หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2

การสร้างคลังข้อมูลแผนที่ฐาน

ถ้าพิจารณาข้อมูลทั้งหมดที่อยู่บนแผนที่เฉพาะกิจแผ่นใดแผ่นหนึ่งจะพบว่า เราสามารถแบ่งข้อมูลเหล่านั้นออกได้เป็น 2 ประเภทใหญ่คือ

1. ข้อมูลแผนที่ฐาน เช่น เส้นชายฝั่ง เกาะ ทะเลสาป เส้นแบ่งพรมแดน

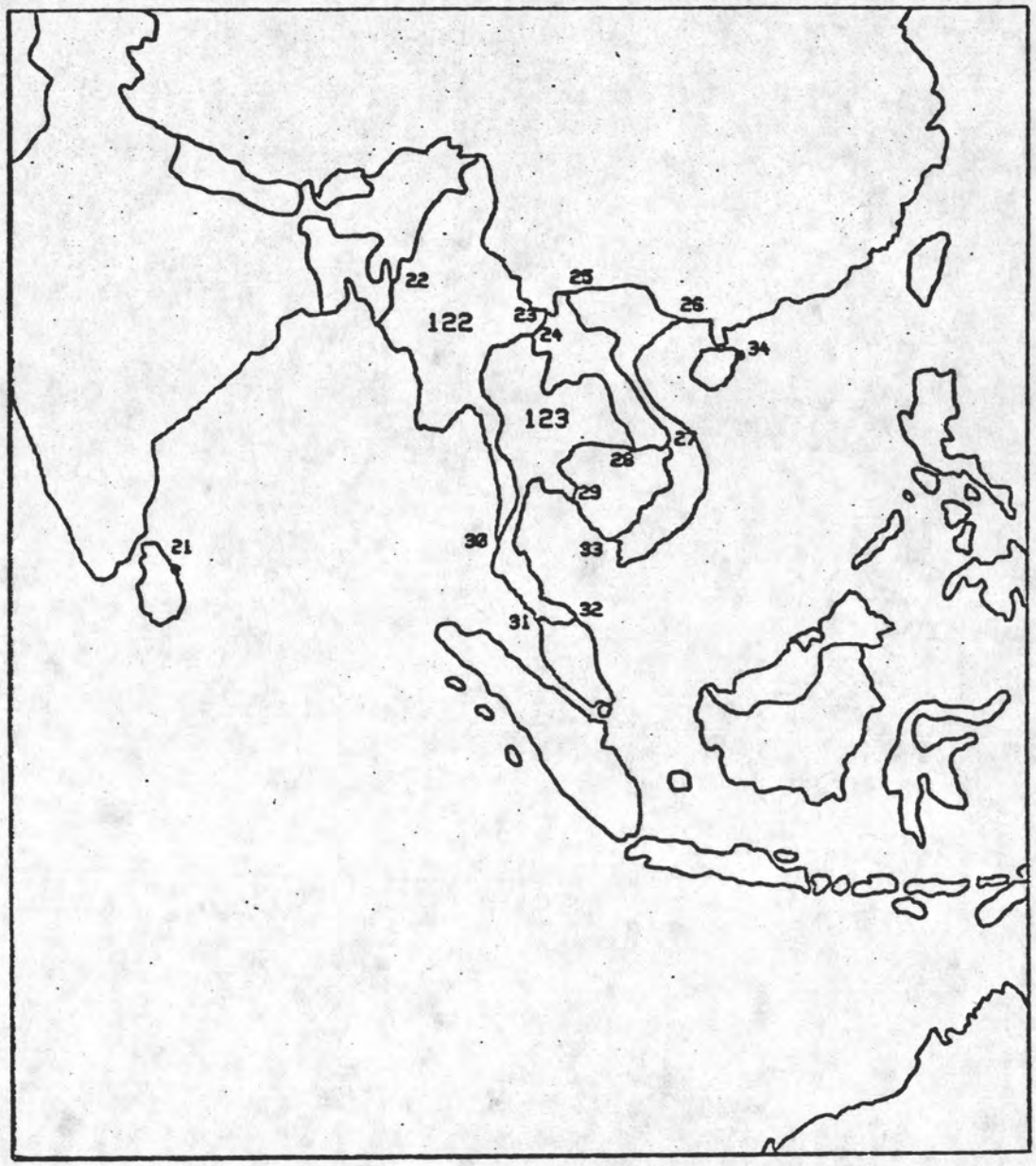
2. ข้อมูลเฉพะกิจ เช่น รายได้เฉลี่ยของประชากร ผลผลิต ฯลฯ ข้อมูลเฉพาะกิจมักอยู่ในรูปสัญญลักษณ์ กราฟ หรือแผนภูมิต่าง ๆ

เนื่องจากข้อมูลแผนที่ฐานเป็นข้อมูลที่ถูกใช้งานอยู่เสมอและไม่ค่อยมีการเปลี่ยนแปลง ในขณะที่ข้อมูลเฉพาะกิจชุดหนึ่ง ๆ จะถูกนำมาใช้เพียงไม่กี่ครั้ง และมีการเปลี่ยนแปลงอยู่ตลอดเวลา ตัวอย่างเช่น แผนที่เฉพาะกิจแสดงผลผลิตดีบุกของกลุ่มประเทศอาเซียน ข้อมูลแผนที่ฐานในกรณีนี้คือ เส้นขอบเขตของแต่ละประเทศ เกาะ หรือลักษณะทางภูมิศาสตร์อื่น ๆ ไม่ว่าเราจะแสดงผลผลิตดีบุกของปีไหนก็ตาม ข้อมูลแผนที่ฐานเหล่านี้ก็ยังเหมือนเดิม ในขณะที่ตัวข้อมูลเฉพาะกิจอันได้แก่ปริมาณดีบุกจะเปลี่ยนแปลงไป ถ้าหากนักทำแผนที่ต้องการสร้างแผนที่แสดงผลผลิตยางพาราของกลุ่มประเทศอาเซียนเพิ่มขึ้นมาอีก ก็ยังใช้ข้อมูลแผนที่ฐานชุดเดียวกับแผนที่แผ่นแรก ส่วนตัวข้อมูลเฉพาะกิจก็จะเปลี่ยนไปเป็นปริมาณยางพารา นอกจากนี้บนแผนที่เฉพาะกิจแผ่นหนึ่ง ๆ ปริมาณข้อมูลของแผนที่ฐานก็มีมากกว่าปริมาณข้อมูลเฉพาะกิจเป็นอย่างมาก ดังนั้นสำหรับการทำแผนที่เฉพาะกิจโดยมีคอมพิวเตอร์ช่วย การนำข้อมูลแผนที่ฐานไปเก็บไว้ในคลังข้อมูล (Data Bank) ซึ่งคอมพิวเตอร์สามารถนำมาใช้งานภายหลังได้จึงมีความเหมาะสมเป็นอย่างมาก

ปัจจุบันมีหลายองค์กรที่ได้สร้างคลังข้อมูลแผนที่ฐานขึ้นมาแล้ว และนักทำแผนที่หรือหน่วยงานต่าง ๆ สามารถหาซื้อมาใช้งานได้ เช่น World Data Bank I (WDB I) และ World Data Bank II (WDB II) ขององค์การ CIA (3) WDB I เก็บค่าพิกัดภูมิศาสตร์ของเส้นชายฝั่งและเส้นพรมแดนที่ได้จากการดิจิไตซ์แผนที่มาตราส่วนประมาณ 1:12,000,000 ส่วน WDB II ได้จากแผนที่มาตราส่วน 1:3,000,000 นอกจากนี้ยังมี Dahlgren File ซึ่งเก็บเฉพาะค่าพิกัดภูมิศาสตร์ของเส้นชายฝั่งทั่วโลกประมาณ 8,000 จุด ฯลฯ อย่างไรก็ตามเพื่อให้ได้ข้อมูลแผนที่ฐานที่เหมาะสม และเพื่อเป็นการศึกษาหาประสบการณ์ ผู้วิจัยนี้จึงได้ทดลองสร้างคลังข้อมูลแผนที่ฐานขึ้นมาใช้เอง โดยดิจิไตซ์เส้นชายฝั่ง เส้นพรมแดนระหว่างประเทศ ทะเลสาป และเกาะที่สำคัญทั่วโลกประมาณ 53,000 จุด เก็บไว้ในไฟล์แบบ ASCII ค่าพิกัดมีความละเอียด Single Precision ข้อมูลทั้งหมดใช้เนื้อที่หน่วยความจำประมาณ 1.5 M bytes

<u>โครงสร้างของข้อมูลแผนที่ฐาน</u>

ก่อนที่จะนำข้อมูลไปเก็บไว้ในคลังข้อมูล จะต้องมีการวางโครงสร้างของข้อมูลแผนที่ฐานและคลังข้อมูล เพื่อให้การจัดเก็บเป็นไปอย่างมีระบบและการนำข้อมูลมาใช้งานภายหลังเป็นไปอย่างมีประสิทธิภาพ สำหรับงานวิจัยนี้ถือว่าข้อมูลแผนที่ฐานถูกแสดงด้วยองค์ประกอบ 3 ส่วนคือ Nodes, Line Segments และ Areas โดยที่ Node คือจุดที่เส้นใด ๆ อย่างน้อย 3 เส้นมาบรรจบกัน ยกเว้นในกรณีของเกาะหรือทะเลสาปภายในประเทศ Line Segment คือเส้นใด ๆ ที่เชื่อมระหว่าง Node สอง Node ยกเว้นเกาะหรือทะเลสาปภายในซึ่ง Line Segment จะบรรจบตัวเองและจุดบรรจบถือเป็น Node (False Node) Line Segment จะทำหน้าที่เป็นโครงสร้างหลักของข้อมูลแผนที่ฐาน ส่วน Area คือพื้นที่ด้านข้างทั้งสองด้านของแต่ละ Line Segment วิธีการแบ่งข้อมูลออกเป็นองค์ประกอบ 3 ส่วนเช่นนี้คล้ายคลึงกับวิธีการที่เรียกว่า DIME (Dual Independent Map Encoding) ซึ่งพัฒนาขึ้นโดย U.S. Bureau of Census (4) รูปที่ 2.1 แสดงให้เห็นถึงการใช้วิธีการดังกล่าวกับข้อมูลแผนที่ฐาน



รูปที่ 2.1 จุดที่มีเส้นอย่างน้อย 3 เส้นมาบรรจบกันคือ Node เช่นจุด 22 จุด 31 ยกเว้นพื้นที่ที่เป็นเกาะหรือทะเลสาป เช่นจุด 21 และจุด 34 เส้นที่เชื่อมระหว่าง Node คือ Line Segment เช่นเส้น 24-30 คือเส้นพรมแดนไทย-พม่า ส่วน Area คือพื้นที่ของแต่ละประเทศ เช่น 122 หมายถึงประเทศพม่า 123 หมายถึงประเทศไทย

Nodes, Line Segments และ Areas จะถูกให้รหัสด้วยระบบตัวเลขที่เหมาะสมเพื่อให้การค้นหา การดึงข้อมูลมาใช้งาน เป็นไปอย่างมีประสิทธิภาพ โดยที่จะต้องไม่มีการซ้ำซ้อนกันด้วย สำหรับ Cadastral Map Node มีความสำคัญมากเนื่องจากมันคือตำแหน่งของหมุดหลักเขตที่ดิน แต่สำหรับแผนที่มาตราส่วนเล็ก ความสำคัญของ Node ก็ลดลงไป โดยที่มันไม่มีความหมายใด ๆ นอกจากทำหน้าที่เพียงแสดงจุดเริ่มต้นและจุดปลายของ Line Segment เท่านั้น การให้รหัสของ Node ในงานวิจัยนี้จึงเป็นตัวเลขใด ๆ ก็ได้ที่ไม่ซ้ำซ้อนกันและมีประโยชน์เฉพาะตอนดิจิไตซ์เท่านั้น หลังจากดิจิไตซ์ Node เสร็จแล้วก็ไม่มีความจำเป็นที่จะต้องเก็บรหัสของ Node เอาไว้อีก สิ่งที่จะถูกนำมาใช้คือค่าพิกัดของ Node เท่านั้น

เนื่องจากความละเอียดของงานวิจัยนี้จำกัดเพียงระดับประเทศ ดังนั้น Area ที่ละเอียดที่สุดก็คือแต่ละประเทศนั่นเอง การให้รหัสของ Area ก็คือการให้รหัสประเทศแต่ละประเทศ รหัสเหล่านี้มีความสำคัญมากเพราะจะเป็นดัชนีที่จะใช้ค้นหาข้อมูลต่าง ๆ ในภายหลัง ตัวอย่างเช่น ถ้านักทำแผนที่ต้องการให้วาดแผนที่ประเทศในกลุ่มลาตินอเมริกา ก็จะต้องให้รหัสของประเทศต่าง ๆ เหล่านั้นอย่างถูกต้อง เพื่อที่โปรแกรมจะได้นำไปใช้ค้นหาค่าพิกัดหรือข้อมูลในคลังข้อมูลแผนที่ เนื่องจากโครงสร้างของคลังข้อมูลแผนที่อิงอยู่กับทวีปต่าง ๆ การให้รหัสประเทศจึงใช้ตัวเลข 3 หลัก หลักแรกแทนรหัสของแต่ละทวีป ส่วนสองหลักที่เหลือแทนรหัสของประเทศในแต่ละทวีป เช่น 401 แทนประเทศแคนาดา 402 แทนประเทศสหรัฐอเมริกา เลข 4 หมายถึงทวีปอเมริกาเหนือ

Line Segments เป็นส่วนที่มีความสำคัญมากที่สุดในองค์ประกอบทั้ง 3 ส่วน เนื่องจากโดยเนื้อแท้แล้ว การวาดแผนที่ของประเทศ เกาะ หรือทะเลสาบใด ๆ ก็ตามก็คือการวาด Line Segment แต่ละเส้นนั่นเอง เนื่องจาก Line Segment แต่ละเส้นมีทั้งลักษณะทางภูมิศาสตร์และลักษณะทางการปกครองอยู่ในตัวเอง เช่น เส้นชายฝั่งของประเทศอินเดีย เกาะ Tasmania ของประเทศออสเตรเลีย หรือบางครั้งก็แต่ลักษณะทางการปกครองล้วน ๆ ซึ่งก็คือเส้นพรมแดนระหว่างประเทศ ระบบการให้รหัสจึงต้องสามารถบอกลักษณะทั้งหมดของ Line Segment ใด ๆ ได้ รหัสของแต่ละ Line Segment ในงานวิจัย

นี้ประกอบด้วยตัวเลขทั้งหมด 10 หลัก 2 หลักแรกบอกลักษณะทางภูมิศาสตร์และการปกครอง 6 หลักถัดมาจะแบ่งเป็นสองส่วน ส่วนละ 3 หลัก แต่ละส่วนคือรหัสของประเทศที่อยู่แต่ละด้านของ Line Segment นั้น ในกรณีที่เป็นเกาะหรือทะเลสาปภายในประเทศ ทั้ง 2 ส่วนจะมีรหัสเหมือนกัน 2 หลักสุดท้ายก็คือเลขเรียงลำดับธรรมดามีได้ตั้งแต่ 01-99 เนื่องจากแต่ละประเทศอาจมีเกาะหลายเกาะหรือทะเลสาปหลายแห่ง หรือสองประเทศใด ๆ อาจมีเส้นพรมแดนหลายช่วงเมื่อเกิดกรณีเช่นนี้ขึ้น ตัวเลข 8 หลักแรกจะซ้ำกันหมด จึงต้องมีส่วนนี้สำหรับป้องกันภาวะซ้ำซ้อนของรหัส

การจัดเก็บข้อมูลแผนที่ฐานไว้ในคลังข้อมูลจะจัดเก็บทีละทวีป แต่ละทวีปจะประกอบด้วยไฟล์ (file) 2 ไฟล์ ไฟล์แรกเป็น Direct Acces file ซึ่งบรรจุข้อมูลของแต่ละ Line Segments รวมทั้งพิกัดของ Node ต้นและปลาย และ pointers ที่จะชี้ไปยังตำแหน่งต่าง ๆ ในอีกไฟล์หนึ่งซึ่งบรรจุค่าพิกัดภูมิศาสตร์ของจุดต่าง ๆ เอาไว้ ดังนั้นกล่าวได้ว่าแต่ละเรคคอร์ดในไฟล์นี้ก็คือแต่ละ Line Segment บนแผนที่นั่นเอง เราอาจเรียกไฟล์นี้ว่า ไฟล์ดัชนี (Index file) เนื่องจากองค์ประกอบทั้ง 3 ส่วนของข้อมูลแผนที่ฐานถูกแสดงไว้ในไฟล์นี้ ส่วนไฟล์ที่สองเป็น Direct Access file เช่นกัน แต่มีเพียงค่าพิกัดภูมิศาสตร์ของจุดต่าง ๆ เท่านั้น ไม่มีข้อมูลอื่นใดอีก ไฟล์นี้จึงเป็นเพียงไฟล์เก็บค่าพิกัด (Coordinate file) ธรรมดา โครงสร้างของไฟล์ดัชนีแสดงไว้ในรูป 2.3 และความสัมพันธ์ระหว่างไฟล์ดัชนีและไฟล์เก็บค่าพิกัดแสดงไว้ในรูป 2.4

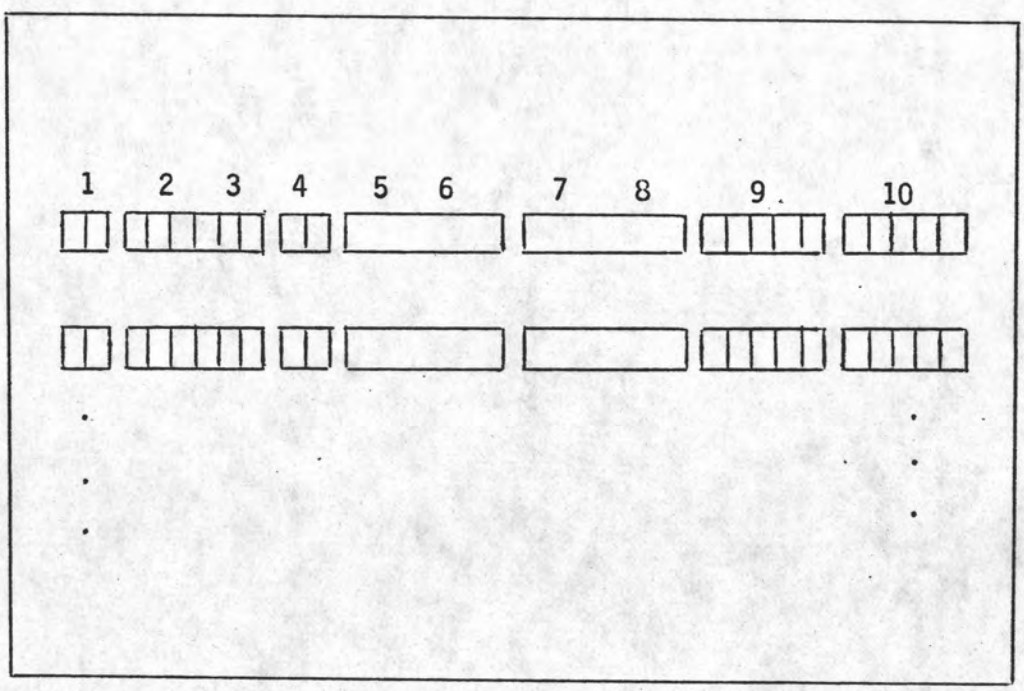
ทวีปต่าง ๆ แบ่งออกเป็น 7 ทวีป ตามหลักของวิชาภูมิศาสตร์ ดังนั้นจะมีไฟล์ดัชนี 7 ไฟล์ และไฟล์เก็บค่าพิกัด 7 ไฟล์ รวมเป็น 14 ไฟล์ โครงสร้างของคลังข้อมูลจึงเป็นดังรูป 2.5

แต่ละเรคคอร์ดของไฟล์ดัชนีประกอบด้วย ฟิลด์ (field) 10 ฟิลด์ แต่ละฟิลด์มีความหมายดังนี้

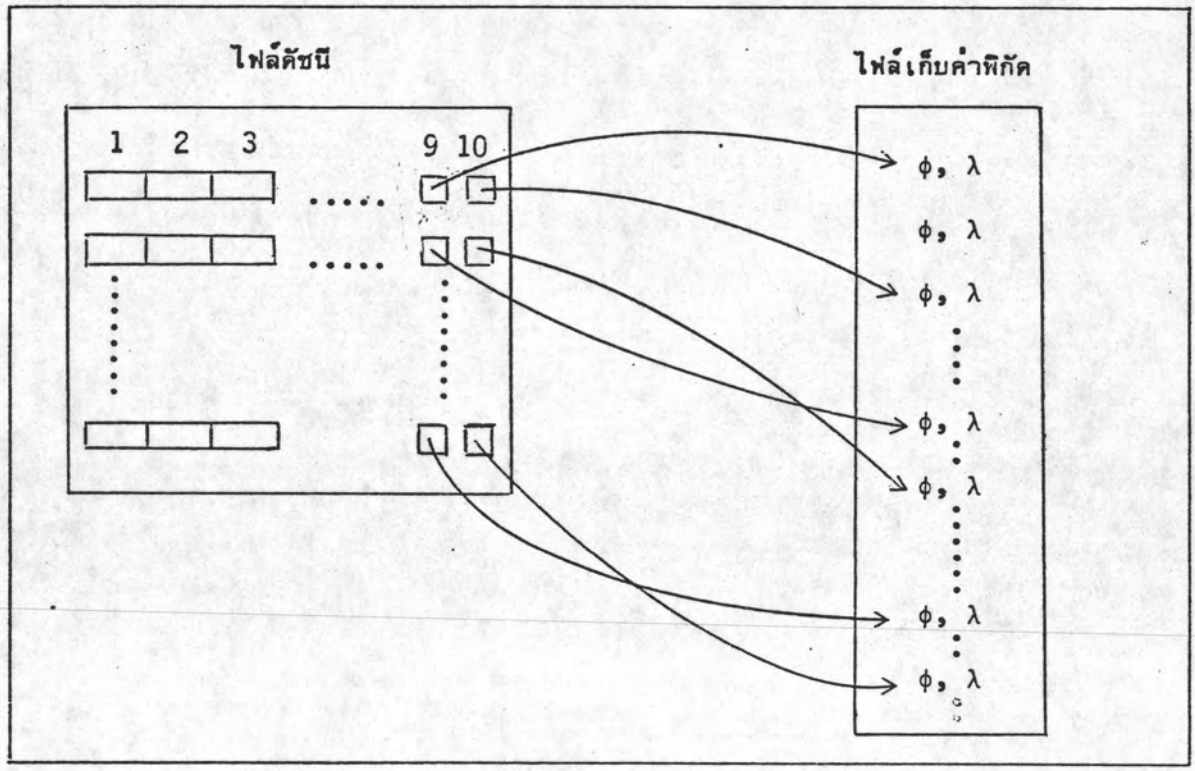
ฟิลด์ 1 ประกอบด้วยตัวเลข 2 ตัว แสดงลักษณะทางภูมิศาสตร์และการ

ลักษณะของ Line Segment	รหัสของ Line Segment
เส้นชายฝั่งพม่า	2012212201
เส้นพรมแดนไทย-พม่า	1012212301
เส้นชายฝั่ง ด้านทะเลอันดามัน	2012312301
เส้นชายฝั่งไทย ด้านอ่าวไทย	2012312302
เกาะศรีลังกา	2012112101
เส้นพรมแดนไทย-มาเลเซีย	1012312701
เส้นชายฝั่งมาเลเซีย	2012712701
เส้นพรมแดนลาว-เวียดนาม	1012412601

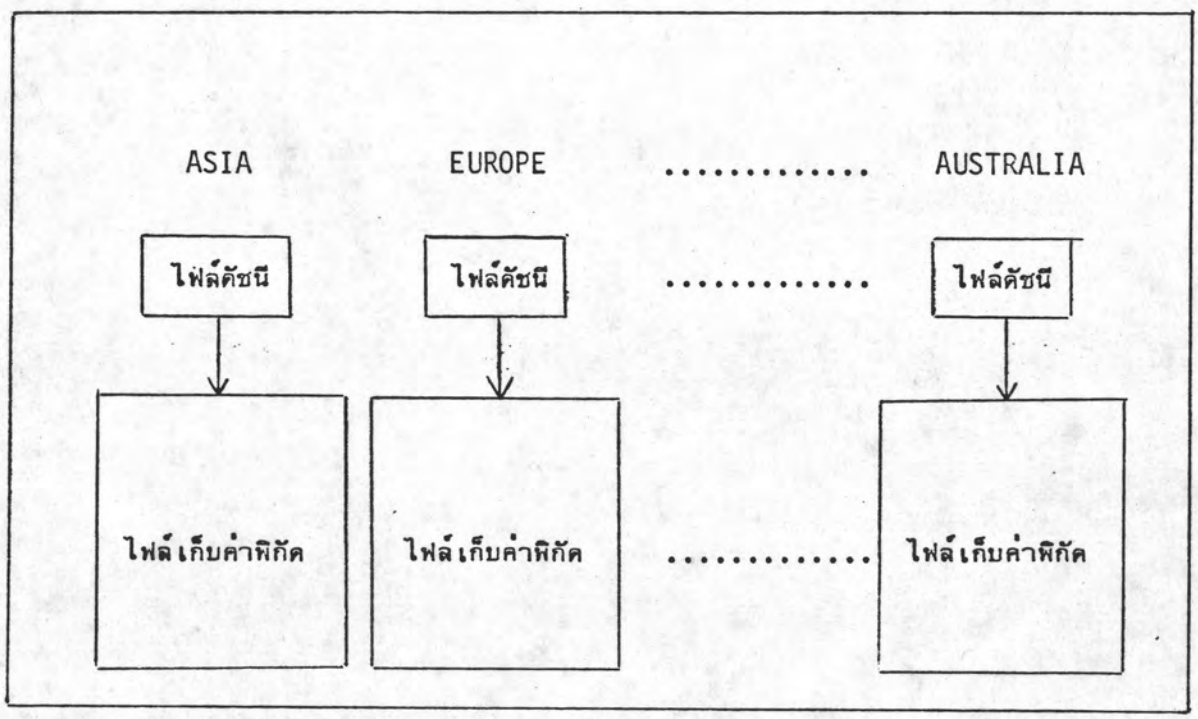
รูปที่ 2.2 การให้รหัสของ Line Segment



รูปที่ 2.3 โครงสร้างของแต่ละเรคคอร์ดในไฟล์ดัชนี



รูปที่ 2.4 ความสัมพันธ์ระหว่างไฟล์ดัชนีและไฟล์เก็บค่าพิกัด



รูปที่ 2.5 แสดงโครงสร้างของคลังข้อมูลแผนที่ฐาน ซึ่งประกอบด้วยไฟล์ทั้งหมด 14 ไฟล์ เป็นไฟล์ดัชนี 7 ไฟล์ และไฟล์เก็บค่าพิกัด 7 ไฟล์ ตั้งแต่ละทวีปจะมีไฟล์ดัชนีและไฟล์เก็บค่าพิกัดอย่างละ 1 ไฟล์

ปกครอง (feature) ของ Line Segment นั้นว่าเป็นอะไร เช่น 20 หมายถึงเส้นชายฝั่ง รหัส 10 หมายถึงเส้นพรมแดน เป็นต้น

ฟิลด์ 2, 3 แต่ละฟิลด์ประกอบด้วยตัวเลข 3 ตัว แทนรหัสประเทศที่อยู่ทางซ้ายและทางขวาของ Line Segment นั้นตามลำดับ ในกรณีที่เป็นเส้นชายฝั่ง เกาะหรือทะเลสาบภายใน ทั้ง 2 ฟิลด์จะเหมือนกัน

ฟิลด์ 4 ประกอบด้วยตัวเลข 2 ตัว เนื่องจากแต่ละประเทศอาจมีเกาะหลายเกาะหรือมีเส้นชายฝั่งหลายเส้น หรือประเทศสองประเทศอาจมีเส้นพรมแดนหลายเส้น เมื่อเกิดกรณีเช่นนี้ ฟิลด์ที่ 1, 2, 3 จะซ้ำกันหมด จึงต้องมีฟิลด์นี้เพื่อไม่ให้ Line Segments มีรหัสซ้ำกัน

ฟิลด์ 5, 6 คือค่าพิกัดภูมิศาสตร์ของ Node ต้นของ Line Segment นั้น

ฟิลด์ 7, 8 คือค่าพิกัดภูมิศาสตร์ของ Node ปลายของ Line Segment นั้น

ฟิลด์ 9, 10 ประกอบด้วยตัวเลขฟิลด์ละ 5 ตัว ชี้ไปยังหมายเลขประจำเรคคอร์ด (Record Number) ของไฟล์เก็บค่าพิกัดของทวีปนั้น ค่าพิกัดภูมิศาสตร์ของแต่ละ Line Segment ก็คือค่าพิกัดที่อ่านได้จากไฟล์เก็บค่าพิกัด โดยที่ฟิลด์ 9 จะชี้ไปยังจุดแรกของ Line Segment ส่วนฟิลด์ 10 ชี้ไปยังจุดสุดท้าย

<u>การนำข้อมูลเข้าสู่คลังข้อมูลแผนที่</u>

หลังจากได้จัดวางโครงสร้างข้อมูลแล้ว ก็เป็นการนำเอาข้อมูลแผนที่ฐานเข้าไปเก็บไว้ในคลังข้อมูลซึ่งอยู่บนจานแม่เหล็ก พิกัดข้อมูลแผนที่ฐานได้จากการดิจิไตซ์แผนที่โลกซึ่งมีมาตราส่วนประมาณ 1:35,000,000 ใช้การฉายแผนที่แบบ Mercator สำหรับทวีป Antarctic ได้มาจากการดิจิไตซ์แผนที่ Modified Mercator มาตราส่วนประมาณ

1:134,000,000 โดยที่ดิจิไตซ์ เส้นชายฝั่ง เส้นพรมแดน ทะเลสาป และเกาะของทุกประเทศที่ปรากฏบนแผนที่ ยกเว้นประเทศที่เล็กมาก ๆ เกาะเล็ก ๆ และทะเลสาปที่ไม่มีความสำคัญ ค่าพิกัดที่จะนำไปเก็บไว้ในคลังข้อมูลอยู่ในรูปพิกัดภูมิศาสตร์ ซึ่งได้มาจากการแปลงพิกัดของดิจิไตเซอร์ โดยมีขั้นตอนดังนี้

1. แปลงพิกัดฉาก X,Y หน่วย count ซึ่งอ้างอิงกับจุดศูนย์กำเนิดของดิจิไตเซอร์ เป็นพิกัดฉาก X,Y CM อ้างอิงกับจุดศูนย์กำเนิดของแผนที่ สำหรับการฉายแผนที่ทั้งสองแบบข้างต้น จุดศูนย์กำเนิดอยู่ที่ละติจูด 0 องศา และลองจิจูด 0 องศา วิธีการแปลงพิกัดใช้ Helmert Transformation มีจุดควบคุม 4 จุดต่อทวีป โดยใช้ Least Square Adjustment ประมาณค่าพารามิเตอร์ออกมา

2. แปลงพิกัดฉากแผนที่ที่ได้จากข้อ 1 เป็นพิกัดภูมิศาสตร์ โดยใช้สมการการฉายแผนที่ Mercator Projection (5)

$$X = R.\lambda \qquad (2\text{-}1)$$

$$Y = R.\ln [\tan(\pi/4 + \phi/2)] \qquad (2\text{-}2)$$

ค่า R คือรัศมีของลูกโลก (Globe Radius) พิกัดภูมิศาสตร์ (ϕ,λ) ได้จากสมการย้อนกลับของ (2-1) และ (2-2) ซึ่งจำเป็นต้องรู้ค่า R ก่อน ค่า R หาได้จากสมการ (2-1) โดยดิจิไตซ์ หา ΔX ของ Segment ต่าง ๆ บนเส้นศูนย์สูตร จำนวน 5 Segments แต่ละ Segment ใช้ค่า $\Delta\lambda$ เท่ากับ 15 องศา หาค่าเฉลี่ยของ ΔX แล้วแทนลงในสมการ (2-1) ก็จะได้ค่า R ออกมา

การดิจิไตซ์ข้อมูลแผนที่ฐานกระทำทีละทวีป แต่ละทวีปจะแบ่งการดิจิไตซ์ออกเป็นสองครั้ง ครั้งแรกเฉพาะ Nodes ส่วนครั้งที่สองดิจิไตซ์จุดต่าง ๆ บนแต่ละ Line Segment ซึ่งทั้งสองครั้งก็จะต้องมีการแปลงค่าพิกัด 2 ขั้นตอนดังกล่าวข้างต้น ลักษณะการดิจิไตซ์ Line Segment เป็นแบบจุดต่อจุด โดยผู้ทำวิจัยเป็นผู้กำหนดเองว่าจะเลือกจุดใดบ้าง โดยพยายามให้ละเอียดที่สุดเท่าที่จะเป็นไปได้ เพื่อรักษาลักษณะภูมิประเทศเอาไว้และเมื่อขยาย

หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

มาตราส่วนจะทำให้ระยะห่างระหว่างจุดไม่มากจนเกินไป

ส่วนข้อมูลแผนที่ฐานที่ไม่ใช่ค่าพิกัด เช่น ลักษณะทางภูมิศาสตร์ ลักษณะทางการปกครอง ชื่อประเทศ จะถูกให้รหัสและป้อนผ่าน keyboard โดยกระทำทีละ Line Segment หลังจากนั้นข้อมูลแผนที่ฐานทั้งหมดจะถูกรวบรวมและจัดการโดยโปรแกรมคอมพิวเตอร์ให้อยู่ในรูปของไฟล์ดัชนีและไฟล์เก็บค่าพิกัดซึ่งเมื่อทำครบทุกทวีปแล้ว ก็จะได้คลังข้อมูลแผนที่ฐานที่สมบูรณ์

<u>ประสิทธิภาพของคลังข้อมูลแผนที่ฐาน</u>

คลังข้อมูลแผนที่ฐานมีโครงสร้างที่แบ่งออกเป็น 2 ระดับ คือ ไฟล์ดัชนีอยู่ระดับบน ส่วนไฟล์เก็บค่าพิกัดอยู่ระดับล่าง การค้นหาข้อมูลต่าง ๆ รวมทั้งการดึงค่าพิกัดมาใช้ต้องผ่านไฟล์ดัชนีเสมอ ประสิทธิภาพของข้อมูลนี้จึงขึ้นอยู่กับโครงสร้างของไฟล์ดัชนีเป็นสำคัญ

ถ้าพิจารณาไฟล์ดัชนีจะเห็นว่าแต่ละเรคคอร์ดบรรจุข้อมูลของแต่ละ Line Segment ไว้อย่างครบถ้วน ดังนั้นจึงสามารถเลือกที่จะแสดงข้อมูลแผนที่ฐานเพียงส่วนหนึ่งบางส่วนหรือแสดงทั้งหมดก็ได้ ตัวอย่างเช่น นักทำแผนที่อาจเลือกที่จะสร้างแผนที่ที่มีแต่เส้นชายฝั่งอย่างเดียวหรืออาจให้มีทั้งเส้นชายฝั่งและเส้นพรมแดนก็ได้ โดยการดึงเอาเฉพาะ Line Segment ที่มีคุณสมบัติตามที่ต้องการมาใช้ ความสามารถเช่นนี้เป็นประโยชน์ต่อนักทำแผนที่ในการที่จะเลือกใช้ข้อมูลแผนที่ฐานเฉพาะส่วนที่คิดว่าจะทำให้การสื่อความหมายเป็นไปอย่างมีประสิทธิภาพมากที่สุด ถ้านักทำแผนที่จะทำแผนที่แสดงผลผลิตน้ำตาลทราย ยางพารา และดีบุก ของกลุ่มอาเซียน โดยคาดว่าผู้ใช้จะสนใจเฉพาะปริมาณรวมของทั้งกลุ่มโดยไม่สนใจว่าแต่ละประเทศจะผลิตได้เท่าไร ดังนั้นอาจจะเลือกไม่ให้มีเส้นพรมแดนปรากฏบนแผนที่ ซึ่งจะเป็นการเน้นความสำคัญของผลผลิตของกลุ่ม ทำให้ประสิทธิภาพในการสื่อความหมายดีขึ้น

ปัจจัยสำคัญอีกประการหนึ่งในการพิจารณาประสิทธิภาพของคลังข้อมูล คือ ความรวดเร็วในการค้นหาข้อมูลที่ต้องการ เนื่องจากแต่ละเรคคอร์ดในไฟล์ดัชนี (ซึ่งแทนแต่ละ Line Segment) ไม่มีความสัมพันธ์กันโดยตรง ไม่มีดัชนีหรือสิ่งอื่น ๆ ที่จะบอกถึงความเกี่ยว

012646

โยงกันของแต่ละเรคคอร์ด หรือกล่าวได้อีกแง่หนึ่งว่า แต่ละเรคคอร์ดอิสระต่อกัน การค้นหา Line Segments ที่ต้องการจะต้องกระทำทีละเรคคอร์ด ตั้งแต่เรคคอร์ดแรกเรื่อยไป จนกว่าจะพบ โดยอาศัยการเปรียบเทียบรหัส เช่น ถ้าต้องการให้วาดประเทศไทยก็สามารถหา Line Segment ของประเทศไทยได้โดยการเปรียบเทียบรหัสประเทศของประเทศไทยกับฟิลด์ที่ 2 และฟิลด์ที่ 3 (รูป 2.3) ของแต่ละเรคคอร์ด ตั้งแต่เรคคอร์ดแรก สมมติว่าพบเส้นแรกที่เรคคอร์ด 21 เป็นเส้นพรมแดนไทย-พม่า ก็ต้องค้นหาเส้นอื่น ๆ ต่อไปอีกโดยวิธีการเดิมตั้งแต่เรคคอร์ดที่ 22 เป็นต้นไป ซึ่งจะได้ครบทั้งประเทศก็ต้องค้นหาตั้งแต่ต้นจนถึงเรคคอร์ดสุดท้ายในไฟล์ดัชนีนั้น ถ้าต้องการวาดประเทศอื่นอีก ก็ต้องเริ่มต้นที่เรคคอร์ดแรกใหม่อีก แล้วอาศัยการเปรียบเทียบรหัสในการหา Line Segments ที่เหลือของประเทศนั้น

จะเห็นได้ว่าการค้นหาข้อมูลเป็นส่วนที่ใช้เวลามาก และเป็นจุดอ่อนของคลังข้อมูลนี้ แต่ทั้งนี้มิได้หมายความว่าจะใช้เวลานานมากจนเกินไป หน่วยของเวลาที่ใช้ในการค้นหาข้อมูลทั้งหมดที่ต้องการยังคงเป็นนาทีเท่านั้น และสามารถลดเวลาอีกโดยการเพิ่มเติมเงื่อนไขบางอย่างเข้าไปในโปรแกรมคอมพิวเตอร์เพื่อป้องกันการอ่านข้อมูลซ้ำซ้อน นอกจากนี้เมื่อได้ข้อมูลที่ต้องการแล้วยังสามารถแยกไปเก็บไว้ในพล้อตไฟล์ (plot file) ซึ่งสามารถนำมาใช้งานภายหลัง ดังนั้นแผนที่แต่ละแผ่นจะเสียเวลาในการค้นหาข้อมูลแผนที่ฐานครั้งแรกเพียงครั้งเดียว ส่วนครั้งต่อไปก็จะวาดจากพล้อตไฟล์