

Matematica pe Web

O introducere în MathML

– Sabin Corneliu Buraga

MathML (Mathematical Markup Language) – începând cu luna martie 2001, versiunea curentă fiind MathML 2.0 – reprezintă o recomandare relativ recentă a Consorțiului World Wide Web pentru descrierea notațiilor din matematică. Scopul acestei recomandări (totodată scopul pentru care a fost creat MathML) este acela de a realiza o metodă prin care elementele matematice să fie procesate pe Web într-un mod asemănător procesării informațiilor textuale realizate cu HTML. MathML se adresează celor care dezvoltă și implementează utilitare de analiză și editare de informații cu specific matematic, putând fi ușor folosite în domenii științifice, ingineresti sau academice. MathML este o aplicație a metalingajului XML.

Problema codificării elementelor matematice pentru procesarea pe calculator a apărut cu mult înainte de dezvoltarea Web-ului. O metodă de adnotare a expresiilor matematice oferită de TeX a fost elaborată de Donald Knuth în anul 1986 și era (și este) utilizată pe scară largă înainte ca Web-ul să devină principalul beneficiar al metodelor de marcare.

Motivul pentru care a devenit necesară specificarea pentru Web a expresiilor matematice este că HTML-ul (sau, mai nou, înlocuitorul acestuia, XHTML-ul) nu este suficient de flexibil pentru a atinge acest deziderat și pune mari probleme la scrierea elementelor complexe din matematică. Până la MathML, metoda utilizată pentru introducerea unei formule complicate într-o pagină Web era folosirea unei imagini care reprezenta la nivel grafic acea formulă.

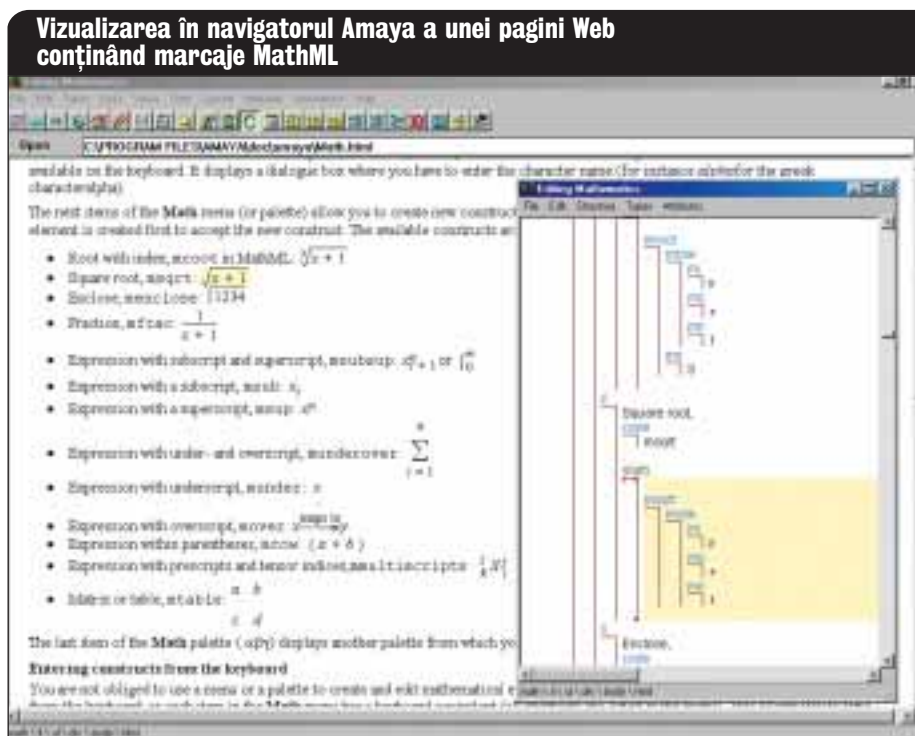
Pentru a veni în întâmpinarea dorințelor comunității științifice, MathML a fost proiectat să satisfacă următoarele scopuri:

- să codifice elementele matematice într-un mod ușor de prezentat și de învățat;
- să codifice notația matematică, dar și înțelesul matematic;
- să fie ușor de convertit în alte metode de prezentare;
- să poată fi utilizat prin intermediul șabloanelor sau altor metode de editare;
- să fie independent de platformă.

Tipuri de elemente MathML

În MathML există *elemente de prezentare*, *elemente de conținut* și *elemente de interfață*.

Vizualizarea în navigatorul Amaya a unei pagini Web conținând marcaje MathML



Elementele de prezentare descriu structura notației matematice. Exemple tipice sunt elementele `<mrow>` care indică o linie orizontală (rând) și `<msup>` specificând o putere. De asemenea, elementele de prezentare acceptă un număr mare de atribute.

Elementele de conținut descriu obiectele matematice în mod direct. Din această categorie fac parte elementul `<plus>` care denotă operatorul uzual de adunare și elementul `<vector>` specificând un vector din algebra liniară.

Toate elementele sunt de tip prezentare sau conținut, exceptând elementul `<math>` care joacă rol de element de cel mai înalt nivel (rădăcină), fiind element de interfață. Scopul acestuia este să trimită către procesorul MathML parametrii care vor afecta întreaga expresie și să comunice vizualizatorului MathML cum să afișeze expresia.

Exemple utilizând elementele de prezentare

Începem printr-un exemplu de reprezentare a unei ecuații de gradul doi, de forma: $x^2+4x+7=0$. Codul de marcaje MathML va fi următorul:

```
<mrow>
  <mrow>
    <msup>
      <mi>x</mi>
      <mn>2</mn>
    </msup>
    <mo>+</mo>
    <mrow>
      <mn>4</mn>
      <mo>&InvisibleTimes;</mo>
      <mi>x</mi>
    </mrow>
    <mo>+</mo>
    <mn>7</mn>
  </mrow>
</mrow>
<mo>=</mo>
<mn>0</mn>
</mrow>
```

Observăm apariția marcărilor `<mi>` (identificator matematic), `<mo>` (operator matematic) și `<mn>` (număr, constantă matematică). De asemenea, în acest exemplu se remarcă utilizarea entității `InvisibleTimes` folosită pentru reprezentarea operatorului de înmulțire care în mod normal este subînțeles. În scrierea

unui document MathML este necesară evidențierea tuturor operatorilor sau a altor elemente care apar într-o notație matematică, chiar dacă aceștia sunt subînțeleși în scrierea uzuală, de mână. Se poate observa că ecuația este descrisă exact cum apare în mod obișnuit, acesta fiind rolul elementelor de prezentare.

În ceea ce urmează, vom descrie în MathML rădăcinile ecuației de gradul doi, adică:

Fragmentul de document MathML va fi:

```
<mrow>
  <mi>x</mi>
  <mo>=</mo>
  <mfrac>
    <mrow>
      <mrow>
        <mo>-</mo>
        <mi>b</mi>
      </mrow>
      <mo>&PlusMinus;</mo>
      <msqrt>
        <mrow>
          <msup>
            <mi>b</mi>
            <mn>2</mn>
          </msup>
          <mo>-</mo>
          <mrow>
            <mn>4</mn>
            <mo>&InvisibleTimes;</mo>
            <mi>a</mi>
            <mo>&InvisibleTimes;</mo>
            <mi>c</mi>
          </mrow>
        </mrow>
      </msqrt>
    </mrow>
  </mfrac>
</mrow>
```

Se observă că s-a folosit o nouă entitate, PlusMinus, care referă notația uzuală din matematică.

Următorul exemplu va specifica o matrice patratică de ordinul doi:

```
<mrow>
  <mi>A</mi>
  <mo>=</mo>
  <mfenced open="[" close="]">
    <mtable>
      <mtr>
        <td>
          <mi>x</mi>
        </td>
        <td>
          <mi>y</mi>
        </td>
      </mtr>
    </mtable>
  </mfenced>
</mrow>
```

Un exemplu de reprezentare a unui text matematic mai complex scris în MathML și procesat de Amaya

5.2 Incoherent Thomson Scattering Theory

When $k_{\perp} \gg 1$, the spectrum of the scattered power for an experiment with input polarization E_i perpendicular to the scattering plane defined by \hat{s} and \hat{i} and selecting scattered light polarized in the E_s direction can be shown [1] to be

$$\frac{d^2 P}{d\Omega_s d\omega_s} = r_e^2 \int_V \langle S_i \rangle d^3 r \left| \frac{1 - \beta_1}{1 - \beta_1} \right|^2 \kappa^2 \delta(k\nu - \omega) d^3 \nu$$

$$= r_e^2 \int_V \langle S_i \rangle d^3 r \left| 1 - \frac{\left(\frac{1 - \beta_1}{1 - \beta_1} \right) \beta_2^2}{(1 - \beta_1)(1 - \beta_2)} \right|^2 \left| \frac{1 - \beta_1}{1 - \beta_2} \right|^2 \times (1 - \beta_2^2) \delta(k\nu - \omega) d^3 \nu, \quad (5.4)$$

where $d^2 P/d\Omega_s d\omega_s$ is the power scattered per solid angle Ω_s and per frequency interval $d\omega_s$ from a volume V , in which the time averaged input wave Poynting vector (power flux) is $\langle S_i \rangle$.

```
<math display="block">\frac{d^2 P}{d\Omega_s d\omega_s} = r_e^2 \int_V \langle S_i \rangle d^3 r \left| \frac{1 - \beta_1}{1 - \beta_1} \right|^2 \kappa^2 \delta(k\nu - \omega) d^3 \nu
<math display="block">= r_e^2 \int_V \langle S_i \rangle d^3 r \left| 1 - \frac{\left( \frac{1 - \beta_1}{1 - \beta_1} \right) \beta_2^2}{(1 - \beta_1)(1 - \beta_2)} \right|^2 \left| \frac{1 - \beta_1}{1 - \beta_2} \right|^2 \times (1 - \beta_2^2) \delta(k\nu - \omega) d^3 \nu, \quad (5.4)
where  $d^2 P/d\Omega_s d\omega_s$  is the power scattered per solid angle  $\Omega_s$  and per frequency interval  $d\omega_s$  from a volume  $V$ , in which the time averaged input wave Poynting vector (power flux) is  $\langle S_i \rangle$ .
```

În limbajul MathML există numeroase elemente posedând atribute utile pentru a controla modul în care acestea apar pe ecran sau cum vor fi tipărite. Marcajul <mfenced> are rolul de a evidenția delimitatorii de la începutul și de la sfârșitul unei expresii. Desigur, în loc de paranteze pătrate puteau fi folosiți alți delimitatori.

Se mai pot folosi marcatorii <mtext> pentru a insera un text în cadrul unei expresii matematice sau <mchar> destinat să adauge anumite simboluri care nu au corespondent în alfabetul latin (i.e. literele grecești).

Exemple utilizând elemente de conținut

Reluând exemplul cu ecuația de gradul doi $x^2+4x+7=0$, putem scrie în MathML următoarele:

```
<math>
  <math display="block">x^2+4x+7=0
</math>
</math>
```

Acest mod de descriere este apropiat celui din limbajele de programare declarative (folosind notația poloneză) și este mult mai ușor de procesat și de analizat decât varianta scrisă cu ajutorul elementelor de prezentare. Se observă că nu este dificil să descriem notații matematice utilizând marcajele predefinite. După cum era de așteptat, elementele <ci> și <cn> sunt similare elementelor de prezentare <mi> și <mn>, respectiv. Rolul marcatorului <mo> este preluat de elementele de conținut desemnând operatorii matematici (e.g. <plus>, <minus> sau <times>).

Pentru specificarea rădăcinilor ecuației de gradul al doilea vom scrie următorul cod MathML:

```
<math>
  <math display="block">x^2+4x+7=0
  <math display="block">x = \frac{-4 \pm \sqrt{4^2 - 4 \cdot 1 \cdot 7}}{2 \cdot 1}
  </math>
</math>
```

```

<apply>
  <root />
  <apply>
    <minus />
    <apply>
      <power />
      <ci>b</ci>
      <cn>2</cn>
    </apply>
    <apply>
      <times />
      <cn>4</cn>
      <ci>a</ci>
      <ci>c</ci>
    </apply>
    </apply>
    <cn>2</cn>
  </apply>
</apply>
<apply>
  <times />
  <cn>2</cn>
  <ci>a</ci>
</apply>
</apply>
</reIn>

```

Inserarea de marcaje MathML în alte documente XML/HTML

Pentru ca documentele MathML să fie utilizate corespunzător, ele trebuie să fie compatibile cu procesoarele, editoarele sau translatorile existente. Deoarece limbajul MathML a fost creat cu scopul de a insera expresii matematice în paginile Web, cel mai important aspect este includerea marcajelor MathML în documente HTML.

În primul rând, documentele MathML trebuie să se integreze semantic în documentele HTML în sensul că vizualizatoarele trebuie să dețină un mecanism de recunoaștere

și de procesare a marcajelor MathML. Mai exact, marcajele MathML trebuie tratate ca marcaje XML, pentru care se face validarea și se respectă modul de afișarea a spațiilor.

Deoarece MathML se dorește a fi cât mai general și cât mai ușor de utilizat, chiar și de către nespecialiști, apare necesitatea dezvoltării unor aplicații care generează corect, în mod automatizat, documente MathML complexe asigurând faptul că documentele sunt corecte (valide în sens XML).

Așa cum am descris mai sus, în cadrul documentelor MathML va exista întotdeauna un element rădăcină denumit math. Atributele acestuia furnizează informații importante despre întreg documentul care va fi tratat unitar sau după valorile atributelor sale. Folosind acest element, putem specifica introducerea de marcaje MathML în cadrul unui document HTML astfel:

```

<math
  xmlns="http://www.w3.org/1998/Math/M
  athML">
  ...
</math>

```

Dacă navigatoarele Web vor putea procesa HTML conținând coduri MathML, atunci va deveni posibilă și procesarea inversă. Deși acest lucru este mai dificil, el constituie unul dintre principalele subiecte de discuție în cadrul Consorțiului Web.

În prezent, există un singur navigator disponibil pentru realizarea și vizualizarea de expresii matematice folosind marcaje MathML – **Amaya** (varianta curentă fiind 5.0), care se conformează standardului MathML 1.0 și variantelor de lucru ale MathML 2.0. Elementele de conținut nu sunt încă suportate, putând fi folosite doar elemente MathML de prezentare. Editarea

de expresii matematice se poate realiza chiar în cadrul navigatorului. Amaya este disponibil gratuit pe diverse platforme, atât în formă de executabil, cât și ca surse (vezi <http://www.w3.org/Amaya/>).

În plus, se pot asocia fragmentelor de expresii matematice legături simple spre alte resurse Web prin intermediul limbajului XLink:

```

<math
  xmlns="http://www.w3.org/1998/Math/M
  athML">
  <mrow>
    <mi>y</mi>
    <mo>=</mo>
    <mi></mi>
    <mfrac
      xmlns:xlink="http://www.w3.org/1999/
      xlink"
      xlink:type="simple"
      xlink:href="http://www.in-
      foiasi.ro/">
    <mi>x</mi>
    <msqrt>
      <mrow>
        <msup>
          <mi>x</mi>
          <mn>2</mn>
        </msup>
        <mo>+</mo>
        <mn>1</mn>
      </mrow>
    </msqrt>
    </mfrac>
  </mrow>
</math>

```

Amaya mai oferă suport pentru documente XHTML și ilustrații SVG (Scalable Vector Graphics), plus pentru adnotări colaborative (folosind RDF, XLink și XPointer), prin intermediul proiectului-aplicație *Annotea*.

De asemenea, se poate utiliza pachetul *WebEQ* pentru editarea și includerea marcajelor MathML în cadrul paginilor Web. Formulele concepute în MathML vor fi afișate prin intermediul unui applet Java. Se pune la dispoziție și un convertor automat al expresiilor scrise într-un dialect TeX în documente MathML. Un alt translator TeX-MathML este programul *Omega* elaborat de *John Plaice*.

O alternativă la MathML este *GELLMU* (*Generalized Extensible Latex-Like Markup*), limbaj de marcare propus de *William Hammond* foarte asemănător LaTeX-ului destinat reprezentării formulelor matematice pe Web. Aplicațiile suportând GELLMU sunt așteptate să apară în cursul acestui an.

Domnul Sabin-Corneliu Buraga este doctorand în Computer Science la Facultatea de Informatică – Universitatea „Al.I. Cuza” din Iași și poate fi contactat pe Web la adresa <http://www.infoiasi.ro/~busaco>. ■ 87

Resurse

MathML – specificații, resurse și informații utile:

- ✓ Miloslav Nic - MathML 2 reference with examples: <http://www.zvon.org/xxl/MathML/Output/>
- ✓ Web Consortium's Math Home Page: <http://www.w3.org/Math/>
- ✓ David Carlisle et al. (eds.) - Mathematical Markup Language (MathML) Version 2.0 W3C Recommendation, 21 February 2001: <http://www.w3.org/TR/MathML2/>
- ✓ MathML Information Center: www.webeq.com/mathml/

MathML – aplicații, implementări și unelte:

- ✓ Amaya: <http://www.w3.org/Amaya/>
- ✓ MathML in Mozilla: <http://www.mozilla.org/projects/mathml/>
- ✓ E-Lite: <http://www.icesoft.no/>
- ✓ IBM techexplorer: <http://www.software.ibm.com/network/techexplorer>
- ✓ Mathcad 2001: <http://www.mathcad.com/>
- ✓ Mathematica: <http://www.wolfram.com/>
- ✓ MathType: <http://www.mathtype.com/>
- ✓ LaTeX2HTML MathML package: <http://www.geom.umn.edu/~ross/webtex/webtex>
- ✓ WebEQ: <http://www.dessci.com/>