Microsoft .NET Speech SDK 1.0 Beta 2:
New Features and Enhancements

Fact Sheet

October 2002

Overview

The Microsoft® .NET Speech SDK is a set of ASP.NET controls, a Microsoft Internet Explorer Speech Add-in, sample applications and documentation that allows Web developers to create, debug and deploy speech-enabled ASP.NET applications. The tools are integrated seamlessly into Microsoft Visual Studio® .NET, allowing developers to leverage the familiar development environment.

Microsoft Corp.'s approach to speech-enabled Web applications is built around a newly emerging standard: Speech Application Language Tags (SALT). SALT has been submitted to the World Wide Web Consortium (W3C) for adoption as a standard for telephony and multimodal applications, which incorporate speech-enabled elements within a visual Web interface.

The beta 2 release of the Microsoft .NET Speech SDK includes a complete tool set for creating and testing SALT-based voice-only telephony applications. It also supports the development of multimodal speech applications on clients such as desktop PCs or Tablet PCs using Internet Explorer browser software.

This fact sheet provides a brief overview of the enhancements built into the .NET Speech SDK 1.0 Beta 2.

Beta 2 Enhancements

The Microsoft .NET Speech SDK consists of ASP.NET controls, a Speech Add-in for Internet Explorer, and numerous libraries and sample applications. In addition, it includes four development tools that integrate natively with Visual Studio .NET. All these elements have been enhanced in the beta 2 release.

ASP.NET Speech Controls

Speech controls are ASP.NET controls that render SALT in a voice-enabled Web application. An add-on to Visual Studio, the Microsoft .NET Speech SDK allows developers to drag and drop these companion ASP.NET controls from an on-screen toolbox onto a page with traditional ASP.NET Web controls. Major enhancements include the following:

- **Application controls.** Application controls encapsulate commonly used interactions in a single control, making tasks such as obtaining a phone number quick and easy. Beta 2 includes numerous prebuilt application controls, including AlphaDigits (obtains a string of letters and numbers), Currency (obtains a dollar amount), Date, Yes No, and Social Security Number.

- **Telephony controls.** This release of the SDK includes call management, dual tone multifrequency (DTMF) and Simple Messaging Extension (SMEX) controls to interact with the telephony world. Telephony controls help developers build and simulate voice-only applications.

- **Semantic Items.** Semantic Items clearly define what information the application requires from the user and the location where the application puts the recognized results, making it easier to design the flow of a dialog. Semantic Items give developers greater control over the binding of spoken answers to the user interface elements of the page, and promote a clean separation of the speech user interface and the visual user interface.

- **Style sheet control.** This new control allows developers to apply global and control-specific settings to an application. For example, it's possible with this control to allow the user to interrupt a long prompt with a spoken command — this is the "barge-in" feature. Using the style sheet control, a developer can configure bargein across the application.

Speech Add-In for Internet Explorer

The SDK comes with an add-in for Internet Explorer that allows the browser to recognize speech tags and execute speech-enabled Web applications. Major enhancements include the following:

- **SALT 1.0 support.** The Microsoft .NET Speech SDK supports SALT 1.0. SALT extends HTML and other markup languages with tags and scriptable objects to perform voice output, spoken-language input, telephony management and messaging. With the new speech add-in for Internet Explorer, developers can type SALT code directly into the browser.

- **Audio meter.** In multimodal applications, a voice meter similar to the display on a stereo equalizer is used to display the audio input. This feature gives users visual verification that they're being heard. The audio meter was an unsupported feature in beta 1.

- **Support for Tablet PC.** With beta 2, the stand-alone Internet Explorer Speech Add-in can be installed on the Tablet PC. This allows developers to build and run .NET Speech applications for the Tablet PC platform.

Libraries and Samples

The .NET Speech SDK includes grammar libraries as well as sample applications that demonstrate the capabilities of the technology:

- **Grammar libraries.** A set of robust grammar libraries, written in the W3C-approved Speech Recognition Grammar Specification (SRGS) format, is included in the SDK, which helps developers obtain abstract, complex concepts from the user. For example, gathering recognizable date or time input is quite complex. If the user says "Thanksgiving" or "Monday the 5th," applications can use grammar libraries to make an intelligent determination of the exact date and year.

- **Samples.** This release of the .NET Speech SDK includes seven new samples, which correspond with the application controls. A mix of multimodal and voice-only samples are included, such as AlphaDigits, Zip Code and Social Security Number for multimodal applications, and SingleItemChoose and YesNo for voice-only applications.

- **New demonstration application.** The SDK includes a finished application called Contacts, which is a valuable user scenario. This voice-only application allows a caller to say a person's name and obtain information such as an e-mail address, phone number or office address from a database of contacts. It's a highly flexible and configurable application, allowing developers to remove or add services within the application.

Speech Controls Editor

The Speech Controls Editor is the primary development environment for building the components of a speech-enabled application. Major enhancements include the following:

- **Redesigned property builders.** The property builders have been completely redesigned to present the most common properties of speech controls in an organized, intuitive format that helps guide the developer through the configuration process. They include links to the Grammar Editor, Prompt Editor and Prompt Function Editor. When a control is dragged from the control toolbox and dropped onto the ASP.NET canvas, the property builder is accessed by right-clicking the control.

- **Support for new speech controls.** The Speech Controls Editor in beta 2 includes full design support — complete with property builders — for the new speech controls, including Validator controls, Semantic Map/Item controls, Application controls and SMEX controls. Developers can now use these helpful controls in their applications as easily as Question and Answer and Command controls.

Grammar Editor

The Grammar Editor provides a graphical view of the grammar incorporated into an application, breaking it down into different paths and displaying them visually. Major enhancements include the following:

- **W3C grammar support.** In beta 1 of the SDK, Microsoft implemented its own grammar format. With this release, the Grammar Editor now opens and saves files only in the W3C-approved SRGS format. All elements, properties and validations have been updated for SRGS.

- **Semantic authoring.** The editing tools now support the W3C specification for adding semantic information to grammar. To simplify this process, two semantic tool windows have been added. One window allows simple assignments for properties and values in one pane, and a script window for more complicated variable setting in another pane. The second window provides an easy interface for monitoring semantic properties.

Prompt Editor

The Prompt Editor is used to record and manage the audio prompts played by the application. It features the following enhancements:

- **Enhanced prompt editing.** Prompt-editing enhancements make it easier to create dynamic prompts. When an application asks for information, the user might mumble, remain silent or say "help." Under these conditions, an application should provide different prompts, such as "I'm sorry, I didn't understand you" for the first error and then, perhaps, "My fault again" when an error happens again. The enhanced Prompt Editor creates a small script block that defines the parameters for each prompt in such strings, and makes them easier to build. In addition, a new file extension, PF, has been created for prompt-function files, with a full editor. Prompt functions in this file type are easier to integrate with speech controls, and allow for more complete and accurate validation.

- **Prompt validation.** Through the Prompt Editor, the validation tool reviews all prompts in an application and indicates which have been recorded and which still need to be completed.

Speech Debugging Console

The Speech Debugging Console allows developers to run through an application's dialogs and test its ability to recognize and act on spoken input. Improvements include the following:

- **Improved telephony simulation.** Enhanced functionality in the Speech Debugging Console enables telephony simulation. It provides developers with an on-screen telephone keypad to simulate telephone input. It also can simulate call transfers and hang-ups to test call management controls.

- **Visible server-side information.** In beta 1, the Speech Debugging Console relayed information purely in terms of the SALT rendered on the client. In beta 2, the console provides information in terms of server-side speech controls. This helps the author relate client behavior directly back to authoring decisions and pinpoint which controls to change on the ASPX page.

- **Replay from file.** This feature allows developers to record a series of interactions with a speech application, and run these log files again through the application and Speech Debugging Console. This functionality allows the developer to fast-forward to specific points in a dialog and to run automated tests on speech applications.

<p align="center">*#########*</p>

For more information, press only:

Rapid Response Team, Waggener Edstrom, (503) 443-7070, rrt@wagged.com