



Manuale Utente

Informazioni Documento

Versione	1.0.0
Redattori	Marco Focchiatti, Giulio Rossetti, Kevin Silvestri, Matteo Rizzo
Verificatori	Manfredi Smaniotto, Samuele Modena
Distribuzione	Prof. Tullio Vardanega Prof. Riccardo Cardin Mivoq S.R.L. Gruppo Graphite
Uso	Esterno
Recapito	graphite.swe@gmail.com



Registro delle modifiche

Versione	Data	Autore	Ruolo	Descrizione
0.2.0	2018-04-12	Samuele Modena	Verificatore	Verifica §5 e §A
0.1.2	2018-04-11	Matteo Rizzo	Amministratore	Stesura §A
0.1.1	2018-04-11	Kevin Silvestri	Progettista	Stesura §5
0.1.0	2018-04-10	Manfredi Smaniotto	Verificatore	Verifica da §1 a §4
0.0.5	2018-04-09	Marco Focchiatti	Programmatore	Stesura §4
0.0.4	2018-04-09	Giulio Rossetti	Programmatore	Stesura §3
0.0.3	2018-04-08	Kevin Silvestri	Programmatore	Stesura §2
0.0.2	2018-04-08	Matteo Rizzo	Amministratore	Stesura §1
0.0.1	2018-04-08	Matteo Rizzo	Amministratore	Creata struttura documento



Indice

1	Introduzione	4
1.1	Scopo del documento	4
1.2	Scopo del prodotto	4
1.3	Informazioni utili	4
1.4	Riferimenti informativi	5
2	Requisiti di sistema	6
2.1	Requisiti hardware	6
2.2	Requisiti software	6
3	Installazione e configurazione	8
4	Guida all'utilizzo	9
4.1	Struttura dell'interfaccia grafica	9
4.1.1	Menù dell'applicazione	10
4.1.1.1	File	11
4.1.1.2	Help	12
4.1.2	Pannello di configurazione	13
4.1.3	Pannello degli utterance processor	15
4.1.4	Pannello delle relations	15
4.1.5	Area del grafo	16
4.1.6	Proprietà del nodo	17
4.1.7	Log errori	17
4.2	Interagire con la voce	17
4.2.1	Caricare la voce	17
4.2.2	Salvare l'audio relativo alla voce	18
4.3	Visualizzare il grafo	19
4.3.1	Visualizzare il grafo step-by-step	21
4.3.2	Visualizzare l'intero grafo	21
4.4	Interagire con il grafo	22
4.4.1	Selezionare gli utterance processor	22



4.4.2	Interagire con le relation	22
4.4.3	Traslare elementi grafici	22
4.4.3.1	Traslare nodi	22
4.4.3.2	Traslare archi	23
5	Risoluzione dei problemi	24
5.1	Errori in DeSpeect	24
5.1.1	Struttura dei codici di errore	24
5.2	Problemi con il reperimento di Speect	25
5.3	Segnalazione di bug	25



1. Introduzione

1.1 Scopo del documento

Il documento ha la finalità di illustrare, a coloro che volessero interfacciarsi con l'applicazione "*DeSpeect: un'interfaccia grafica per Speect*", i requisiti necessari per poterlo utilizzare e le modalità di installazione e di utilizzo. Nonostante la versione attuale rappresenti una prima bozza del documento, una volta concluso esso rappresenterà sia una guida che un riferimento completo per l'utilizzo del prodotto da parte di un utente.

1.2 Scopo del prodotto

Lo scopo del progetto è la realizzazione di un'interfaccia grafica per *Speect_G* [Meraka Institute(2008-2013)], una libreria per la creazione di sistemi di sintesi vocale, che agevoli l'ispezione del suo stato interno durante il funzionamento e la scrittura di test per le sue funzionalità.

1.3 Informazioni utili

La stesura di questo documento assume come utente target del prodotto un programmatore esperto nell'utilizzo di *Speect*.

Per completezza, viene riportato in appendice §A un glossario comprensivo di termini tecnici o riguardanti particolari funzionalità di *DeSpeect*. Per identificare i termini presenti nel glossario, la loro prima occorrenza all'interno del documento è riportata in corsivo e marcata con una G al pedice.



1.4 Riferimenti informativi

- **Documentazione Speect:**

<http://speect.sourceforge.net/contents.html>;

Documentazione ufficiale della libreria di *Text-To-Speech* di riferimento per il progetto.

- **Documentazione Qt:**

<http://doc.qt.io/>;

Documentazione ufficiale del *framework_G* utilizzato per lo sviluppo dell'interfaccia grafica.

- **Documentazione CMake:**

<https://cmake.org/documentation/>.

Documentazione ufficiale del framework utilizzato per la build del prodotto.



2. Requisiti di sistema

2.1 Requisiti hardware

L'installazione ed esecuzione del software DeSpeect richiede una macchina con i seguenti requisiti:

Hardware	Requisiti minimi	Requisiti consigliati
CPU	2 GHz dual core x86	3GHz quad core x64
RAM	4 GB	8GB
Spazio libero su disco	256MB	1GB
GPU	accelerazione video 3D risoluzione 1280x1024 256 MiB di memoria	accelerazione video 3D risoluzione 1920x1080 512 MiB di memoria

2.2 Requisiti software

L'installazione ed esecuzione del software DeSpeect richiede i seguenti programmi preinstallati:

- Sistema operativo Unix / Unix-like (il software è stato testato solo per piattaforma Ubuntu 16.04 LTS)
<https://www.ubuntu.com/download/desktop>
- CMake (versione minima 2.8)
<https://cmake.org/download/>
- Compilatore ANSI C/ISO C90 *GCC*_G (versione minima 5.0)
<https://gcc.gnu.org/install/binaries.html>



- Qt_G 5.9.0
<https://www.qt.io/download>
- Git
<https://git-scm.com/>
- Curl
<https://curl.haxx.se/>
- Swig
<http://www.swig.org/>
- libxml2-dev
<https://packages.debian.org/stretch/libxml2-dev>
- python-dev
<https://pypi.python.org/pypi/dev/0.4.0>



3. Installazione e configurazione

DeSpeect è reperibile su GitHub al seguente link:

<https://github.com/graphiteSWE/DeSpeect>

Una volta soddisfatti i prerequisiti descritti in §2 "Requisiti di sistema" di questo documento, per installare ed eseguire il software è necessario seguire la seguente procedura:

1. Clonare o scaricare il repository sulla propria macchina;
2. Entrare nella cartella clonata/scaricata ed eseguire lo script `Installer.sh` presente al suo interno;
3. Selezionare la cartella di destinazione per l'installazione: lo script installerà Speect ed effettuerà una build del progetto all'interno di tale directory;
4. Per eseguire DeSpeect, entrare nella directory di installazione ed avviare l'eseguibile `DeSpeect`.



4. Guida all'utilizzo

Viene di seguito illustrata la guida all'utilizzo del software DeSpect.
Si fa notare che i termini evidenziati in *corsivo* corrispondono a pulsanti specifici presenti all'interno dell'applicazione.

4.1 Struttura dell'interfaccia grafica

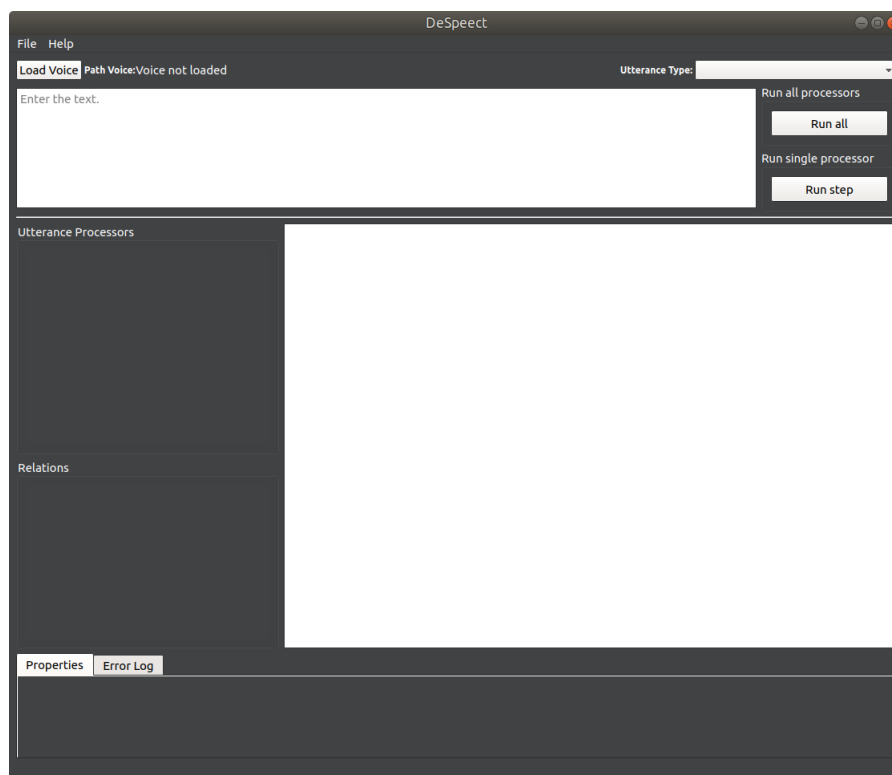


Figura 4.1: Interfaccia grafica - Schermata iniziale



All'avvio dell'applicazione viene presentata la schermata riportata in Figura 4.1. L'interfaccia grafica è costituita dalle componenti illustrate nelle sezioni seguenti.

4.1.1 Menù dell'applicazione

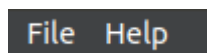


Figura 4.2: Menù dell'applicazione - Visione generale

Menù situato nella parte superiore della schermata. Tramite il menù *File* è possibile interagire con alcune funzionalità offerte dal sistema, mentre tramite il menù *Help* è possibile visualizzare manuale utente e licenza del prodotto (vedi Figura 4.2). Il menù è sempre disponibile in qualunque posizione all'interno dell'applicazione e al suo interno è possibile selezionare le seguenti voci:



4.1.1.1 File

- **Load Voice JSon:** carica il *file Voice_G JSon_G*;
- **Save Audio file:** salva il file audio prodotto in seguito all'esecuzione di *Speect*;
- **Search path:** cerca un nodo inserendo come input un percorso specifico (es. Words.parent.n);
- **Exit:** chiude l'interfaccia *DeSpeect*.

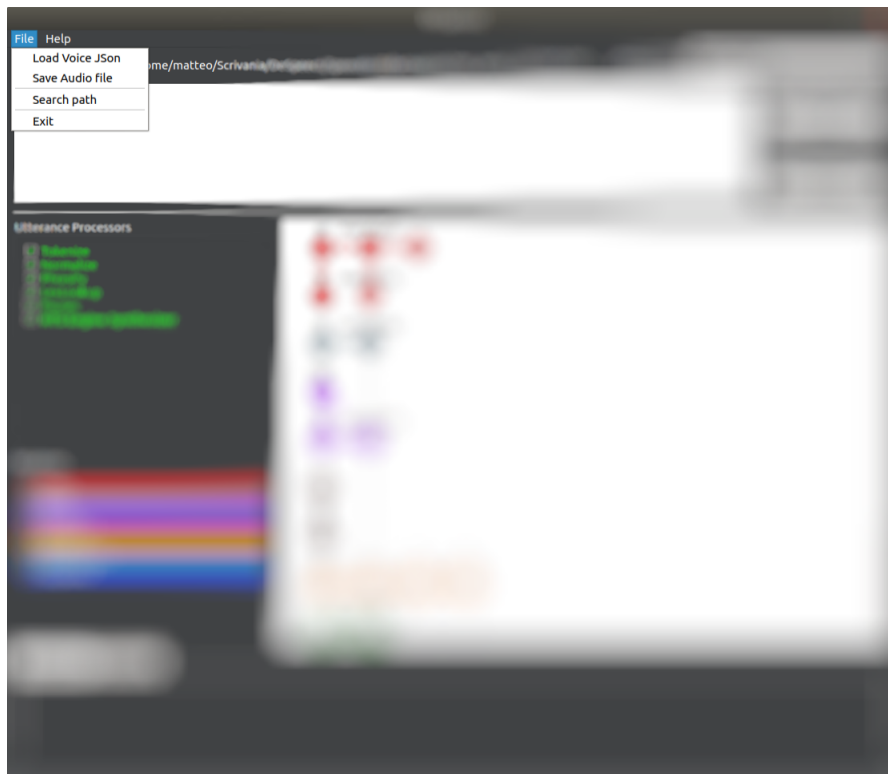


Figura 4.3: Menù dell'applicazione - Sezione "File"



4.1.1.2 Help

- **Manual:** apre il manuale utente;
- **Licence:** visualizza la licenza di *DeSpeech*.

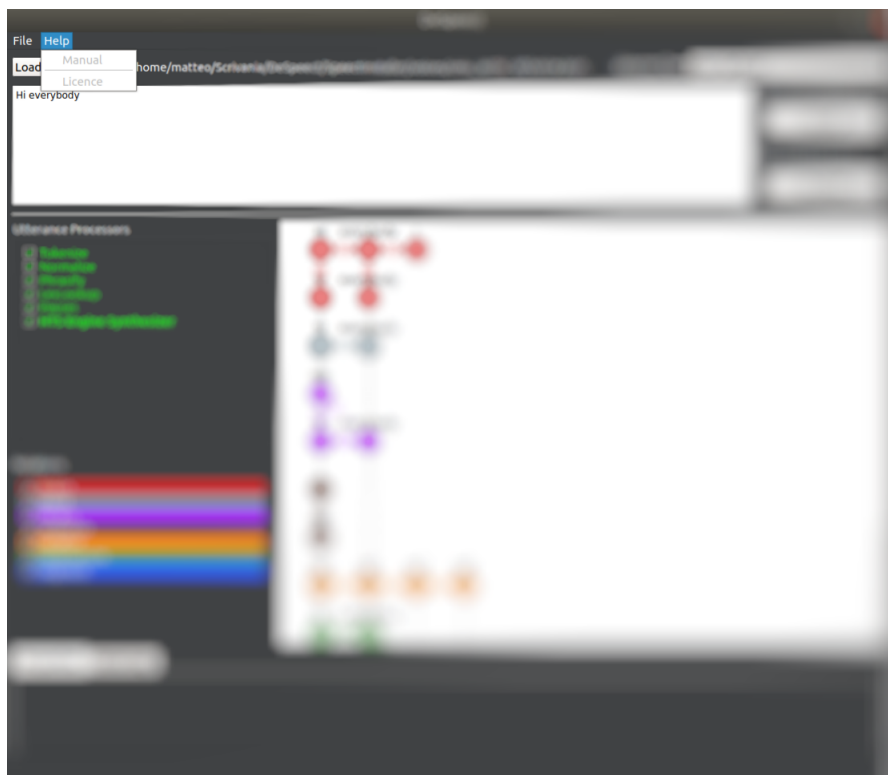


Figura 4.4: Menù dell'applicazione - Sezione "Help"



4.1.2 Pannello di configurazione

Situato nella parte superiore della schermata sotto il menù citato al punto precedente. Qui è possibile caricare un file Voice (tramite pulsante *Load Voice*), inserire input testuale (tramite area di testo dedicata), selezionare l'*utterance_G* type e avviare l'esecuzione di *Speect* eseguendo singolarmente ogni utterance processor (pulsante *Run step*) o tutti assieme in sequenza (pulsante *Run all*).

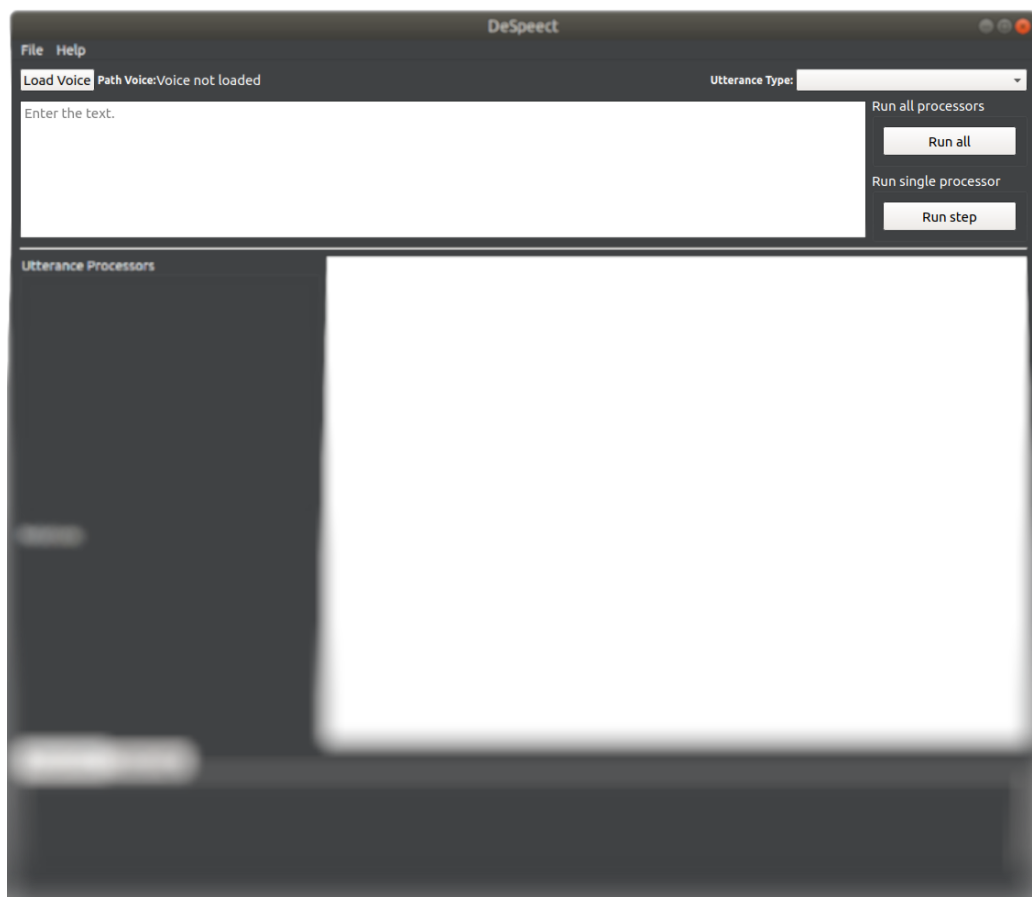


Figura 4.5: Interfaccia grafica - Pannello di configurazione

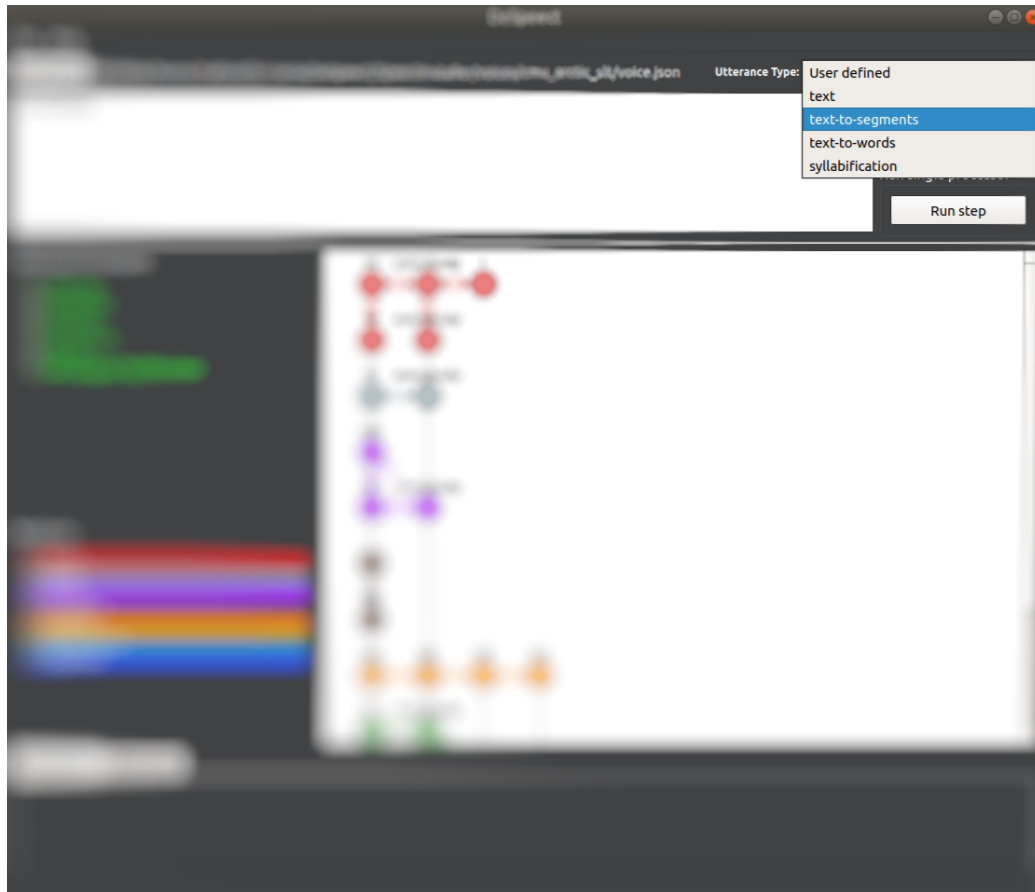


Figura 4.6: Interfaccia grafica - Selezione utterance type



4.1.3 Pannello degli utterance processor

Situato sulla sinistra, sotto al pannello di configurazione. Qui è presente una lista degli utterance processor relativi alla *voice_G* caricata e all'utterance type selezionata, a ognuno dei quali è assegnata una checkbox. Spuntando le checkbox l'utente seleziona gli utterance processor che desidera eseguire per un dato input tramite i pulsanti dedicati presenti nel pannello di configurazione (pulsanti *Run all* e *Run step*). Avviata l'esecuzione, i processor eseguiti verranno evidenziati dal colore verde, mentre quello corrente è riportato in grassetto (vedi Figura 4.4);

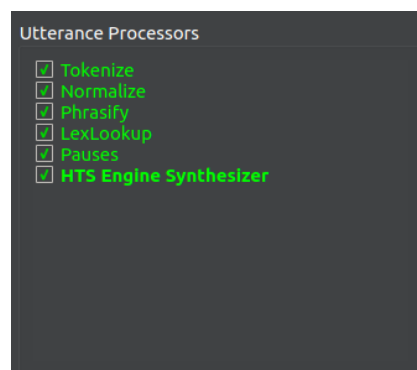


Figura 4.7: Interfaccia grafica - Pannello degli utterance processors

4.1.4 Pannello delle relations

Situato sulla sinistra, sotto al pannello degli utterance processor. Qui è possibile selezionare quali *relation_G* visualizzare nel grafo. Di default, il grafo mostra ogni relation disponibile, rimuovendo la spunta dalla checkbox della relation desiderata le componenti del grafo relative vengono rimosse dalla visualizzazione (vedi Figura 4.5);

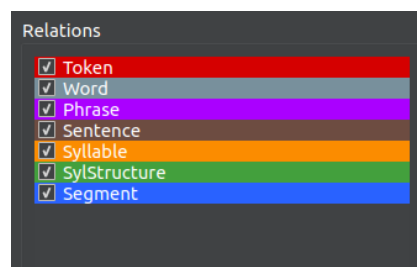


Figura 4.8: Interfaccia grafica - Pannello delle Relations



4.1.5 Area del grafo

Situata sulla destra, sotto al pannello di configurazione. Qui viene visualizzato il grafo relativo ad una data voce, input testuale e configurazione di utterance processor. I nodi sono selezionabili (nel cui caso, le informazioni relative al nodo vengono visualizzate nell'apposita barra approfondita al punto seguente) e spostabili (vedi Figura 4.6);

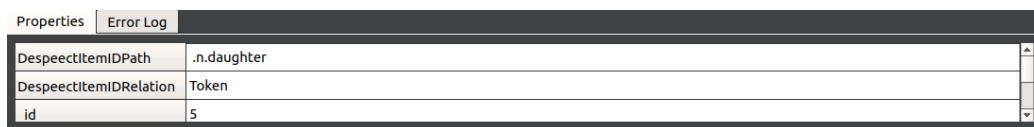


Figura 4.9: Interfaccia grafica - Area del grafo



4.1.6 Proprietà del nodo

Situato nella parte inferiore dell'interfaccia, in corrispondenza del primo tab. Qui vengono visualizzate le informazioni specifiche relative al nodo selezionato (vedi Figura 4.7).

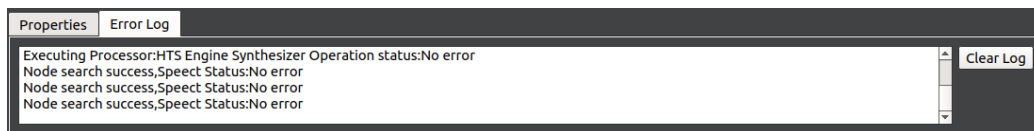


Properties	Error Log
DespectItemIDPath	.n.daughter
DespectItemIDRelation	Token
id	5

Figura 4.10: Interfaccia grafica - Proprietà del nodo

4.1.7 Log errori

Situato nella parte inferiore dell'interfaccia, in corrispondenza del secondo tab. Qui vengono visualizzate le informazioni relative allo stato delle operazioni eseguite, ed in particolare degli errori di Speect (vedi Figura 4.7).



Properties	Error Log
Executing Processor:HTS Engine Synthesizer Operation status:No error	
Node search success,Speect Status:No error	
Node search success,Speect Status:No error	
Node search success,Speect Status:No error	
Clear Log	

Figura 4.11: Interfaccia grafica - Log errori

4.2 Interagire con la voce

Viene di seguito trattata l'interazione con la voce (rappresentata da un file `.json`), dal suo caricamento alla generazione dell'audio prodotto da *Speect*.

4.2.1 Caricare la voce

Per caricare la voce è sufficiente cliccare il pulsante *Load Voice*, situato nell'angolo in alto a sinistra del pannello di configurazione (vedi §4.1.2) o, in alternativa, cliccare sul menù *File* e, successivamente, sulla voce *Load Voice*. Una volta fatto ciò si aprirà il file browser del sistema operativo, che permetterà la ricerca di un file `.json` corrispondente alla voce desiderata. Una volta reperito il file, sarà sufficiente aprirlo mediante file browser.

A questo punto, se verrà visualizzato il percorso del file `.json` subito dopo l'etichetta "Path Voice", vicino al pulsante di caricamento della voce, allora il caricamento della stessa avrà avuto successo.



Nel caso in cui il caricamento del file abbia avuto esito negativo, è necessario ripetere la procedura assicurandosi si averla eseguita correttamente.

Esempio

In questo esempio si mostra come caricare una delle voice open source fornite con Speect. Sarà necessario attenersi alla seguente procedura:

1. Cliccare il pulsante *Load Voice* del pannello di configurazione (si noti che accanto a tale pulsante, se non è ancora stata caricata alcuna voice, è riportata un'etichetta "Voice not loaded") o sulla voce *Load Voice* del menù *File*;
2. Navigare tramite file browser secondo il path `SpeectInstaller/voices/cmu_arctic_slt/`;
3. Aprire il file `voice.json` presente all'interno della directory raggiunta al punto precedente.

Al termine della procedura, si noterà che accanto al pulsante *Load Voice* viene mostrato il path corrispondente alla voice caricata, a conferma della buona riuscita dell'operazione.

4.2.2 Salvare l'audio relativo alla voice

Uno volta eseguito *Speect*, per salvare l'audio è sufficiente cliccare sul pulsante *File* del menù e selezionare la voce *Save Audio file*.

Fatto ciò, si aprirà una finestra di ricerca del file browser dove cercare la cartella in cui salvare il file. Una volta trovata basterà selezionarla, cliccandoci sopra o scrivere il nome del nuovo file nella barra di testo dedicata e confermare, con attenzione di riportare l'estensione `.wav`.

Il file audio prodotto avrà l'estensione `.wav`. Questa operazione controllerà se è possibile salvare l'audio e come ogni altra operazione darà un report nella log di errore.

Esempio

In questo esempio si mostra come salvare l'audio relativo generato da Speect in relazione alla voice caricata. Sarà necessario attenersi alla seguente procedura:

1. Caricare la voice desiderata (vedi §4.2.1);



2. Avviare un'esecuzione di Speect, è possibile fare ciò inserendo un input testuale nell'apposita area di testo del pannello di configurazione e cliccare su *Run all* o *Run step* (vedi §4.3.1 - §4.3.2);
3. Cliccare sul pulsante *File* del menù e selezionare la voce *Save Audio file*;
4. Scegliere una cartella di destinazione per il file tramite file browser;
5. Digitare un nome per il file assicurandosi che abbia estensione *.wav*;
6. Confermare il salvataggio.

Al termine della procedura, il file audio generato da Speect (in formato *.wav*) sarà reperibile nella directory selezionata al punto 4 della precedente procedura.

4.3 Visualizzare il grafo

Per visualizzare il grafo è necessario aver caricato una voce, inserito un input testuale ed eseguito *Speect* tramite pulsante *Run all* o *Run step*.

Se sono state compiute tali azioni, viene visualizzato il grafo nell'area dedicata (§4.1.5), nonché la lista delle relation (§4.1.4). Il grafo è composto da:

- **Nodi:** visualizzati tramite cerchi con relativo name, i nodi hanno colori diversi a seconda della relazione a cui appartengono (la lista delle relation funge da legenda, vedi §4.1.4);
- **Archi:** visualizzati tramite frecce direzionali che collegano due nodi, anche gli archi sono di colori diversi a seconda della relazione a cui fanno riferimento. Gli archi sono differenziati in base al rapporto tra i nodi che collegano, come evidenziato dagli schemi seguenti.

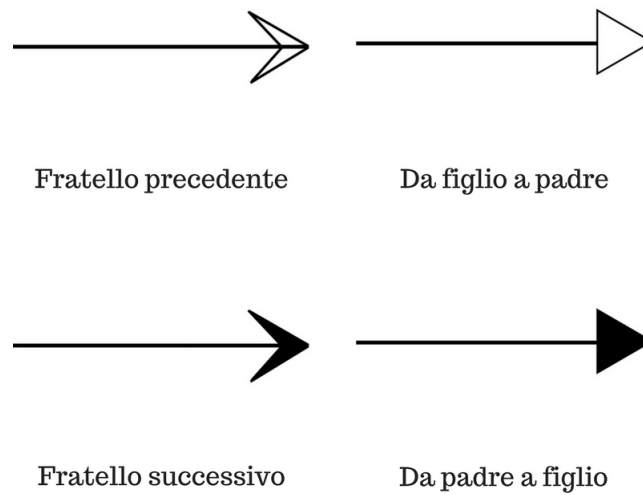


Figura 4.12: Legenda frecce grafo

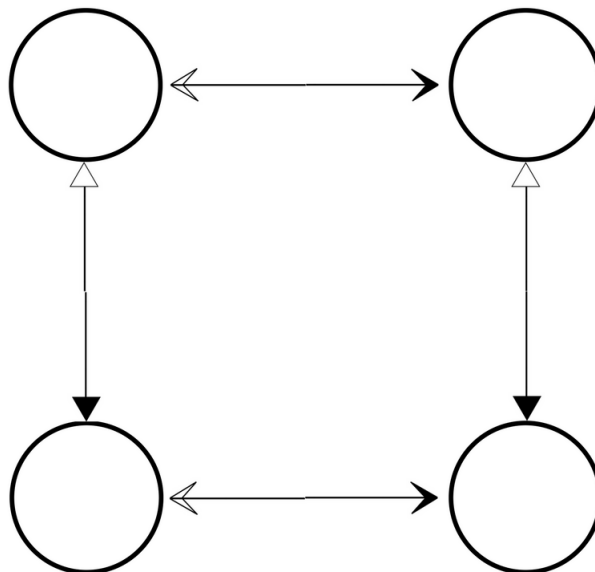


Figura 4.13: Schema frecce grafo



4.3.1 Visualizzare il grafo step-by-step

Eseguendo *Speect* tramite pulsante *Run step*, è possibile eseguire un utterance processor alla volta e vederne il risultato sul grafo, che ad ogni step si aggiorna aggiungendo nuove relazioni.

Esempio

In questo esempio viene mostrato come visualizzare il grafo step-by-step, ovvero eseguendo gli utterance processor selezionati uno alla volta sequenzialmente. La procedura è la seguente:

1. Caricare una voice (vedi §4.2.1);
2. Inserire un input testuale nell'apposita area di testo del pannello di configurazione (§4.1.2). Si inserisca ad esempio l'input "Hi everybody!";
3. Selezionare un utterance type dall'apposito menù a tendina del del pannello di configurazione (§4.1.2 - Figura 4.6). Si selezioni ad esempio "text-to-segments";
4. Cliccare ripetutamente il pulsante *Run step* del pannello di configurazione (§4.1.2), ad ogni click l'area del grafo (§4.1.5) verrà arricchita con le relazioni appropriate secondo la lista d'esecuzione degli utterance processor (§4.1.3) che si aggiornerà man mano. Seguendo gli input inseriti, verrà eseguito prima "Tokenize", successivamente "Normalize" e così via.

4.3.2 Visualizzare l'intero grafo

Eseguendo *Speect* tramite pulsante *Run all* è possibile eseguire tutti gli utterance processor in una sola volta e vedere subito il grafo completo.

Esempio

In questo esempio viene mostrato come visualizzare direttamente l'intero grafo, ovvero eseguendo tutti gli utterance processor selezionati. La procedura è la seguente:

1. Caricare una voice (vedi §4.2.1);



2. Inserire un input testuale nell'apposita area di testo del pannello di configurazione (§4.1.2). Si inserisca ad esempio l'input "Hi everybody!";
3. Selezionare un utterance type dall'apposito menù a tendina del pannello di configurazione (§4.1.2 - Figura 4.6). Si selezioni ad esempio "text-to-segments";
4. Cliccare il pulsante *Run all* del pannello di configurazione (§4.1.2): l'area del grafo (§4.1.5) mostrerà immediatamente tutte le relazioni appropriate secondo la lista d'esecuzione degli utterance processor (§4.1.3).

4.4 Interagire con il grafo

Questa sezione spiega come poter interagire col grafo spostando nodi e archi, rimuovendo nodi e archi relativi a determinate relation dalla visualizzazione.

4.4.1 Selezionare gli utterance processor

Nel pannello degli utterance processor (§4.1.3) è possibile interagire con i processor tramite le caselle di spunta a lato.

Togliendo la spunta da un processor, esso non verrà eseguito da *Speect*, e viceversa.

4.4.2 Interagire con le relation

Nel pannello delle relation è possibile interagire con esse tramite le caselle di spunta a lato.

Togliendo la spunta da una relation, i relativi nodi verranno rimossi dalla visualizzazione del grafo, viceversa se si aggiunge la spunta, verranno aggiunti i relativi nodi al grafo.

4.4.3 Traslare elementi grafici

Questa sezione spiega come traslare gli elementi grafici che compongono il grafo, ovvero nodi e archi.

4.4.3.1 Traslare nodi

Per traslare un nodo è sufficiente cliccare su di esso tenendo premuto e, spostando il cursore, trascinarlo dove si desidera all'interno dell'area del grafo.



Inoltre è possibile selezionare più nodi tenendo premuto CTRL durante la selezione

4.4.3.2 Traslare archi

Quando si trasla un nodo, gli archi che sono coinvolti direttamente con quel nodo si adatteranno alla nuova posizione del nodo.



5. Risoluzione dei problemi

5.1 Errori in DeSpeect

Gli errori che si possono riscontrare utilizzando l'applicazione verranno stampati nella log di errore, nel caso l'errore sia dell'engine Speect il suo errore verrà riportato nel file `ErrorLog.txt` *DeSpeect*.

5.1.1 Struttura dei codici di errore

Gli errori rispettano la seguente notazione:

Operazione in esecuzione, Operation Status: stato di Speect in seguito all'operazione. Gli stati sono:

- No error Successo, nessun errore;
- Failure Una failure non specificata è avvenuta;
- Memory allocation failed Errore di allocazione di memoria;
- Function argument(s) invalid Errore negli argomenti della funzione;
- Class/object method does not exist Classe/Oggetto non esistono;
- End of file/stream Fine del file o dello stream raggiunta;
- Warning, possible error Attenzione possibile errore;
- Error context continued Errore il contesto continua;
- Undefined error Errore non definito.

Se l'operazione finisce in uno stato diverso da No error, nel file `ErrorLog.txt` verrà descritto in maniera più dettagliata.



5.2 Problemi con il reperimento di Speect

Il portale che ospita documentazione e download di *Speect* può occasionalmente risultare offline per dei giorni.

Si consiglia di scaricare in locale tutto il materiale reperibile online così da tutelarsi nel caso il problema dovesse occorrere.

5.3 Segnalazione di bug

DeSpeect potrebbe contenere bug o potrebbe essere desiderabile apportare modifiche e ampliamenti alle sue funzionalità.

È possibile segnalare malfunzionamenti o richieste di nuove funzionalità sotto forma di GitHub *issue*_G all'indirizzo:

<https://github.com/graphiteSWE/DeSpeect>

oppure scrivendo direttamente all'indirizzo e-mail:

graphite.swe@gmail.com