

Ginga-NCL com objetos de mídia SSML embutidos

Relatório Técnico: Requisitos

Rafael Diniz
Matrícula: 1312398
5 de agosto de 2014

Sumário

1	Introdução	2
1.1	Propósito	2
1.2	Público Alvo	2
1.3	Escopo	2
1.4	Definições, Acrônimos e Abreviações	2
1.5	Referências	2
1.6	Visão geral do documento	2
2	Descrição do projeto	2
3	Descrição dos usuários	3
4	Requisitos Funcionais	3
4.1	Suporte à sintetização de objetos de mídia SSML embutidos em código NCL	3
4.2	Suporte à execução de objetos de mídia de texto puro ou XML mal formado	3
4.3	Sintetizar o objeto de mídia SSML mesmo quando o atributo <i>type</i> não estiver declarado na mídia	3
4.4	Suporte à definição de âncoras de conteúdo no objetos de mídia	3
4.5	Suporte às propriedades de mídias de áudio pelo reprodutor SSML	4
4.6	Suporte à propriedade que ative a redução automática do áudio principal (Talk-Over)	4
4.7	Suporte à redefinição da língua e velocidade da fala através de propriedade da mídia	4
5	Requisitos Não-Funcionais	4
5.1	Baixo consumo de CPU e memória de forma a permitir a sintetização de documentos SSML em sistemas embarcados	5

1 Introdução

1.1 Propósito

Este documento especifica os requisitos do projeto final de programação no âmbito da disciplina INF2102. O projeto final de programação consiste da adição de nova funcionalidade ao middleware Ginga na sua versão em C++, tema escolhido devido à relação com o assunto da dissertação de mestrado do autor.

1.2 Público Alvo

Este documento se destina aos arquitetos de software, engenheiros de software e testadores interessados no sistema, assim como ao professor avaliador da disciplina.

1.3 Escopo

Este documento lista os requisitos funcionais e não-funcionais do Ginga-NCL com objetos SSML embutidos.

1.4 Definições, Acrônimos e Abreviações

Ginga	Middleware padrão do Sistema Brasileiro de TV Digital
Ginga-NCL	Subsistema Declarativo do middleware Ginga
NCL	Nested Context Language
XML	eXtensible Markup Language
SSML	Speech Synthesis Markup Language

1.5 Referências

- [1] Soares, L. F. G., Rodrigues, R. F., Moreno, M. F. (2007). Ginga-NCL: the declarative environment of the Brazilian digital TV system. *Journal of the Brazilian Computer Society*, 12(4), 37-46.
- [2] Soares, L. F. G. S. (2009). Programando em NCL 3.0: desenvolvimento de aplicações para middleware Ginga: TV digital e Web. Elsevier.
- [3] Garrido, J., Bofias, E., Laplaza, Y., Marquina, M., Aylett, M., & Pidcock, C. (2008). The Cerevoice speech synthesiser. *Actas de las V Jornadas de Tecnología del Habla (Bilbao)*.
- [4] W3C Recommendation (2010). Speech Synthesis Markup Language (SSML) Version 1.1.

1.6 Visão geral do documento

A seção 2 apresenta uma descrição geral do projeto, seu escopo, objetivos e usuários. A seção 3 lista os requisitos funcionais e na seção 4 são apresentados os requisitos não-funcionais do sistema.

2 Descrição do projeto

O software desenvolvido no escopo deste trabalho estende a implementação do middleware Ginga-NCL para adicionar suporte ao tipo de mídia SSML em sua versão 1.1 assim como definido pela W3C. A mídia SSML será controlada no ambiente declarativo provido pela linguagem NCL. O código fonte original do Ginga está disponível no endereço <http://git.telemidia.puc-rio.br>. O framework para sintetização de voz utilizado no projeto é o CereVoice¹, que provê suporte quase completo da norma SSML, incluído

¹<https://www.cereproc.com>

chamada de função de *callback* no momento em que a sintetização do documento passa por uma âncora declarada no documento SSML.

A adição desse novo recurso ao middleware Ginga é parte do trabalho de mestrado do autor, cujo tema é a definição de um middleware para uso no Sistema Brasileiro de Rádio Digital.

O foco do suporte à sintetização de voz no Ginga é seu uso no rádio digital, de modo a permitir uma interação vocal da aplicação com o ouvinte de rádio, no entanto seu uso se estende também à TV digital e Web.

3 Descrição dos usuários

Os usuários do Ginga-NCL com suporte à mídias SSML embarcadas serão os desenvolvedores de aplicações multimídia interativas, produtores de conteúdo para rádio e TV digitais e desenvolvedores Web. Um interesse especial da tecnologia é a autoria de aplicações que não dependam somente da interação visual da aplicação, como no caso em que o usuário não pode ficar olhando para a tela, por exemplo, quando está dirigindo, ou no caso em que o usuário não vê a tela devido à restrições de visão.

Os usuários dos serviços de rádio, TV ou Web que venham a executar as aplicações NCL com mídias SSML se beneficiarão de uma interatividade vocal ampliada no Ginga.

4 Requisitos Funcionais

Nesta seção são descritos os requisitos funcionais do sistema, separados nas subseções a seguir.

4.1 Suporte à sintetização de objetos de mídia SSML embutidos em código NCL

Um documento NCL poderá referenciar uma mídia SSML. O middleware deverá ser capaz de corretamente sintetizar o texto presente no documento, respeitando as definições da norma SSML e as regras de sincronismo temporal do NCL.

Exemplo de como um documento NCL deve referenciar uma mídia SSML:

```
<media id="ssml1" src="fala.ssml" type="application/ssml+xml"/>
```

4.2 Suporte à execução de objetos de mídia de texto puro ou XML mal formado

Uma mídia SSML deverá ser reproduzida mesmo quando houver elementos XML não conformes à norma ou quando o documento não for um XML válido. Nesses casos o texto presente no documento, excluindo-se as *tags* XML, deverá ser sintetizado.

Nos casos em que a língua não estiver explicitada no documento ou não puder ser inferida, o middleware deve assumir a língua definida como padrão do sistema.

4.3 Sintetizar o objeto de mídia SSML mesmo quando o atributo *type* não estiver declarado na mídia

Na linguagem NCL, o atributo *type* não é obrigatório, visto que é possível deduzir qual o tipo da mídia a partir da extensão do nome do arquivo indicado no atributo *src*. Assim sendo, o sistema aqui desenvolvido deve estar apto a deduzir o tipo de mídia SSML apenas a partir do atributo *src* das mídias.

Caso o atributo *type* seja especificado, ele tem precedência na escolha do *player* que deve reproduzir aquela mídia. Por exemplo, um arquivo de texto *.txt* pode ser reproduzido pelo tocador SSML atribuindo-se o valor "application/ssml+xml" ao atributo *type*, permitindo o reúso da mídia de texto para apresentação tanto visual como em forma de áudio sintetizado.

4.4 Suporte à definição de âncoras de conteúdo no objetos de mídia

Deverá ser possível definir âncoras aos objetos de mídia SSML que indiquem o momento em que determinada marcação (elemento *<mark>*, atributo *name*) no documento SSML é atingida durante a sintetização. Dessa forma é possível sincronizar ações *via* NCL no momento em que determinada parte

do texto que está sendo sintetizado é atingida. A âncora deve ser especificada utilizando o elemento *area* da NCL da seguinte forma:

```
<media id="ssml1" src="fala.ssml" type="application/ssml+xml">
  <area id="ssmlmark1" label="mark1"/>
</media>
```

Onde o *label mark1* é igual ao atributo *name* de um elemento *<mark>* definido em um documento SSML.

4.5 Suporte às propriedades de mídias de áudio pelo reprodutor SSML

Além das propriedades comuns à todos os objetos de mídia do NCL, as seguintes propriedades para mídias de áudio do Ginga serão suportadas pelo *player* de mídias SSML: *soundLevel*, *balanceLevel*, *trebleLevel*, *bassLevel* e *freeze*.

4.6 Suporte à propriedade que ative a redução automática do áudio principal (Talk-Over)

No contexto do rádio e TV digitais, no qual existe sempre um fluxo de áudio principal sendo reproduzido, uma propriedade da mídia pode ser especificada, de nome *talkOverReduction*. O valor dessa propriedade é uma porcentagem do volume do áudio principal, indicando que o volume do áudio principal será reduzido para a porcentagem indicada. Exemplo de redução do áudio principal para 50%:

```
<media id="ssml1" src="fala.ssml" type="application/ssml+xml">
  <property name="talkOverReduction" value="50%"/>
</media>
```

4.7 Suporte à redefinição da língua e velocidade da fala através de propriedade da mídia

Serão adicionadas novas propriedades de mídia que definem ou redefinem, no caso do documento SSML especificar uma língua, a língua e velocidade da fala. Abaixo um exemplo que especifica a língua para português brasileiro e a velocidade da fala para 60% do normal.

```
<media id="ssml1" src="fala.ssml" type="application/ssml+xml">
  <property name="speechLanguage" value="pt-br"/>
  <property name="speechRate" value="60%"/>
</media>
```

5 Requisitos Não-Funcionais

Dentre os requisitos não funcionais do sistema estão:

- Segurança: Associada à integridade dos dados, privacidade, como o sistema trata de informação confidencial, liberação de acesso aos usuários do sistema.
- Performance: Relacionada ao tempo de resposta do sistema durante o uso dos recursos disponibilizados.
- Usabilidade: Associada à facilidade de uso do sistema.
- Confiabilidade: Deve-se buscar baixa frequência de falhas e robustez do sistema na recuperação destas falhas.
- Padrões: Conformidade com os padrões e normas a serem seguidas no desenvolvimento do sistema.

A subseção a seguir descreve o requisito não-funcional específico ao recurso introduzido ao Ginga.

5.1 Baixo consumo de CPU e memória de forma a permitir a sintetização de documentos SSML em sistemas embarcados

É um pré-requisito para um *player* NCL que ele possa ser embarcado em receptores de TV e rádio digitais. Para isso um baixo consumo de CPU e memória deve ser levado em conta, visto que essas plataformas possuem recursos computacionais limitados.