



Conferenza  
GARR 2021  
sostenibile  
digitale

# Riconoscimento vocale di formule

Matteo Amore, Tiziana Armano, Cristian Bernareggi, Anna Capietto,  
Sandro Coriasco, Roberta Crespan, Mattia Ducci, Maria Luisa Gabrielli,  
Mariafrancesca Guadalupi, Alessandro Mazzei, Antonio Mazzei, Adriano  
Sofia, Francesco Tarasconi

Dipartimento di Matematica, Dipartimento di Informatica – Università di Torino  
CELI s.r.l.



# LABORATORIO "S. POLIN "

PER LA RICERCA E LA SPERIMENTAZIONE DI  
NUOVE TECNOLOGIE ASSISTIVE PER LE STEM

Responsabile scientifica Prof.ssa Anna Capietto



DIPARTIMENTO  
DI MATEMATICA  
GIUSEPPE PEANO  
UNIVERSITÀ DI TORINO

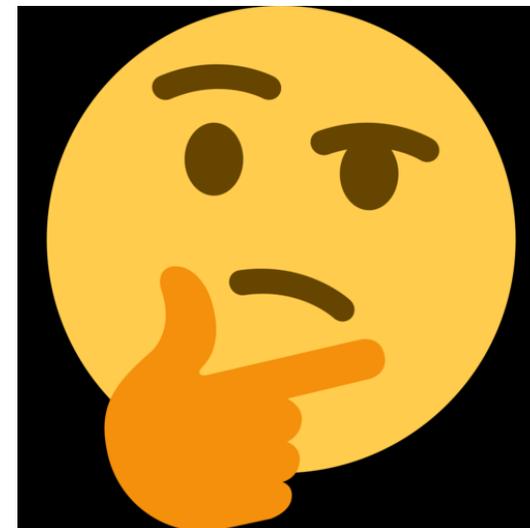
- Dipartimento di Matematica "G.Peano" Università di Torino

# Il problema

Sviluppo elevato e rapido di tecnologie

Elevata richiesta di competenze scientifiche e tecnologiche

Le tecnologie, le tecnologie assistive e gli strumenti compensativi non gestiscono in modo opportuno i contenuti scientifici (formule e grafici)



# Le formule

Sono oggetti non lineari e contengono simboli e notazioni particolari

$$\phi = \frac{1 + \sqrt{5}}{2}$$

# Matematica a voce

Gli attuali sistemi di riconoscimento e di interazione vocale hanno problemi con le formule

Progetto *Matematica a Voce*

cofinanziato da Fondazione CRT,  
in collaborazione con CELI s.r.l.

# Matematica a voce

1

SpeechMathE

Dettatura di  
formule

2

VoiceMath

Trascrizione di  
formule

# SpeechMatE

Le persone con disabilità motorie traggono numerosi vantaggi dalle applicazioni con interazione vocale

## SpeechMatE per :

- dettare espressioni matematiche in italiano;
- semplificare le espressioni matematiche attraverso la voce.

# SpeechMatE

Trasforma espressioni matematiche lette a voce in LaTeX, un linguaggio per rappresentare la matematica in forma digitale;

- mostra l'espressione matematica dettata in un editor;
- permette la modifica dell'espressione nell'editor mediante semplici comandi vocali.

# SpeechMatE

The screenshot displays the PolinSpeech interface. On the left, a green bar says "Speak now. Math mode!". Below it, the text "ICS alla terza fine" is visible. At the bottom, there are dropdown menus for "Italiano" and "Italia". The central part of the interface shows a LaTeX document editor with the following code:

```
%opening  
\title{  
\author{  
\begin{document}  
\maketitle  
\begin{abstract}  
|[10x^{3}y^{3}]/(-\frac{5y^{4}}{0})|  
\end{abstract}  
\section{  
\end{document}
```

A blue highlight is under the abstract code. To the right, a mobile device preview shows a document page with the number "1" at the top right.

April 7, 2020

Abstract

$$10x^3y^3 / (-\frac{5y^4}{0})$$

1

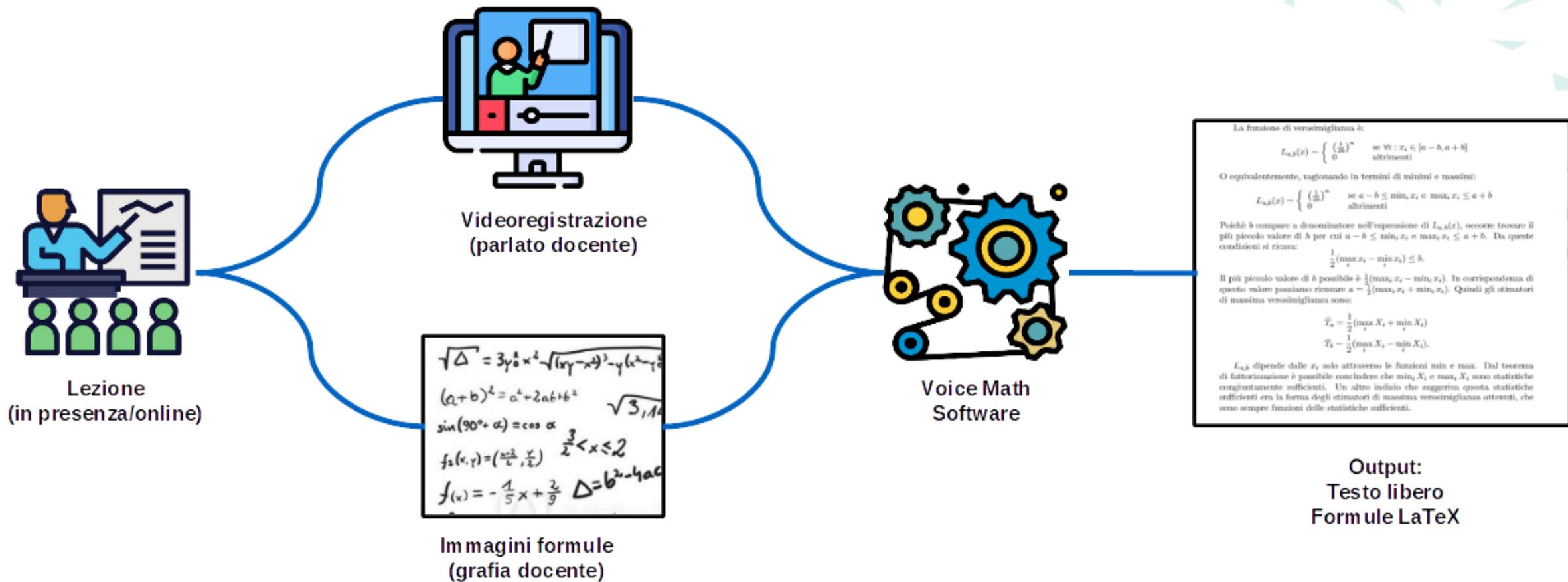
# VoiceMath

Trascrizione e sottotitolazione di lezioni con formule a vantaggio di persone con disabilità e DSA.

Trascrizione e sottotitoli con formule in linguaggio naturale o in LaTeX.



# VoiceMath



La funzione di verosimiglianza è:

$$L_{a,b}(x) = \begin{cases} \left(\frac{x}{b}\right)^a & \text{se } \forall i: x_i \in [a-b, a+b] \\ 0 & \text{altrimenti} \end{cases}$$

O equivalentemente, ragionando in termini di minimi e massimi:

$$L_{a,b}(x) = \begin{cases} \left(\frac{x}{b}\right)^a & \text{se } a-b \leq \min_i x_i \text{ e } \max_i x_i \leq a+b \\ 0 & \text{altrimenti} \end{cases}$$

Poiché è comparso a denominatore nell'espressione di  $L_{a,b}(x)$ , occorre trovare il più piccolo valore di  $b$  per cui  $a-b \leq \min_i x_i$  o  $\max_i x_i \leq a+b$ . Da queste condizioni si ricava:

$$\frac{1}{2}(\max x_i - \min x_i) \leq b.$$

Il più piccolo valore di  $b$  possibile è  $\frac{1}{2}(\max x_i - \min x_i)$ . In corrispondenza di questo valore possiamo ricavare  $a = \frac{1}{2}(\max x_i + \min x_i)$ . Quindi gli stimatori di massima verosimiglianza sono:

$$\hat{T}_a = \frac{1}{2}(\max X_i + \min X_i)$$
$$\hat{T}_b = \frac{1}{2}(\max X_i - \min X_i).$$

$L_{a,b}$  dipende dalle  $x_i$  solo attraverso le funzioni min e max. Dal teorema di fattorizzazione è possibile concludere che  $\min X_i$  e  $\max X_i$  sono statistiche congiuntamente sufficienti. Un altro indizio che suggeriva questa statistiche sufficienti era la forma degli stimatori di massima verosimiglianza ottenuti, che sono sempre funzioni delle statistiche sufficienti.

# VoiceMath

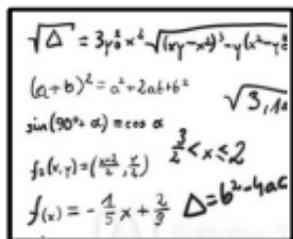


Videoregistrazione

Preprocessing  
audio

Trascrizione parlato  
(Speech-To-Text)

Riconciliazione  
flussi



Immagini Formule

Preprocessing  
immagini

Conversione formule  
(Computer Vision)



Output (Visualizzatore)



# VoiceMath - Sfide

- Qualità dell'input (audio e video)
  - Dipendente dallo strumento
  - Dipendente dalla persona
- Linguaggio di dominio

# GRAZIE!!

<http://www.integr-abile.unito.it/>

tiziana.armano@unito.it

mariafrancesca.guadalupi@h-farm.com

