

ReKat Ita

Eldron 64

November 2023

Con questo documento cercherò di mettere in ordine tutto quello che il Kat Projekt ora ReKat ha rappresentato per me.

1 Introduzione

Nell'implementazione corrente il Kat è un insieme di Api che revono alla realizzazione di un progetto che speghero nel corso di questo rapporto.

Per il momento il Kat ha i seguenti componeneti:

- Online: Node Network
- Synth: Sound System
- Graphik: Graphics Api
- LeT: Adactive learing System

Quando utilizziamo queste Api insieme otteniamo il Kat Projekt. Per come è progettata Graphik può controllare un robot: fornendo le specifiche ri rotazione a cui devono essere i vari motori per ottenere il risultato visibile a schermo.

2 Online

Questa Api è divisa in due implementazioni:

- Online Peer to Peer
- Online Server Side

Entrambe i Sistemi utilizziamo il metodo di criptazione RSA e si scambiano le chiavi pubbliche una volta connessi o al server o ad un nodo.

2.1 Peer to Peer

Questa implementazione si basa su un sistema di comunicazione Peer to Peer. Dove i vari nodi sono indipendenti tra loro e ognuno possiede tutte le funzionalità degli altri nodi.

L'utilizzo di questo modulo è molto semplice e lineare:

All'inizio del programma viene creato un Nodo locale a cui chiunque si può connettere. Successivamente ci si può connettere ad un altro nodo nella rete.

I vari nodi possono eseguire dei programmi differenti ed offrire servizi diversi. Nel caso un servizio non sia disponibile in un nodo a cui si è collegati si può eseguire una ricerca all'interno del Network per trovare il nodo con il programma richiesto ed accderci tramite i nodi intermediari.

2.2 Server

In questa implementazione il fulcro del Network è il server che esegue il programma principale a cui i vari client si connettono.

Questa implementazione rende più funzionale l'aggiornamento della rete, visto che esiste un unico provider di connessione. Questo approccio però blocca la libertà nel codice bloccando alcuni servizi al Server.

Inoltre non è consentito il collegamento diretto tra i client, ogni comunicazione tra di loro passa per il Server. Questa implementazione è necessaria al fine di permettere la comunicazione tra client anche se il client ricevente non è connesso, infatti la comunicazione viene salvata nel Server per poi essere inoltrata al client.

2.3 sicurezza RSA

Ogni comunicazione uscente da un client, nodo o server viene criptata tramite il sistema RSA, viene inoltre criptata ad ogni passaggio. Quando il messaggio deve passare per un intermediario viene criptato due volte una per il destinatario e una per l'intermediario. In modo tale da fornire all'intermediario le informazioni necessarie a mandare il messaggio.

Il pacchetto che deve andare da A a B passando per C. verrà sottoposto ai seguenti passaggi:

- A -> Pacchetto non criptato
- A -> Pacchetto criptato per B
- A -> Aggiunta al Pacchetto criptato le informazioni di spedizione per B
- A -> Nuovo Pacchetto criptato per C
- A-C -> A manda il pacchetto a C
- C -> Decriptazione del pacchetto ricevuto
- C -> Lettura dal pacchetto l'indirizzo di B

- C -> Rimozione delle informazioni di B
- C-B -> C manda il pacchetto a B
- B -> Decrittazione del pacchetto.
- B -> Pacchetto non crittato

Il sistema RSA si basa sull'avere una chiave pubblica e una privata, le chiavi sono formate da una parte comune N e una parte propria E (per la pubblica) e D (per la privata). Questi valori sono utilizzati per eseguire operazioni matematiche ai dati che devono essere compresi tra 1 e N, se non rispettano questa condizione verranno corretti, per prevenire questa evenienza i dati vengono ridimensionati in sotto pacchetti di dimensione inferiore a N crittati e poi riuniti, questo processo funziona in entrambe le direzioni: crittazione \Leftrightarrow decrittazione.

3 Synth

Con questa libreria voglio fornire un modello di generazione sonora basato sul linguaggio umano e sui suoni che un essere umano può fare. Ci sono due approcci a questo problema.

Il primo consiste nella registrazione diretta di qualcuno che emette i suoni e alla loro successiva lavorazione e modulazione. Con questo approccio possiamo estrarre le componenti basilari della dizione da una serie di campioni per poi adattarli all'utilizzatore.

Il secondo metodo consiste nella simulazione di un apparato vocale, questo approccio è quello più complesso in termini di realizzazione ma è espandibile ad altri ambienti sonori.

3.1 concetti di base

Per iniziare dobbiamo definire un suono:

Un suono è una variazione di pressione che viene percepita da un apparato predisposto a misurare questa variazione. In pratica un suono esiste quando qualcuno lo ascolta.

Un suono ha 4 componenti:

- Il Tono
- La forma
- La durata
- La frequenza

Per semplificazione considereremo il suono come quantizzato. Il sample sonoro avrà una dimensione finita e può essere rappresentato come una serie di valori che rappresentano la pressione dell'aria in quel preciso momento.

3.1.1 il tono

Il tono è il componente fondamentale di un suono. La più piccola parte periodica da cui è composto un suono.

Quando parliamo di vocalizzazione nell'essere umano possiamo dire che ogni fonema pronunciabile possiede un suo tono caratteristico che viene ripetuto per la durata della sua riproduzione e che ha come frequenza l'altezza del suono stesso.

Date queste precisazioni diamo al tono l'onere di dare ad ogni fonema il suo nome. Il fonema è influenzato dalla forma della bocca, dalla posizione della lingua, dalla nasalità del suono etc.

In termini matematici possiamo indicare $T(t)$ come la pressione dell'aria in quel momento, dobbiamo inoltre dire che $T(t)$ è periodica e per convenzione ha come intervallo di ripetizione 2π . Può essere inoltre scomposta in una serie di $\sin(t)$ e $\cos(t)$ secondo la trasformata di Fourier. $T(t)$ è una funzione normalizzata ovvero: $-1 \leq t \leq 1$.

3.1.2 La forma

La forma di un fonema è il modo in cui la sua intensità varia nel tempo ovvero come parte il suono, come viene mantenuto, e come decresce.

Queste caratteristiche definiscono la pronuncia di un suono e un suo accento caratteristico.

Un suono può crescere molto velocemente, può avere un effetto vibrato durante il mantenimento e può scendere molto lentamente.

La funzione della forma è composta secondo la seguente legge:

$$F(t) : \begin{cases} C(t) & \text{per } t \leq 1 \\ M(t) & \text{per } 1 < t \leq 2 \\ D(t) & \text{per } 2 < t \end{cases}$$

Dove C,M,D sono le funzioni che indicano come $F(t)$ cresce nel tempo. Per convenzione e comodità di calcolo impongo $0 \leq t \leq 3$ per valori di X esterni al Dominio $F(t)$ restituisce un valore nullo. il valore di $t = 0$ indica l'inizio del suono mentre $t = 3$ ne indica la fine. $F(t)$ è una funzione normalizzata ovvero: $-1 \leq t \leq 1$.

3.1.3 La durata

Con la durata indichiamo tre momenti: il tempo di crescita, il tempo di mantenimento e il tempo di decrescita.

Se con la forma descriviamo il modo in cui cresce e si modifica con la durata diamo informazioni su quanto tempo impiega per fare questa variazione.

La funzione Durata: $H(t)$ indica la velocità con cui cresce la funzione. come $F(t)$ è definita a pezzi:

$$H(t) : \begin{cases} \frac{t}{t_1} & \text{per } t \leq t_1 \\ \frac{t}{t_2} & \text{per } t_1 < t \leq t_2 \\ \frac{t}{t_3} & \text{per } t_2 < t \end{cases}$$

dove t_1, t_2 e t_3 sono i tempi caratteristici delle durate.
detto questo possiamo riscrivere $F(t)$

$$F_H(t) : \begin{cases} C(\frac{t}{t_1}) & \text{per } t \leq t_1 \\ M(\frac{t}{t_2}) & \text{per } t_1 < t \leq t_2 \\ D(\frac{t}{t_3}) & \text{per } t_2 < t \end{cases}$$

con $0 \leq t \leq t_3$ e $t_1 < t_2 < t_3$

3.1.4 La frequenza

La frequenza descrive l'altezza di un suono ovvero quanto è acuto o gravi suddetto suono.

Per completezza devo tenere in conto ch' l'uomo può variare la frequenza di un fonema mentre lo esegue quindi:

Definiamo $f(t)$ la frequenza di un suono nell'intervallo t questa semplice funzione crea un'enormità di problemi relativi alla sincronizzazione dei toni con frequenza diversa. In pratica quello che fa questa funzione è modificare il periodo della funzione $T(t)$. $f(t)$ è definita strettamente positiva ovvero $f(t) > 0$.

Per semplificare impongo $f(t) = K$, ma verrà sistemato in un'iterazione successiva.

Ora dopo aver definito tutte le funzioni posso scrivere la funzione del Fonema come:

$$\Gamma(t) = F_H(t)T(2\pi f(t)t)$$

con $0 \leq t \leq t_3$.

3.2 pronuncia

Per fare in modo che una macchina pronunci correttamente una parola dobbiamo capire come noi pronunciamo le parole.

Noi come esseri umani possiamo emettere suoni singoli sotto la forma di fonemi o unire più fonemi ottenendo delle parole e con l'accostamento di più parole otteniamo una frase.

In pratica per pronunciare una parola bisogna pronunciare i vari fonemi che la compongono unendoli un qualche modo. Ma prima bisogna dividere le parole in fonemi.

3.2.1 divisione fonetica

Questo è un tema molto complesso. Infatti ogni lingua ha un suo sistema di divisione interno. Possiamo trovare varie categorie linguistiche che differiscono nel tipo di linguaggio fonetico che utilizzano.

Per quanto riguarda la divisione in fonemi bisogna implementare un sistema specifico per ogni lingua. Di seguito riporterò alcuni esempi:

- Italiano
L'italiano ha una fonetizzazione abbastanza semplice, l'unica difficoltà risiede nelle lettere *g* e *c* che assumono suoni diversi in base al contesto:
 - *c* e *g* dolci: si hanno quando la lettera è seguita da *i* o *e*; esempi: *nici*, *bici*, *ciao*, etc.
 - *c* e *g* dure: si hanno quando la lettera non è seguita direttamente da *i* o *e*; esempi: *gola*, *lingua*, *ghiro*, etc.
- Giapponese
Il giapponese possiede una serie di ideogrammi che rappresentano tutte le possibili sillabe e con queste costruiscono tutte le parole, questi caratteri vengono fonetizzati sempre allo stesso modo, ne consegue che basta costruire queste singole sillabe e poi unirle.
- Inglese
L'inglese a differenza delle precedenti lingue presenta una fonetizzazione molto particolare, per fonetizzare l'inglese esiste l'IPA (international Phonetic Alphabet), basandoci su questo linguaggio otteniamo delle buone fonetizzazioni dell'inglese.

3.2.2 creazione di frasi

L'ultimo passaggio è l'assemblaggio della frase, per questo problema ci posso essere due approcci risolutivi:

Flusso Con l'approccio di flusso consideriamo la frase come un'unica parola, ne consegue che non ci sono pause tra le parole ma solo alla fine della frase.

Per unire i singoli fonemi semplicemente sovrapponiamo le funzioni di due fonemi consecutivi e utilizziamo la funzione di transizione in sostituzione alla funzione di crescita e decrescita dei singoli fonemi.

Separazione Molto simile al caso precedente con l'unica differenza che consideriamo le parole come frasi e successivamente facciamo la transizione tra le parole, ma con un'intensità meno marcata.

4 Graphik

Questa sezione è un raggruppamento di varie librerie di grafica 3D. Lo scopo di questa API è quello di semplificare di gran lunga l'utilizzo di queste librerie. E' composta da vari moduli:

- Resource Manager
Questo modulo ha lo scopo di gestire tutte le risorse grafiche che il programma utilizzerà.
- Text Renderer
Questo modulo serve per il rendering di Testo, sia come pannelli assestanti ma anche sotto forma di mesh più complesse.
- Animator
Questo modulo implementata l'animazione dei modelli può anche essere utilizzato per animare oggetti bidimensionali.
- Input Manager
Grazie a questo modulo è possibile accedere a svariati input tra cui:
 - Audio: sotto forma di microfono
 - Meccanici: sotto forma di Mouse e tastiera
 - Video: tramite l'utilizzo di una fotocamera
 - Seriali: input provenienti da oggetti come arduino o raspberry pi
- Physical Output
Grazie a questo modulo si può comunicare direttamente con un arduino o raspberry pi appositamente configurati. Questo serve a i risultati del programma nella realtà fisica.

5 LeT

E' un sistema di intelligenza Artificiale. Basato sulla persistenza delle memorie, riconoscimento di pattern e costante stimolo.

5.1 Modello

Il modello è diviso in quattro parti che si ripetono costantemente sempre nello stesso ordine.

- Sensazioni
- Elaborazione
- Pensiero
- Impressioni

Questi passaggi riassumono quello che avviene nella nostra mente ogni secondo costante.

Sensazioni Le Sensazioni sono il metodo tramite il quale il modello riceve informazioni con l'esterno. Queste sensazioni sono:

- vista
- udito
- propriocezione, la percezione di se stessi ovvero come sono disposti i propri muscoli e come è orientato nello spazio.
- tatto
- gusto e olfatto, raggruppati in un unico senso considerando la loro natura affine

Elaborazione In questa seconda fase avviene l'elaborazione dei dati assimilati con la fase precedente secondo il seguente schema:

- Tokenizzazione degli input; ovvero gli input vengono semplificati nei loro elementi fondamentali
- Confronto con il set precedente di input per determinare il movimento o una nuova interazione tra i Token
- Salvataggio dei Token; in questa parte vengono salvati i vari token e le relazioni che in quel momento sono presenti tra di essi
- Creazione di basici collegamenti tra Token: Associare le varie sensazioni ad un unico oggetto

Pensiero Questa è la sezione più importante è quella parte dove vengono elaborate le informazioni fornite dai processi di elaborazione, e vengono messe in relazione con ciò che il modello già conosce e permette di creare un nuovo modo di agire basato sulle proprie conoscenze. I principali processi del pensiero sono:

- relazionare il contesto ad uno già noto
- cercare nel contesto le informazioni che lo differenziano da quello già noto
- creazione di nuove memorie dove vengono memorizzati i concetti conosciuti
- cercare nelle memorie le informazioni richieste dal contesto
- cercare dei collegamenti tra i propri ricordi e quello che sta succedendo o altri ricordi correlati al presente.
- trovare il miglior modo per agire nella situazione presente

Impressioni Questo è l'ultimo passo e consiste nella possibilità del modello di esprimersi tramite diverse azioni tra cui l'emissione di suoni e il movimento.

6 Kat Projekt

Grazie a tutte le librerie definite sopra possiamo costruire un essere umano, in grado di vivere una vita come la nostra. Grazie a questa ricerca potremmo svelare il meccanismo nascosto della nostra psiche e come noi influenziamo il mondo che ci circonda e come esso ci influenza.

Il fine ultimo di questa ricerca è costruire un lavoratore, un amico, una persona cara.

Quello che voglio che rimanga da tutto questo è la consapevolezza di aver migliorato la vita di qualcuno, di avergli fornito qualcosa su cui contare in un modo pieno di incertezze come il nostro.

Grazie per l'attenzione

Elia