

Особенности построения диалоговых систем в комплексах компьютерной телефонии

Давид Петросян

Институт проблем информатики и автоматизации НАН РА

Аннотация

Обсуждаются возможные принципы и механизмы оптимального построения диалоговых систем компьютерной телефонии.

В любой интерактивной системе компьютерной телефонии (СТТ) связующим звеном между системой и телефонным пользователем являются диалоговые системы, которые служат своеобразным "интерфейсом" между компьютером и телефонным абонентом. Голосовая информация от компьютера передается абоненту в виде:

- заранее подготовленных голосовых файлов (информационные /справочные системы, расписания, реклама, "подсказки" абоненту, запросы системы абоненту на выполнение тех или иных действий).
- "сконструированных" файлов, содержащих также и переменные, быстро меняющиеся данные (например, информация о погоде "... в городе xxxx температура уу градусов ...").
- голосовой почты, которая может содержать кроме принятого голосового сообщения также и дополнительные голосовые файлы "сконструированные" компьютером (например, дату и время поступления сообщения).

Информация от телефонного абонента на компьютер (команды на установку того или другого режима, запросы определенной информации из предложенного меню, ответы на запросы компьютера в виде ожидаемой программой цифры набора (или их комбинации), поставленной в соответствие с возможными вариантами ответа передаются абонентом с помощью тактильного набора телефона в тоновом режиме. В последние годы в системах СТТ начали использовать аппарат распознавания речи.

Специализированные системы распознавания речи, ориентированные на использование в СТТ, еще не стали массовой технологией, хотя перспективы их использования очевидны, учитывая, что подобные системы являются наиболее естественным способом общения человека с компьютером. Возвращаясь к рассмотренной выше другой технологией взаимодействия человек-компьютер, при которой используются голосовые запросы или команды компьютера пользователю и ответные действия телефонного пользователя в виде набора запрошенных данных (или заранее оговоренных кодов-команд)

Построение диалоговой системы, несмотря на кажущуюся простоту, является непростой задачей, т.к. при этом в полной мере должна быть учтена специфика взаимодействия "человек-машина".

Приведем основные положения, которые должны учитываться при проектировании диалоговых систем компьютерной телефонии.

- Система диалога должна быть предельно лаконичной. "Болтливые" диалоги вызывают у пользователя раздражение и, как следствие, неприятие системы в целом.

- Вопросы, задаваемые системой пользователю, должны предполагать однозначный ответ ("да", "нет", цифры).

- В начальной фазе диалога желательна подсказка ответных действий пользователя (...при ответе "да" - нажмите кнопку "1", при ответе "нет" - кнопку "3").

- Кнопки, наиболее часто используемые программой, должны легко запоминаться пользователем (например, "да"-кнопка "1", "нет" - кнопка "3", но никак не "да"-кнопка 2, "нет"-кнопка "7").

- Назначаемые программой диалога кнопки, соответствующие тем или иным функциям, должны располагаться на клавиатуре телефонного аппарата "тематически" (например, верхний ряд кнопок телефона - "1", "да", "2" - повтор запроса, "3" - "нет").

Второй ряд кнопок - управление функциями почтового ящика (в случае системы голосовой почты) - кнопка "4" - "листать голосовые письма в почтовом ящике вперед", кнопка "5" - воспроизвести письмо повторно, кнопка "6" - листать письма назад" и т.д.

- В особо ответственных случаях последовательность запросов системы должна быть построена таким образом, чтобы система имела возможность логического контроля возможных ошибок при ответах пользователя. Приведем условный фрагмент возможного диалога, иллюстрирующего приведенное выше положение.

```
система > "удалить данное сообщение ?"
пользователь < "да"
система > "сохранить данное сообщение ?"
пользователь < "да" (?)
```

При вводе пользователем цифровой информации система в ответственных случаях может запросить повторно ввести требуемую информацию (например, # телефона)

Очевидно, что в обоих случаях система должна производить анализ введенных данных. В первом случае недостоверным является два взаимоисключающих ответа, во втором - несовпадение введенных номеров телефонов. При запросе цифровой информации заранее известного формата, система должна анализировать соответствие введенных данных ожидаемому формату (например, при запросе идентификационного кода (или # телефона) известна разрядность вводимого кода (# телефона).

При построении диалоговых систем следует учитывать "человеческий фактор", вводя программные механизмы "защиты от дурака" и "защиты от зависания системы". (например, при работе с информационными файлами удаление информации должно задаваться усложненным кодом (* + X) или в два этапа диалога, как это делается в компьютерных программах - ".. Вы уверены, что хотите удалить этот файл?").

При ответе на запрос системы абонент может прервать ввод данных, оставаясь на связи с системой (возможны также сбои или помехи в телефонном канале). Должны быть предусмотрены механизмы контроля времени ожидания ответа на запрос и времени исполнения запроса. Если реакция абонента на запрос системы отсутствует в течении заданного времени - сеанс связи прерывается. При заведомо неправильном ответе на запрос система может повторно запросить данные; при повторном неправильном ответе - прервать сеанс связи.

Большое значение при построении диалога имеет "дружественность" голосовых запросов системы.

Приведем некоторые рекомендации фирмы Berkley Production по построению дружественного голосового интерфейса. "Выбор наилучшей системы голосового интерфейса не заканчивается выбором аппаратно-программного решения. Исследователи обнаружили, что 84% впечатлений о системе формируются ощущением от голоса, которым озвучиваются голосовые сообщения.

- Избегайте слишком громкого или слишком тихого голоса. Если голос для абонента звучит на повышенных тонах или, наоборот, слишком мягко, это может быть и проблемой оборудования.

Но наиболее вероятно, что диктор не умеет модулировать оттенки своего голоса и говорить с интонацией, приятной для своего собеседника.

- Убедитесь, что Ваш диктор улыбается, когда говорит. Это слышно по голосу.
- Не переусердствуйте в четкости. В попытке быть ясными некоторые дикторы забывают о плавности и естественности речи.
- Говорите в размеренном темпе. Звонящих раздражает слишком медленная речь. С другой стороны, абонент может упустить нужную информацию, если диктор говорит слишком быстро.
- Звонящие предпочитают краткие меню. Устраните лишние слова типа "для", "пожалуйста".
- Имейте в виду, что хотя на телефоне имеется целых 12 управляющих клавиш, это не значит, что Вы должны использовать их все одновременно ! "

В отличие от привычных графических средств "общения" с компьютером, в которых реакция компьютера на действия человека или выдаваемая ею информация выводится на экран монитора и пользователь, при необходимости, может вновь прочитать и осмыслить переданное сообщение, голосовая форма представления информации является "одноразовой". Очевидно, что при проектировании голосовых диалоговых и информационных систем должны предусматриваться механизмы, позволяющие пользователю "переспросить" непонятый фрагмент переданной ему голосовой информации или запроса компьютера.

Из этого следует, что голосовые информационные материалы и запросы диалога должны быть представлены в виде последовательности логически завершенных фрагментов, а система управления диалогом (или информационной системой) должна допускать прерывание пользователем передаваемой ему информации с повторным воспроизведением прослушанного фрагмента. Сформулируем приведенные выше соображения в виде следующих положений:

- При построении информационных, информационно-справочных систем компьютерной телефонии, предоставляющих по запросу информационные блоки необходимо структуризация массива передаваемых голосовых данных в виде последовательности логически законченных фрагментов.

- При разработке программ управления диалогом необходимо предусмотреть возможность прерывания текущего воспроизводимого фрагмента и его повторного воспроизведения.

- Те же принципы должны быть использованы при голосовых запросах системы, т.е. должна быть реализована функция "повтор"

Литература

[1] <http://www.comptek.ru/telephony/board/recognition.html>

[2] <http://www.comptek.ru/telephony>

**Կոմպյուտերային հեռախոսակապի համալիրներում երկխոսակցական
համակարգերի կառուցման առանձնահատկությունները**

Դ. Պետրոսյան

Ամփոփում

Զննարկվում են կոմպյուտերային հեռախոսակապի երկխոսակցական համակարգերի օպտիմալ կառուցման հնարավոր սկզբունքները և կառուցվածքները: