

**Speech recognition
helps visually impaired people
writing mathematical formulas**

Tomáš Hanakovič, Marek Nagy
Comenius University, Bratislava

Introduction

- Mathematical expressions are common in many science domains – important
- We can effectively use computer to write them
- There are many WYSIWYG editors, but are these tools adaptable for disabled people?
- Non-WYSIWYG solution can have too complicated syntax

Motivation - multimodal approach

- Most editors use only one modality (mainly visual). Why the restriction?
- Multimodal (visual and aural) application offers more freedom
- Each modality has it's own advantages:
 - Using mouse can speed up navigation
 - Using voice can speed up entering new data
- Voice helps visually impaired people

Dictation and mathematical formulas

Ambiguity of formula dictation

- Utterance „*a i plus b j*“ has at least two possible interpretations:

$$ai + bj$$

or

$$a_i + b_j$$

- Which one is correct? Intuition: the second one. Can we count on intuition? No.
- What is the main problem here?

2D to 1D – linearisation

- Mathematical expressions have 2D spatial structure (over, under, upper left index,...)
- Dictation: producing sequence of words – one dimensional structure (vector)
- Linearisation – lossless transformation of 2D structure to 1D using syntactical mark-up
- Must be intuitive and still exact

Analyse of formula structure

- Several structure categories (function, fraction, index,...)
- Recurrence (fraction in fraction)
- Logical separation to groups – three blocks:

$$a^2 + 2ab + b^2$$

Navigation

- How to mark (navigate to) specific element easily? (note error marked red):

$$\sin(u + v) = \sin u \cos b + \cos u \sin v$$

- Solution: hierarchical numbering:
 - allows to fast and intuitively mark any element of formula
 - Procedural nature – consists from block selections and expansions

Hierarchical numbering – example

$$\begin{array}{|c|} \hline \textit{sine} \\ \hline \sin(u+v) \\ \hline 1 \\ \hline \end{array} = \begin{array}{|c|c|} \hline \textit{sine} & \textit{cosine} \\ \hline \sin u & \cos b \\ \hline 1 & 2 \\ \hline \end{array} + \begin{array}{|c|c|} \hline \textit{cosine} & \textit{sine} \\ \hline \cos u & \sin v \\ \hline 3 & \\ \hline \end{array}$$

$$\sin(u+v) = \begin{array}{|c|c|} \hline \textit{sine} & \textit{cosine} \\ \hline \sin u & \cos b \\ \hline 1 & 2 \\ \hline \end{array} + \cos u \sin v$$

$$\sin(u+v) = \sin u \cos \begin{array}{|c|} \hline \textit{value} \\ \hline b \\ \hline 1 \\ \hline \end{array} + \cos u \sin v$$

Application – key facts

- Application for entering mathematical formulas
- Multimodal – modalities are independent, but can cooperate.
- Simple, easy to use, good orientation in formula
- Clear and colourful visual and voice output
- Highly configurable, customizable

Resume

- 1.Dictation can be ambiguous, use good linearisation
- 2.Hierarchical approach to navigation with block selections/expansions
- 3.Rich command set with synonyms and optional parameters
- 4.Colourful output

Thank you for your attention!