

広域大規模分散環境における音声情報の利用

2S-10

大野 浩之¹⁾村井 純²⁾

1) 東京工業大学 理学部 2) 慶應義塾大学 環境情報学部

1 はじめに

分散処理に基づいた計算機環境を相互に接続することにより構築される広域大規模分散環境においては、利用者の位置に依存しないサービスを提供することが可能となる。たとえば WIDE プロジェクト [1] では、64Kbps 以上のデータ回線と UNIX ワークステーションによって構成される WIDE インターネット上で広域大規模分散環境の実験を行っており、WIDE インターネットの各ノードにおいて同一の作業環境を提供するための試みが続けられている [2]。

このように計算機利用者の行動範囲は、従来に比して大きく拡大されたが、実際には回線およびその回線を利用して実現されるマンマシンインターフェースの機能によって制限を受けている。しかし、計算機利用者の実際の活動に対する計算機環境の役割の拡大を考慮すると、柔軟なマンマシンインターフェースと多様な回線への対応によるサービス提供可能性の拡張に関する研究開発は急務であり、例えば音声は文字や図形とともに、利用価値の高いメディアとして期待されている。特に、従来文字が利用されていた計算機システムからのメッセージに音声による表現を追加することは、分散システム上での作業環境を考慮すると、非常に効果が高いと考えられる。

そこで、WIDE プロジェクトでは計算機システムからのメッセージに限定して音声表現に関する実験を行なった。この実験において着目したのは、限られたリソース（ディスク、通信回線、音声合成装置の品質）の中で、いかにして音声の特徴を生かしたサービスを提供できるかという点である。

さらに、音声は公衆電話回線で伝搬できるので、電話のタッチトーン (DTMF) 信号を利用した入力系について検討を行ない、DTMF 信号と音声を利用した shell である PhoneShell [3] を試作した。これにより、音声を活用した双方向のマンマシンインターフェースが実現し、さらに実際の運用を通じてその有効性を確認した。

2 音声情報の利用

計算機システムが音声情報を利用可能になると、さまざまなサービスが可能になる。

例えば文書ファイルを読み上げるといったサービスが考えられる。これは、視覚に障害を持つ人々にとっては非常に有効なサービスであろう。利用者の希望に応じてチュートリアルを音声で行なうといったこともできる。また、計算機システムに何らかの障害が生じた時、音声によってこれをシステム管理者や利用者に伝えることができれば、混乱の回避に役立つ。

この例からも明らかなように、音声情報を有効に利用できる場面は多岐に及ぶが、本研究ではその中でも、システムの稼働状況に関する大量の情報を常に正確に把握している必要があるシステム管理者や、計算機システムを使いこなしている一般の利用者に焦点を合わせ、彼らにとって音声による情報の提供が必要になるのはどのような場合かをまず最初に検討した。

その結果、文書ファイルの中身を朗読や、音声によるチュートリアルといったサービスよりも計算機システムが出力するメッセージを音声で的確に伝達するサービスの方がより期待されていることがわかった。

2.1 計算機システムのメッセージの分析

計算機システムからのメッセージには、ユーザが作成したプログラムから出力されるものと、システム側（OS および、標準的に用意されているアプリケーション）から出力されるものがあるが、ここでは、後者に着目する。そこで試みに筆者らの利用している UNIX

システムの /bin, /usr/bin, /usr/ucb の下にある実行可能なファイルを 100 個無作為に取り出して、そのソースファイルにおけるメッセージの構成を分析をしたところ、出力されるメッセージは以下のように分類されることがわかった [4]。

	メッセージの構成	割合 (%)
1	特定の固定フレーズからなるもの	45
2	特定のフレーズ + 変数 (OS の名前空間に対応する名詞)	23
3	特定のフレーズ + 変数 (数字)	15
4	2,3 の組合せ およびそれ以外	17

図 1: システムからのメッセージの構成

この結果から、特定のフレーズ、数字に加え、login 名や host 名といった、分散環境におけるリソースの名前を発声できれば、システムからのメッセージの大半を読み上げることが可能であり、必要となる語彙の数は、OS の名前空間の大きさで決まることがわかった。

2.2 音声情報の利用

計算機に音声を発声させるには、さまざまな方法がある。その中でも、テキスト音声合成はもっとも有望な手法であり、将来的にはこの方法が主流になると一般的には考えられているが、今回は採用を見送った。前項で示したように、本実験で必要となる語彙は大半が OS の名前空間の読み、すなわち固有名詞であるので、コストの点も考慮すると現時点ではテキスト音声合成が必ずしも最適な手段にはならないからである。

結局今回は、単語単位で音声を録音して保持し、再生時にはそれを単につなげて再生するという単純な方法を採用したが、当然今後さまざまな方式の音声合成装置が使われることが予想されるので、音声情報は、音声合成装置の種類や方式とは独立に決めた標準的な形式で保持する必要がある。

さて、音声情報を単語単位で保持する場合、データは分散して所有し必要に応じて融通できれば効率が良い。この場合ネットワークに対しては以下のような要求が生じる

1. 必要な音声情報をすべて蓄積してから発声する方法なら伝送路にリアルタイム性は不要
2. 転送を受けながら順次発声する場合にはデータ転送にリアルタイム性が望まれる

いずれにせよ、計算機からのメッセージを音声で伝えるという目的のもとでは、発声中に単語間で不自然な間があいてしまっても目的は達せられるので、転送方法を工夫すれば、現状の 64kbps の回線を介して音声データをリアルタイムにやりとりすることも不可能ではない。

なお、現行のシステムの改造をを最小限に押えて音声をサポートするようにするには以下のような方法がある。

1. ライブラリを改造する。たとえば vprintf といったものを作り、すべてのツールをコンパイルし直す
2. 音声発声のための専用のコマンドを作りそれを明示的に呼び出す
3. pty を使い、stdout, stderr に出力されるメッセージを監視し必要に応じて発声する

3 応用可能性の検討

音声情報が利用可能になった広域大規模分散計算機環境における音声発声機構の導入形態としては、まず

- (1) 既存のアプリケーション中に音声を用いたサービスを追加する。

が考えられる。一例として、phone (ネットワークを介して利用者同士が文字ベースで会話を行なうためのツール) の着信メッセージの音声化を挙げることができる。

前述のように、現在では計算機利用者は、ワークステーション上で多数のウィンドウを開き複数の仕事を同時進行させることが多いので、phone が着信した際、当該ウィンドウが他のウィンドウの下に隠れていたり、iconify されていると phone の着信を見逃してしまうことがある。また当該ウィンドウが表示されていても利用者が画面に注目していなければ着信を見落とすことになる。もちろん phone は着信時には文字によるメッセージだけでなく、ベルによりユーザの注意を促すようになっているが、ベルの鳴るアプリケーションは他にも数多くありどこかで何かが起こったことはわかるが、どこで何が起こったかを知ることはできない。このようなケースでは、音声によりイベントの発生を伝えるのが効果的で、この例ではどの(あるいは誰が利用している)ワークステーションの、どのウィンドウに着信があったかの2点が明確になれば十分である。ここで重要なことは、これら2つの情報が、phone の着信したマシンの OS の名前空間に必ず対応しているということである。すなわちシステムの名前空間に対応させて音声情報を事前に収集しておけば、確実に読み上げることができるのである。他にも shutdown メッセージなど、同様の観点から音声によるサービスの追加が効果的であるアプリケーションは多い。

さて、発声機構のもうひとつの導入形態として、

(2) 音声を積極的に利用するアプリケーションを新規に開発する

というアプローチもある。すなわち、計算機の出力はディスプレイに表示されるという概念に拘らず、文字ではなく音声を主に利用するツールを充実させるのである。

この方式のツールの利用者は、ディスプレイの前で作業をする必要はなく、例えば電話器を介して計算機を操作するといった新たな可能性が開ける。

この方式においても、出力されるメッセージがシステムの名前空間に対応するものであれば、発声は容易である。

いずれにせよ、発声するメッセージを計算機システムが出力するものに絞らざることで、音声化のための作業が不必要に大きくなるのを避けることができ、音声化を効率良く行なうことができる。このことは現在の計算機環境の改善に大きく寄与する。

4 試験系の構成と運用

前節の応用可能性の検討を踏まえ、音声情報を扱うための試験系を構成し、以下の2つのアプリケーションを作成した。

1. 音声によるメッセージ機能付の phone
2. phone shell と呼ばれる、音声と DTMF 信号を併用した shell

4.1 音声メッセージ機能つき phone の構成

計算機利用者に注意や警告を促すために音声を利用する一例として、音声メッセージ機能を phone に付加した。phone には、外部から着信があった場合その旨のメッセージを表示する代わりに、ある shell スクリプト (.busy) を実行する機能があるが、ここではそれを用いずに、phone のソースを変更した。具体的には着信を伝えるメッセージを表示する部分の printf ライブラリを音声対応型の vprintf に差し替えてコンパイルし直した。vprintf は自前のライブラリで音声合成装置が稼働していたら、画面に文字列を表示するとともに発声も行なうという仕様になっている。

これにより、ウィンドウが iconify されていても、多数のウィンドウの下に埋もれていても、phone の着信を見落とすことがなくなった。なお、このためのハードウェアは、音声合成装置を搭載したワークステーションや X 端末であれば何でもよく、現在 SONY NEWS、X MiNT などで利用可能な状態になっている。

4.2 phone shell の構成

phone shell は、通常の電話から、モデムや小型のパソコンなどを使うことなしにワークステーションを操作することを目標につくられた一種の shell である。

phone shell を稼働するためのハードウェアはワークステーションとそれに接続された NCU (回線制御装置) およびオーディオミキサーによって構成されており、それらは

- ワークステーションからの音声を電話回線に送り出す部分
- 電話回線から送られてきた DTMF 信号を文字列に変換する部分

の2つに分けられる。

phone shell の利用者は、計算機に対しては DTMF 信号を使って情報を送り、計算機からは音声によって情報を受け取る。phone shell を用いたもっとも単純なアプリケーションは、テレホンサービスである。すなわち、プッシュホンで操作し、計算機からの音声による指示にしたがって、あらかじめ決められたメニューを選択し必要な情報を取り出すというものである。現在この方法により、任意のマシンに対する ping、あらかじめ設定したマシンのユーザ数、ロードアベレージなどを知ることができる。もちろん、ユーザがプログラムを事前に用意することで、より高度なサービスを受けることもできる。

また、障害発生時には計算機が電話を通じて音声による障害報告を管理者に対して行なうことも可能である。これらの機能により、大規模分散システムの管理者は、障害発生を検知する機会が従来より大幅に増えるとともに、障害発生時には、もよりのプッシュホンから応急措置が講じられるようになり、システムの安定稼働に貢献している。

5 おわりに

広域大規模分散環境における音声情報の有効利用について検討した。その結果、音声によるサービスを計算機システムからのメッセージの伝達に限定すると、必要となる音声情報のおよその規模を見積もることができることがわかった。この点に着目して試験系を作って運用した。

今後早急に対処する必要がある課題としては、音声情報の分散管理の体系の整備、現在のネットワーク環境のもとでの音声データの相互交換サービスの実現、phone shell の機能の強化などがある。

謝辞

WIDE インターネットへの音声情報の導入にあたり、さまざまな議論に参加してくれた WIDE プロジェクトのメンバーに感謝する。

参考文献

- [1] Murai, J., Kato, A., Kusumoto, H., Yamaguchi, S., and Sato, T.: Construction of the Widely Integrated Distributed Environment. In *Proceedings of TENCON '89* (November 1989), IEEE. Bombay, India.
- [2] Murai, J., Kusumoto, H., Yamaguchi, S., and Kato, A.: Construction of Internet for Japanese Academic Communities. In *Proceedings of Supercomputing '89* (November 1989), ACM. Reno, Nevada.
- [3] 大野浩之, 村井純: WIDE プロジェクトにおける音声情報の活用について — phone shell の設計とネットワーク管理への応用 —, JUN 16th UNIX SYMPOSIUM(1990), pp164-173
- [4] 村井純 他: WIDE プロジェクト 1989 年度研究報告書, WIDE プロジェクト (1990)