Adaptive Management of the Answering Process for a Call Center System

Federica Cena and Ilaria Torre
Department of Computer Sciences – University of Torino
Corso Svizzera 185 - 10149 Torino (Italy)
fede.cena@tiscali.it, ilaria.torre@unito.it

Abstract. This paper describes the development of an Adaptive Call Center which personalizes the management of the answering process. The system is composed by an adaptive response system with a speech recognition engine and by an operator support structure, which is dynamically involved in the answer when the sentence is not recognized or the question belongs to those classified as complex. In this case the call is routed to the operator which best fits the caller features.

1 Introduction

Over the last few years the world of Call Center Systems is being involved in a deep transformation. Moving from the evidence of cost differential between live contact handling and automated transactions¹, many companies (but also non-profit organizations) are switching from fully operator-based solutions to self-service ones. IVR (Interactive Voice Response) and web contact are the main self-service tools. Besides reducing costs of work and costs of training, self-service allows also to create a common way of presentation and to extend the time of access to the entire day. Anyway, self-service cannot entirely substitute the role of human agents: it can automate some operations, lighten human agents’ work, but it difficultly succeeds in managing complex inquiries and special cases, in facing unexpected situations and in overcoming the resistance of users toward non-usual and non-human interactions.

Given that, a solution which i) deploys IVRs, particularly with automatic speech recognition – ASR -, and furthermore ii) integrates IVR events within an operator support structure, seems very interesting: it allows companies to combine the advantages of flexible self-services with those of employing human agents, supporting the user incrementally. Indeed the initial interaction and the management of mechanical operations are automated and operators are dynamically drawn in for accomplishing complex operations or in case of misunderstanding or problems with the ASR.

¹ Live contact handlings run from $3 to $10 per contact and up, while IVR (Interactive Voice Response) transactions generally cost tens of cents and Web contacts may be mere pennies [1].
However, the mentioned solution is still not satisfactory: speech-enabled-IVRs offer standardized conversations, they are not as flexible as human operators and they cannot cope with the differences between users and their different needs (e.g. disabled people, novice users, etc.); moreover, unless it is well integrated, the step of switching the call to an operator can be experienced as a system failure and can decrease the trust of users, especially if the final answering process is not successful.

The project described in this paper aims at finding out a solution for solving these problems, inside a speech-recognition-enabled IVR structure, collaboratively supported by an operator-based structure. The idea is that of using adaptivity to manage the workflow of the answering process. In particular, for the goal of integration, it will have to be applied to both the main components, that means:

- Personalizing the interaction with the automatic voice response system,
- Routing of the call toward the operator that best fits the caller needs, when the first solution is not applicable or was not successful.

The advantages coming from this approach can be summarized in the following list:

(i) efficiency: using a language fitting the user features (IVR) and/or routing the call toward an operator fitting the user features (Routing), decreases the possibility of user’s misunderstanding (see [3] for a comparison on a different approach) and, consequently, increases the probability user’s problems are solved and shortens calls.

(ii) positive experience of the user: a positive and friendly interaction, which satisfies the user requests, and improves over time, makes the interaction enjoyable,

(iii) decreased frustration of the operator: on the one hand, (s)he can leave the boring and repeated operations to IVR and, on the other hand, the adaptive routing allows her/him to manage situations that are adequate for her/his competences and skills.

2 Architecture of the system and flow of an inbound call

The main components of a typical Call Center infrastructure running on standard PC servers are represented by a Gateway, which interconnects the Call Center applications with the Public Switched Telephone Network (PSTN), and a Communication Server, which provides the basic functionalities for IP Telephony. In figure 1 they are not included, in order to focus the schema on the logical architecture that carries out the personalized interaction with the user (caller).

As it can be seen in the figure, the core of our application is the Response Manager toward which the Communication server routes the traffic. It controls the flow of the call and the dialog among the modules which manage:

i) the Voice User Interface - VUI -, composed by the Automatic Speech Recognition engine, in charge of understanding natural language user input, and the Response Generation agent, in charge of prompting, providing menus and answering to the user;

ii) the routing of the call, accomplished by the Routing agent.

Whenever a call is received, the system checks the calling number. If it is not recognized, the user is taken into account as a new one, otherwise, there are two options: the caller is a customer or (s)he is using the phone of another customer.
To manage this second possibility, in order to avoid wrong forms of personalization, the system asks if the calling number is the usual one. Consequently, there are two types of dialog interactions: standard, for non-customer callers; personalized, for customer callers, based on her/his model. During the interaction, the user makes her/his request, which is understood by the Speech Recognition Engine (using a keyword spotting technique) and analyzed by the Response Manager. If the request is identified as a simple one, the IVR is charged for supplying the service requested and the Response Generation Agent produces the answer according to the features of the caller. The TextToSpeech Engine, inside of it, will generate the final output. If the request is complex, or the ASR engine does not recognize the user input, the Response Manager switches the call to the Routing Agent. As it will be described in the following, this agent routes the call toward an operator, taking into account both caller’s and operator’s models.

3 Personalized Automatic voice response

In this section, and in the next one, we will focus on the components which couple voice response systems and routing systems with adaptation rules, in order to carry out the adaptive management of the call. Our prototype regards the Call Center of a bank, thus the adaptation rules, the features of the users models and also the grammars for automatic speech recognition are specific for this application domain.

As seen, the VUI of the system includes two kinds of complementary technologies: speech input and speech output. Similarly to a GUI, personalizing a VUI can regard the adaptation of the contents, of the form of presentation and of the
structure of navigation, which, for an IVR, means the menu commands. Given the
goals discussed in the introduction, we limited the adaptation to the first two aspects.

To accomplish that, we integrated the automatic response system with an agent that
decides which is the right answer for the specific caller, on the basis of a set of rules
(see the Knowledge Base - KB, in the figure). Then the system uses TextToSpeech for dynamic generation of words that are specific and variable (e.g. user name, user features, data from DB, etc.), adding them to previously recorded message chunks (e.g. welcome formulas, the questions to ask, pieces of phrases for asking details about the user question, etc.). The reason is that people do not like TextToSpeech because of its unnatural prosody and poor intelligibility (see the experience of British Telecom [5]).

USER MODELING. As it can be seen in figure 1, the Response Generation agent
accomplishes its task inquiring the model of the caller. The system stores a user
customer model for each caller, building and updating it on the basis of the Customer DB of the
bank. The model is structured in a set of dimensions. Both the Response Generation
agent and the Routing agent access the same user model, but use different dimensions.
In particular, those used by the Response Generation agent are the following: age,
skill level (ability in the use of the application, which basically depends on the
number of calls), knowledge of the domain (which is deduced with secondary
inferences from the user’s school level, job and kind of question), satisfaction
(inferred from the lack of complaints and problems during the previous interactions),
cost (it is related with the time subtracted to others calls and it is also a monetary cost
if the number is free for the caller. It is estimated on the basis of the number of calls
to the Call Center in a period of time).

A customer DB stores the basic data, typically coming from the operational systems
of the bank, and data regarding calls. The user model is processed starting from these
data and according with the above-sketched criteria. The dimensions depending on
calls are updated at each interaction, while the others are updated periodically.

PERSONALIZATION. According to usability researches [2], and the preliminary tests we
performed on the use of IVR vs. traditional call centers, we concluded that the typical
forms human operators modify their dialog with the caller are varying the detail level
of explanation, adding some hints, and changing the way and the style of utterance.
We used the following techniques [4] to reproduce this behavior:

- Variation of the amount of information (text addition or removal): if the caller is a
  novice in the use of the application, we add some hints to the system prompts (in
  the form of answer’s examples), hypnotizing that (s)he probably needs help.

- Variation of the type of the information provided (changing a part of text
  according to the user’s features):
  a) prompt variant, we produce different versions of the same prompt on the basis
  of user’s cost (if high, the system answers are shorter to save time), of user’s
  satisfaction (if low, the system answers are longer as they contain incentives) and
  of user’s experience (if the user is very expert, we replace generic prompts with
  her/his last request to advance her/his needs and avoid to annoy her/him);

---

2 In the figure it corresponds to some modules inside the Response Generation Agent
b) adaptive natural-language generation: we formulate phrases with a degree of complexity which varies on the basis of the user’s levels of knowledge of domain;

- **Style** variation: generation of natural-language sentences characterized by different levels of formalism according to the user’s age.

### 4 Adaptive Routing

The switch of the call from the IVR to the operator (when it is necessary) has to be experienced as a soft and natural transition, namely as an additional service and not as a deficiency. For this reason it is important to route the call toward an operator which fits the user features and keeps on a registry of interaction homogeneous with the previous one.

Common Call Routing systems route the calls following techniques like FIFO (arrival order) or considering some kinds of priority on groups of operators; instead, the requirement for an Adaptive Routing System is that of selecting the operator that is more adequate for answering a specific call of a specific customer. To implement that, we integrated the routing system with an *agent* which, starting from caller’s features, and taking into account the models of all the operators, selects the operator that is best fitting and is idle at that time. The criteria for the match are defined in a set of business rules, stored in a KB (see the *Knowledge Base* – KB₂ -, in the figure).

**USER MODELING.** As explained, in this second step of the answering process, the models taken into account and matched are those of the caller and of the operators. As regards the first one, we have already briefly described its structure and here we just mention the specific features taken into account by the Adaptive Routing agent: value (customer’s profitability for the bank, estimated from data such as account balance, volume of operations, etc.), risk of churn (probability that the customer closes her/his account/s), region of birth, knowledge, satisfaction, age, gender. The first two dimensions of the model are processed according to some formulas provided by experts (the marketing CEO, in our prototype). As regards the operator, the system stores a model for each one in the Call Center and updates their data periodically through evaluation tests. The main dimensions of the operator models are: skill, knowledge, area of birth, communicative ability, expertise, age, gender. As clear, these features rarely change.

**PERSONALIZATION.** A set of business rules specifies the match caller/operator. Some of them are simple correspondences between users’ features (e.g. age, region of birth) with some aggregations of ranges of the operators’ features. E.g. “same area of birth” is due to the fact that people are pleased when the unknown voice which answers is somehow familiar to them (e.g., immediately perceivable elements are dialectal inflection, age and scholarship). Other rules are combined evaluations of utility, which consider, for each caller, his/her economic value, her/his performance required in the response (inferred by the frequency of operations and her/his expertise) and the risk of churn. For the operator, the dimensions taken into account are the skill level, the ability of communication and the rate and speed of well ended answers. Anyway,
in any case, the agent must have domain knowledge higher than the caller and expertise and communication ability adapted to manage risk of churn.

On the basis of these rules, the inference engine dynamically assigns a score to each agents’ characteristic. Consequently, the call is routed to the operator with the higher score if (s)he is idle, otherwise to the one with second higher score and so on.

5 Conclusions

The paper presents a prototype for the adaptive management of calls toward a Call Center system, integrating technologies from different fields. Important applications can be built upon this infrastructure. Just for example we can cite those for disabled people and for CRM purposes. A natural evolution of this system is to apply adaptivity also to the agent’s screen, to supply an adaptive support to manage calls.

Implementation remarks. The goal we pursued was to develop modules which could be integrated into commercial CTIs in an open architecture. For the prototype, we implemented our agents on Cisco platform –Customer Response Application v3-, which is supported by IP networks and runs in a Java environment. Others components are: ICM (Cisco Intelligent Contact Manager), which routes the call; TTS Nuance server, that translates text into voice, ASR Nuance server, which contains the voice recognition engine (based on GSL language) and JESS shell to implement the routing agent.

References