บทที่ 4

การออกแบบฐานความรู้และกลไกการหาเหตุผล

เมื่อได้ทำการศึกษาลักษณะทางสัณฐานวิทยาของอ้อยและความรู้เกี่ยวกับอ้อย ซึ่งเป็นประกอบส่วนสำคัญที่จะนำมาใช้สร้างฐานความรู้แล้ว ได้ทำการรวบรวมข้อมูลลักษณะ (attribute) ต่างๆ ของอ้อยจำนวน 50 สายพันธุ์จากแปลงรวบรวมพันธุ์ของศูนย์เกษตรอ้อยภาคกลาง อำเภอท่าม่วง จังหวัดกาญจนบุรี ภายใต้คำแนะนำของผู้เชี่ยวชาญ แล้วนำมาสร้างฐานความรู้และกลไกการหาเหตุผลตามขั้นตอนดังต่อไปนี้

4.1 วิเคราะห์สภาพปัญหาและแนวทางแก้ไข

จากการศึกษาลักษณะของปัญหาและแนวทางการแก้ปัญหาของผู้เชี่ยวชาญแล้ว สามารถวิเคราะห์องค์ความรู้ที่จะนำมาสร้างระบบได้ดังนี้

4.1.1 สภาพของปัญหา

4.1.1.1 ปัญหาลักษณะนี้ เป็นปัญหาที่จะต้องอาศัยผู้ที่มีประสบการณ์ มีความชำนาญจริงๆในการแก้ปัญหา เพื่อให้ได้มาซึ่งคำตอบที่เหมาะสม

4.1.1.2 ผู้เชี่ยวชาญที่สามารถแก้ไขปัญหาได้ดี จะทำหน้าที่ 2 ส่วน คือ ส่วนหนึ่งจะต้องจดจำลักษณะต่างๆ ของอ้อยแต่ละพันธุ์ ซึ่งจะรวมถึงค่าที่เป็นไปได้ของลักษณะนั้นๆ ด้วย และอีกส่วนหนึ่งจะเป็นขั้นตอนการวินิจฉัย โดยการพิจารณาลักษณะของอ้อยที่พบ ซึ่งจะอาศัยความชำนาญและประสบการณ์ในการพิจารณาและค้นหาคำตอบ

4.1.1.3 ลักษณะต่างๆ ของอ้อย ประกอบด้วยลักษณะเชิงปริมาณ (quantitative characteristic) เป็นลักษณะที่สามารถวัดได้ บอกออกมาเป็นค่าตัวเลขที่มีความต่อเนื่อง ได้แก่ ขนาดของลำ ความยาวใบ ความกว้างของบริเวณเกิดราก เป็นต้น และลักษณะเชิง

คุณภาพ (quanlitative characteristic) เป็นลักษณะที่ไม่สามารถวัดเป็นตัวเลขได้ ได้แก่ รูปร่าง ปล้อง สีลำต้น ร่องเหนือตา เป็นต้น ลักษณะทั้งหมดนี้เป็นลักษณะที่ปรากฏออกมาภายนอก ซึ่งถูกควบคุมโดยทางพันธุกรรมและสภาพแวดล้อม ดังนั้นการวินิจฉัยจึงค่อนข้างที่จะลำบากในการหาเหตุผลมาอธิบายคำตอบ รวมถึงขั้นตอนที่ใช้ในการพิจารณาหาคำตอบ

4.1.1.4 ผู้เชี่ยวชาญแต่ละคน อาจจะมีวิธีการหรือขั้นตอนในการพิจารณาที่ไม่เหมือนกัน และลักษณะในการจดจำค่าของคุณสมบัติต่างๆ ของอ้อยจะแตกต่างกัน ขึ้นอยู่กับประสบการณ์ของแต่ละคน

4.1.1.5 ลักษณะของความรู้ในปัญหานี้ ค่อนข้างที่จะเปลี่ยนแปลงได้ง่ายกว่าความรู้ประเภทอื่นๆ เนื่องจากมีปัจจัยที่ทำให้เกิดความแปรปรวนที่ไม่สามารถควบคุมได้ นั่นคือ สารพันธุกรรมและสภาพแวดล้อม นอกจากนี้อ้อยซึ่งเป็นสิ่งมีชีวิตอาจจะมีการปรับตัวให้เข้ากับสภาพแวดล้อมได้ หรือมีการพัฒนาการเพื่อความอยู่รอดได้เมื่อเวลาผ่านไประยะหนึ่ง

4.1.2. <u>แนวทางการแก้ปัญหาในระบบผู้เชี่ยวชาญ</u>

4.1.2.1 ลักษณะการทำงานของระบบผู้เชี่ยวชาญมีความสอดคล้องกับการแก้ปัญหาของผู้เชี่ยวชาญ กล่าวคือ ในส่วนของความจำลักษณะของอ้อยพันธุ์ต่างๆ ระบบจะสามารถจดจำโดยการเก็บรวบรวมความรู้ทั้งหมดไว้ในฐานความรู้ ในกรณีนี้จะค่อนข้างได้เปรียบกว่าผู้เชี่ยวชาญคือระบบจะสามารถเก็บความรู้ในรายละเอียดได้มากและแม่นยำกว่า ส่วนด้านการวินิจฉัยเพื่อหาคำตอบ ระบบจะทำงานตามขั้นตอนของกลไกการหาเหตุผล สำหรับกรณีนี้ระบบจะมีความยืดหยุ่นในการหาเหตุผลได้ดีเท่ากับผู้เชี่ยวชาญหรือไม่นั้นขึ้นอยู่กับความสามารถของระบบและความยากในการออกแบบและพัฒนาระบบ ในการที่จะทำให้ระบบผู้เชี่ยวชาญมีความสามารถคล้ายผู้เชี่ยวชาญจริงๆ นั้นคงจะต้องอาศัยเวลาในการสร้างและปรับปรุงระบบเพื่อให้มีความยืดหยุ่นสูงคล้ายกับมนุษย์

4.1.2.2 ในขั้นต้นของการพัฒนาระบบผู้เชี่ยวชาญนี้ ต้องพยายามที่จะหาหลักการในการใช้เหตุผลของผู้เชี่ยวชาญ เพื่อที่จะได้มาซึ่งคำตอบที่ถูกต้องหรือใกล้เคียงความจริงมากที่สุด จึงได้กำหนดให้ลักษณะของอ้อยที่สภาพแวดล้อมมีอิทธิพลมากเป็นลักษณะที่ใช้

ประกอบการตัดสินใจ หรือให้สภาพแวดล้อมมีส่วนในการพิจารณาน้อยที่สุดก่อน โดยจะยังไม่คำนึงถึงสถานที่ปลูก สภาพการให้น้ำ ดิน อุณหภูมิ ช่วงฤดูปลูก การรับแสง การระบาดและทำลายของโรคและแมลง ฯ เหล่านี้เป็นต้น

4.2 การออกแบบฐานความรู้

4.2.1 การพิจารณาเลือกวิธีการแทนค่าความรู้

จากการศึกษาลักษณะของพันธุ์อ้อย พบว่าอ้อยแต่ละพันธุ์มีลักษณะทางพฤกษศาสตร์เหมือนกัน คือมี ลำต้น ข้อ ปล้อง ใบ ดอก ในลักษณะเดียวกัน แต่จะแตกต่างกันที่ค่าของลักษณะ เช่น รูปร่างปล้อง พันธุ์ F 140 อาจจะเป็นแบบทรงกระบอก ส่วนพันธุ์อู่ทอง 1 เป็นแบบกลางคอด เป็นต้น ดังนั้นลักษณะของความรู้จะเป็นในรูปแบบ object - attribute - value ในการจำแนกพันธุ์อ้อยจะต้องพิจารณาค่าของลักษณะ (attribute value) ของแต่ละลักษณะ (attribute) ซึ่งมีจำนวนหลายๆ ลักษณะดังนั้นวัตถุจึงมีขนาดค่อนข้างใหญ่เพราะมีคุณสมบัติมาก สำหรับการออกแบบฐานความรู้ในการศึกษาครั้งนี้เลือกวิธีการแทนค่าความรูปแบบกรอบและใช้ตารางประกอบ ด้วยเหตุผลที่ว่า กรอบเป็นการแทนค่าความรู้ที่เน้นวัตถุ เน้นสิ่งที่ต้องการแสดงเป็นหลัก กรอบสามารถใช้อธิบายวัตถุได้ดี เหมาะสำหรับการพัฒนาเป็นระบบต้นแบบ ซึ่งจะสามารถเพิ่มวัตถุที่เกี่ยวข้องในระบบได้ภายหลัง ส่วนการใช้ตารางช่วยเก็บความรู้ที่นอกเหนือจากคุณสมบัติที่สำคัญจะทำให้ขนาดของกรอบเล็กลง และการทำงานว่องไวมากขึ้นในกรณีที่ใช้ภาษาวิชวล เบสิค ซึ่งจะถูกเลือกเป็นเครื่องมือที่ใช้พัฒนาโปรแกรม (รายละเอียดในบทที่ 5) และการเพิ่มเติมความรู้ใหม่ทำได้ไม่ยาก

4.2.2 คุณสมบัติของกรอบ

การแทนค่าความรู้แบบกรอบ ใช้ในการรวบรวมสิ่งต่างๆ ที่มีส่วนเกี่ยวข้องกันให้มาอยู่รวมกัน เช่น ความคิดที่มีความสัมพันธ์กัน แนวความคิดหลักการ ความจริง นิพจน์ เป็นต้น

กรอบเป็นการแสดงความรู้ที่เน้นวัตถุ เหตุการณ์หรือสภาพโดยที่ความรู้ต่างๆ จะถูกจัดเป็นสัดส่วนเรียกว่ากรอบ มีข้อมูลแสดงความสัมพันธ์ระหว่างกรอบ (วิลาศ วูวงศ์ และบุญเจริญ ศิริเนาวกุล, 2535)

กลุ่มของกรอบจะประกอบเป็นฐานความรู้ ซึ่งเมื่อรวมกับกลไกการอนุมานจะเป็นระบบฐานกรอบ (frame base system) หรือระบบกรอบ (frame system)

4.2.2.1 โครงสร้างของกรอบ (frame structure) กรอบเป็นโครงสร้างที่มีคุณสมบัติมากกว่าโครงสร้างของภาษาบางภาษา เช่น ภาษาซี หรือ ปาสคาล (PASCAL) ทั้งนี้เพราะในกรอบนอกจากจะประกอบด้วย Slot ซึ่งใช้เก็บค่าแสดงคุณสมบัติของวัตถุหรือสภาพเหตุการณ์ สารสนเทศ (information) และค่าต่างๆ หรือ pointer ที่ชี้ไปยังกรอบอื่นๆ แล้ว ยังสามารถประกอบด้วยชุดคำสั่ง (procedure) ที่จะถูกใช้งานเมื่อมีการเปลี่ยนแปลงค่าของ slot

4.2.2.2 โครงสร้างของระบบกรอบประกอบด้วยกรอบและความสัมพันธ์ระหว่างกรอบจะถูกจัดเป็นลำดับชั้น (hierarchy) โดยที่กรอบในระดับที่ต่ำกว่าจะเก็บความรู้ที่มีความจำเพาะ (specific) มากกว่ากรอบที่อยู่ในระดับสูง

4.2.2.3 คุณสมบัติการถ่ายทอดและการกำหนดค่า (inheritance and default value) เนื่องจากกรอบถูกจัดเป็นลำดับชั้น ทำให้มีคุณสมบัติในการถ่ายทอด โดยที่กรอบที่อยู่ในระดับต่ำกว่าจะได้รับการถ่ายทอดคุณสมบัติของกรอบที่อยู่ในระดับสูงกว่า

4.2.2.4 กรอบเป็นการแทนค่าความรู้แบบรวบยอด โดยใช้โครงสร้างข้อมูลชนิดเดียวสามารถแทนความรู้ได้หลายประเภท

4.2.2.5 การอนุมานที่ใช้ในระบบกรอบ จะสามารถเป็นไปตามความคาดหมายและมีความยืดหยุ่นในการอนุมาน ไม่มีการกำหนดวิธีที่ตายตัว เลือกใช้การอนุมานที่เหมาะสมกับปัญหาได้ และได้ผลลัพธ์ตามความคาดหมาย

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

4.2.3 <u>คุณสมบัติของตาราง</u>

ตาราง (table) เป็นลักษณะของการจัดเก็บข้อมูลประเภทหนึ่งที่มีประสิทธิภาพสูง สามารถใช้ได้ง่ายในหลายๆ กรณี การจัดเก็บข้อมูลจะอยู่ในลักษณะของระเบียน (record) และเขตข้อมูล (field) ทำให้การจัดเก็บเป็นระเบียบ ปัจจุบันเทคโนโลยีเกี่ยวกับการใช้ตารางในด้านฐานข้อมูลได้พัฒนาไปไกลมาก ทำให้รูปแบบการใช้ทำได้ง่ายและขีดความสามารถสูง

4.2.4 <u>การสร้างฐานความรู้</u>

จากคุณสมบัติของกรอบและตารางดังกล่าว สามารถนำมาเป็นแนวทางในกำหนดคุณลักษณะและสร้างเป็นกรอบและตารางที่จะใช้แทนค่าความรู้พันธุ์อ้อยได้ ตามขั้นตอนดังนี้

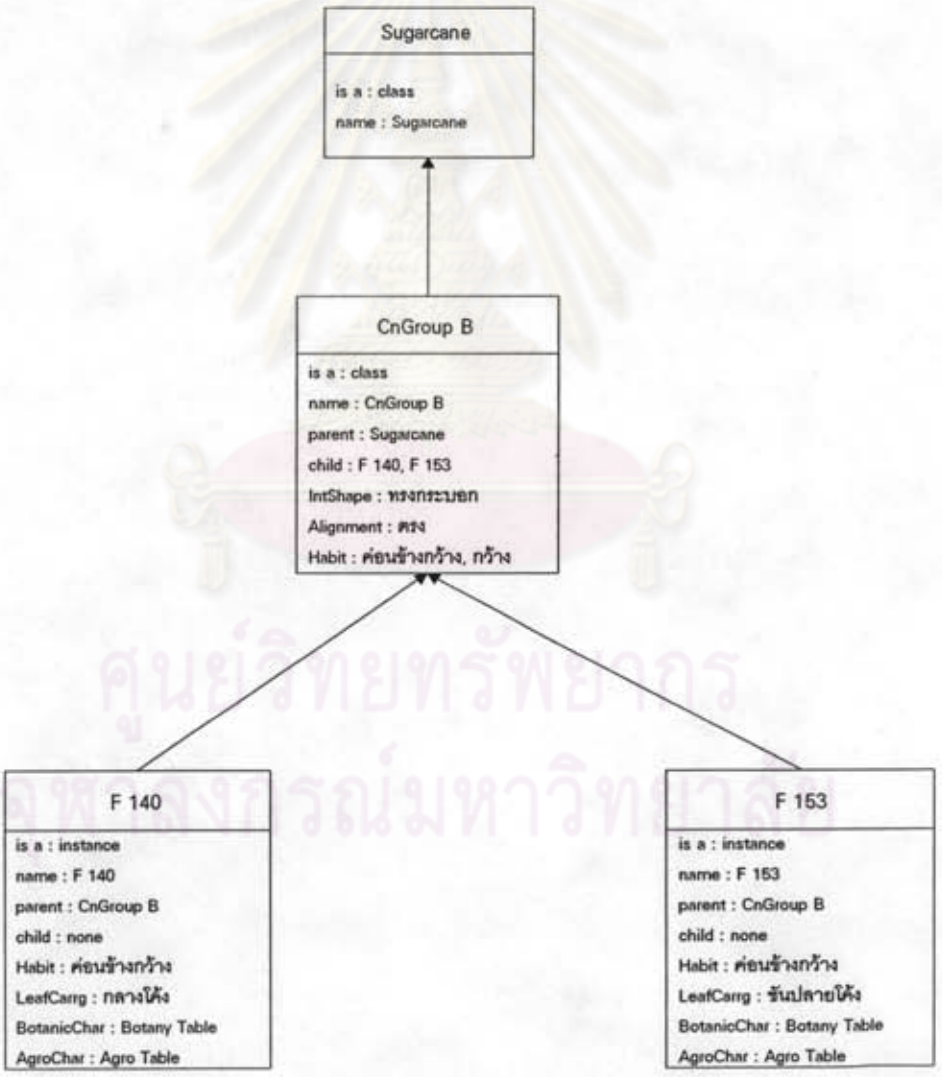
4.2.4.1 กำหนดวัตถุ เนื่องจากกรอบสามารถเขียนแทนด้วยภาษาโปรแกรมเชิงวัตถุ (object oreinted programming) ซึ่งจะมีวิธีสร้างที่คล้ายกับการสร้างคลาส (class) เริ่มต้นด้วยการหาวัตถุ อธิบายลักษณะของวัตถุ โดยพิจารณาปัญหาแล้วหาว่ามีวัตถุใดบ้างที่เกี่ยวข้องกับปัญหานั้น สำหรับการจำแนกพันธุ์อ้อยในที่นี้ มีวัตถุที่เกี่ยวข้องและอยู่ในขอบเขตที่พิจารณาเพียงอย่างเดียว คือ อ้อย 50 พันธุ์ ในกรณีที่สนใจปัจจัยอื่นๆ ที่มีผลต่อพันธุ์อ้อยจะสามารถเพิ่มเติมวัตถุภายหลังได้ง่าย

4.2.4.2 กำหนดคุณสมบัติของวัตถุ เมื่อสามารถหาวัตถุได้แล้วขั้นตอนต่อไปคือ การกำหนดคุณสมบัติของวัตถุแต่ละชนิด อธิบายคุณสมบัติตั้งแต่ทั่วไป จนถึง คุณสมบัติเฉพาะเจาะจง ตั้งแต่คุณสมบัติทางรูปธรรมจนถึงคุณสมบัติประจำที่เป็นนามธรรม ได้แก่ หน้าที่หรือการกระทำ ของวัตถุนั้นๆ ในขั้นนี้เป็นการกำหนดช่อง (slot) ของกรอบด้วย ในการศึกษาครั้งนี้ได้ออกแบบกรอบไว้ดังตัวอย่างรูปที่ 4.1

4.2.4.3 กำหนดความสัมพันธ์ เมื่อได้ชนิดของวัตถุและคุณสมบัติประจำของวัตถุแล้ว ขั้นตอนต่อไป จะเป็นการหาความสัมพันธ์ของวัตถุเพื่อสร้างเป็นระบบกรอบ สำหรับการศึกษาในครั้งนี้มีวัตถุเพียงอย่างเดียว แต่มีจำนวนตัวอย่าง (instance) 50 ตัวอย่าง ซึ่งตัวอย่างเหล่านี้จะมีความสัมพันธ์ที่สามารถจัดเป็นโครงสร้างลำดับชั้น (hierachy) โดยการจัดเป็นคลาสย่อยซึ่ง

จะเป็นคลาสนามธรรม เป็นการจัดกลุ่มด้วยการนำตัวอย่างที่มีคุณสมบัติหลักเหมือนกันจัดรวมไว้เป็นกลุ่มเดียวกัน ดังแสดงในรูปที่ 4.4

4.2.4.4 การจัดเก็บคุณสมบัติในตาราง เมื่อสร้างกรอบได้เรียบร้อยแล้ว เปลี่ยนรูปคุณสมบัติที่จำเพาะของอ้อยแต่ละพันธุ์ คุณสมบัติที่ไม่ได้ใช้ในการจำแนก รวมถึงลักษณะด้านพฤกษศาสตร์และด้านการเกษตร เก็บไว้ในตารางโดยให้ชื่อพันธุ์เป็นระเบียน และคุณสมบัติเป็นเขตข้อมูล และค่าของคุณสมบัติเป็นค่าของเขตข้อมูล ซึ่งในกรณีนี้ 1 ระเบียนจะมีค่าเท่ากับ 1 กรอบ



รูปที่ 4.1 ตัวอย่างแสดงกรอบและช่อง (slot) ของอ้อยพันธุ์ F 140 และ F 153

4.3 การออกแบบกลไกการหาเหตุผล

4.3.1 การพิจารณาเลือกกลไกการหาเหตุผล

กลไกการหาเหตุผลเป็นสิ่งที่สำคัญส่วนหนึ่งที่จะทำให้การทำงานของระบบผู้เชี่ยวชาญมีประสิทธิภาพ การเลือกวิธีการหาเหตุผลหรือกลไกการทำงานจะขึ้นอยู่กับลักษณะของปัญหา เมื่อพิจารณาหากลไกการหาเหตุผลในการจำแนกพันธุ์อ้อยตามความเหมาะสมของปัญหา สมควรที่จะใช้วิธีกลไกการหาเหตุผลที่ค่อนข้างไปทางข่ายความหมาย และใช้หลักการที่เลียนแบบผู้เชี่ยวชาญที่ใช้เหตุผลในการพิจารณาลักษณะต่างๆ โดยสร้างเป็นต้นไม้การตัดสินใจ (decision tree) กำหนดให้โหนดของต้นไม้แทนลักษณะของอ้อย ส่วนใบเป็นพันธุ์อ้อยที่จำแนกได้ การค้นหาใช้วิธีค้นตามแนวลึก (depth first search) ซึ่งในกรณีที่พบโหนดที่ต้องการจะเป็นวิธีการที่ใช้เวลาน้อยและระยะทางในการค้นหาสั้น และยังเหมาะสมกับการออกแบบฐานความรู้ที่จัดกลุ่มพันธุ์ที่มีลักษณะหลักเหมือนกันไว้เป็นกลุ่มเดียวกัน

4.3.2 การเลือกลักษณะที่ใช้ในการวินิจฉัยพันธุ์

ในการพิจารณาเลือกลักษณะที่จะใช้ในการวินิจฉัย (decision factors) เป็นเรื่องที่ยากถ้าเลือกลักษณะที่ไม่ถูกต้องหรือไม่เหมาะสม จะทำให้ได้คำตอบที่ไม่ถูกต้อง ลักษณะของอ้อยที่รวบรวมไว้ทั้งหมด 40 ลักษณะ ในการจำแนกพันธุ์ 50 พันธุ์ ไม่จำเป็นที่จะต้องใช้ทุกลักษณะที่มีในฐานความรู้ เพื่อให้ระบบทำงานได้รวดเร็วควรใช้ลักษณะที่สำคัญและสามารถจำแนกพันธุ์ได้อย่างถูกต้อง Quinlan (1986) กล่าวว่า ความสำคัญอย่างหนึ่งของการวิจัยด้านการเรียนรู้ (learning) ก็เพื่อหาลักษณะที่ใช้ในการตัดสินใจที่ดีได้ ดังนั้นในขั้นตอนนี้อาจใช้วิธีการ induction ซึ่งเป็นการสร้างกฎขึ้นจากตัวอย่างที่ดีของความรู้ (Durkin, 1994) โดยใช้วิธีการของ ID3 เข้าช่วยจะสามารถลดปัญหายุ่งยากบางประการในการเลือกลักษณะที่ใช้ตัดสินใจได้ และทำให้ระบบผู้เชี่ยวชาญทำงานได้อย่างมีประสิทธิภาพ

วิธีการทำงานของ ID3 จะคำนวณหาค่าเอนโทรปี (entropy) ของแต่ละลักษณะ จากทฤษฎีความรู้ (information theory) โดยใช้ฟังก์ชัน (Ignizio, 1991)

$$H(C/A_k)=\sum_{j=1}^{Mk} p(a_{k,j})\cdot[-\sum_{i=1}^{N} p(c_i/a_{k,j})\cdot \log_2 p(c_i/a_{k,j})]$$

โดยที่ $H(C/A_k)$ เป็นค่า entropy ของการแยกคลาสของลักษณะ A_k

$p(a_{k,j})$ เป็นค่าความเป็นไปได้ของลักษณะ k ที่ค่าเป็น j

$p(c_i/a_{k,j})$ เป็นค่าความเป็นไปได้ที่ค่าของคลาสเป็น c_i เมื่อลักษณะ k มีค่าเป็น j

M_k จำนวนทั้งหมดของค่าที่เป็นไปได้ของลักษณะ Ak ที่ j = 1, 2,...,M_k

N เป็นจำนวนทั้งหมดของคลาสอื่น มีค่าตั้งแต่ I = 1, 2,...,N

K เป็นจำนวนทั้งหมดของลักษณะ มีค่าตั้งแต่ k = 1, 2,...,K

ค่าเอนโทรปีที่น้อยแสดงว่า ลักษณะนั้นสามารถแยกตัวอย่างออกเป็นคลาสได้ดี ควรเลือกลักษณะนั้นเป็นลักษณะเริ่มต้น แล้วนำลักษณะที่เหลือมาคำนวณค่าเอนโทรปีต่อไป แล้วนำลักษณะเหล่านี้มาต่อเป็นต้นไม้การตัดสินใจ ทำซ้ำจนกระทั่งสามารถแยกคลาสได้ตามต้องการ ก็จะได้ต้นไม้ขนาดเล็กและมีประสิทธิภาพมาก อย่างไรก็ตามการทำ induction ก็ยังมีข้อจำกัดที่ความถูกต้องของกฎขึ้นอยู่กับคุณภาพของข้อมูล Durkin (1994) ได้กล่าวถึงข้อจำกัดของ ID3 ว่า ในบางกรณีกฎบางข้อไม่สามารถเป็นไปได้ ในกรณีที่มีตัวอย่างเหมือนกันหลายตัวอย่างจะไม่มีผลมากกว่าคลาสที่มีเพียงหนึ่งตัวอย่าง และไม่สามารถใช้ได้กับตัวอย่างที่ขัดแย้งกัน ถ้าจำนวนตัวอย่างมีน้อย อาจมีผลกระทบต่อการสร้างต้นไม้ที่มีประสิทธิภาพได้ ปัจจุบัน ID3 ถูกปรับปรุงให้มีประสิทธิภาพการทำงานสูงขึ้นจากการศึกษาของ Ignizio (1991) พบว่า ID3 ไม่เพียงแต่มีวิธีการที่จะลดขนาดของต้นไม้ลงเท่านั้น แต่ยังมีประสิทธิภาพการทำงานถูกต้องสูงขึ้นอีกด้วย โดยที่ C4.5 เป็นโปรแกรมที่สร้างกฎและต้นไม้การตัดสินใจโดยใช้ขั้นตอนวิธีของ ID3 สามารถทำงานได้โดยมีความถูกต้องมากกว่า 85 เปอร์เซนต์

จากการศึกษาในครั้งนี้ไม่ได้ใช้ ID3 algorithm ในการสร้างต้นไม้การตัดสินใจ เนื่องจากจำนวนตัวอย่างที่ใช้ทั้งหมดมีเพียง 50 ตัวอย่าง และแบ่งเป็นคลาส 50 คลาส นับว่าเป็นตัวอย่างที่มีขนาดเล็กมากในการเรียนรู้ เมื่อเทียบกับระบบ AQ11 ที่ใช้ 290 ตัวอย่างในการเรียนรู้ และใช้ 340 ตัวอย่างสำหรับการทดสอบ เพื่อวินิจฉัยโรคถั่วเหลือง 15 โรค โดยใช้ลักษณะที่พิจารณา

35 ลักษณะ (Durkin, 1994) และจากการค้นคว้าข้อมูลจากเอกสารหลายแหล่งพบว่า ลักษณะสำคัญของอ้อยที่มีการบันทึกไว้ขัดแย้งกัน ทั้งนี้เนื่องมาจากเวลาและสถานที่เก็บข้อมูลต่างกัน จำนวนตัวอย่างที่ได้ก็ต่างกัน บางพันธุ์มีมากกว่า 5 ตัวอย่าง ในขณะที่บางพันธุ์ไม่มีตัวอย่างเลย จากปัจจัยในเรื่องเวลาและสถานที่ อาจทำให้อ้อยมีการเปลี่ยนแปลง เช่น พันธุ์เสื่อมลง ปรับตัวตามสภาพแวดล้อมแตกต่างกัน และด้วยสาเหตุซึ่งเกิดจากปริมาณและคุณภาพข้อมูลดังกล่าว ซึ่งมีผลกระทบต่อความถูกต้องในการสร้างกฎฉะนั้นในการพัฒนาระบบเพื่อให้ได้โปรแกรมต้นแบบในขั้นต้น จึงได้เลือกใช้วิธีการหลักของผู้เชี่ยวชาญ ประกอบกับหลักการของ Induction โดยคัดเลือกลักษณะที่มีความสอดคล้องกับเอกสารอ้างอิงต่างๆ มาเป็นส่วนหนึ่งของลักษณะที่ใช้วินิจฉัย เพื่อให้ได้ต้นไม้ที่มีขนาดเหมาะสมและทำงานได้ใกล้เคียงกับผู้เชี่ยวชาญมากที่สุด

จากการกำหนดวัตถุและคุณสมบัติของวัตถุเพื่อสร้างกรอบในฐานความรู้ทำให้ทราบว่าลักษณะของอ้อยแต่ละพันธุ์เป็นอย่างไร การที่ผู้เชี่ยวชาญพิจารณาพันธุ์อ้อยจะดูที่ลักษณะใดบ้าง เมื่อทำการศึกษาเรื่องพันธุ์อ้อยพบว่า ลักษณะภายนอกของอ้อยจะขึ้นอยู่กับพันธุ์และสภาพแวดล้อม ในการทำต้นแบบระบบระบบผู้เชี่ยวชาญในขั้นต้นจะพยายามหลีกเลี่ยงปัจจัยที่มีผลกระทบต่อการให้เหตุผลก่อน เพื่อให้การพัฒนาระบบเป็นไปอย่างมีขั้นตอนและสามารถทำงานได้อย่างมีประสิทธิภาพ ได้แบ่งกลุ่มของลักษณะประจำพันธุ์ออกเป็น 4 กลุ่ม คือ

กลุ่มที่ 1 เป็นกลุ่มลักษณะเห็นได้ชัดเจนและสภาพแวดล้อมมีผลน้อย ได้แก่ ลักษณะรูปร่างปล้อง การจัดเรียงปล้อง ทรงกอ และทรงใบ

กลุ่มที่ 2 เป็นกลุ่มลักษณะเห็นได้ค่อนข้างชัดเจนและสภาพแวดล้อมมีผลกระทบต่อการแสดงออกของลักษณะดังกล่าวมากกว่ากลุ่มแรก ลักษณะที่จัดอยู่ในกลุ่มนี้ ได้แก่ ร่องเหนือตา รอยแตกตื้น รอยแตกลึก ตำแหน่งของวงเจริญเมื่อเทียบกับตา และรูปร่างตา

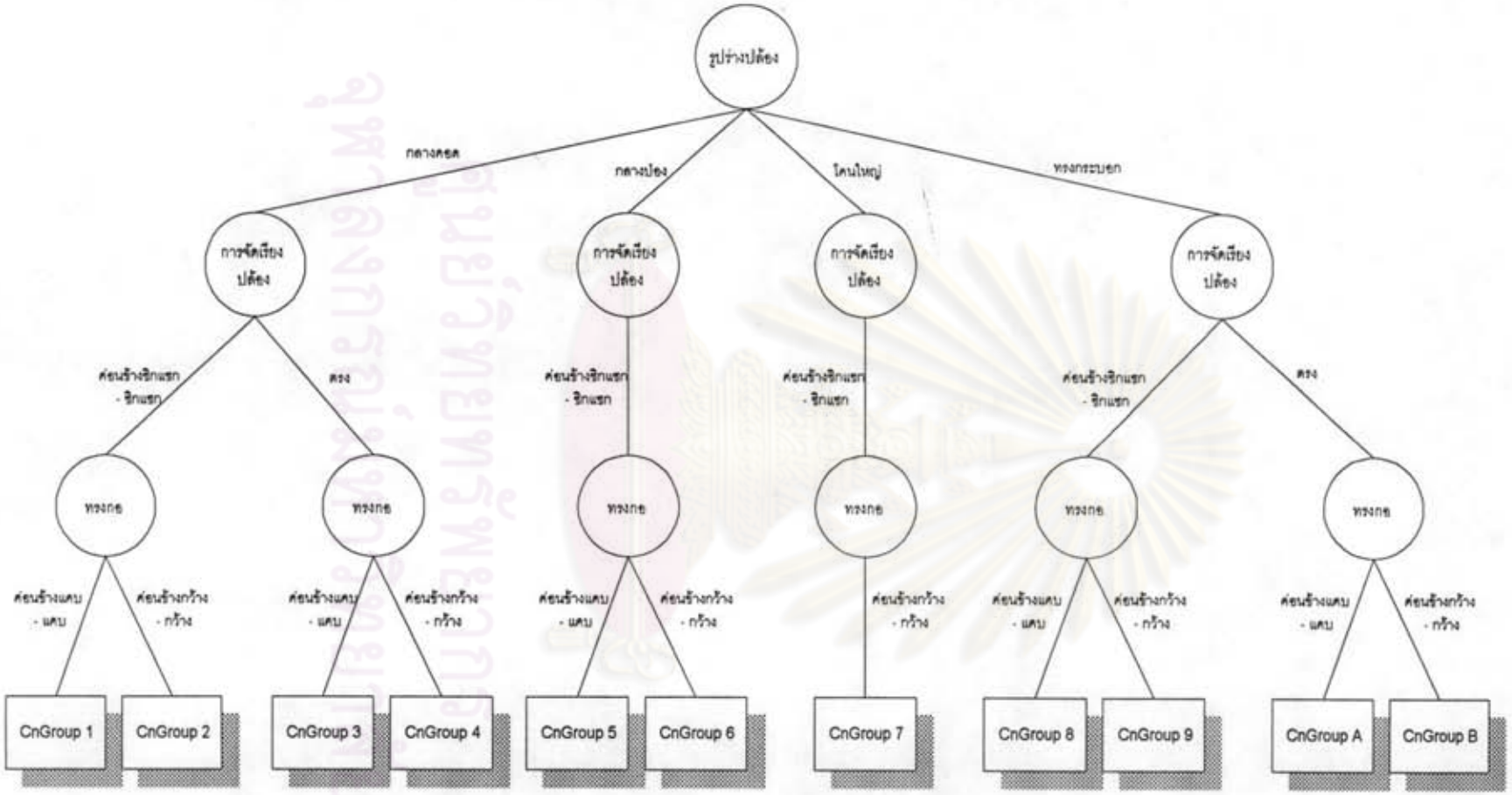
กลุ่มที่ 3 เป็นกลุ่มลักษณะที่พิจารณาได้ค่อนข้างยากและมีความแปรปรวนมากกว่ากลุ่มที่ 2 แต่ก็ยังสามารถใช้พิจารณาได้ตามความชำนาญของผู้เชี่ยวชาญ ลักษณะที่จัดอยู่ในกลุ่มนี้ ได้แก่ วงไข หูใบ คอใบ

กลุ่มที่ 4 เป็นกลุ่มลักษณะที่มีความแปรปรวนมากกว่ากลุ่มอื่นๆ การพิจารณาลักษณะในกลุ่มนี้ทำได้ยาก อย่างไรก็ตาม ลักษณะที่อยู่ในกลุ่มนี้มักเป็นลักษณะเด่นที่เป็นเอกลักษณ์ของพันธุ์ ลักษณะในกลุ่มนี้ ได้แก่ สีของลำต้นส่วนที่กาบหลุด การจัดเรียงตัวของปุ่มราก ขนเบอร์ 57 สีกาบใบ สีคอใบ ขนาดลำ ขนาดตา ความยาวร่องตา สีใบ รวมถึงลักษณะอื่นๆ ที่เป็นรายละเอียดนอกเหนือจากที่กล่าวมาข้างต้นทั้งหมด

เมื่อได้ลักษณะที่ใช้พิจารณาแล้ว ในขั้นต่อไปก็จะกำหนดค่าที่เป็นไปได้ของแต่ละลักษณะโดยใช้หลักการของ Artschwager และ Brandes (1958) กับ กรมวิชาการเกษตร (2533) เช่น รูปร่างปล้อง มีลักษณะที่เป็นได้ 6 แบบ คือ กลางคอด กลางโค้ง โคนใหญ่ ปลายใหญ่ กลางป่อง และทรงกระบอก เป็นต้น รายละเอียดส่วนใหญ่ได้อธิบายไว้ในบทที่ 3 แล้ว

4.3.3 <u>การสร้างต้นไม้การตัดสินใจ (decision tree)</u>

เมื่อเลือกลักษณะที่ใช้ในการวินิจฉัยได้แล้ว นำลักษณะเหล่านั้นมาสร้างเป็นต้นไม้การตัดสินใจโดยให้ความสำคัญของลักษณะในกลุ่มที่ 1 ก่อนโดยเรียงลำดับจากรูปร่างปล้อง การจัดเรียงปล้อง ทรงกอ และทรงใบ ทั้ง 4 ลักษณะจะเป็นโหนด ส่วนกลุ่มของพันธุ์ที่ถูกจัดในฐานความรู้จะเป็นใบ ดังแสดงในตัวอย่างรูปที่ 4.2 เมื่อแยกได้คลาสย่อย (กลุ่ม) ต่างๆ แล้ว ในคลาสย่อยก็จะประกอบด้วยต้นไม้เล็กๆ อีก โดยมีลักษณะในกลุ่มที่ 2 และ 3 เป็นโหนดที่ใช้ในการค้นหา ส่วนลักษณะในกลุ่มที่ 4 เป็นโหนดสุดท้ายที่จะถามเกี่ยวกับลักษณะเด่นของพันธุ์เพื่อเพิ่มความมั่นใจในการตัดสินใจก่อนถึงโหนดสุดท้ายซึ่งจะเป็นใบแสดงค่าชื่อพันธุ์ที่วินิจฉัยได้ กรณีที่โหนดปลายไม่มีพันธุ์อ้อยก็จะให้คำตอบคืนแก่ผู้ใช้ สิ่งสำคัญที่ใช้เป็นหลักการในการสร้างโครงสร้างต้นไม้นี้ Ignizio (1991) ได้กล่าวว่า การใช้ลักษณะที่พิจารณาจำนวนมากเกินไปจะทำให้ฐานความรู้ทำงานได้เชื่องช้าและดูเหมือนไม่ฉลาด ควรพิจารณาแต่ลักษณะที่สำคัญๆ และการใช้ลักษณะที่พิจารณาน้อยเกินไปจะเป็นการจำกัดประโยชน์ของลักษณะนั้น และในอนาคตจะทำให้การปรับปรุงหรือเปลี่ยนแปลงระบบทำได้ยาก



รูปที่ 4.2 แสดงต้นไม้หลักของการตัดสินใจจำแนกพันธุ์อ้อย

4.3.4 การสร้างวิธีการค้นหา

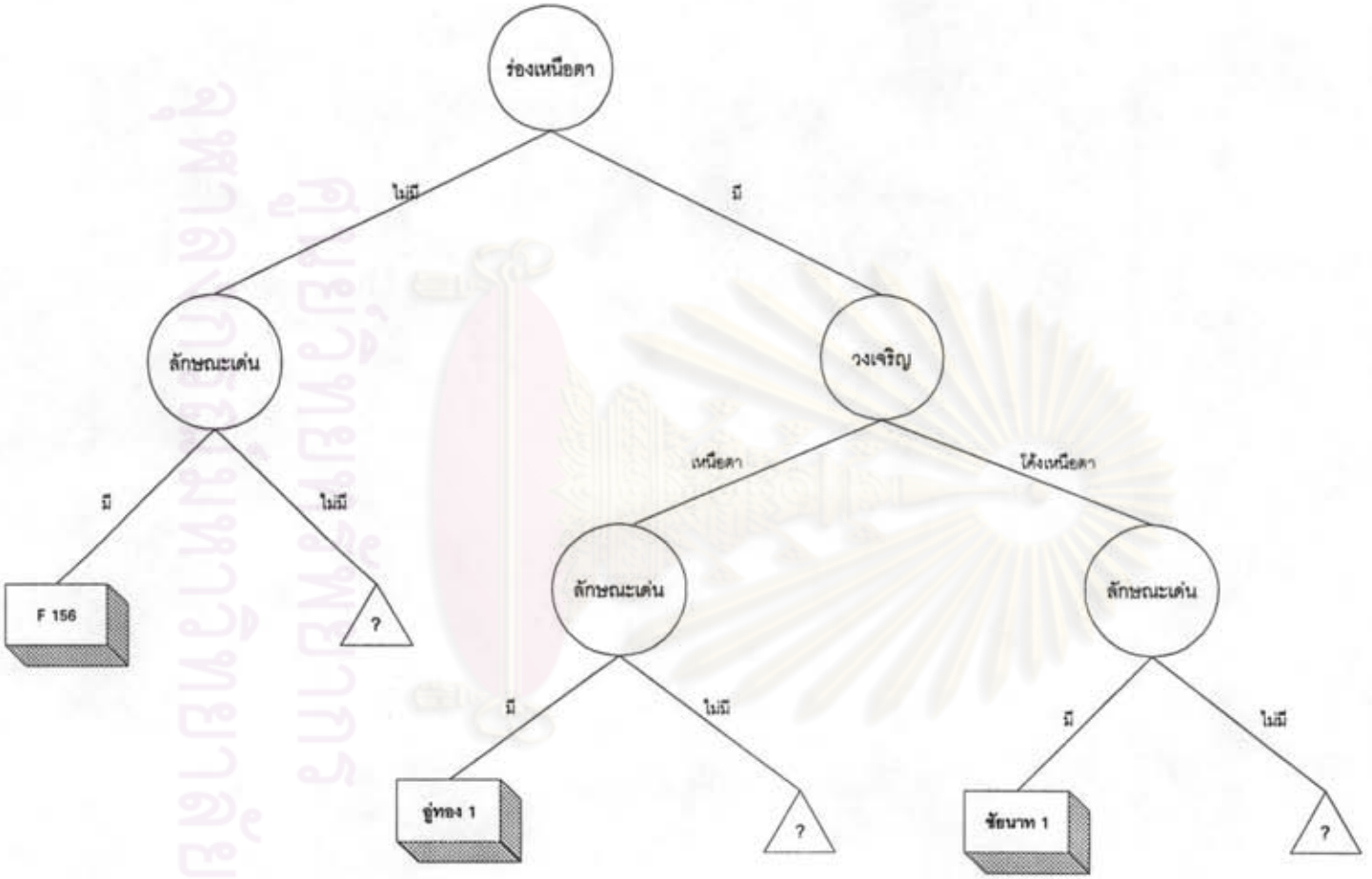
เมื่อได้โครงสร้างต้นไม้แล้ว สามารถเปลี่ยนโครงสร้างต้นไม้ให้เป็นกฎได้โดยการให้โหนดเป็นชื่อลักษณะ ส่วนกิ่งของต้นไม้เป็นค่าของลักษณะ และใบเป็นพันธุ์อ้อยที่เป็นคำตอบ แต่เนื่องจากกฎในที่นี้ไม่ได้ทำหน้าที่เป็นฐานความรู้ หากแต่ทำหน้าที่เป็นกลไกการหาเหตุผล ดังนั้นโหนดที่แสดงชื่อลักษณะจะถูกแทนด้วยคำถามจากระบบและเชื่อมโยงด้วยวิธีการค้นหาตามแนวลึกจากการใช้คำถาม if-then-else จากตัวอย่างในรูปที่ 4.3 สามารถสร้างเป็นกฎได้ดังนี้

```
Ask for ร่องเหนือตา ?
if ร่องเหนือตา = yes then
    Ask for ลักษณะเด่น?
    if ลักษณะเด่น = yes then
        Answer = F 156
else ร่องเหนือตา = No
    Ask for ตำแหน่งวงเจริญ ?
    if ตำแหน่งวงเจริญ = เหนือตา then
        Ask for ลักษณะเด่น ?
        if ลักษณะเด่น = yes then
            Answer = อู่ทอง 1
    else ตำแหน่งวงเจริญ = โค้งเหนือตา
        Ask for ลักษณะเด่น ?
        if ลักษณะเด่น = yes then
            Answer = ชัยนาท 1
```

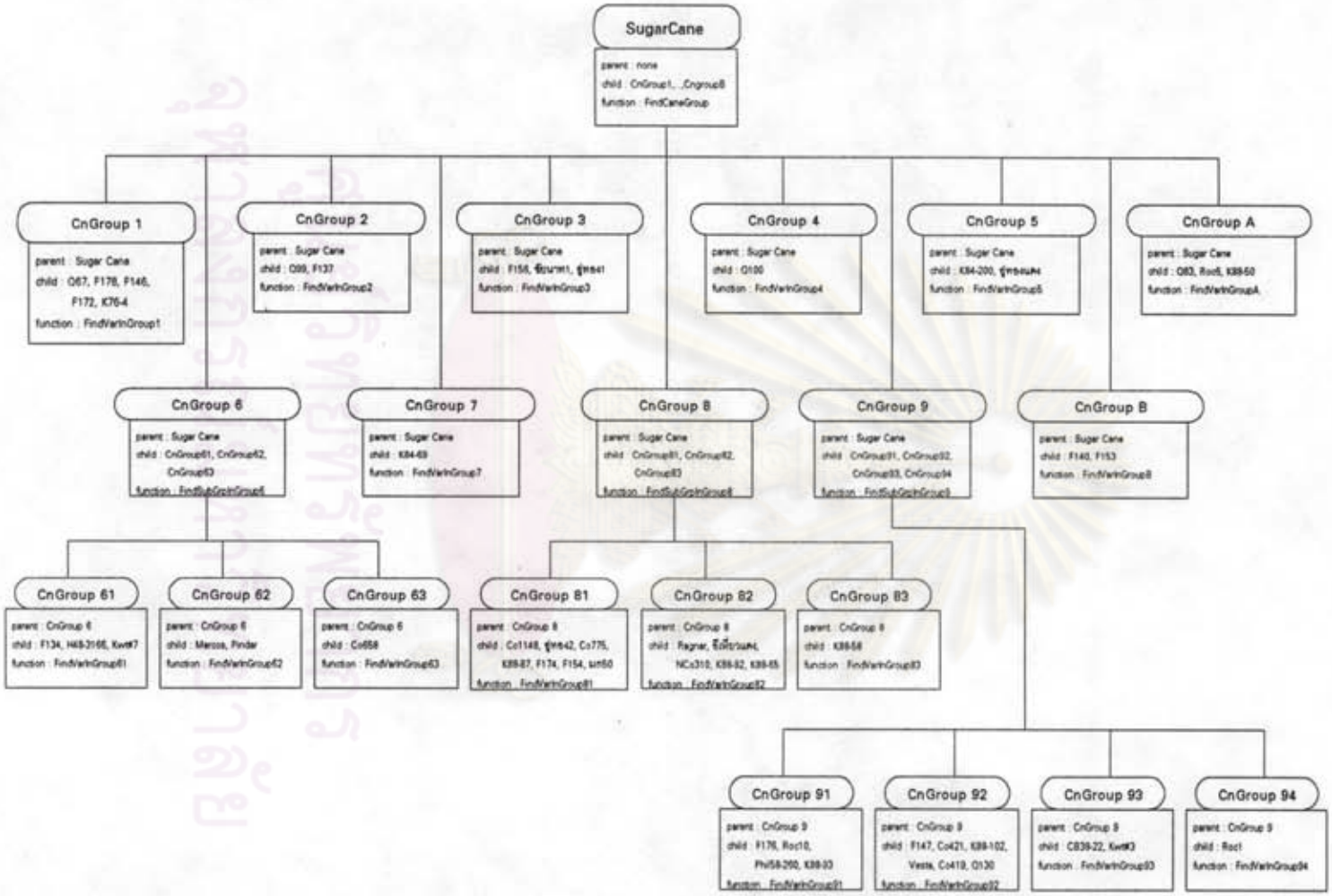
4.5 ฐานความรู้พันธุ์อ้อย

จากการออกแบบฐานความรู้และกลไกการหาเหตุผลเมื่อนำมารวมกันเป็นระบบกรอบในระบบผู้เชี่ยวชาญจะสามารถเขียนเป็นแผนภาพได้ดังรูปที่ 4.4

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 4.3 แสดงต้นไม้ย่อยในกลุ่มย่อย CnGroup 3



รูปที่ 4.4 แสดงโครงสร้างลำดับชั้นของกรอบความรู้พันธุ์อ้อย