ORCCA

Ontario Research Centre for Computer Algebra

Νοτατιον Σελεκτιον ιν Μαθεματικαλ Κομπυτινγ Ενωιρονμεντσ

Ελενα Σμιρνοωα ανδ Στεθεν Μ. Ωαττ

Ονταριο Ρεσεαρχ Χεντρε φορ Κομπυτερ Αλγεβρα

Της Υνιωερσιτψ οφ Ωεστερν Ονταριο

ε–μαιλ: {ελενα, ωαττ}@ορχχα.ον.χα

ORCCA
Ontario Research Centre for Computer Algebra

Notation Selection in Mathematical Computing Environments

Elena Smirnova and **Stephen M. Watt**

Ontario Research Centre for Computer Algebra

The University of Western Ontario

e-mail: {elena, watt}@orcca.on.ca

▪ Simple test 1

How do we write the following?

1. the open interval with end points a and b:

2. the base 10 logarithm of x:

3. the inner product of p and q:

4. the arctangent of 2:

▪ Simple test 1

How do we write the following?

1. the open interval with end points *a* and *b*: $]a, b[$
2. the base 10 logarithm of x: $\lg x$
3. the inner product of *p* and *q*: $p \llcorner q$
4. the arctangent of 2: $\mathrm{tg}^{-1} 2$

▪ Simple test 2

What does it mean for you?

1. lg x

2. $\tan^{-1} a$

3. $\left(\frac{p}{q}\right)$

4. u'

▪ Simple test 2

What does it mean for you?

1. lg x
   - ○ $\log_{10} x$
   - ○ $\log_2 x$

2. $\tan^{-1} a$
   - ○ cotangent a
   - ○ arctangent a

3. $\left(\frac{p}{q}\right)$
   - ○ Legendre symbol
   - ○ Parenthesized fraction

4. u'
   - ○ derivative of u
   - ○ u minutes
   - ○ transformation performed on an original u

▪ Origin of notational ambiguity

- **Mathematical context**
  - an ordinary derivative can be denoted as

    f', f_x, Df , $\frac{df}{dx}$ etc.

- **Area of application**
  - *i* for $\sqrt{-1}$ (complex analysis) vs.

    j (electrical engineering).
  - the integral

    $\int f(x)\mathrm{d}x$ (pure math) vs. $\int \mathrm{d}x\, f(x)$ (physics)

▪ Origin of notational ambiguity (cont-d)

- **National and cultural conventions**
  - the tangent function:

    tan (England, Canada) vs. *tg* (Russia, China);

  - the open interval:

    (*a, b*) (United States) vs.]*a, b*[(France).

- **Historical period**
  - e.g. ancient $3\overline{a+b}$ vs. the modern 3(*a* + *b*).

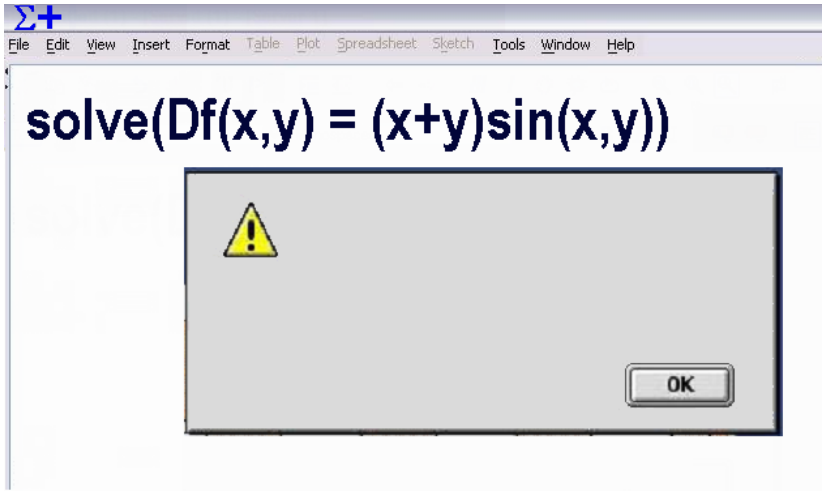
- **Level of mathematical sophistication**
  - e.g. $a \div b$ vs. $b\overline{)a}$ vs. $\frac{a}{b}$ vs. *a/b* vs. a/b.

▪ Who needs to understand written math?

- Human readers
- Software tools:
  - CAS,
  - theorem provers,
  - format converters,
  - pen-based applications

▪ What happens if math is misunderstood?

- Human readers
  - confusion
  - frustration
  - headache



- Software tools
  - unexpected results
  - incorrect calculation
  - hidden errors

*Maple 10 - Untitled (1) - [Server 1]

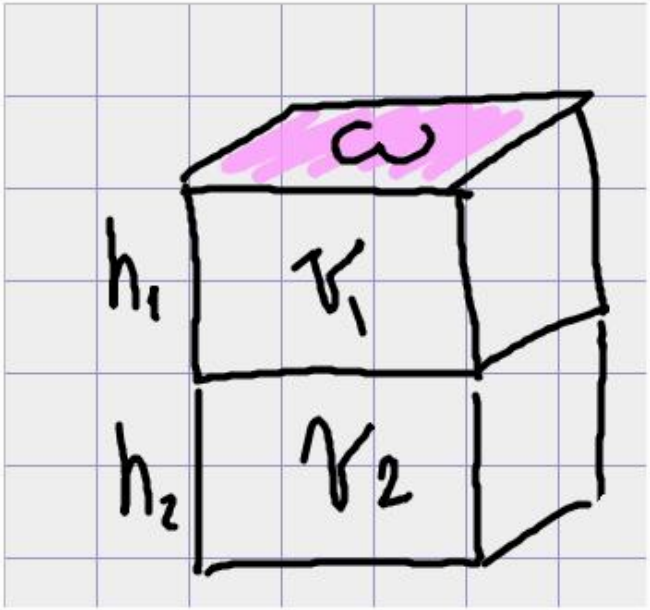
File Edit View Insert Format Table Plot Spreadsheet Sketch Tools Window Help

Text Math Error Monospaced 24

$$h_1 := 5'7"$$

$$h2 := 8'10"$$

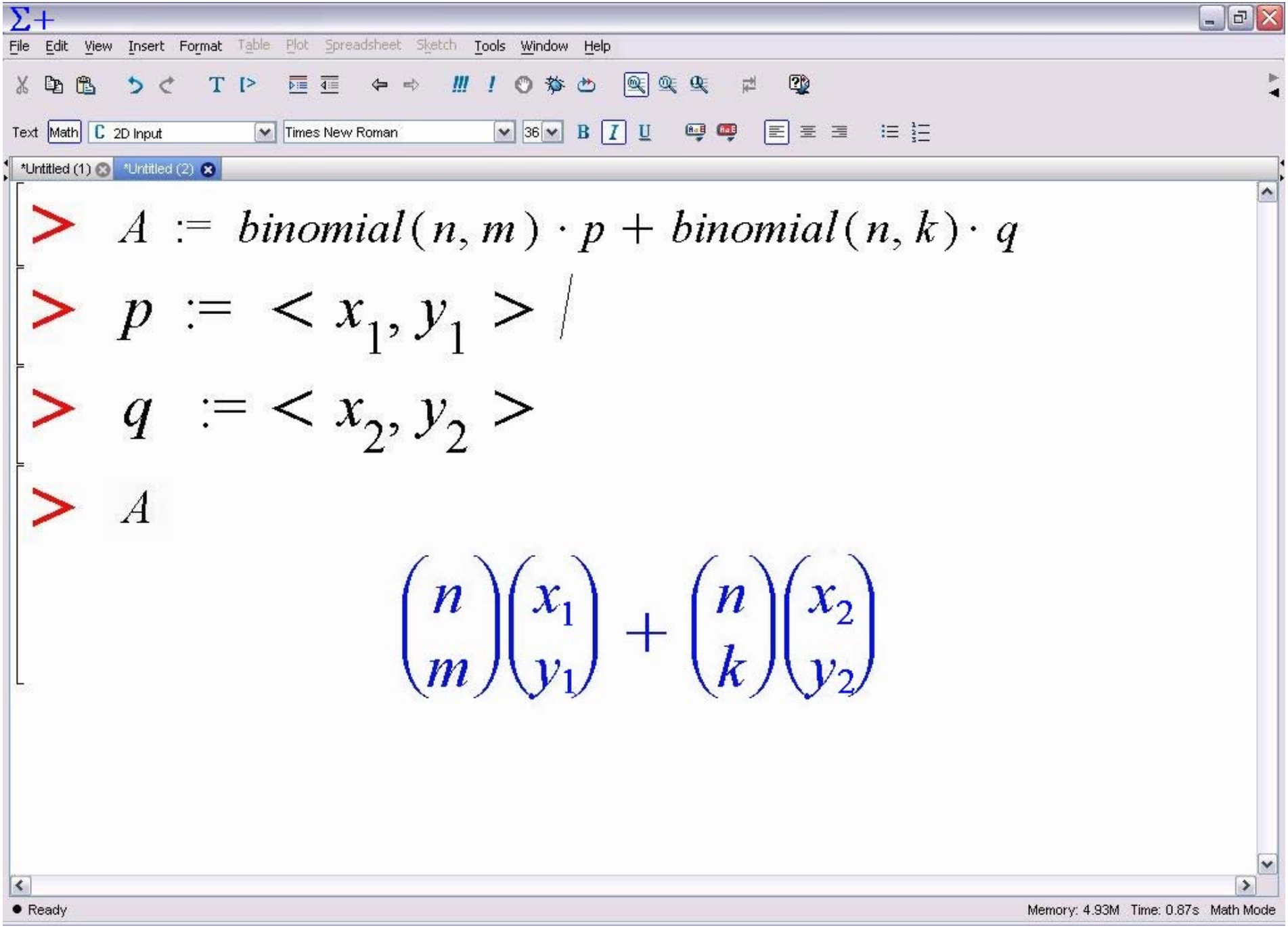
$$S_w = \frac{V_1 + V_2}{(h_1 + h_2)}$$



Error, numeric exception: division by zero

Ready

Memory: 5.62M Time: 9.53s Text Mode


File Edit View Insert Format Table Plot Spreadsheet Sketch Tools Window Help
Text Math 2D Input Times New Roman 36
*Untitled (1) *Untitled (2)
> A := binomial(n, m) · p + binomial(n, k) · q
> p := < x1, y1 >
> q := < x2, y2 >
> A
Ready
Memory: 4.93M Time: 0.87s Math Mode

▪ Mathematical content vs. Notation

Software often confuses

mathematics with ***notation***

▪ Separating notation from content

1. To choose which of several

   different mathematical notations to use

   for *the same concept*.

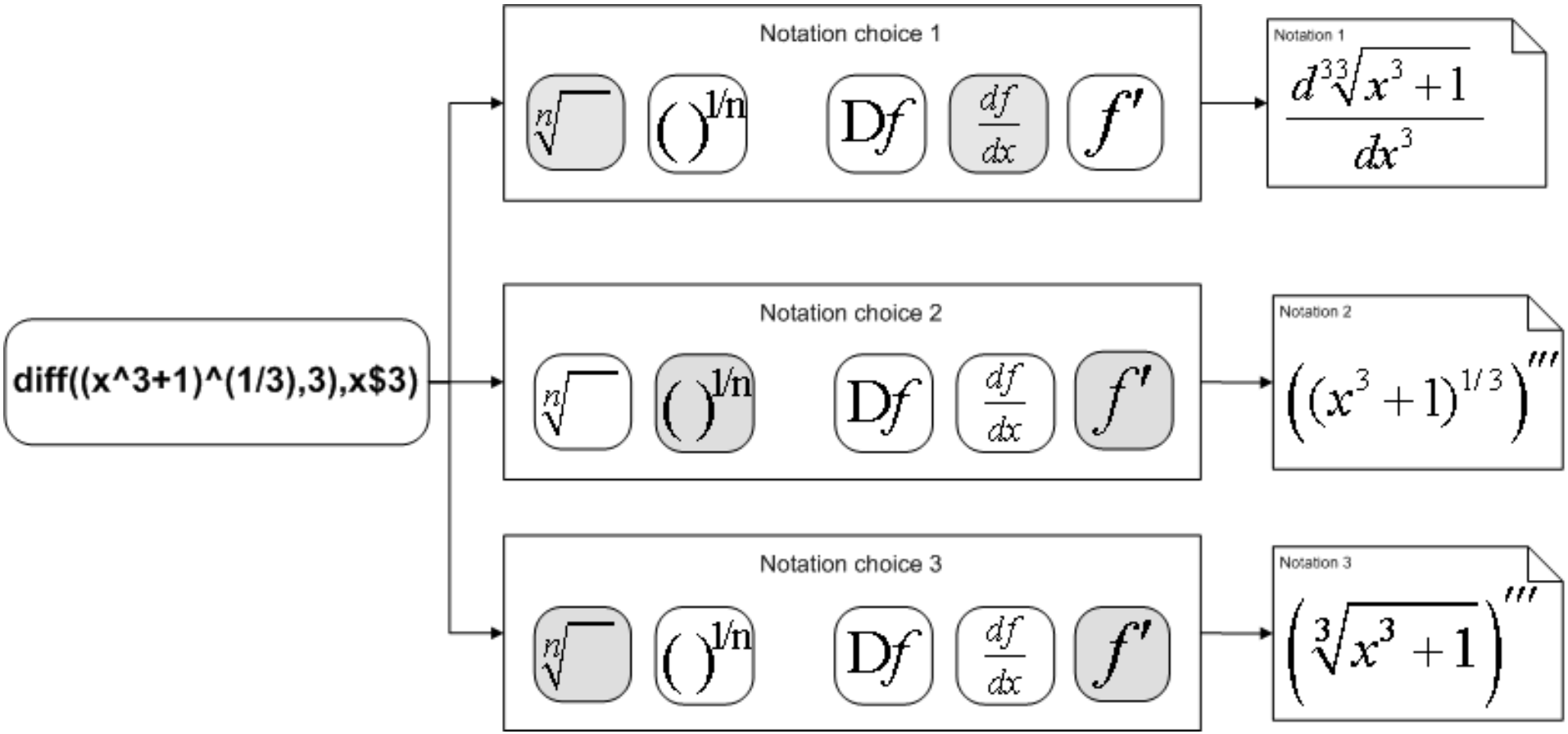
2. To disambiguate where

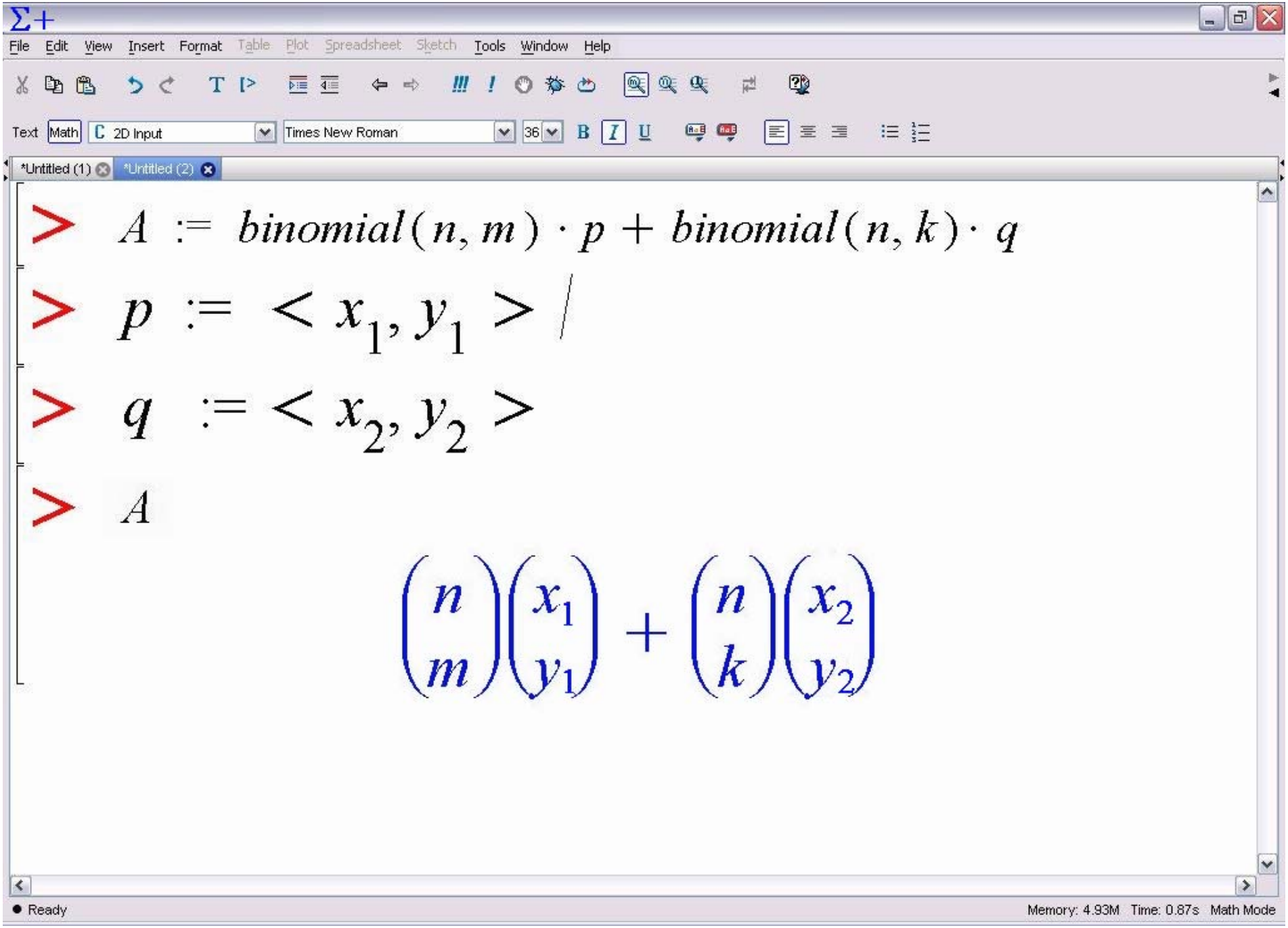
   the *same notation* could be used

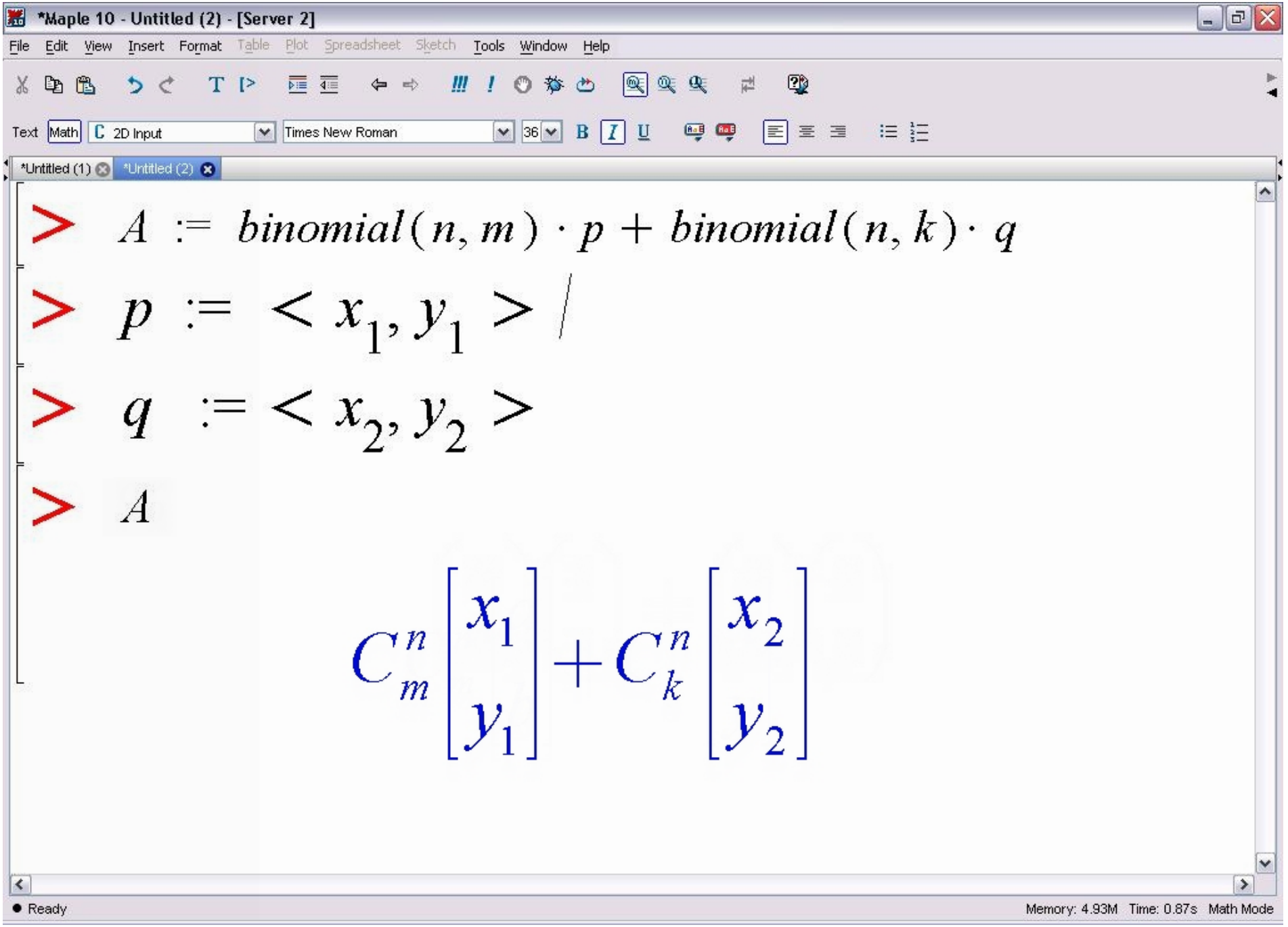
   for *different concepts*.

▪ Rendering math with different notations

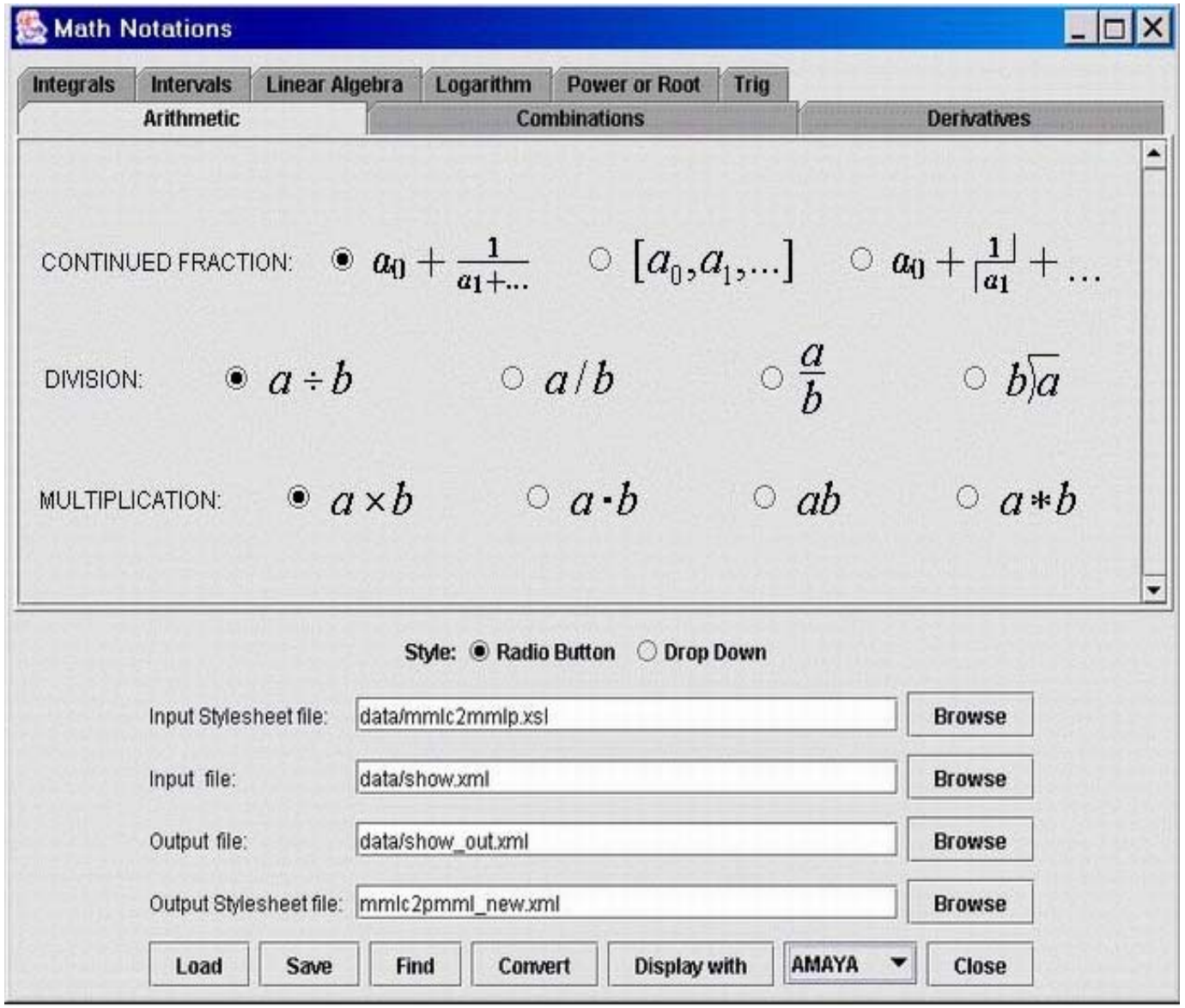
- Making your math look the way you want




File Edit View Insert Format Table Plot Spreadsheet Sketch Tools Window Help
Text Math 2D Input Times New Roman 36
*Untitled (1) *Untitled (2)
> A := binomial(n, m) · p + binomial(n, k) · q
> p := < x1, y1 >
> q := < x2, y2 >
> A
Ready
Memory: 4.93M Time: 0.87s Math Mode


*Maple 10 - Untitled (2) - [Server 2]
File Edit View Insert Format Table Plot Spreadsheet Sketch Tools Window Help
Text Math 2D Input Times New Roman 36
*Untitled (1) *Untitled (2)
> A := binomial(n, m) · p + binomial(n, k) · q
> p := < x1, y1 >
> q := < x2, y2 >
> A
C_m^n [x1, y1] + C_k^n [x2, y2]
Ready
Memory: 4.93M Time: 0.87s Math Mode

▪ A Notation Selection Tool outlook



▪ A Notation Selection Tool spec

- Designed to drive the conversion from *XML-based* ***conceptually*-oriented** *math documents* into ***notationally-oriented*** formats

- It **generates rules** to translate between math expressions in different XML formats

- Main feature: **extensibility –**
  user can introduce additional math concepts along with the set of possible notations

A Notation Selection Tool anatomy

- Java program + web service
- Set of target domains
- Configuration file
- Initial XSLT stylesheet

▪ Notation Selection Tool configuration

- Notation Selection Tool is just a software engine. Its actual behaviour depends on the configuration.

- Configuration defines
  - target mathematical domains,
  - concepts in each domain,
  - set of alternative notations for each concept
  - transformation rules for each notational choice

- Core of the tool is the *configuration file*

The configuration file

```
<catalog>
  <name> LINEAR ALGEBRA </name>
  <itemlist>
    <item>
      <keyword> INNER PRODUCT </keyword>
      <choicelist>
        <choice>
          <image src = "inn_prod1.gif"/>
          <keyvalue> 1 </keyvalue>
          <presentation>
              <converter input="C MathML" output="P MathML>
                ... <!-- XSLT template for this notation-->
            </converter>
          </presentation>
        </choice>
       %other choices,
     %other items,
%other catalogs
```

▪ Features of the Notation Selection Tool

Advantages our this approach:

flexibility and **extensibility**.

- To *introduce new notations* for existing math concepts the user needs simply update Notation Selection Tool configuration file.

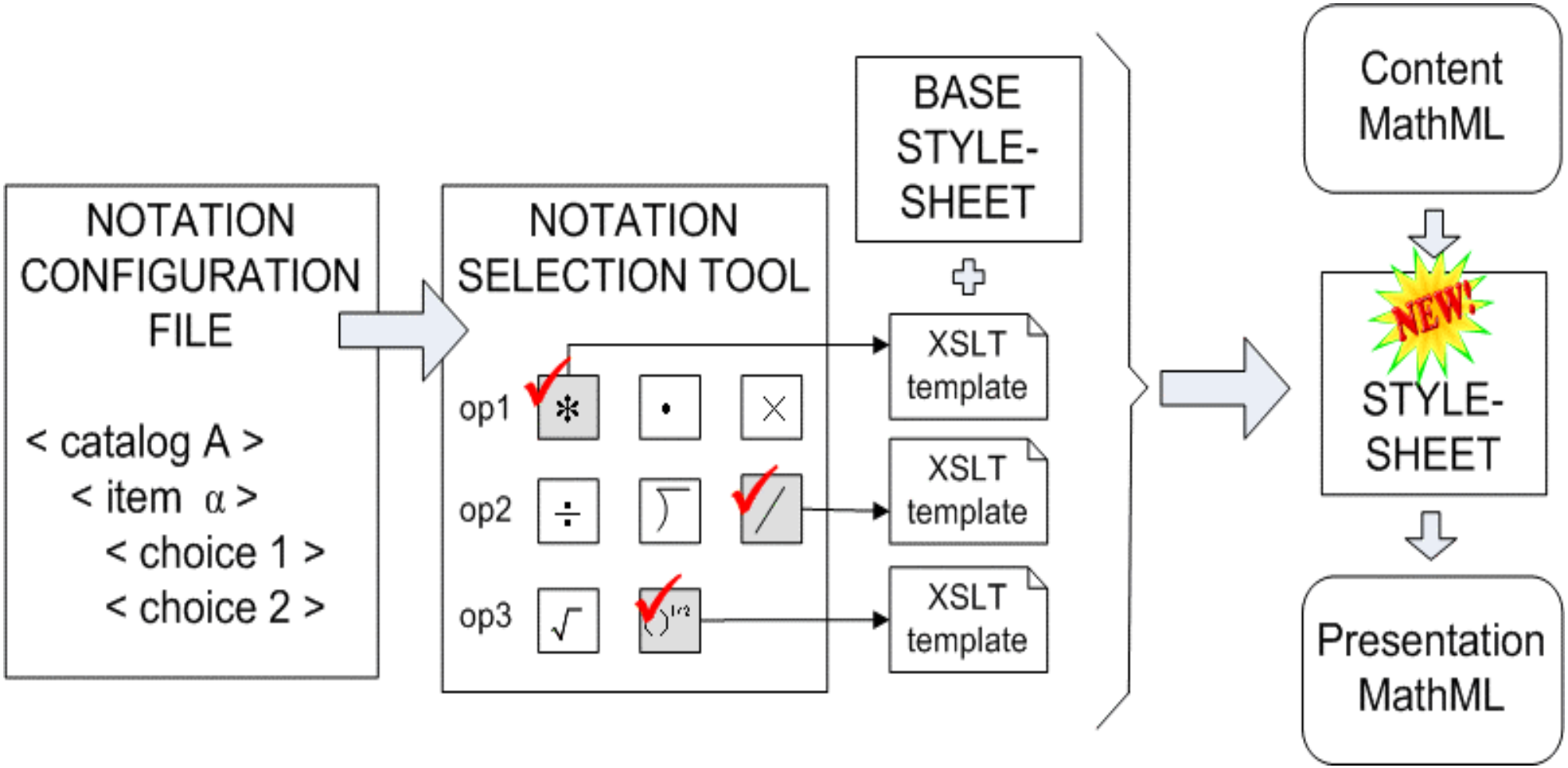
▪ Features of the Notation Selection Tool (2)

Advantages our this approach:

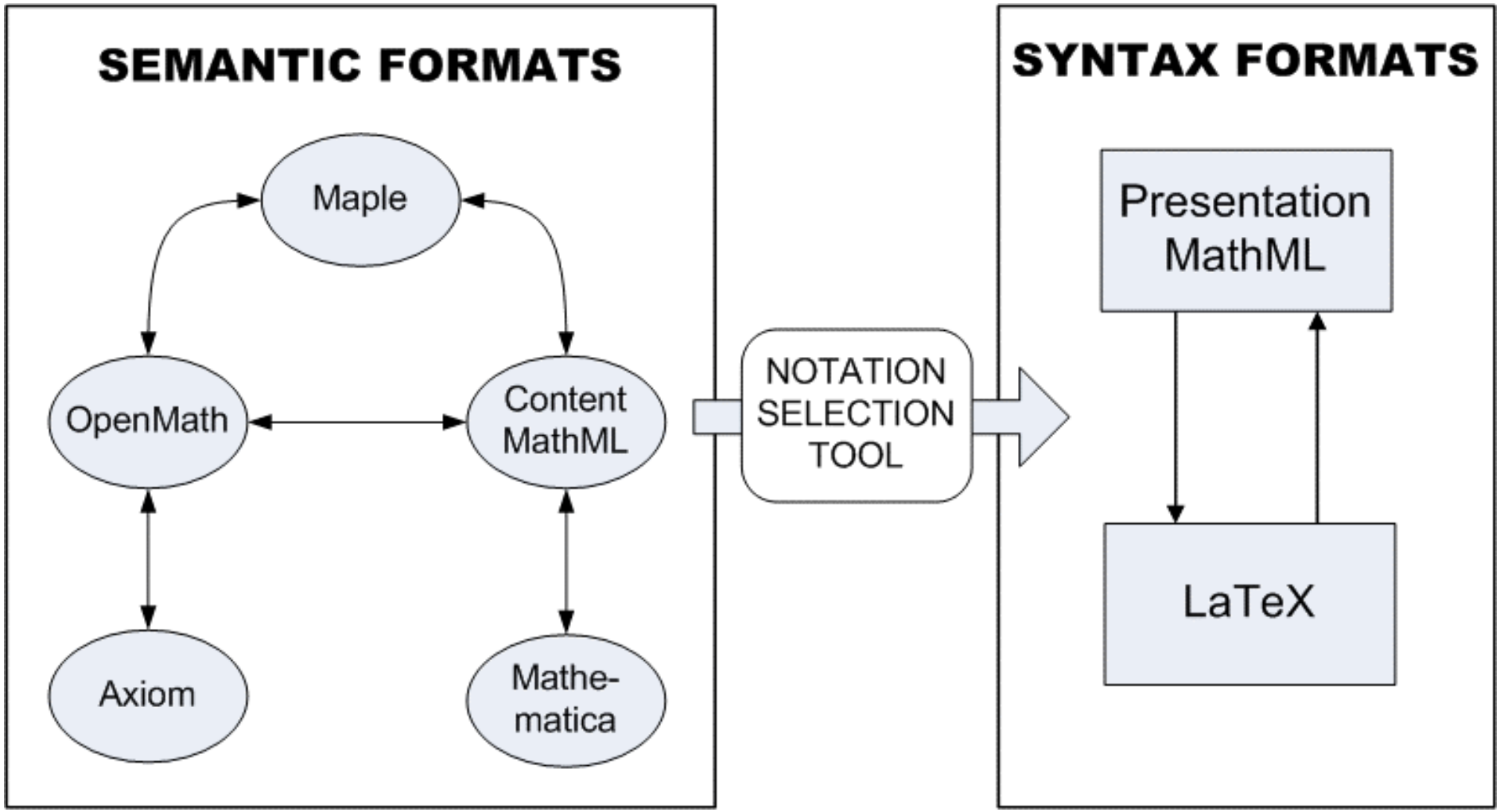
flexibility and **extensibility**.

- *New mathematical concepts* can be introduced in existing settings.
  - e.g. binomial or continued fractions are defined neither in Content MathML, nor in Presentation MathML, but they can be introduced as additional stylesheet templates.
  - The same approach allows to set preferred rendering for OpenMath CDs

▪ Notation Selection Tool in action

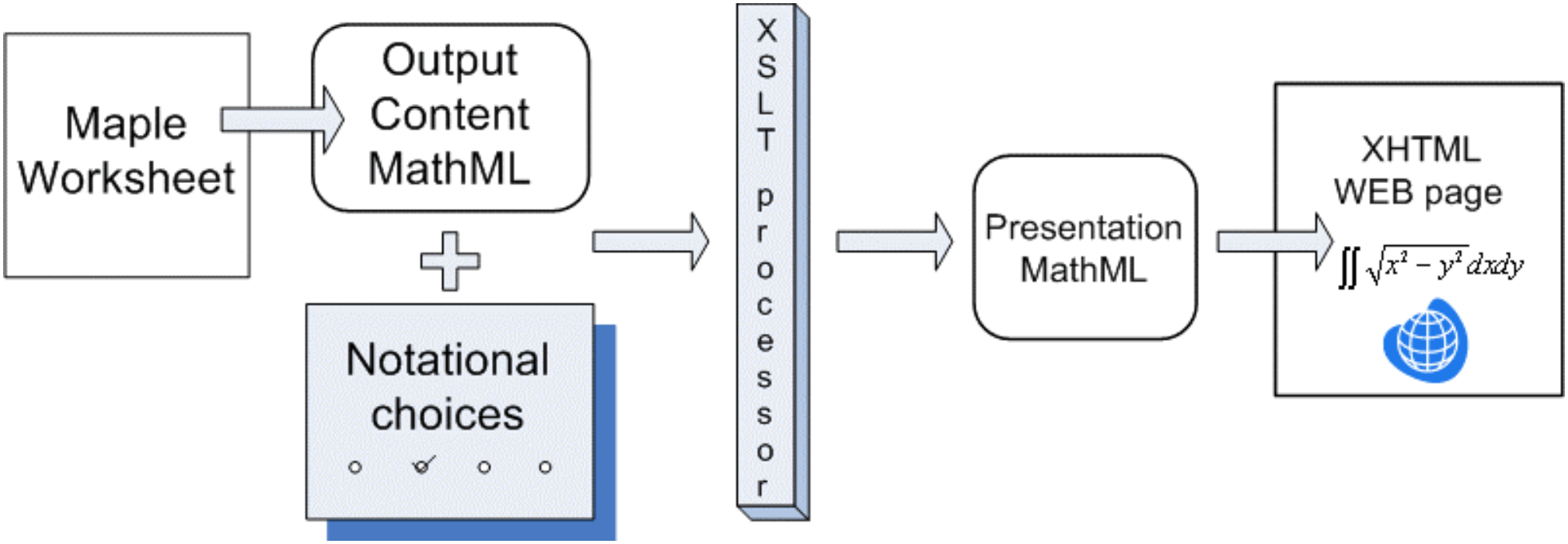


▪ Mathematical Data formats involved



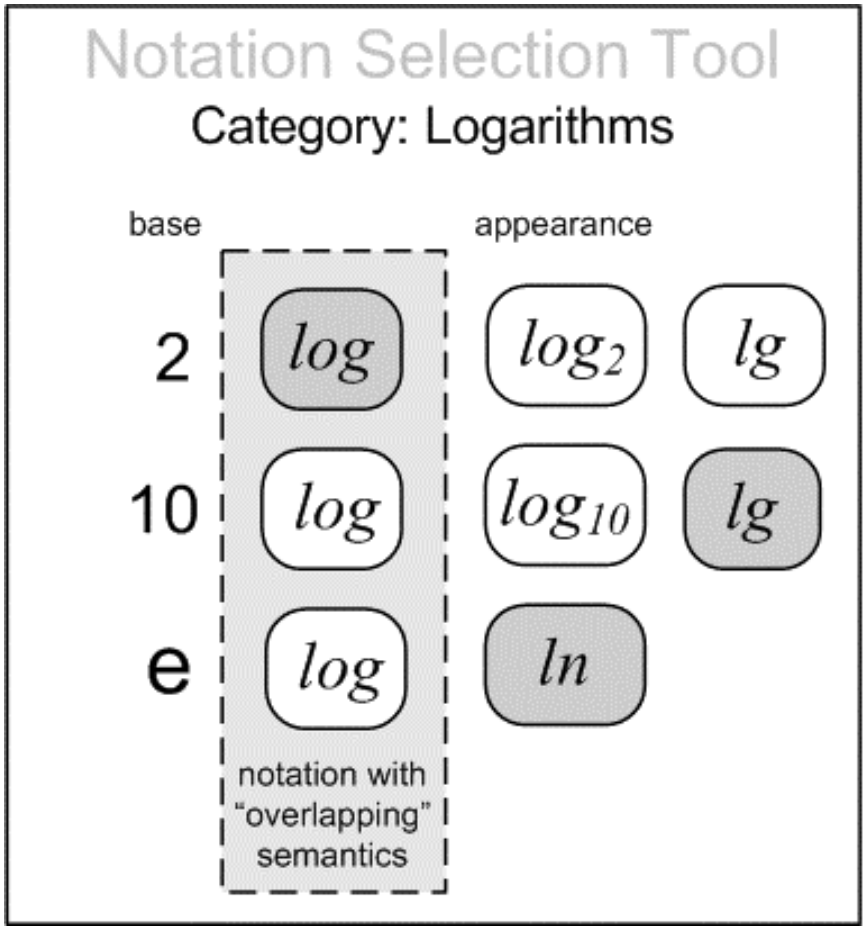
▪ Application of the notation selection (1)

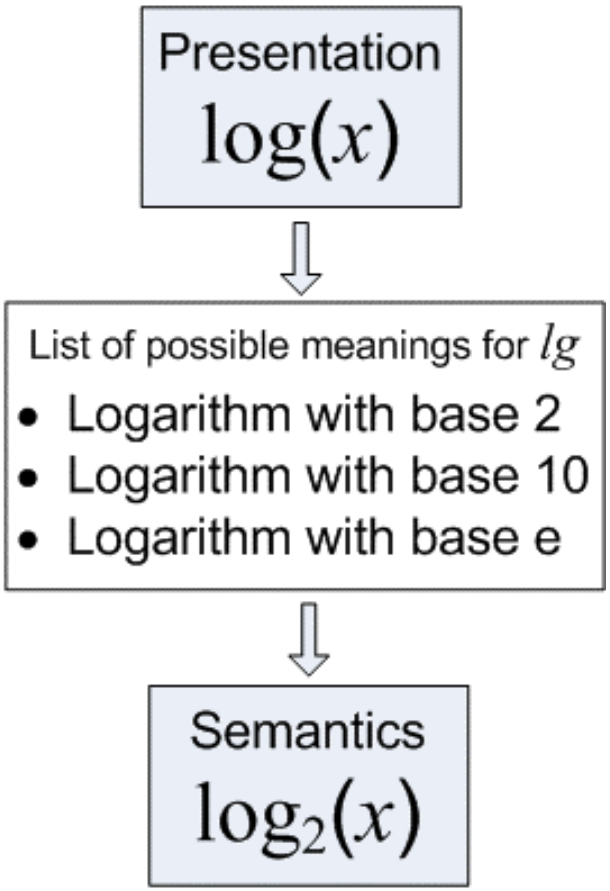
- Rendering mathematical content produced by CAS



▪ Application of the notation selection (2)

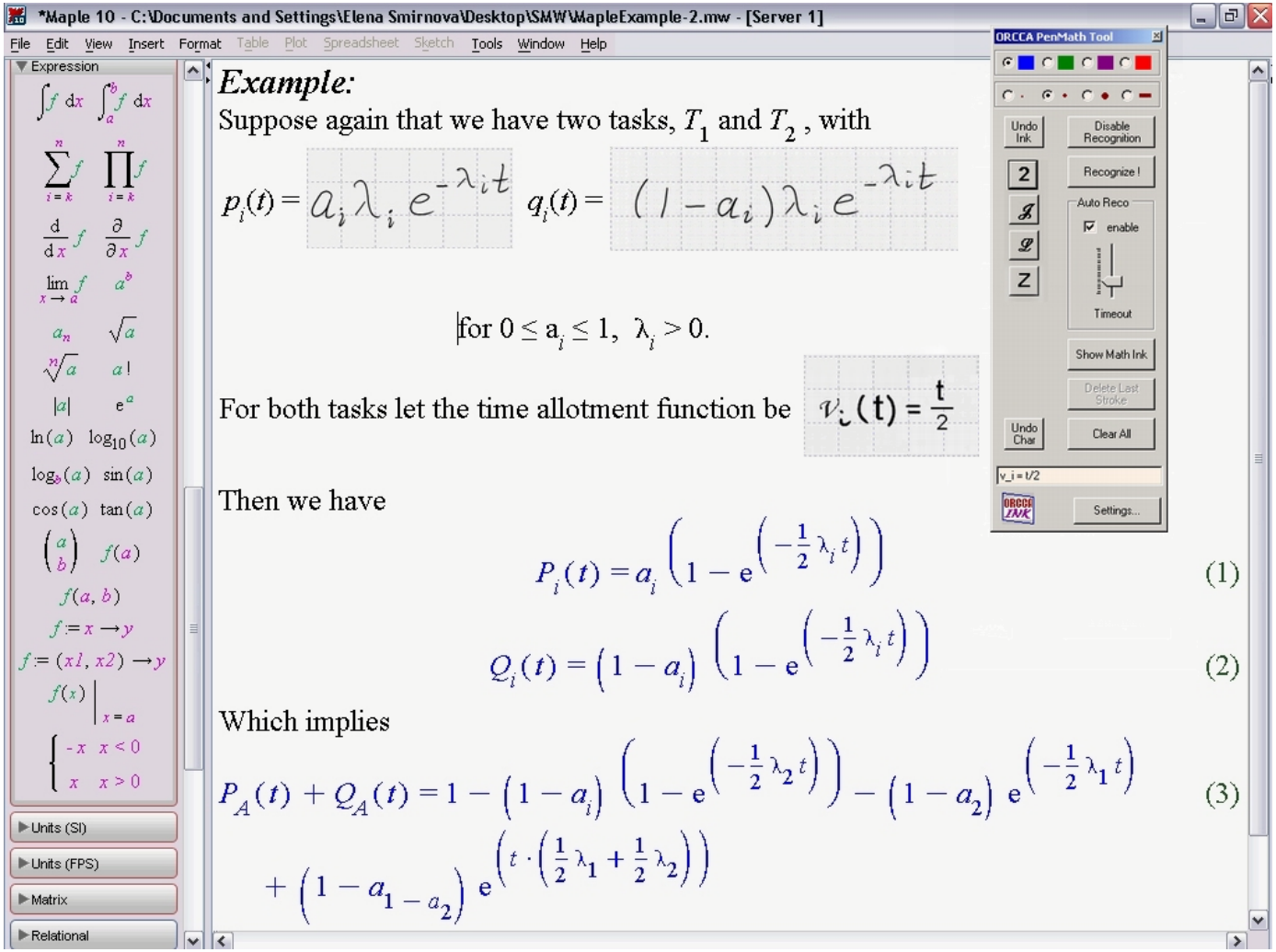
- Syntax disambiguation





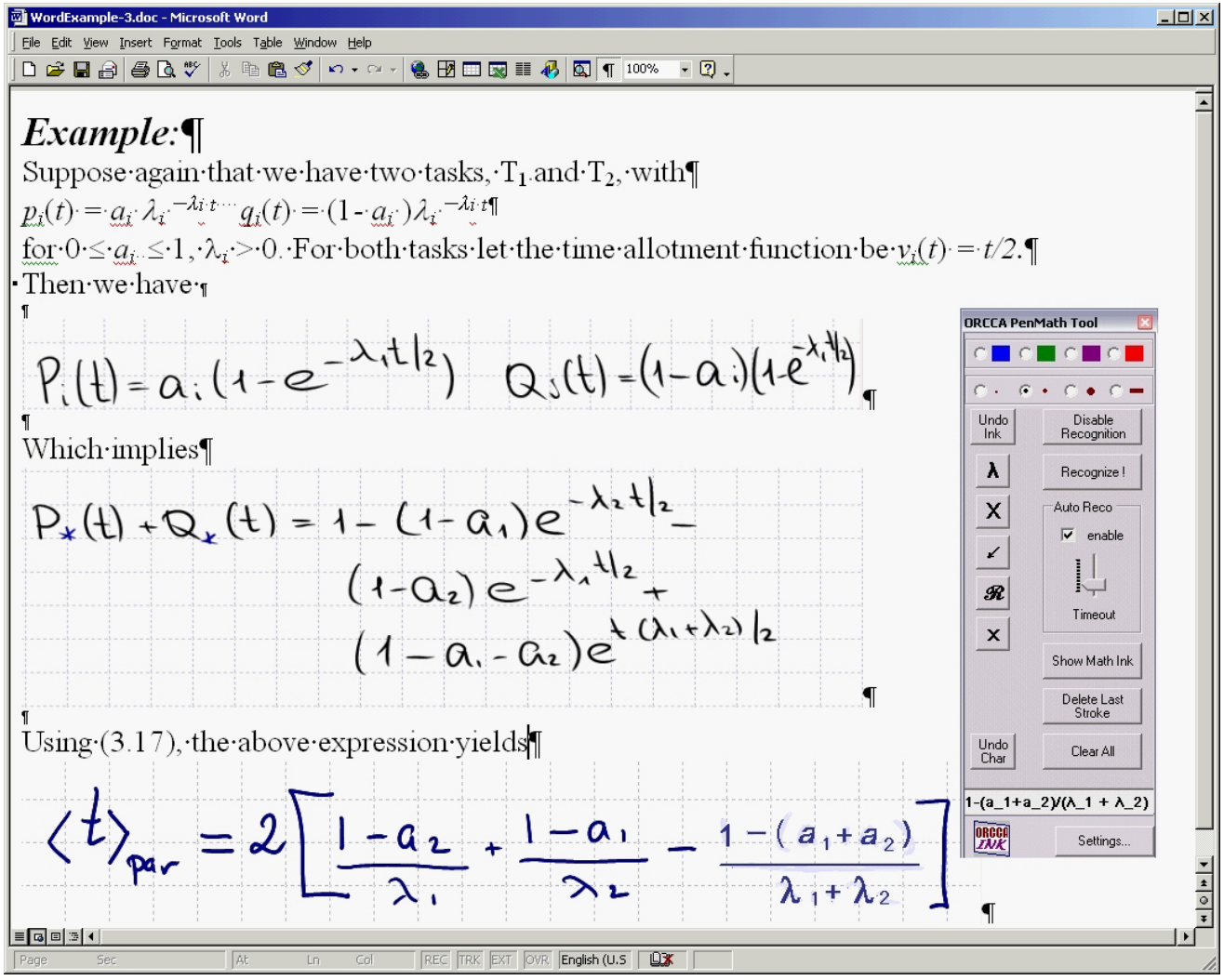
▪ Application of the notation selection (3)

- Mathematical handwriting recognition
  - Computer Algebra Systems



▪ Application of the notation selection (3)

- Mathematical handwriting recognition
  - Document-processing software



▪ Application of the notation selection (3)

- Mathematical handwriting recognition

Helps to narrow down a set of recognition candidates without restricting user's writing habits:

- to eliminate un-used notations.

  Ex. if $\sim$ is chosen for proportionality, then

  $\propto$ will not be suggested as a candidate for α or ∞

- to activate/deactivate various rules in structures recognition.

  Ex.]a,b] or $\int \mathrm{d}x\, f(x)$

▪ Application of the notation selection (4)

- Mathematical knowledge management

  NS used as a front end for mathematical knowledge databases could help to

  - avoid storing duplicated information
  - allow queries without forcing a particular notation

Ex. sinh(x) and *sh(x)*

Should *arcsin(x)* – $sin^{-1}(x)$ simplify to 0?

▪ Application of the notation selection (5)

- Computer Algebra Systems

  The most widely used math software packages are designed with *input* and *output* systems **separate** from one another.

  - NS is aimed to ensure consistency in *both* input and output notations

  - It allows to set the appearance of new math expressions, generated by CAS, such as $J_0(z)$ instead of bulky `BesselJ[0](z)`

▪ Application of the notation selection (sum)

- Rendering mathematical content
- Syntax disambiguation
- Mathematical handwriting recognition
- Mathematical knowledge management
- Computer Algebra Systems

▪ Conclusions

- Notation ambiguity often takes place in mathematics
- Software often confuses mathematics with notation
- By forcing too early the choice of notations, flexibility is lost and work is restricted to too narrow context
- Special tools can be used to translate meaningful math constructs both **to** and **from** a wide range of notations
- This will allow math documents to be deployed in a wide range of settings with notation customized by country, field, level of sophistication etc.

Conclusions

- Notation Selection Tool is an interactive stylesheet constructor dealing with Content and Presentation MathML

- By using it in conjunction with other translators NST covers the following conversions:

Content MathML OpenMath, Maple Axiom Mathematica	→	Presentation MathML LaTeX

- NS is promising to be useful in application to
  - Context of Computer Algebra
  - Syntax disambiguation
  - Pen-based mathematical interfaces
  - Mathematical knowledge management

Related projects at ORCCA

Software for mathematical interfaces

math communication and math format conversions:

- MathML $\leftrightarrow$ LaTeX
- Content MathML $\rightarrow$ Presentation MathML
- OpenMath $\leftrightarrow$ Content MathML
- Maple $\leftrightarrow$ OpenMath
- Pen-Based Interfaces for Mathematics

http://www.orcca.on.ca