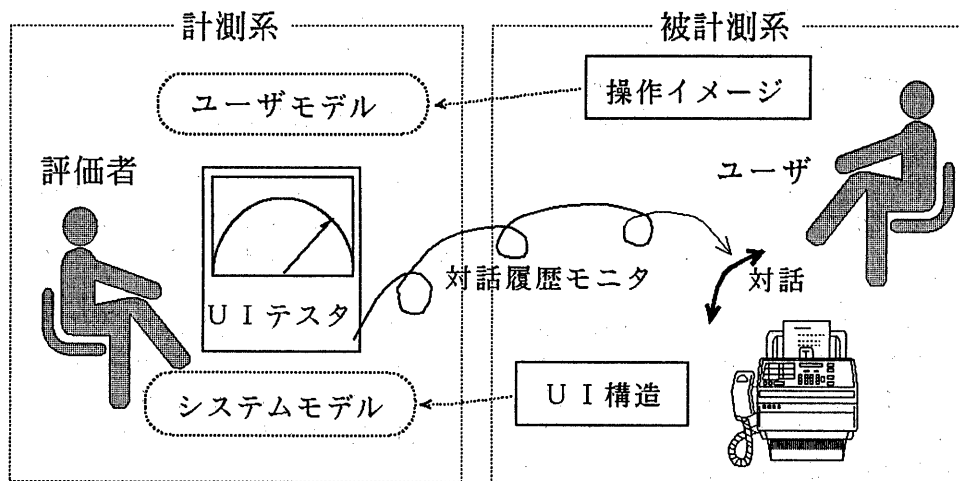


図2 自動評価システムの概念

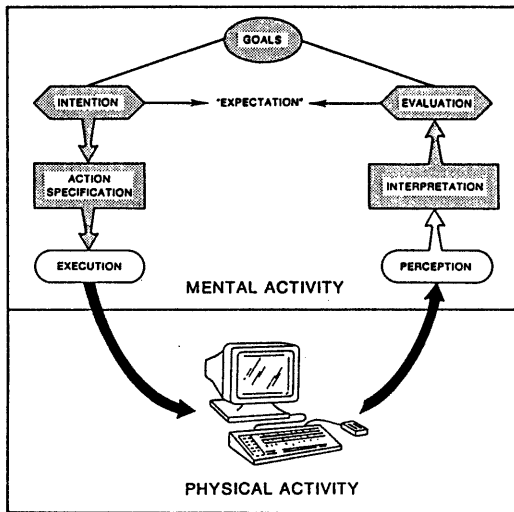


図3 対話進行のモデル<sup>7)</sup>

UI テスタでは、ユーザと機器との対話シーケンスを計測対象とし、これを適当な形式に変換することでユーザがイメージした操作手順をモデル化することを考える(図2)。これは、操作イメージという抽象的事象を計算機処理可能な中間表現に落とす処理を見なすことができる。装置側のUI 構造についてもこれと同じ形式で表現できれば、両者の一致度(相違度)を判定することで使いやすさ(使いにくさ)計測を実現することができる。

### 3.2 対話履歴分析

3.1で述べた評価システムを実現するためには、何らかの方法で対話状況をモニタし、これを処理可能なデータ形式に変換する必要がある。従来、ユーザの操作履歴を記録し、これをベースに使いやすさ評価を試みた例は少なくない。ただし、この場合の記録は多くが打鍵レベルの操作履歴であり、人間がみて意味のある対話状況の再現は難しい。したがって、キーボードの各キーの打鍵頻度など、定量的かつ局所的な評価指標として利用されることが多かった<sup>8)</sup>。

一方、ユーザの操作イメージと装置側UI の認知的な整合性を議論する際には、「ユーザと装置との対話(のコンテキスト)」を解析対象とする必要がある。この場合、ユーザの操作履歴に関しては、少なくともコマンドレベル以上の意味を持ったデータをベースにして分析を進める必要がある。さらに、対話の進行状況を規定するためには、装置側の応答履歴も不可欠である。

したがってUI テスタにおいては、

- ①コマンドレベルの操作履歴
- ②装置側応答内容

の2種類の時系列データを計測/分析の対象とする(図4)。

### 状態遷移表現

応答内容が装置側の状態を表わすとし、ユーザ操作はその状態の変化を惹起する事象であるとする、対話の進行状況を状態遷移と見なして表現することができる。すなわち、応答履歴の各内容をノード、操作内容をアークとした状態遷移図を作成することができる(図5)。

状態遷移表現は、時系列の事象を直感的/視覚的に把握することに適しており<sup>9)</sup>、従来テスト方式では難しかったユーザの操作イメージとUI 構造の大局的な比較検討が可能となる。

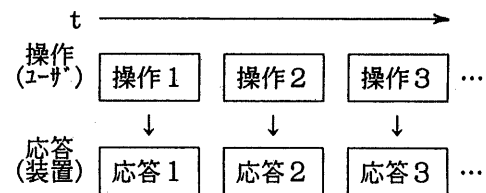


図4 UI テスタの計測データ

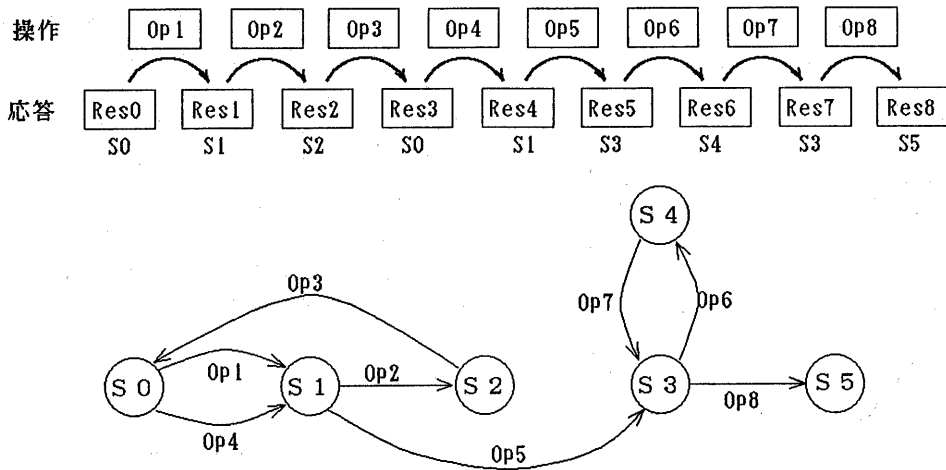


図5 対話履歴の状態遷移表現

一般に、ユーザ(特に初心者)の操作イメージは固定的なものではなく、対話が進行するにつれて次々に変化する。ユーザは原則として、対話中に示される装置側の応答とタスクの実現イメージ(ゴール)を比較評価し、次段階の操作方針を決定する(図3)。したがってUIの設計において、応答内容のデザインはユーザのガイダンスとしての意味が大きい。様々な局面における装置の応答とその後のユーザ操作を分析することで、応答内容やUI構造上の問題点などを評価することができる。特にユーザの誤り操作に着目し、この原因と考えられる応答内容を検討することで、UI設計時のデザイナー予測と操作イメージの不一致を抽出することができる。

#### 4. 対話履歴表示機能の試作

前述したアイデアを具体化し今後のUIテスト開発のベースとするため、ファクシミリを対象とした対話履歴モニタ及び状態遷移表示機能をパソコン上に実現した。本試作機は、ファクシミリ内の信号を検出/解析することにより、ユーザと装置との対話履歴を表示するものである(図6)。対話履歴とその状態遷移表現例をそれぞれ図7、8に示す。

本試作機を用いていくつかの試行テストを

実施した。ユーザの示した対話履歴と熟練者による正解操作の対話履歴を比較し、両者が異なる部分(誤操作や冗長操作)の装置側応答内容を抽出することで、要検討部分を抽出するというレベルでは評価が可能であることがわかった。

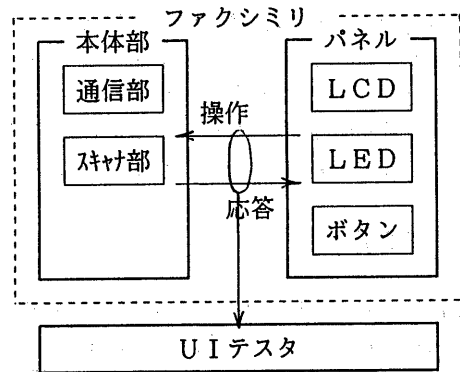


図6 試作機の構成

#### 5. 検討

以上、多面的な評価が必要となる背景と各種評価方式を紹介するとともに、従来方式とは違った側面からの評価を指向したUIテストの概念について説明した。

UIテストは、最終的には製品評価テスト

の自動化を目的としている。人間の判定を介さずに使いやすさの評価が実現できるシステムが理想であるが、そこに至るまでのステップとして、

- 1) 評価が容易な形式で、データを提示する。
- 2) 評価の一部を代行する、あるいは評価作業を支援する。
- 3) 自動評価（問題点指摘）。

を考えている。筆者らはこの概念をCAUT (Computer Aided Usability Testing) と命名して、段階的に研究を進めていく予定である。現在検討中のいくつかのアイデアを以下に述べる。

### UIモデルとの比較

デザイナーが意図した対話手順と実際にユーザが実行した手順の違いは、ユーザインタフェースの問題点を示唆していることが多く、使いやすさ評価では特に詳細に検討する必要がある。UIテストにおいても、正常な対話手順 (UIモデル) とユーザの対話履歴を比較し、その相違点が明瞭になる表示形式を考える必要がある。本論文で示した対話履歴の例において、熟練者の示した履歴はUIモデルの一形態であると考えられることができる。したがって、これとユーザ対話履歴との違いを明確に提示できれば、上記1)のステップを達成することになる。例えば、2つの履歴を同時表示し各状態の対応関係を提示する方式や、UIモデルにしたがってユーザ対話履歴を變形する方式などを検討している。

### ユーザモデル評価

UI評価において、テストに参加するユーザの個人差は常に難しい問題である。当部で行ってきたプロトコル解析テストでは、小人数のユーザで抽出された「明瞭な」問題点に検討を絞ることで、この問題に対処してきた。この理由として、時間と手間の関係で多数のユーザでテストすることと、評価データを複数のユーザ間で比較することが困難であったことが挙げられる。UIテストでは、テストそのものはユーザだけで実行可能であり、さ

らに最初からデータが計算機に入っていることから、多数データの収集と比較が容易であると考えられる。この特性を生かし、複数ユーザ間の異常操作傾向を抽出し、これを個人差の影響が少ないユーザモデルとして検討対象とする方式が考えられる。計算機による評価作業支援であると考ええる。

### シミュレータ評価

今回試作したテストは、電氣的に評価対象装置に接続することで対話履歴をモニタしているため、他の製品に適用することが一般的には難しい。これを解決する方式としては、評価対象 (の入出力部) の動作をシミュレータ上で実現することが考えられる。すなわち、シミュレートされた製品をユーザに操作してもらい、その対話履歴を評価するという方式である。これが実現すれば、評価対象の制限がなくなるだけでなく、製品試作以前の段階でユーザビリティテストが実現できることになる。CAUTシステムの1つの実現形態として目標にする。

以上いくつかのアイデアを述べたが、もちろんこれだけでは自動評価が達成できるものではない。最終的には、プロトコル解析などユーザビリティテストを積み重ねることでノウハウを蓄積し、対話履歴の特徴から製品問題点の候補を推定できるまでに発展させていきたい。

### 参考文献

- 1) 稲葉他：「使いやすさ」の設計手法が実用に、日経エレクトロニクス, no. 529, pp. 107-124 (1991)。
- 2) 神場他：マルチユーザ環境に基づくユーザインタフェース設計システム "U-face", 情報処理学会研究会報告 ヒューマンインタフェース 32-2 (1990)。
- 3) Asahi, T., Miyai, H. (1990): A Usability Testing Method Employing the "Trouble Model", Proc. Human Factor Society 34th Annual Meeting, pp. 1233- 1237.
- 4) Mills, C. (1987): Usability Testing in the Real World, SIGCHI Bulletin, vol. 19, no. 1, pp. 43-46.

5) Montford, S. (1989): User Observation: Guidelines for Apple Developers, ACM SIGGRAPH '89 tutorial text#15.

6) Nielsen, J., Molich, R. (1990): Heuristic Evaluation of User Interfaces, CHI '90 Proceedings, pp. 250-256.

7) Norman, D. (1986): User Centered

System Design, LEA Publishers.

8) Brown, C. (1986): Human-Computer Interface Design Guidelines, Ablex Publishing Company, Norwood, NJ.

9) Myers, B.: Tools for Creating User Interfaces: An Introduction and Survey, CMU technical Report, CMU-CS-88-107.

操作履歴(キー)	応答履歴(LED)	応答履歴(液晶)	状態
(1) オープン	01000000010100	1991.06.06(MON) 17:09 ゲンコウヲ セット シテタサイ	S 1
(2) 登録	01000000010100	カクシュ データ トウロク セツテイ ジキョク テンワハンゴウ	S 2
(3) カーソル下	01000000010100	カクシュ データ トウロク セツテイ ジキョク タンシュカハンゴウ	S 3
(4) カーソル下	01000000010100	カクシュ データ トウロク セツテイ ハースワート トウロク	S 4
⋮	⋮	⋮	⋮
(24) 入力	01000000010100	カクシュ データ トウロク セツテイ ジキョク テンワハンゴウ	S 2
(25) 完了	01000000010100	1991.06.06(MON) 17:30 ゲンコウヲ セット シテタサイ	S 1

図7 ファクシミリ対話履歴の例

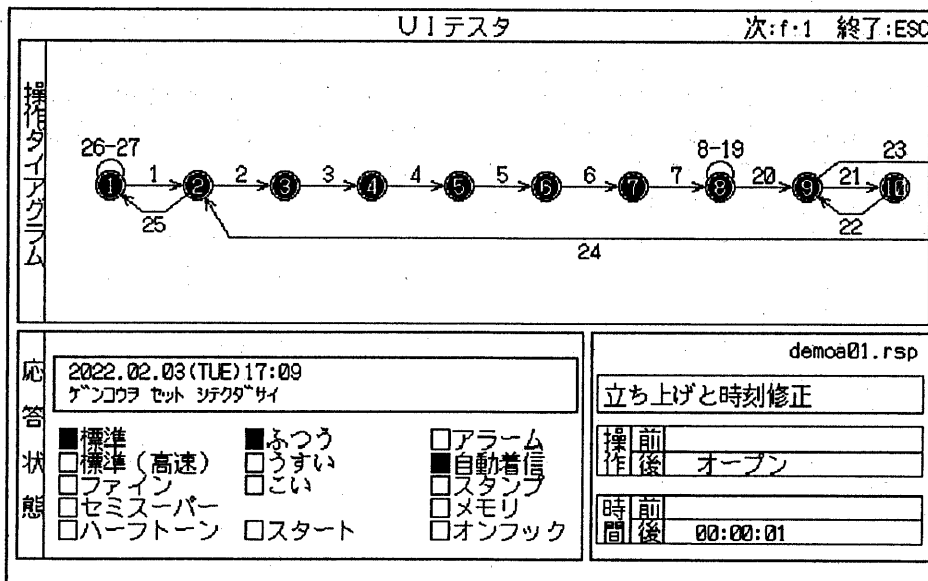


図8 UIテストの画面例