

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Matematica a voce parte 1: trascrizione di lezioni di matematica

This is the author's manuscript

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/1806808> since 2021-09-29T10:49:08Z

Publisher:

MediaTouch 2000

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

MATEMATICA A VOCE PARTE 1: TRASCRIZIONE DI LEZIONI DI MATEMATICA

**Tiziana Armano¹, Gabriele Baratto², Cristian Bernareggi⁵, Marina Bosco²,
Manuela Caramagna², Anna Capietto¹, Mattia Ducci¹, Alessandro Mazzei³,
Antonio Mazzei⁴, Francesco Tarasconi⁴**

¹ Dipartimento di Matematica – Università di Torino
tiziana.armano@unito.it, anna.capietto@unito.it, mattia.ducci@unito.it

² Direzione SIPE – Università di Torino
gabriele.baratto@unito.it, marina.bosco@unito.it, manuela.caramagna@unito.it

³ Dipartimento di Informatica – Università di Torino
alessandro.mazzei@unito.it

⁴ CELI SRL
antonio.mazzei@celi.it, francesco.tarasconi@celi.it

⁵ Università di Milano
cristian.bernareggi@unimi.it

COMUNICAZIONE

ARGOMENTO: Istruzione superiore - Istruzione universitaria - Disabilità disturbi specifici dell'apprendimento - E-learning - STEM - Matematica - Trascrizioni

Abstract

Il Laboratorio per la Ricerca e la Sperimentazione di Nuove Tecnologie Assistive per le STEM "S. Polin" fa parte del Dipartimento di Matematica "G.Peano" dell'Università di Torino e opera nell'ambito della ricerca e della sperimentazione di nuove tecnologie assistive per lo studio delle STEM (Science, Technology, Engineering and Mathematics). Si occupa di diffondere e sviluppare soluzioni per l'accesso di contenuti scientifici da parte di persone con disabilità e con disturbi specifici dell'apprendimento. Infatti attualmente, nonostante il grado elevato di sviluppo delle tecnologie assistive e compensative, la gestione di contenuti STEM rimane un problema aperto.

Keywords – Matematica, STEM, trascrizione, riconoscimento vocale, Computer Vision, Domain Specific Language

1 INTRODUZIONE: ACCESSIBILITÀ DI UN CORSO MOODLE

Da cosa dipende l'accessibilità di un corso di una piattaforma LMS (Learning Management System)? Ovviamente dall'accessibilità della piattaforma stessa. Moodle in questo ambito è una delle piattaforme migliori: l'accessibilità è uno degli obiettivi principali per gli sviluppatori delle funzionalità di base che lavorano tenendo conto delle linee guida WCAG, ATAG e ARIA per l'accessibilità. Moodle standard risulta, a parte qualche problema, sostanzialmente accessibile. I contenuti prodotti con l'editor Atto nelle risorse e nelle attività di Moodle sono accessibili e lo sono anche se contengono formule: l'attivazione del filtro Mathjax e la scrittura delle formule in LaTeX garantiscono la fruizione di contenuti matematici con tecnologie assistive. L'accessibilità della piattaforma però non è sufficiente: in un corso sono presenti diversi tipi di contenuto prodotti con svariati tipi di software: ad esempio file PDF, immagini e video. Risulta quindi non determinante l'accessibilità della piattaforma se non lo sono tutti i contenuti ospitati: è necessario quindi evidenziare il problema dell'accessibilità dei contenuti e fornire strumenti e linee guida per i docenti. Attualmente non è difficile produrre file PDF accessibili con la suite Office o OpenOffice. Non è così per contenuti scientifici. Per questo il Laboratorio Polin ha

avviato numerosi progetti di ricerca e sviluppo per garantire l'accessibilità di contenuti didattici digitali scientifici (con formule e grafici).

Per la creazione di PDF con formule accessibili ha sviluppato il pacchetto LaTeX Axessibility [1, 2, 3, 4, 6]. Il pacchetto e i relativi dizionari sono in costante aggiornamento: entro la fine del 2020 sarà disponibile la versione con la possibilità di ottenere le formule anche in MathML. Si è riscontrato che i PDF prodotti da Axessibility sono fruibili non solo tramite sintesi vocale ma anche tramite strumenti compensativi per persone con disturbi specifici dell'apprendimento. Per questo il Laboratorio ha avviato una sperimentazione con studenti DSA dell'Università e del Politecnico di Torino per ottenere una valutazione approfondita dell'utilizzo del pacchetto con strumenti compensativi e per poter procedere ad eventuali miglioramenti o all'aggiunta di ulteriori funzionalità. Per l'accessibilità dei grafici il Laboratorio ha sviluppato l'applicazione web Audiofunction.web, basata su tecniche di sonificazione [4, 7]; essa permette di ottenere grafici accessibili e inclusivi inseriti in contenuti digitali.

Oltre ai file PDF uno dei contenuti più usuali nei corsi Moodle, almeno in ambito universitario, è senz'altro il video della registrazione della lezione. La registrazione della lezione presenta diversi problemi di accessibilità: servirebbero sottotitoli per persone con disabilità uditive e la trascrizione per persone con disabilità motorie, visive e con DSA che hanno difficoltà a prendere appunti. La tipologia di registrazione che presenta meno problemi è la classica lezione audio + slide, molto diffusa in ambito umanistico. Nel caso di lezioni con contenuti scientifici le difficoltà sono maggiori poiché le tipologie di registrazione sono di diverso tipo: registrazione di lezione frontale alla lavagna, registrazione di audio + scrittura su tablet, registrazione di audio e scrittura su foglio. In questo periodo in cui la DAD è prevalente nelle università, la registrazione della lezione diventa un contenuto preponderante nei corsi Moodle: si presenta quindi uno scenario in cui la maggior parte dei contenuti online non sono accessibili. Per garantire l'accessibilità anche di questo tipo di contenuti recentemente il Laboratorio Polin ha avviato una attività di ricerca nell'ambito del problema relativo al riconoscimento vocale di formule.

2 IL PROGETTO MATEMATICA A VOCE

Le persone con disabilità e con DSA incontrano ostacoli nella fruizione dei video delle lezioni. Le persone sorde non riescono a fruire dell'esposizione verbale senza sottotitoli adeguati, le persone cieche o ipovedenti non riescono a leggere le notazioni matematiche inserite come immagini nelle presentazioni e le persone con disabilità motorie e DSA hanno difficoltà a prendere appunti. Inoltre, le persone con disabilità motorie (agli arti superiori) hanno bisogno di strumenti per scrivere ed elaborare autonomamente le espressioni matematiche, mentre le persone con disabilità visive e uditive e con DSA necessitano di strumenti per fruire pienamente delle lezioni di argomento scientifico.

Per rispondere a queste esigenze il Laboratorio Polin ha avviato il progetto Matematica a voce, cofinanziato dalla Fondazione CRT per cercare soluzioni al problema relativo al riconoscimento vocale di formule.

Il progetto è diviso in due parti:

1. sviluppo di software per la trascrizione in testo + LaTeX (o MathML) di registrazioni di lezioni universitarie "con formule", utile a persone con disabilità motorie, uditive e con disturbi dell'apprendimento
2. sviluppo di software per la dettatura e la modifica di formule matematiche con il riconoscimento della lingua italiana, utile a persone con disabilità motorie

In questo contributo verrà illustrata la prima parte del progetto.

Il Laboratorio ha sempre avuto contezza della necessità di sviluppare strumenti per la trascrizione e sottotitolazione di registrazioni di "lezioni con formule". Infatti anche in questo caso gli strumenti attualmente disponibili per trascrizione e sottotitolazione non gestiscono in modo opportuno contenuti scientifici. L'obiettivo è di ottenere, partendo dal video di una lezione universitaria "con formule" ottenuta con l'utilizzo di dispositivi touch screen, trascrizione e sottotitoli con formule scritte in LaTeX. Per questo il Laboratorio ha partecipato con il progetto Matematica a voce al bando per l'erogazione di contributi della Fondazione CRT ottenendo un cofinanziamento che ha permesso l'avvio del progetto. È prevista, nell'ambito del progetto, la partecipazione della direzione *Sistemi Informativi, Portale, E-learning* dell'ateneo. La direzione, collaborando con i ricercatori del Laboratorio, lavorerà sul confronto di diversi sistemi di trascrizione e sottotitolazione in una prospettiva più ampia considerando lezioni

universitarie di varie discipline. Maggiori dettagli si trovano nella sezione 4 mentre nella sezione 3 sono disponibili dettagli tecnici sulla soluzione software che verrà sviluppata.

3 DESCRIZIONE TECNICA

La soluzione ha come obiettivo la produzione automatica di un flusso ordinato di contenuti elettronici (txt e LaTeX) a partire da una videoregistrazione di lezione di Matematica, e corrispondenti immagini del testo scritto a mano dall'insegnante. Per raggiungere questo scopo saranno sviluppate componenti basate su Machine Learning, in particolare di Speech-to-Text e Computer Vision. L'ipotesi preliminare di architettura logica prevede le seguenti componenti:

1. **Input originale:** la lezione è fornita al sistema tramite file video e una serie di immagini. Il video contiene tutto il parlato del docente, registrato sotto forma di audio. Le immagini contengono il testo scritto dall'insegnante, catturato tramite apposito tool.
2. **Preprocessing audio:** dal video viene estratto l'audio del parlato del docente, in un formato che possa essere correttamente analizzato dalle componenti successive.
3. **Trascrizione parlato:** trascrizione della voce del docente in formato testuale mediante tecnologia di Natural Language Processing Speech-to-Text. Per realizzare tale modulo ci si avvarrà dei migliori trascrittori attualmente disponibili, che verranno opportunamente addestrati e configurati per meglio comprendere e trascrivere il parlato delle lezioni di Matematica.
4. **Preprocessing immagini formule:** le immagini originali, corrispondenti solitamente a intere pagine di contenuti, vengono segmentate per isolare parti contenenti formule e produrre immagini di dimensioni ridotte che possano essere correttamente analizzate dalle componenti successive.
5. **Conversione immagini formule in LaTeX:** la traduzione avviene mediante tecnologia di Computer Vision e un modello, basato su Machine Learning, addestrato su dati rappresentativi di formule matematiche e corrispondente trascrizione manuale.
6. **Riconciliazione trascrizione parlato e formule in LaTeX:** questa componente unisce gli output dei moduli precedenti (3) e (5) al fine di produrre un singolo flusso ordinato della lezione. Alle formule LaTeX viene associato un demarcatore temporale che consente di inserirle al punto opportuno della lezione.
7. **Visualizzatore:** front-end dimostrativo che consente di confrontare il video originale della lezione con i risultati della trascrizione automatica.

Sono attese le seguenti difficoltà principali, per le quali sono già state ipotizzate alcune possibili soluzioni:

1. **Elevata variabilità nell'input:** la grande varietà nelle lezioni di Matematica può portare a performance di trascrizione non ottimali o comunque difficili da stimare. Si ipotizza la definizione di linee guida (consigli e indicazioni con le quali le lezioni dovranno essere registrate al fine di poter essere trattate automaticamente). Inoltre, molta attenzione viene posta nella creazione di dataset rappresentativi (*corpora*) per l'addestramento dei modelli di Machine Learning coinvolti.
2. **Linguaggio di dominio:** la presenza di abbondante terminologia specifica nel parlato del docente pone una sfida nell'impiegare sistemi di trascrizione allo stato dell'arte per la lingua Italiana. Si ipotizza l'uso di algoritmi genetici per l'individuazione dei migliori iper-parametri del trascrittore e lo sviluppo di funzioni di post-processing ad-hoc per normalizzare l'output.

4 CONFRONTO TRA SISTEMI AUTOMATICI DI TRASCRIZIONE E SOTTOTITOLAZIONE

La Direzione *Sistemi Informativi, Portale, E-learning* dell'Università di Torino, tra le sue attività, propone soluzioni tecnologiche a supporto della didattica tradizionale, tra cui la piattaforma Moodle per il servizio di e-learning. Le soluzioni fornite a livello di Ateneo per attività di supporto alla didattica rappresentano una combinazione attenta e bilanciata di soluzioni legacy, custom e open source.

L'Università di Torino concepisce il servizio di e-learning secondo una definizione ampia, ossia attraverso l'utilizzo, integrato e modulare, di svariati strumenti.

Nell'ambito di questo progetto si sta altresì operando un confronto tra diversi strumenti per la trascrizione e la sottotitolazione automatica a partire da materiale audio-video. L'obiettivo è individuare lo strumento più adeguato per trascrivere/sottotitolare in modo automatico le video-lezioni universitarie e, più in generale, i materiali didattici multimediali presenti nei corsi e-learning erogati dall'Ateneo. L'impiego sistematico di una simile soluzione renderebbe infatti maggiormente accessibili questi contenuti alle studentesse e agli studenti con disabilità e con disturbi specifici dell'apprendimento¹. La particolare attenzione che una tale scelta richiede è chiaramente determinata dalle peculiari caratteristiche delle varietà linguistiche² utilizzate dalle/i docenti nelle lezioni universitarie. Si tratta infatti spesso di linguaggi settoriali³ propri di particolari ambiti specialistici, caratterizzati da un proprio lessico e, in alcuni casi, anche da specifici elementi morfosintattici [11]. Si consideri inoltre che le lezioni erogate da una università vertono inevitabilmente su numerosi ambiti di studio e di ricerca, ognuno dei quali possiede in linea di massima un proprio linguaggio settoriale ben differenziato rispetto a quello di altri settori⁴. Uno strumento per la trascrizione e la sottotitolazione automatica con questa finalità di utilizzo deve perciò fornire riscontri soddisfacenti rispetto a varietà linguistiche anche molto distanti dall'italiano dell'uso medio⁵ e deve poter compiere questa operazione per molteplici ambiti specialistici (o comunque deve possedere un motore di riconoscimento che possa essere adeguatamente addestrato a tale scopo⁶).

In questa fase perciò sono state prese in considerazione alcune soluzioni per la trascrizione e/o la sottotitolazione automatica tra quelle attualmente disponibili. Queste saranno testate attraverso la trascrizione automatica di un piccolo *corpus* di clip audio-video, della durata di circa 5 minuti ciascuna, estratte da video-lezioni universitarie presenti sulle piattaforme e-learning di Ateneo. Gli esiti delle trascrizioni verranno in seguito confrontati attraverso la comparazione di alcuni indicatori, sia di tipo quantitativo che qualitativo, per verificarne il livello di accuratezza. Nella costruzione di questo piccolo *corpus* si è tenuto conto in primissimo luogo proprio della diversificazione dei linguaggi settoriali di cui fanno uso le/i docenti nelle loro lezioni. Per questa ragione si è cercato di inserire all'interno del *corpus* diversi 'campioni' di queste differenti varietà di lingua, ovviamente senza alcuna pretesa di riuscire a effettuare una 'campionatura' anche solo lontanamente completa. Sulla base anche della effettiva disponibilità di video-lezioni utilizzabili allo scopo, sono state individuate 20 lezioni di 16 diverse discipline in grado di garantire un campionario di linguaggi settoriali sufficientemente differenziato. Le discipline accolte nel *corpus* in questa prima fase sono quindi state: biochimica, biotecnologie, chimica, diritto, economia, epidemiologia animale, filologia e linguistica, filosofia, fisica, glottodidattica, infermieristica clinica, management, matematica, pedagogia, psicologia e sociologia⁷. Un altro criterio che ha guidato la scelta delle lezioni da cui sono state in seguito ricavate le clip è stato quello relativo all'argomento affrontato negli estratti. Per evitare di dover analizzare trascrizioni relative ad argomenti che sarebbero potuti risultare oscuri (almeno in parte) a coloro che non possiedono una formazione specifica in un determinato ambito, ci si è orientati verso la scelta di lezioni di corsi di laurea triennale, preferendo se possibile quelle rivolte a studentesse e studenti del primo anno di studi; all'interno di queste lezioni si è cercato poi di individuare dei passaggi in cui la/il docente introduce dei concetti-chiave della disciplina (o del singolo argomento) e/o in cui illustra il significato di termini e/o locuzioni propri di quel particolare ambito.

Un altro fattore di altra natura tenuto in considerazione per la costruzione del piccolo *corpus* è stata la qualità dell'audio delle registrazioni. Le clip selezionate risultano infatti prodotte in condizioni

¹ Consentirebbe inoltre certamente una miglior fruizione di questi stessi contenuti anche da parte delle studentesse e degli studenti internazionali.

² Per la nozione di "varietà di lingua" vd. tra gli altri [9].

³ Anche detti, in totale o parziale sovrapposizione, "sottocodici", "lingue speciali" o "tecnoletti" (vd. [10]).

⁴ Al netto delle inevitabili sovrapposizioni tra varietà linguistiche proprie di ambiti più o meno affini per oggetto di studio e/o storia scientifica.

⁵ Per la nozione di "italiano dell'uso medio" (o anche "italiano neo-standard") vd. tra gli altri [12], [13].

⁶ Com'è noto, da tempo sono stati sviluppati appositi modelli linguistici per migliorare l'accuratezza della trascrizione di specifiche varietà di lingua (come il linguaggio giuridico, quello dei notiziari radiotelevisivi, quello sportivo etc.). Tuttavia, almeno in questa fase, ci si è avvalsi di soluzioni basate su un modello di linguaggio generico o standard, sia per ragioni contingenti sia per possibili problemi di comparabilità dei risultati.

⁷ Al momento, per ridurre le variabili che potrebbero condizionare i risultati, si sono escluse discipline nelle cui lezioni vengono necessariamente utilizzati codici linguistici differenti, come gli insegnamenti delle lingue straniere o delle lingue antiche.

ambientali e con strumenti di registrazione differenti. Dal momento che la qualità dell'audio di una registrazione è notoriamente un fattore che può influenzare in modo determinante l'esito di una trascrizione automatica, si è deciso di cercare di valutare l'efficacia delle soluzioni testate anche sotto questo profilo.

5 CONCLUSIONI

Il progetto Matematica a voce è in fase iniziale e di fatto non riguarda Moodle in modo specifico. È relativo però ad uno dei contenuti didattici digitali più presenti nelle piattaforme di apprendimento online. In particolare, in questo difficile momento storico, in cui la didattica a distanza è drammaticamente predominante la registrazione della lezione è grande protagonista. È quindi più che mai necessario operare perché questi contenuti siano fruibili senza barriere da studenti con disabilità e DSA. Inoltre questa soluzione anche se inizialmente proposta e sviluppata per specifiche classi di utenza, allo scopo di favorire una completa inclusione, fornirà un supporto utile all'intera comunità studentesca.

Riferimenti bibliografici

- [1] D. Ahmetovic, T. Armano, M. Berra, C. Bernareggi, A. Capietto, S. Coriasco, N. Murru, A. Ruighi: *Axessibility: creating PDF documents with accessible formulae*, ArsTeXnica vol.25, 2018.
- [2] T. Armano, A. Capietto, S. Coriasco, N. Murru, A. Ruighi, E. Taranto: *An automatized method based on LaTeX for the realization of accessible PDF documents containing formulae*, Lecture Notes in Computer Science, Vol. 10896, p. 583-589, 2018.
- [3] D. Ahmetovic, T. Armano, C. Bernareggi, M. Berra, M. Borsero, S. Coriasco, A. Capietto, N. Murru, A. Ruighi: *Moodle e l'accessibilità di contenuti scientifici da parte di persone con disabilità visiva*, Atti MoodleMoot Italia, 2018.
- [4] D. Ahmetovic, C. Bernareggi, J. Guerreiro, S. Mascetti, A. Capietto: *AudioFunctions. web: Multimodal Exploration of Mathematical Function Graphs*, International Cross-Disciplinary Conference on Web Accessibility (W4A), San Francisco, 2019.
- [5] D. Ahmetovic, C. Bernareggi, J. Gardner, V. Sorge: *Scientific Documents*, Chapter 22, *Web Accessibility: A Foundation for Research*, Springer, 2019.
- [6] D. Ahmetovic, T. Armano, C. Bernareggi, A. Capietto, S. Coriasco, B. Doubrov, A. Kozlovskiy and N. Murru: *Axessibility 2.0: creating tagged PDF documents with accessible formulae*, Guit Meeting 2019, Ars TeXnica, vol. 27/28, p. 138-145, 2019
- [7] D. Ahmetovic, T. Armano, C. Bernareggi, A. Capietto: *Utilizzare grafici di funzione accessibili nelle attività di Moodle con AUDIOFUNCTIONS.WEB*, MoodleMoot Italia 2019, Atti del convegno, p. 65-70, 2020.
- [8] Berruto, G., *Fondamenti di sociolinguistica*. Laterza (1995), pp. 74-79.
- [9] Berruto, G., *Fondamenti di sociolinguistica*. Laterza (1995), p. 148.
- [10] Berruto, G., *Sociolinguistica dell'italiano contemporaneo*. La Nuova Italia Scientifica (1987) pp. 62-99.
- [11] D'Achille, P., *L'italiano contemporaneo*. Il Mulino (2003), pp. 79-88.
- [12] Sabatini, F., "*L'italiano dell'uso medio*": *una realtà tra le varietà linguistiche italiane* (1985), in *Gesprochenes Italienisch in Geschichte und Gegenwart*, Holtus, G. & Radtke, E. (a cura di). Gunter Narr Verlag, pp. 154-184.