

ระบบแปลงนิพจน์คณิตศาสตร์เป็นข้อความภาษาไทย
Mathematical Expressions Conversion System into Thai Text

อัจจรา ธานีกุล (Achara Thaneekul)* ดร.พฤษดี ศิริแสงตระกูล (Dr.Pusadee Seresangtakul)**

บทคัดย่อ

บทความนี้นำเสนอการพัฒนา “ระบบอ่านนิพจน์คณิตศาสตร์” ซึ่งเป็นระบบที่ทำการแปลงนิพจน์คณิตศาสตร์เป็นข้อความภาษาไทย โดยระบบที่พัฒนาขึ้นนอกจากสามารถนำข้อมูลเข้าทางแป้นพิมพ์แล้วยังสามารถนำเข้าข้อมูลไฟล์ที่สร้างโดยโปรแกรมไมโครซอฟต์เวิร์ด 2007 (Microsoft Office Word 2007) โปรแกรมไมโครซอฟต์โอเพนเอกเซล (Microsoft Office Open XML) และโปรแกรมลาเทก (LaTeX) ในการพัฒนาระบบเริ่มจากวิเคราะห์ชนิดของไฟล์ข้อมูลนำเข้าและแปลงโครงสร้างข้อมูลให้อยู่ในรูปแบบโครงสร้างเอกเซล (XML) ซึ่งเป็นโครงสร้างแม่แบบของระบบ หลังจากนั้นทำการแปลงข้อมูลจากแท็กเอกเซล (XML tag) เป็นข้อความภาษาไทย ระบบที่พัฒนาขึ้นสามารถแปลงนิพจน์คณิตศาสตร์ ที่ใช้ในการเรียนจากระดับชั้นประถมศึกษาปีที่ 1 ถึง ระดับชั้นมัธยมศึกษาปีที่ 6 โดยอ้างอิงนิพจน์คณิตศาสตร์ตามหนังสือเรียนรายวิชาพื้นฐานคณิตศาสตร์ กระทรวงศึกษาธิการและพจนานุกรมศัพท์คณิตศาสตร์ ฉบับราชบัณฑิตยสถาน เพื่อทดสอบประสิทธิภาพของระบบที่พัฒนาขึ้น ผู้วิจัยได้คัดเลือกนิพจน์คณิตศาสตร์จากหนังสือเรียนรายวิชาพื้นฐานคณิตศาสตร์ ครอบคลุมเนื้อหาจากชั้นประถมศึกษาปีที่ 1 ถึง ระดับมัธยมศึกษาปีที่ 6 จำนวน 232 นิพจน์ ทดสอบด้วยวิธีการสร้างไฟล์อิเล็กทรอนิกส์นำเข้าระบบและข้อความภาษาไทยจากระบบ หลังจากนั้นให้อาจารย์ผู้สอนรายวิชาคณิตศาสตร์จำนวน 12 คน เป็นผู้ตรวจสอบประสิทธิภาพความถูกต้องของผลลัพธ์ข้อความภาษาไทย ผลการทดสอบความถูกต้องคิดเป็นร้อยละ 91.81

ABSTRACT

This paper presents the conversion of mathematical expressions to Thai text, named “Mathematical Expressions Reading System”. The system not only supports input data via keyboard, but it also supports imported data file which created by Microsoft Office Word 2007, Microsoft Office Open XML, and LaTeX. In the system development process, the input file-type is analyzed. The structure of input data is converted to XML structure, which is the template structure of the system. Next, the XML tag is converted to Thai text. The mathematical expressions system is able to convert the mathematics expression that study in Grade 1 to Grade 12. The expressions that use in this system is referred from mathematics fundamental textbook of the ministry of education and mathematics dictionary of the Royal Institute. In order to evaluate the system, 232 mathematics expressions (covers grade 1 to Grade 12) were converted to Thai text. The results were examined by 12 mathematics teachers. The correctness of the proposed system is 91.81%

คำสำคัญ : นิพจน์คณิตศาสตร์ ข้อความภาษาไทย ไมโครซอฟท์ออฟฟิศเอกเซล ลาเทก

Key Words : Mathematical expressions, Thai text, Microsoft office XML formats, LaTeX

* นักศึกษา หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

** ผู้ช่วยศาสตราจารย์ ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

บทนำ

รายวิชาคณิตศาสตร์เป็นกลุ่มสาระการเรียนรู้ใน 8 กลุ่มสาระการเรียนรู้ของหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 กระทรวงศึกษาธิการ (กระทรวงศึกษาธิการ, 2551) ซึ่งจะประกอบด้วยสาระหลักที่จำเป็น ได้แก่ จำนวนและการดำเนินการ การวัด เรขาคณิต พีชคณิต การวิเคราะห์ข้อมูลและความน่าจะเป็น ทักษะและกระบวนการทางคณิตศาสตร์ โดยแบ่งเนื้อหาสำหรับการเรียนการสอนจากระดับชั้นประถมศึกษาถึงระดับชั้นมัธยมศึกษา ปัจจุบันรายวิชาคณิตศาสตร์บรรจุในหลักสูตรการเรียนการสอนของการศึกษาทุกระดับชั้นในประเทศไทย เนื่องจากเป็นรายวิชาพื้นฐานและเป็นรายวิชาที่สามารถนำมาประยุกต์ใช้ในการเรียนการสอนวิชาอื่น อีกทั้งสามารถประยุกต์ใช้สำหรับการประกอบอาชีพอีกด้วย ดังนั้นจึงเป็นรายวิชาหนึ่งที่มีความจำเป็นอย่างยิ่ง

ระบบแปลงนิพจน์คณิตศาสตร์เป็นข้อความภาษาไทยเป็นระบบที่ช่วยเสริมสร้างทักษะการเรียนรู้จดจำสัญลักษณ์ ทราบถึงการนำไปใช้งานในเบื้องต้น กล่าวคือระบบจะแปลงนิพจน์คณิตศาสตร์จากไฟล์อิเล็กทรอนิกส์เป็นข้อความคำอ่านของนิพจน์นั้นๆ สามารถกระตุ้นการเรียนรู้และจดจำได้ทั้งระบบประสาทสัมผัสทั้งทางหูและทางตา ด้วยกระบวนการทำงานที่มีความรวดเร็ว นั้น ระบบแปลงนิพจน์คณิตศาสตร์จึงเป็นอีกหนึ่งทางเลือกที่อาจารย์ผู้สอนนำมาประกอบกับกระบวนการเรียนรู้ในชั้นเรียน และผู้ที่สนใจสามารถศึกษาวิธีการใช้งานและนำไปปรับใช้ในชีวิตประจำวันได้อย่างสะดวก แต่เดิมนั้นระบบแปลงนิพจน์คณิตศาสตร์เป็นข้อความนั้นมีหลากหลายระบบ โดยส่วนใหญ่จะเป็นระบบที่แปลงนิพจน์คณิตศาสตร์เป็นข้อความภาษาอังกฤษ ตัวอย่างงานวิจัย SWIMS (Speech-based Web Interface for Mathematics using Statistical language models) (Attanayake et al., 2012) เป็นระบบ Web-based ที่ผู้ใช้สามารถพิมพ์นิพจน์คณิตศาสตร์ผ่านแป้นพิมพ์

หลังจากนั้นระบบจะประมวลและแสดงผลพีชคณิตในรูปแบบข้อความภาษาอังกฤษ ระบบนี้มีชื่อว่า "Talkmaths" นอกจากนี้ มีงานวิจัยภายใต้ความร่วมมือจากหลายสถาบันเพื่อสร้างมาตรฐานการใช้งานคณิตศาสตร์ ได้แก่ UMA: A System for Universal Mathematics Accessibility (Karshmer et al., 2004) รองรับเอกสารอักษรเบรลล์ที่ใช้รหัส Nemeth และ Marburg โดยแปลงเอกสารเป็นรหัสเฉพาะทางด้านคณิตศาสตร์ ได้แก่ LaTeX, MathML, OpenMath

จากการศึกษาในงานวิจัยในครั้งนี้พบว่ามียุคที่พัฒนาขึ้นสำหรับอ่านไฟล์เอกสารอัตโนมัติ ซึ่งสามารถอ่านนิพจน์คณิตศาสตร์เป็นข้อความภาษาไทย ชื่อระบบ i-Math (Wongkia et al., 2012) พัฒนาขึ้นสำหรับผู้พิการทางสายตา รองรับการอ่านนิพจน์คณิตศาสตร์ถึงระดับชั้นมัธยมศึกษาตอนต้น

ในบทความนี้จะนำเสนอวิธีการแปลงนิพจน์คณิตศาสตร์เป็นข้อความภาษาไทย และกฎการแปลงนิพจน์คณิตศาสตร์ รองรับการพิมพ์นิพจน์คณิตศาสตร์ผ่านคีย์บอร์ด โดยมีกระบวนการทำงานทั้งหมด 4 ส่วน ประกอบด้วย 1) File type analysis 2) Extract file to XML format 3) Mathematical Expressions Generator 4) Thai Text Generator งานวิจัยนี้ รองรับนิพจน์คณิตศาสตร์ในระดับประถมศึกษาปีที่ 1 ถึงระดับมัธยมศึกษาปีที่ 6 โดยอ้างอิงเนื้อหาหลักจากหนังสือเรียนวิชาพื้นฐานคณิตศาสตร์ กลุ่มสาระการเรียนรู้คณิตศาสตร์ ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

การอ่านออกเสียงนิพจน์คณิตศาสตร์เป็นภาษาอังกฤษ

Chang et al. (1983) ได้จัดทำคู่มือสำหรับอ่านออกเสียงคณิตศาสตร์เป็นภาษาอังกฤษ ซึ่งเป็กลุ่มมือเพื่อช่วยให้ผู้อ่านสามารถอ่านสูตรหรือสมการคณิตศาสตร์ได้อย่างถูกต้อง กลุ่มมือจะแบ่งรายละเอียดของเนื้อหาเป็นบท ได้แก่ Alphabets, Basic Symbols,

Algebra, Trigonometric and Hyperbolic Expressions, Logic and Set Theory, Elementary and Analytic Geometry โดยจะครอบคลุมเนื้อหาคณิตศาสตร์ทั้งหมด คู่มือนี้จะแสดงแนวทางในการอ่านที่สามารถอ่านได้ในแต่ละนิพจน์นั้นๆ เช่น a.b อ่านได้ว่า “a dot b”, “the multiply of a and b” และ “a times by b” เป็นต้น เพื่อเป็นแนวทางในการเลือกใช้คำอ่านได้ เนื่องจากบางนิพจน์สามารถอ่านได้หลายรูปแบบ ขึ้นอยู่กับผู้ใช้และลักษณะการใช้งาน

Extensible Markup Language (XML)

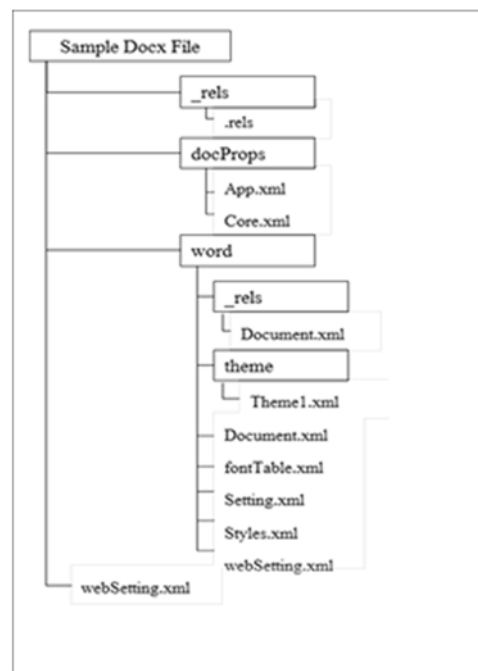
The World Wide Web Consortium (W3C, 2014) ได้พัฒนา Extensible Markup Language (XML) เป็นภาษา markup มีวัตถุประสงค์เพื่อใช้สำหรับระบบที่แตกต่างกัน โดยแต่ละระบบสามารถเข้าถึงข้อมูลและใช้งานร่วมกันได้ ภาษา XML เป็นภาษาพื้นฐานสำหรับภาษาอื่น เช่น Geography Markup Language (GML), RDF/XML, RSS, MathML, XHTML เป็นต้น การใช้งานภาษา XML สามารถเขียนโดยการใช้แท็กหรืออิลิเมนต์ (Element) มีหลักการเขียนแท็กดังนี้ <Tag> เป็นแท็กเปิดและใช้ </Tag> เป็นแท็กปิด ซึ่งแต่ละแท็กสามารถสร้าง Attribute เพื่อเก็บรายละเอียดข้อมูลเพิ่มเติมของแท็กนั้นๆ ได้ ตัวอย่างการเขียน Attribute เช่น <Tag id= “123”> สำหรับการใช้งานสามารถเขียนเอกสารในรูปแบบ DTD และ Schema เพื่อความสะดวกในการจัดรูปแบบเอกสาร ดังในรูปที่ 1 จะแสดงตัวอย่างการเก็บข้อมูลนักศึกษาโดยใช้โครงสร้างการเขียนภาษา XML ดังต่อไปนี้

```
<?xml version="1.0" ?>
<student id="123456789">
  <fname>Achara</fname>
  <lname>Thaneeekul</lname>
  <nickname>Aon</nickname>
  <univ>Khon Kaen University</univ>
  <education>Master Degree</education>
</student>
```

รูปที่ 1 ตัวอย่างคำสั่ง XML

Microsoft Office XML formats

Gouyaun et al. (2011) ได้นำเสนอข้อมูลรูปแบบไฟล์ Microsoft Office Word 2007 ซึ่งเป็นเอกสารของ Microsoft Office Word ที่จัดเก็บรายละเอียดของเอกสารในรูปแบบ XML โดยที่ไฟล์เหล่านี้จะถูกบีบอัดเป็นเอกสาร Microsoft Office Word ผ่านเครื่องมือมาตรฐานในการบีบอัดไฟล์ เพื่อจัดเก็บไฟล์ในรูปแบบไฟล์ ZIP โดยที่สามารถนำไฟล์นั้นๆ มาแตกไฟล์เพื่อแสดงไฟล์ข้อมูลที่ถูกบีบอัดอยู่ด้านในได้ วิธีการแสดงไฟล์ XML แบบง่ายสามารถทำได้โดยการบันทึกไฟล์เอกสารเป็น Word Office 2007 แบบทั่วไป (.docx) หลังจากนั้นเปลี่ยนนามสกุลไฟล์เป็น .zip และทำการแตกไฟล์ จะสามารถเห็นไฟล์ในรูปแบบ XML ที่ซ่อนอยู่ด้านใน โดยมีโครงสร้างของไฟล์ทั้งหมดดังรูปที่ 2 ต่อไปนี้



รูปที่ 2 โครงสร้างไฟล์ Microsoft Office XML

ไฟล์ XML แต่ละไฟล์จะเก็บข้อมูลที่แตกต่างกัน ตัวอย่างเช่น app.xml จะเก็บลักษณะทั่วไปของไฟล์ เช่น จำนวนคำ จำนวนบรรทัด เป็นต้น ส่วนไฟล์ cored.xml เก็บข้อมูลเกี่ยวกับชื่อผู้เขียน และไฟล์ document.xml เป็นไฟล์ข้อมูลเอกสารที่แสดงบน

หน้าจอ พร้อมคุณลักษณะสี ขนาดตัวอักษร การเว้นวรรค ระยะห่างตัวอักษร รูปแบบที่สำคัญของเอกสาร และเนื้อหาทั้งหมดของไฟล์นั้นๆ

LaTeX

LaTeX (Lamport et al., 1985) เป็นชุดคำสั่งสำหรับการจัดเรียงเอกสารที่มีคุณภาพสูง ใช้สำหรับสร้างเอกสารหลายประเภทที่มีคุณลักษณะเฉพาะ ตัวอย่าง คณิตศาสตร์ ฟิสิกส์ วิทยาการคอมพิวเตอร์ สถิติ เศรษฐศาสตร์และวิทยาศาสตร์ เอกสารจะประกอบด้วยอักขระพิเศษเป็นส่วนหนึ่งของเอกสาร LaTeX เป็นชุดคำสั่งฟรีที่สามารถใช้งานได้หลาย platforms เป็นชุดคำสั่งที่เข้าใจง่าย พัฒนาขึ้นโดย Leslie Lamport โดยใช้ TeX เป็นตัวประมวลผลหลัก คัดตัวอย่างการเขียนคำสั่ง LaTeX ดังต่อไปนี้

```
\begin{displaymath}
\sum_{i=1}^n \frac{1}{i} = \frac{n(n+1)}{2}
\end{displaymath}
```

จากคำสั่ง LaTeX ข้างต้น เมื่อแสดงผลบนหน้าจอดังสมการที่ 1

$$\sum_{i=1}^n i = \frac{n(n+1)}{2} \quad (1)$$

จะเห็นได้ว่าการเขียนนิพจน์หนึ่งๆ จะประกอบด้วยคำสั่งหลายส่วน ซึ่งคำสั่ง LaTeX แบ่งออกตามลักษณะการใช้งาน ได้แก่ เครื่องหมายประกอบตัวแปร เช่น $\bar{X} : \backslash \text{bar}\{X\}$, $\hat{a} : \backslash \text{hat}\{a\}$, $\ddot{a} : \backslash \text{ddot}\{a\}$ เป็นต้น ตัวอักษรกรีกตัวเล็ก เช่น $\alpha : \backslash \text{alpha}$, $\beta : \backslash \text{beta}$, $\gamma : \backslash \text{gamma}$, $\delta : \backslash \text{delta}$ เป็นต้น ตัวอักษรกรีกใหญ่ เช่น $\Delta : \backslash \text{Delta}$, $\Pi : \backslash \text{Pi}$, $\Omega : \backslash \text{Omega}$ เป็นต้น ลูกศรต่าง เช่น $\leftarrow : \backslash \text{leftarrow}$ or $\backslash \text{gets}$, $\rightarrow : \backslash \text{rightarrow}$ or $\backslash \text{to}$, $\leftrightarrow : \backslash \text{leftrightarrow}$ เป็นต้น จากตัวอย่างชุดคำสั่งข้างต้น คำสั่งจะขึ้นต้นด้วย \ เสมอ

งานวิจัยที่เกี่ยวข้อง

งานวิจัยนี้ได้ศึกษางานวิจัยที่เกี่ยวข้องกับการกับการอ่านนิพจน์คณิตศาสตร์เป็นข้อความดังต่อไปนี้

Attanayake et al. (2012) ได้ศึกษาและพัฒนา

ระบบ Speech-based Web Interface for Mathematics using Statistical Language models (SWIMS) : An intelligent editing assistant for mathematical text ขึ้นซึ่งเป็นระบบที่ผู้ใช้สามารถพิมพ์นิพจน์คณิตศาสตร์ผ่านหน้าจอ หลังจากนั้นระบบจะประมวลและแสดงผลลัพธ์ในรูปแบบข้อความภาษาอังกฤษ พร้อมทั้งเสนอทางเลือกการใช้คำในภาษาอังกฤษให้ผู้ใช้สามารถเลือกใช้งานตามความเหมาะสมและถูกต้อง ซึ่งข้อความที่นำเสนอเป็นผลลัพธ์นั้นจะจัดเรียงตามอัตราส่วนความถูกต้องและความสมบูรณ์ของข้อความ โดยเรียงลำดับจากข้อความที่ถูกเลือกใช้งานจากผู้ใ้มากที่สุด ระบบนี้มีชื่อว่า Talkmaths เป็นระบบที่ทำงานในรูปแบบ Web-based

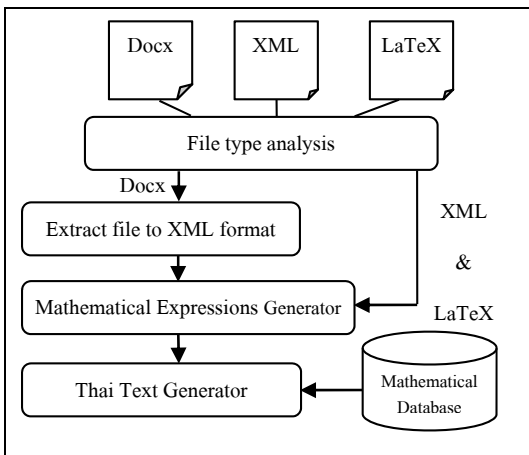
Karshmer et al. (2004) ได้นำเสนอระบบ A System for Universal Mathematics Accessibility (UMA) ซึ่งเป็นระบบที่พัฒนาขึ้นภายใต้ความร่วมมือจากหลายสถาบัน เพื่อสร้างมาตรฐานการเข้าถึงคณิตศาสตร์ ระบบ UMA นำส่วนการแปลงเอกสารที่ประกอบด้วยเนื้อหาทางคณิตศาสตร์เฉพาะด้าน เช่น Nemeth, Marburg โดยการแปลงเป็นรหัสสัญลักษณ์ทางคณิตศาสตร์ ได้แก่ LaTeX, MathML และ OpenMath ทั้งนี้ระบบมีโปรแกรมช่วยสำหรับอ่านออกเสียงเอกสารทางคณิตศาสตร์เพื่อเป็นส่วนช่วยสำหรับการได้ยินเสียงของผู้ใช้งานระบบอีกทางหนึ่ง

Wongkia et al. (2012) ได้ศึกษาและพัฒนา ระบบการอ่านนิพจน์คณิตศาสตร์อัตโนมัติสำหรับนักเรียนพิการทางสายตา ชื่อระบบว่า i-Math ซึ่งระบบจะประมวลผลจากเอกสารที่มีส่วนประกอบของเนื้อหาทางด้านคณิตศาสตร์ หรือเอกสารเฉพาะด้าน เช่น เอกสารทางวิทยาศาสตร์ ระบบจะประมวลผลด้วยโปรแกรมช่วยช่วยอ่านออกเสียงข้อความทางคณิตศาสตร์เป็นภาษาไทย

ปัจจุบันนี้ระบบแปลงนิพจน์คณิตศาสตร์เป็นข้อความภาษาไทยนั้นมีปริมาณน้อย จากการศึกษาพบว่ามียุทธวิธีที่พัฒนาขึ้นสำหรับอ่านเอกสารอัตโนมัติและรองรับในระดับมัธยมศึกษาตอนต้นเท่านั้น

วิธีดำเนินงานวิจัย

งานวิจัยนี้แบ่งส่วนการทำงานของระบบ ออกเป็น 4 ขั้นตอนหลักๆ คือ 1) File type analysis 2) Extract file to XML format 3) Mathematical Expressions Generator 4) Thai Text Generator โดยสามารถเขียนเป็นแผนภาพขั้นตอนการทำงานของระบบ ได้ดังรูปที่ 3 ต่อไปนี้ ซึ่งในแต่ละขั้นตอนการทำงานจะอธิบายถึงรายละเอียดการประมวลผลที่แตกต่างกัน ซึ่งการทำงานของระบบจะแบ่งออกตามชนิดของไฟล์นำเข้า



รูปที่ 3 ขั้นตอนการทำงานของระบบแปลงนิพจน์คณิตศาสตร์เป็นข้อความภาษาไทย

File type analysis

ระบบที่พัฒนารองรับไฟล์นำเข้าที่แตกต่างกัน 3 ประเภท ได้แก่ Microsoft Office Word 2007, Open Office XML และ LaTeX ซึ่งแต่ละประเภทจะมีโครงสร้างภายในที่แตกต่างกัน ระบบสามารถโหลดไฟล์ที่สร้างไว้แล้วหรือเขียนนิพจน์ใหม่ผ่านหน้าจอระบบได้ เพื่ออำนวยความสะดวกในการใช้งานระบบแบ่งส่วนการแสดงผลออกเป็น 2 ส่วน ประกอบด้วยส่วนเขียนคำสั่งและส่วนแสดงนิพจน์ หลังจากนั้นระบบจะวิเคราะห์ชนิดของไฟล์นำเข้าว่าเป็นไฟล์ชนิดใด หลังจากนั้นระบบจะตรวจสอบว่าเป็นไฟล์ Microsoft Office Word 2007 หรือไม่ ถ้าเป็นไฟล์

Microsoft Office Word 2007 ระบบจะทำการแตกไฟล์ในลำดับถัดไป แต่ถ้าไม่ใช่ ระบบจะทำงานในส่วน Mathematical Expressions Generator เพื่ออ่านค่าข้อมูลจากไฟล์นั้นๆ

Extract file to XML format

โมดูลนี้ใช้งานสำหรับไฟล์ Microsoft Office Word 2007 (.docx) โดยเฉพาะ ระบบจะแตกไฟล์ .docx เพื่อแสดงไฟล์โครงสร้างที่เก็บในรูปแบบ XML ซึ่งถูกบีบอัดในรูปแบบไฟล์ zip หลังจากนั้นระบบจะเลือกใช้ไฟล์ document.xml ซึ่งเป็นไฟล์สำหรับเก็บข้อมูลเนื้อหาทั้งหมดของไฟล์นำเข้า ระบบจะดึงข้อมูลโครงสร้าง XML เพื่อเลือกแท็ก (Tag) ที่เกี่ยวข้องกับนิพจน์คณิตศาสตร์ ซึ่งแท็กเหล่านั้นจะขึ้นต้นด้วย <m:oMathPara> และนำแท็กที่มีคุณลักษณะเป็นแท็ก Child Node ทั้งหมดมาประมวล โดยแต่ละแท็ก จะมีคุณลักษณะที่แตกต่างกัน เช่น <m:num> และ <m:den> เป็นแท็กสำหรับนิพจน์เศษส่วน, <m:rad> เป็นแท็กสำหรับนิพจน์แสดควรรุท และ <m:deg> ใช้สำหรับบอกจำนวน degree ของ squaroot เป็นต้น ดังตัวอย่างต่อไปนี้เป็นตัวอย่างการเขียนคำสั่ง XML ภายในเอกสาร .docx จากสมการที่ 2

$$c = \sqrt{a^2 + b^2} \tag{2}$$

เมื่อระบบแตกไฟล์เสร็จแล้วนั้นจะแสดงคำสั่ง XML ดังรูปที่ 4 โดยโครงสร้างจะขึ้นต้นคำสั่งด้วยแท็ก <m:oMathPara> เป็นการเริ่มต้น Paragraph ของนิพจน์คณิตศาสตร์ ถัดมาจะเป็นแท็ก <m:oMath> หมายถึงการเริ่มต้นคำสั่ง Office Math และแท็ก <m:rad> เป็นจุดเริ่มต้นของแสดควรรุทและแท็ก <m:deg> เป็นแท็กบอกจำนวน degree ของ square root สำหรับตัวอักษรทั้งหมดในสมการจะเขียนด้วยแท็ก <m:t> ซึ่งเป็นแท็กสำหรับการเก็บค่าตัวอักษร โดยเฉพาะ จากชุดคำสั่งข้างต้น ระบบจะนำชุดคำสั่งทั้งหมดใช้ในการประมวลผลสำหรับการสร้างคำอ่านของนิพจน์นั้นๆ ในส่วนการทำงานถัดไป

```

1. <m:oMathPara>
2.   <m:oMath>
3.     <m:r>
4.       <m:t>c</m:t>
5.     <m:r>
6.       <m:rad>
7.         <m:radPr>
8.           <m:degHide m:val="on"/>
9.         </m:radPr>
10.        <m:deg/>
11.       <m:e>
12.         <m:sSup>
13.           <m:e>
14.             <m:r>
15.               <m:t>a</m:t>
16.             </m:r>
17.           </m:e>
18.         <m:sup>
19.           <m:r>
20.             <m:t>2</m:t>
21.           </m:r>
22.         </m:sup>
23.       </m:sSup>
24.     <m:r>
25.       <m:t>+</m:t>
26.     </m:r>
27.     <m:sSup>
28.       <m:e>
29.         <m:r>
30.           <m:t>b</m:t>
31.         </m:r>
32.       </m:e>
33.     <m:sup>
34.       <m:r>
35.         <m:t>2</m:t>
36.       </m:r>
37.     </m:sup>
38.   </m:oMath>
39. </m:oMathPara>

```

รูปที่ 4 คำสั่ง XML ในเอกสาร document.xml

Mathematical Expressions Generator

โมดูลนี้เป็นส่วนสำหรับแปลงโครงสร้างคำสั่ง XML จากโมดูลก่อนหน้านี้เป็นข้อความนิพจน์คณิตศาสตร์ โดยการวิเคราะห์โครงสร้างและจำแนกชนิดของคำสั่ง XML ซึ่งสามารถจำแนกได้เป็น 3 กลุ่มได้แก่ 1) คำสั่งพิเศษ เช่น <m:oMathPara> <m:oMath> <m:rad> <m:deg> เป็นต้น 2) ฟังก์ชันคณิตศาสตร์ เช่น ฟังก์ชัน sin, cos, tan, cosec, log, lim เป็นต้น และ 3) ตัวเลขและตัวอักษร โดยถ้าเป็นคำสั่งพิเศษระบบจะเลือกคำอ่านที่สอดคล้องกับนิพจน์จากกฎการแปลงนิพจน์คณิตศาสตร์ที่สร้างไว้ดังตารางที่ 2 หากเป็นคำสั่งชนิดฟังก์ชันคณิตศาสตร์ ระบบจะทำการแปลงเป็นคำอ่านของฟังก์ชันนั้นๆ เป็นข้อความได้ทันที และกลุ่มสุดท้ายได้แก่ตัวเลขและตัวอักษร ในส่วนนี้

ระบบจะดึงข้อมูลคำอ่านจากฐานข้อมูลเพื่อแสดงคำอ่านของตัวเลขหรือข้อความนั้นๆ ภายหลังจากการทำงานของคำสั่งแต่ละกลุ่มเสร็จสิ้น ลำดับถัดไปคือการจัดเรียงคำอ่าน สำหรับคำสั่งบางคำสั่งที่จะต้องจัดเรียง เช่น \bar{X} ถ้าอ่านจาก แท็ก โดยตรงจะอ่านว่า “บาร์เอ็กซ์” จึงจำเป็นต้องมาจัดเรียงใหม่อีกครั้งเพื่อความถูกต้อง และในลำดับสุดท้ายจะตรวจสอบคำอ่านของตัวเลขซึ่งจะประกอบด้วยชุดคำอ่านจำนวนเต็ม เช่น 125 อ่านว่า “หนึ่งร้อยยี่สิบห้า : one hundred twenty-five” และชุดคำอ่านที่อ่านเรียงตัวเลข เช่น 125 อ่านว่า “หนึ่งสองห้า : one, two, five” โดยสองส่วนจะใช้งานต่างกันสำหรับตัวเลขทศนิยม

ในขั้นตอนนี้ผู้วิจัยได้สร้างกฎการแปลงนิพจน์คณิตศาสตร์เป็นข้อความภาษาไทย ดังตารางที่ 2 เพื่อเก็บรูปแบบการอ่านนิพจน์คณิตศาสตร์

ตารางที่ 2 แสดงกฎการแปลงนิพจน์คณิตศาสตร์

กฎ	ตัวอย่าง	คำอ่าน
จุด	$a.b$	เอ ภูณ บี
ตัวอักษรขีด	$4ab$	สี่ ภูณ เอ ภูณ บี
ในวงเล็บ...	$(a + b)$	ในวงเล็บ เอ บวก บี
ในวงเล็บ...ทั้งหมด	$(a + b)^2$	ในวงเล็บ เอ บวก บี ทั้งหมด ยกกำลังสอง
หารด้วย	$a \div b$	เอ หารด้วย บี
เศษ...ส่วน...	$\frac{a}{b}$	เศษ เอ ส่วน บี
เศษ...ส่วน...ของ	$\frac{a}{b}c$	เศษ เอ ส่วน บี ของ ซี
เศษ...ทั้งหมดส่วน...	$\frac{a + b}{c}$	เศษ เอ บวก บี ทั้งหมดส่วน ซี
อนุพันธ์ของ...เทียบกับ..	$\frac{dy}{dx}$	อนุพันธ์ของ วาย เทียบกับ เอ็กซ์
ความน่าจะเป็น...	$P(E_2)$	ความน่าจะเป็นที่จะเกิด เหตุการณ์ E_2

จากตารางที่ 2 จะเห็นว่า กฎที่สร้างขึ้นจะเป็นส่วนในการควบคุมโครงสร้างการอ่านนิพจน์ แต่อีกส่วนที่สำคัญคือคำอ่านสำหรับตัวเลขและตัวอักษร ซึ่งในงานวิจัยนี้ได้ทำการออกแบบฐานข้อมูลนิพจน์

คณิตศาสตร์ โดยมีโครงสร้างฐานข้อมูลประกอบด้วย แท็กเอกเอมแอล แท็กลาเทก คำภาษาไทย ชนิดของ แท็ก ดังแสดงในตารางที่ 3

ตารางที่ 3 โครงสร้างการเก็บข้อมูล

NO	Entity	Type	Caption
1	ID	AutoNumber	ไอดีอัตโนมัติ
2	Tagname	TEXT(100)	แท็กเอกเอมแอล
3	Latextag	TEXT(100)	แท็กลาเทก
4	Thai	TEXT(150)	คำภาษาไทย
5	Chktype	TEXT(20)	ชนิดของแท็ก

Thai Text Generator

จากการทำงานของระบบในโมดูลก่อนหน้า จะเห็นได้ว่านิพจน์คณิตศาสตร์มีการจัดเรียงและแบ่ง ลักษณะของนิพจน์เป็นที่เรียบร้อย ในส่วนโมดูลนี้ ระบบจะแสดงข้อความภาษาไทยที่ตรงกับนิพจน์ที่ นำเข้าจากระบบและผ่านการประมวลผลจาก โมดูล ก่อนหน้า จากสมการที่ 2 เมื่อประมวลผลเสร็จแล้วนั้น จะได้ข้อความภาษาไทยดังต่อไปนี้

“สี่ เท่ากับ รากที่สองของ เอ ยกกำลัง สอง บวก บี ยกกำลัง สอง”

ผลการวิจัย

งานวิจัยนี้ทดสอบประสิทธิภาพของระบบที่ พัฒนาขึ้น โดยใช้เลือกนิพจน์คณิตศาสตร์จากหนังสือ เรียนวิชาพื้นฐานคณิตศาสตร์ กลุ่มสาระการเรียนรู้ คณิตศาสตร์ ตามหลักสูตรแกนกลางการศึกษาขั้น พื้นฐาน พุทธศักราช 2551 จากระดับชั้นประถมศึกษา ปีที่ 1 ถึง ชั้นมัธยมศึกษาปีที่ 6 จำนวน 232 นิพจน์ เพื่อ ใช้ในการแปลงนิพจน์คณิตศาสตร์เป็นข้อความ ภาษาไทย และวัดประสิทธิภาพของการแปลงนิพจน์ นั้นๆ โดยใช้ตัวชี้วัดร้อยละความถูกต้อง คือจำนวน นิพจน์ที่แปลงถูกต้องต่อจำนวนนิพจน์ทั้งหมดคิดเป็น เปอร์เซ็นต์แสดงได้ดังสมการที่ 3

$$\text{ประสิทธิภาพ (\%)} = \frac{X}{N} \times 100 \quad (3)$$

โดยที่ X คือ จำนวนนิพจน์ที่แปลงถูกต้อง
N คือ จำนวนนิพจน์ทั้งหมด

ผู้วิจัยตรวจสอบความถูกต้องของการแปลง นิพจน์คณิตศาสตร์เป็นข้อความภาษาไทย โดยนำคำที่ ได้จากระบบนั้น ให้อาจารย์ผู้สอนรายวิชาคณิตศาสตร์ จำนวน 12 ท่าน ตรวจสอบความถูกต้องของการแปลง นิพจน์คณิตศาสตร์ โดยอาจารย์ผู้สอน 1 ท่านต่อการ ตรวจสอบการแปลง 1 ระดับชั้น ซึ่งจำนวนนิพจน์ที่ นำมาทดสอบแต่ละระดับชั้นจะไม่เท่ากัน เนื่องจาก เลือกนิพจน์จากลักษณะการใช้งานที่มีในหนังสือเรียน ของแต่ละระดับชั้น แสดงตัวอย่างผลลัพธ์ข้อความ ภาษาไทยจากนิพจน์ที่ผู้วิจัยนำมาทดสอบระบบ โดย เริ่มต้นนิพจน์คณิตศาสตร์ตามชนิดของไฟล์นำเข้า ได้แก่ Open Office XML Format, LaTeX และผลลัพธ์ ข้อความภาษาไทยตัวอย่างต่อไปนี้ โดยแสดงเริ่มต้น จากนิพจน์คณิตศาสตร์ที่นำเข้าสู่ระบบ และชุดคำสั่งที่ เป็นชุดคำสั่ง LaTeX และแสดงข้อมูล คำสั่ง Open Office XML และผลลัพธ์จากระบบเป็นข้อความ ภาษาไทย

นิพจน์คณิตศาสตร์ :

$$3x + 5y^2 = 5$$

LaTeX format :

$$3x+5y^{2}=5$$

Open Office XML Format :

```
<m:oMathPara>
  <m:oMath>
    <m:t>3x+</m:t>
    <m:sSup>
      <m:e>
        <m:t>5y</m:t>
      </m:e>
    <m:sup>
      <m:t>2</m:t>
    </m:sup>
  </m:sSup>
  <m:t>=5</m:t>
</m:oMath>
</m:oMathPara>
```

ข้อความภาษาไทย :

“สามคูณเอ็กซ์บวกห้าคูณวายยกกำลังสองเท่ากับห้า”

จากตัวอย่างข้างต้น ผู้วิจัยนำผลการทดสอบประสิทธิภาพของการแปลงนิพจน์คณิตศาสตร์เป็นข้อความภาษาไทย ซึ่งผลการทดสอบแบ่งตามระดับชั้น โดยแบ่งจำนวนนิพจน์และคำอ่านที่อ่าน ได้ถูกต้อง แสดงเป็นเปอร์เซ็นต์ ดังแสดงในตารางที่ 4 ต่อไปนี้

ตารางที่ 4 แสดงผลการประเมินความถูกต้องระบบ

ระดับชั้น	จำนวนนิพจน์	อ่านถูกต้อง	เปอร์เซ็นต์
ประถมศึกษาปีที่ 1-3	33	31	93.93
ประถมศึกษาปีที่ 4-6	53	50	94.33
มัธยมศึกษาปีที่ 1-3	82	75	91.46
มัธยมศึกษาปีที่ 4-6	64	57	89.06
รวมทั้งสิ้น	232	213	91.81

อภิปรายและสรุปผลการวิจัย

งานวิจัยนี้นำเสนอการพัฒนาระบบแปลงนิพจน์คณิตศาสตร์เป็นข้อความภาษาไทย โดยใช้หนังสือเรียนวิชาพื้นฐานคณิตศาสตร์ กลุ่มสาระการเรียนรู้คณิตศาสตร์ ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 จากระดับชั้นประถมศึกษาปีที่ 1 ถึงชั้นมัธยมศึกษาปีที่ 6 เป็นเอกสารอ้างอิงสำหรับเลือกทดสอบนิพจน์คณิตศาสตร์ ผลการทดสอบเพื่อประเมินประสิทธิภาพความถูกต้องของระบบมีความถูกต้องที่ร้อยละ 91.81

ผลการวิเคราะห์พบว่า ผลลัพธ์ที่ได้สามารถประยุกต์ใช้งานต่อไปได้ เช่น ใช้เป็นข้อมูลนำเข้าของระบบอ่านออกเสียง เช่น โปรแกรมวาจา (VAJA) (NECTEC, 2557) และสามารถพัฒนาเป็นแอปพลิเคชันสำหรับคนตาบอดได้ ดังนั้นเพื่อให้ระบบมีความสมบูรณ์ครบถ้วนเหมาะสำหรับการใช้ประกอบการเรียนการสอนต่อไป ควรพัฒนาในส่วนของการอ่านออกเสียง และพัฒนาทิวเพิ่มเติมสำหรับการอ่านสมการคณิตศาสตร์ในระดับชั้นที่สูงขึ้น เพื่ออำนวยความสะดวกให้ผู้ใช้งานได้เข้าใจคำอ่านนิพจน์ทุกนิพจน์ที่มีในสาขาวิชาคณิตศาสตร์ได้อย่างถูกต้องและครบถ้วน

กิตติกรรมประกาศ

งานวิจัยนี้สามารถดำเนินจนสำเร็จลุล่วงไปได้ด้วยความช่วยเหลืออย่างดียิ่งของอาจารย์ทัศนีย์ อารยะตระกูลลิขิต และอาจารย์ประจำภาควิชาคณิตศาสตร์ มหาวิทยาลัยขอนแก่น ที่ให้คำแนะนำและข้อคิดเห็นต่างๆ มาโดยตลอด ทั้งด้านคำอ่านภาษาไทย การเลือกนิพจน์สำหรับทดสอบระบบ ขอขอบพระคุณอาจารย์ผู้สอนรายวิชาคณิตศาสตร์ทุกท่าน ที่ให้ความอนุเคราะห์ในการตรวจประสิทธิภาพความถูกต้องทุกท่าน พร้อมผู้ที่มีส่วนเกี่ยวข้องกับการทำวิจัยในครั้งนี้ทุกท่าน

เอกสารอ้างอิง

กระทรวงศึกษาธิการ. หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551. กรุงเทพมหานคร: กระทรวงศึกษาธิการ; 2551.

Attanayake D, Pfluegel E, Hunter G, Denholm-Price J. SWIMS (Speech-based Web Interface for Mathematics using Statistical language models): An intelligent editing assistant for mathematical text. Proceedings of 2012 8th International Conference on the Intelligent Environments (IE); 2012 June 26-29; Mexico: Guanajuato. p. 327 - 330.

Karshmer AI, Gupta G, Pontelli E, Miesenberger K, Ammalai N, Gopal D. et al. UMA: A System for Universal Mathematics Accessibility. Proceedings of the ACM SIGACCESS Conference on Computers And Accessibility, ASSETS'04; 2004 October 18 – 20; USA. : Georgia. p. 55 - 62.

Wongkia W, Naruedomkul K, Cercone N. i-Math : Automatic math reader for Thai blind an visually impaired students. Proceeding of Computer and Mathematics with Application; 2012 April 11; vol. 64; p. 2128-2140.

Chang LA, White CM. Handbook for Spoken Mathematics. USA.: California. 1983.

The World Wide Web Consortium (W3C). Extensible markup Language (XML) [online] 2014 [cited 2014 Jan 12]. Available from: <http://www.w3.org/XML/>

Gouyaun L, Gouhui W. A new information hiding method based on word 2007. Proceedings of 2011 IEEE 2nd International Conference on Digital Object, Software Engineering and Service Science (ICSESS); 2011 July 15-17; China:Beijing. p. 709-711.

Lamport L. LaTeX : A Document Preparation System User's Guide and Manual. Addison-Wesley Company, 1985.

National Electronic and Computer Technology Center (NECTEC), Ministry of Science Technology and Environment. VAJA Thai text-to-speech [online] 2014 [cited 2014 Sep 12]. Available from: <http://vaja.nectec.or.th/>