29 JUL 2002

LIBRARY AND INFORMATION CENTER
MAHIDOL UNIVERSITY

COMPUTER SOFTWARE FOR ASSISTING BIOLOGICAL IDENTIFICATION

ATTASIT THIPSONGKHROH

อภิธันพนาคาร
จาก
บัณฑิตวิทยาลัย มหาวิทยาลัยมหิดล

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR
THE DEGREE OF MASTER OF SCIENCE
(TECHNOLOGY OF INFORMATION SYSTEM MANAGEMENT)
MAJOR IN INFORMATION MANAGEMENT ON
ENVIRONMENTS AND NATURAL RESOURCES
FACULTY OF GRADUATE STUDIES
MAHIDOL UNIVERSITY
2002

TH
A883c
2002
c.2.

ISBN 974-04-1761-2

COPYRIGHT OF MAHIDOL UNIVERSITY
Copyright by Mahidol University

4036784 ENIM/M : MAJOR : INFORMATION MANAGEMENT ON ENVIRONMENTS AND NATURAL RESOURCES ; M.Sc. (TECHNOLOGY OF INFORMATION SYSTEM MANAGEMENT)
KEY WORDS : BIOLOGICAL IDENTIFICATION, IDENTIFICATION SOFTWARE, NAMEID

ATTASIT THIPSONGKHROH : COMPUTER SOFTWARE FOR ASSISTING BIOLOGICAL IDENTIFICATION. THESIS ADVISORS : RAYWADEE ROACHANAKANAN, M.Sc., SARANYA SUTJARITKUL, M.Sc., VACHAROBON THIRAKHUPT, Ph.D. 128 P. ISBN 974-04-1761-2.

Biological identification is a main channel to retrieve systematics information for other biological-related sciences. This study develops a computer program for assisting the identification called NameID, which separates identification data from the identification program.

The NameID consists of two important modules, the NameID Editor and the NameID Viewer. The Editor was used by experts to create identification data packages that could be sent to the Viewer's users for interactive identification. It used two relational databases as data storages; one to store main identification data, the other is internally used to store vector data for lines that can be drawn over graphic illustrations to help identification or to illustrate certain features. Its data structure had a limitation in that it cannot store character dependency information if one child character had more than one parent characters. The NameID assists identification via LucID score matrix encoding, character dependency information and comparison function calculation.

Two sample identification data were developed to demonstrate the program's usefulness, a key to families of turtles and tortoises of Thailand and a key to the Durian varieties sold in markets in Thailand. These two sets of data were developed by external experts. Testings were conducted by questionnaire with 10 users divided into two groups. The two experts tested both the NameID Editor and the NameID Viewer while the eight graduate students tested the NameID Viewer only. The NameID obtained 83.00% of overall score. In the user's opinion, the NameID Editor often crashed (5.50 points of 10 maximum points). The NameID Viewer could efficiently improve identification process in the expert's opinions when comparing to conventional identification (time saving = 8.50 points and accuracy = 9.00 points of 10 maximum points). However, the effectiveness of the NameID highly depends on the quality of data set being used. It should be noted that the data set of local market Durian was later found to contain incorrect data and this may have affected the results.

The identification software which identification data separated from the programs like the NameID will leverage the data creators from concerning about identification software development and enable them to more concentrate on the identification data development instead.

Copyright by Mahidol University

4036784 ENIM/M : สาขาวิชา : การจัดการสารสนเทศสิ่งแวดล้อมและทรัพยากร ; วท.ม.
(เทคโนโลยีการจัดการระบบสารสนเทศ)

อรรถสิทธิ์ ทิพย์สงเคราะห์ : โปรแกรมคอมพิวเตอร์เพื่อช่วยในการระบุชนิดสิ่งมีชีวิต (COMPUTER SOFTWARE FOR ASSISTING BIOLOGICAL IDENTIFICATION). คณะกรรมการควบคุมวิทยานิพนธ์ : เรวดี โรจนกนันท์, M.Sc., ศรัณยา สุจริตกุล, M.Sc., วัชโรบล ธีรคุปต์, Ph.D. 128 หน้า. ISBN 974-04-1761-2.

การระบุชนิดสิ่งมีชีวิตเป็นช่องทางสำคัญในการสืบค้นข้อมูลทางด้านอนุกรมวิธานสำหรับใช้ในการศึกษาในสาขาที่เกี่ยวข้องกับชีววิทยา ในการศึกษานี้เป็นการพัฒนาโปรแกรมคอมพิวเตอร์ซึ่งมีชื่อว่า NameID เพื่อช่วยในการระบุชนิดสิ่งมีชีวิตโดยที่มีข้อมูลสำหรับการระบุชนิดนั้นแยกออกจากตัวโปรแกรม โปรแกรม NameID มีสองส่วนหลักคือโปรแกรม NameID Editor และโปรแกรม NameID Viewer โดยที่ NameID Editor เป็นโปรแกรมสำหรับให้ผู้ที่มีความรู้หรือเชี่ยวชาญใช้ในการสร้างแฟ้มข้อมูลการระบุชนิด (identification data package) เพื่อส่งต่อให้กับผู้ใช้ทั่วไปซึ่งจะใช้โปรแกรม NameID Viewer ในการเปิดแฟ้มข้อมูลการระบุชนิดขึ้นมาเพื่อใช้ในการระบุชนิดแบบปฏิสัมพันธ์ (interactive identification) ต่อไป โปรแกรม NameID มีการใช้ฐานข้อมูลจำนวนสองฐานข้อมูลในการเก็บข้อมูลหลักสำหรับการระบุชนิดและสำหรับการเก็บข้อมูลเส้นตรงที่ลากเพื่อชี้ตำแหน่งต่างๆ บนภาพประกอบ โครงสร้างข้อมูลของโปรแกรมมีข้อจำกัดที่ไม่สามารถจะเก็บข้อมูลความสัมพันธ์ระหว่างลักษณะ (characters) ต่างๆ ที่มีลักษณะแม่ (parent characters) ที่มากกว่าหนึ่งลักษณะได้ โปรแกรม NameID สามารถช่วยในการระบุชนิดได้โดยการใช้เทคนิคต่างๆ คือ LucID score matrix encoding, character dependency และ comparison function calculation

ในการศึกษานี้ยังได้พัฒนาแฟ้มข้อมูลการระบุชนิดขึ้นจำนวนสองชุดเพื่อสาธิตถึงประสิทธิภาพของโปรแกรม คือ แฟ้มข้อมูลการระบุชนิดในระดับแฟมิลี (family) ของเต่าและตะพาบน้ำในไทย และแฟ้มข้อมูลการระบุชนิดของทุเรียนพันธุ์การค้าในไทย โดยที่แฟ้มข้อมูลทั้งสองชุดนี้พัฒนาโดยผู้เชี่ยวชาญในแต่ละสาขา ผลการทดสอบการใช้งานโปรแกรมกับกลุ่มตัวอย่างจำนวน 10 คน ซึ่งแบ่งออกเป็นสองกลุ่มอันประกอบด้วยผู้เชี่ยวชาญ 2 คน (ทดสอบ NameID Editor และ NameID Viewer) และนักศึกษาระดับปริญญาโท 8 คน (ทดสอบ NameID Viewer เพียงอย่างเดียว) พบว่า โปรแกรม NameID ได้รับคะแนนรวมคิดเป็นร้อยละ 83.00% ผู้ใช้มีความเห็นว่า NameID Editor มีปัญหาหยุดการทำงาน (crash) บ่อยครั้ง (5.50 จาก 10 คะแนน) ในขณะที่ผู้เชี่ยวชาญมีความเห็นว่า NameID Viewer สามารถเพิ่มประสิทธิภาพในการระบุชนิดได้ดีเมื่อเทียบกับการระบุชนิดแบบวิธีดั้งเดิม (การทุ่นเวลา = 8.50 คะแนน และความแม่นยำ = 9.00 คะแนนจากคะแนนเต็ม 10 คะแนน) ทั้งนี้ประสิทธิภาพของการระบุชนิดนั้นขึ้นกับคุณภาพของแฟ้มข้อมูลที่นำมาใช้ด้วยเป็นอย่างยิ่ง ยิ่งไปกว่านั้นแฟ้มข้อมูลของทุเรียนยังมีข้อผิดพลาดซึ่งอาจส่งผลต่อผลการประเมิน อย่างไรก็ดีโปรแกรมช่วยในการระบุชนิดสิ่งมีชีวิตแบบที่ข้อมูลแยกออกจากตัวโปรแกรมในลักษณะเดียวกันกับโปรแกรม NameID นี้จะช่วยให้ผู้สร้างข้อมูลการระบุชนิดไม่ต้องกังวลกับการพัฒนาโปรแกรมและหันมาให้ความสนใจกับการพัฒนาข้อมูลการระบุชนิดได้

Copyright by Mahidol University