



ELSEVIER

Speech Communication 31 (2000) 279–291

SPEECH
COMMUNICATION

www.elsevier.nl/locate/specom

Practical issues in the application of speech technology to network and customer service applications

David Attwater^{a,*}, Mike Edgington^{a,b}, Peter Durston^a, Steve Whittaker^a

^a BT laboratories, Mobility and Network Services, Martlesham Heath, Ipswich, Suffolk, IP5 3RE, UK

^b SRI International, Menlo Park, CA, USA

Received 26 April 1999; received in revised form 14 June 1999; accepted 25 November 1999

Abstract

This paper proposes a simple model to characterise the different stages of short telephone transactions. It also discusses the impact of the context of the caller when entering an automated service. Three different styles of service were then identified, namely, large vocabulary information gathering, spoken language command and natural language task identification for helpdesks. By considering human dialogue equivalents, the requirements for each style are considered. Consequently, it is shown that each style pushes different technological limits. Three case studies, selected from current project from BT laboratories, are presented to highlight the practical design issues in these different styles. The styles and case studies presented are:

- Information gathering – UK name and address recognition.
- Spoken language command – network service configuration.
- Natural language helpdesks – BT operator services.

It is shown that large vocabulary information gathering systems require high accuracy, careful data modelling and well-designed strategies to boost confidence and accuracy. Spoken language command requires dialogue and grammar design and test complexity to be managed. Natural language task identification requires large volumes of training data, good learning algorithms and good data generalisation techniques. These styles can be mixed into a single interaction meaning that design frameworks of the future will have to address all of the aspects of the different interaction styles. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Speech recognition; Dialogue modelling; Network service automation; Address recognition; Natural language processing; Semantic classification

1. Introduction

1.1. The model

1.1.1. Caller context

The context of a call to an automated service is very important. We note two important related dimensions:

* Corresponding author.

E-mail address: david.attwater@bt.com (D. Attwater).

- Victim or volunteer – was the caller expecting automation or were they unsuspecting victims?
- Frequent or infrequent – is the caller well primed and experienced or do they rarely call the service?

It is the clear experience of the authors that these two dimensions strongly dictate what can be achieved, and in what style, for a given service. It is also extremely common for these two dimensions to pair-up into *frequent volunteers* and *infrequent victims*. By definition, frequent callers to a service will quickly come to expect automation and become volunteers if they continue to call.

The term *victim* is deliberately emotive. In the UK IVR services, especially those based on touch tone, are widely disliked when callers are not expecting them (Attwater et al., 1998a). Early indications are that acceptance of dialogue-based speech recognition systems is higher, but there are currently no well-established norms for talking with machines. Consequently, spoken language behaviour from callers who have not been primed for a service can be difficult to predict.

1.1.2. Four-layer call handling model

There are typically four phases during a transaction with a service:

- *Problem specification* – in which the problem to be solved is identified.
- *Task identification* – in which the customer intent is identified within the framework of available services.
- *Information gathering* – in which all details necessary to achieve the task are gathered from the customer
- *Task completion* – in which the customer receives the service or information they require.

In practice, when a customer calls a human agent there is often significant overlap between these various phases. For example, there may be several stages of negotiation in order to discover the actual problem experienced by the customer, during which several potential services may be offered to the customer. Fig. 1 shows a real call to a BT international operator, annotated into these four phases.

This model is helpful for analysing operator-based and automatic interactions. It is important to note that the first two phases of a transaction

may also be implicitly satisfied. For example, the BT directory enquiries service on “192” uses a human operator to achieve information gathering and then automates the task completion phase by use of recorded number announcement. Since the directory enquiry service is very well known, the first two phases are implicitly fulfilled when the customer dials the “192” access number.

1.2. Dialogue styles

This paper suggests that the point in the model at which a caller engages a dialogue system will be a deciding factor in the style of dialogue that the caller and agent conduct. The style may even change as the caller advances through the stages of the four layer model. We propose a progression of dialogue styles, based on the patterns of dialogue which have been observed to be successful. These are:

| | |
|-----------------------|--------------------------------|
| Information gathering | “ <i>answer the question</i> ” |
| Command | “ <i>tell me what to do</i> ” |
| Helpdesk | “ <i>what’s the problem?</i> ” |

In information gathering dialogues it is often satisfactory for the agent to take the initiative and follow a structured question and response style (Bennacef et al., 1995). Once the information is gathered, the task may be completed. Automation of this style of dialogue lends itself to large vocabulary isolated word speech recognition and highly structured dialogues.

In command dialogues the caller will take the initiative and give a clear direction initially to the agent, which is often followed by information gathering. A key element to this style of dialogue is that callers may spontaneously give the whole command and information in a single utterance. Automation of this style lends itself to hand-coded finite state speech recognition grammars and slot-filling natural language style dialogues.

Finally with assistance dialogues, callers tend to take the initiative and describe the problem which they are experiencing to the agent with the expectation that the agent will propose a potential solution. Once agreed, subsequent information gathering may occur. An important element of

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات