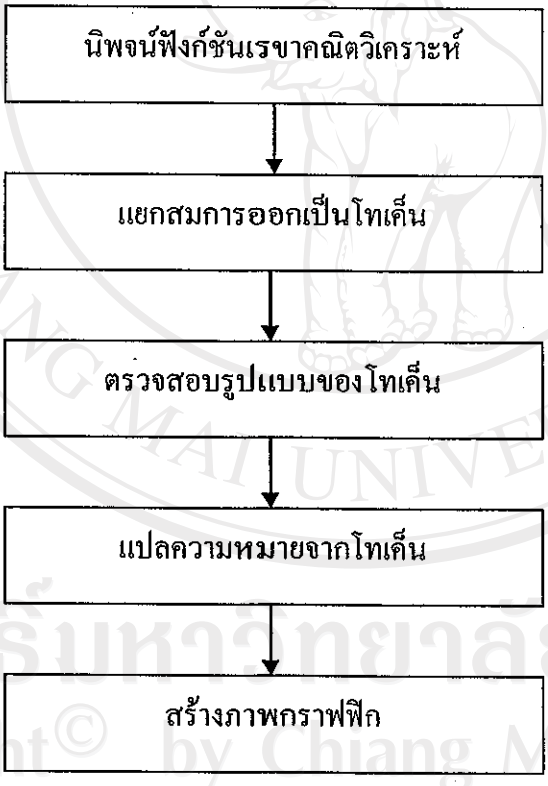
บทที่ 3
การออกแบบโครงสร้างและการตรวจสอบไวยากรณ์

ระบบการตรวจสอบไวยากรณ์ของสมการเรขาคณิตวิเคราะห์แบบ 2 มิติ ในการค้นคว้าอิสระครั้งนี้ ระบบเริ่มจากกระบวนการนำเข้าข้อมูลสมการเรขาคณิตวิเคราะห์เพื่อที่จะทำการตัดคำ เพื่อกราดตรวจทำการแยกสมการที่นำเข้าออกเป็นโทเค็น และทำการแทนค่าของแต่ละโทเค็นในรูปของสัญลักษณ์ที่ใช้แทนโทเค็นนั้น ๆ ซึ่งแต่ละโทเค็นจะมีการกำหนดคุณลักษณะของโทเค็นนั้น ๆ

ระบบการวาดภาพของฟังก์ชันเรขาคณิตวิเคราะห์ในการพัฒนาระบบมีขั้นตอนการดำเนินงานดังรูป 3.1



รูป 3.1 ขั้นตอนการดำเนินงาน

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

3.1 นิพจน์ปกติสำหรับโทเค็น

3.1.1 ชนิดของโทเค็น แบ่งออกเป็นหลายกลุ่มดังนี้

3.1.1.1 คำสงวน

คำสงวน เป็นโทเค็นของสายอักขระที่มีขนาดและรูปแบบที่ตายตัว และมีความหมายเฉพาะ เช่น sin , cos , tan , arcsin , arcos , arctan , ln เป็นต้น

3.1.1.2 อักขระพิเศษ

อักขระพิเศษเป็นกลุ่มของโทเค็นที่เป็นเครื่องหมายทางคณิตศาสตร์ต่าง ๆ อยู่ในรูปแบบของอักขระเดี่ยว เช่น "+" , "-" , "*" , "/" , "^" ,"#" เป็นต้น

3.1.1.3 ค่าคงที่

ค่าคงที่เป็นโทเค็นที่ใช้แทนค่าคงที่ของจำนวน เช่น π มีค่าเท่ากับ 3.1415926535897931 หรือ ค่าคงที่ของอักขระ เช่น "x" ในกรณีที่โทเค็นเป็นค่าตัวเองที่ใช้ในการแสดงค่าจำนวนเต็ม จำนวนจริง ซึ่งอยู่ในรูปของเลขฐานสิบ

3.1.1.4 พื้นที่ว่าง

ที่ว่าง (การเว้นวรรค) เป็นโทเค็นพื้นที่ว่าง

3.2 การตรวจสอบรูปแบบของสายอักขระ

นิพจน์ปกติเป็นข้อกำหนดของรูปแบบสายอักขระที่ต้องการตรวจสอบ โดยในการตรวจสอบรูปแบบของสายอักขระดังกล่าวนั้น จะทำการแยกสายอักขระเป็นกลุ่ม ๆ โดยมีรูปแบบตามที่กำหนดในนิพจน์ปกติ เช่น

(15+20)

จากนิพจน์ (15+20) สามารถตรวจสอบอักขระที่เป็นองค์ประกอบของไวยากรณ์ ได้ดังนี้

นิพจน์ปกติ	โทเค็น
(	(
)	)
+	+
[0-9]+	number

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

โดยในการทำงานของตัววิเคราะห์ศัพท์จะทำการแยกโดยอาศัยการกำหนดสถานะของสมการเพื่อตรวจสอบรูปแบบที่คาดหวังของสายอักขระที่กำลังตรวจสอบ โดยจะพบโทเค็น "(" , ")" และ "+" ซึ่งสามารถตรวจสอบได้ง่าย แต่สำหรับโทเค็นที่เป็น number นั้นจำเป็นต้องมีการตรวจสอบองค์ประกอบของโทเค็นทั้งหมด (เนื่องจากอาจจะประกอบด้วยสัญลักษณ์มากกว่า 1 ตัว) โดยมีการอ่านข้อมูลนำเข้าตัวแรก (ซึ่งเป็นตัวเลข) ตัววิเคราะห์ศัพท์จะย้ายสถานะของการตรวจสอบไปยังสถานะของการตรวจสอบตัวเลข และในสถานะดังกล่าวหากข้อมูลนำเข้าตัวถัดไปนั้นเป็นตัวเลขอีก ตัวเลขดังกล่าวจะถูกนำไปรวมกับตัวเลขก่อนหน้า แล้วทำการอ่านข้อมูลตัวถัดไป ในกรณีที่ข้อมูลตัวถัดไปไม่ใช้ตัวเลข ตัววิเคราะห์ศัพท์ก็จะออกจากสถานะของการตรวจสอบตัวเลข และจะรู้ได้ทันทีว่าข้อมูลที่มีอยู่เดิมในขณะนั้นเป็นโทเค็นที่เป็นตัวเลข และส่งคืนค่าโทเค็นและชนิดของโทเค็นกลับไปยังผู้เรียกใช้ จากนั้นจะทำการตรวจสอบข้อมูลนำเข้าตัวต่อไป ดังตารางที่ 3.1

ตารางที่ 3.1 การตรวจสอบโทเค็น (15+20)

ข้อมูลนำเข้า	สถานการณ์ตรวจสอบ	การทำงาน
(	สถานะเริ่มต้น	ส่งค่าโทเค็น "("
1	สถานะเริ่มต้น	จัดเก็บค่า 1 และเข้าไปทำงานในสถานะตรวจสอบตัวเลข
5	ตรวจสอบตัวเลข	จัดเก็บ 5
+	ตรวจสอบตัวเลข	พบเครื่องหมายบวก ออกจากสถานะตรวจสอบตัวเลข ส่งโทเค็น 15 กลับไปยังผู้เรียกใช้
+	สถานะเริ่มต้น	ส่งค่าโทเค็น "+" กลับไปยังผู้เรียกใช้
2	สถานะเริ่มต้น	จัดเก็บค่า 2 และเข้าไปทำงานในสถานะตรวจสอบตัวเลข
0	ตรวจสอบตัวเลข	จัดเก็บค่า 0
)	ตรวจสอบตัวเลข	พบเครื่องหมาย ")" ออกจากสถานะตรวจสอบตัวเลข ส่งโทเค็น 20 กลับไปยังผู้เรียกใช้
)	สถานะเริ่มต้น	ส่งค่าโทเค็น ")" กลับไปยังผู้เรียกใช้

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

กระบวนการตรวจสอบค่าที่แท้จริงของสายอักขระ สามารถแสดงให้เห็นชัดเจนยิ่งขึ้นโดยแสดงในรูปแบบของลำดับของสถานะที่มีการทำงาน ซึ่งอธิบายข้อกำหนดของจำนวนจริง ดังต่อไปนี้

$$Digit = [0-9]$$
$$NaturalNumber = Digit +$$
$$SingNaturalNumber = (+|-)?\, NaturalNumber$$
$$Number = SingNaturalNumber("."\, NaturalNumber)$$

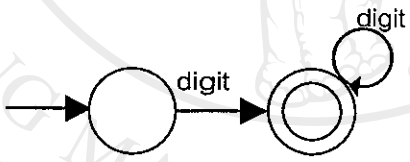
หมายเหตุ $Digit +$ หมายถึง มี Digit กี่ Digit ก็ได้

$(+|-)$ จะมีเครื่องหมาย "+" หรือ "-" อยู่ข้างหน้า NaturalNumber เครื่องหมายใดเครื่องหมายหนึ่ง หรือ ไม่มี ซึ่งมีความหมายเป็น "+"

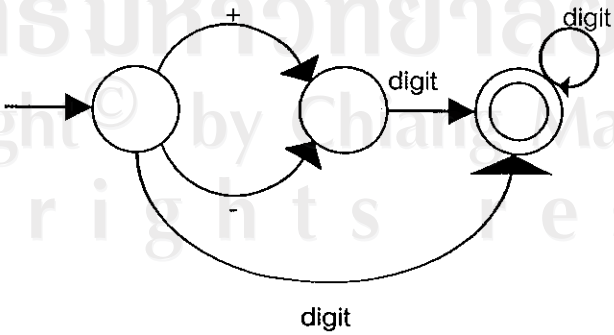
$("."\, NaturalNumber)$ ถ้ามีเป็นจำนวนจริง ถ้าไม่มีเป็นจำนวนเต็ม

จะทำการสร้างกำหนดสถานะโดยทำทีละขั้นตอนดังนี้

1) เครื่องจักรสถานะจำกัดที่กำหนดสถานะได้สำหรับ NaturalNumber

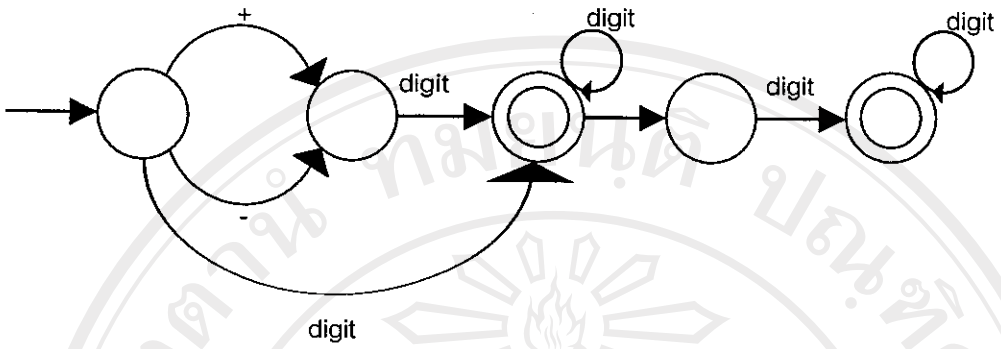


2) เครื่องจักรสถานะจำกัดที่กำหนดสถานะได้สำหรับ SignNaturalNumber



ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

3) เครื่องจักรสถานะจำกัดที่กำหนดสถานะได้สำหรับ Number ที่มีทศนิยม



3.3 ข้อกำหนดของนิพจน์

ข้อกำหนดของนิพจน์ เมื่อมีการตรวจสอบรูปแบบตัววิเคราะห์ศัพท์สามารถตรวจสอบรูปแบบที่ถูกต้องตามข้อกำหนดที่ประกาศค่าไว้ได้ โดยจะทำการตรวจสอบว่าโทเค็นที่พบนั้นเป็นโทเค็นชื่อหรือไม่ หากไม่ใช่ตัวอักษรก็แสดงว่าโทเค็นนั้นไม่ใช่ชื่อ แต่ในกรณีที่อักขระแรกเป็นตัวอักษร ก็จะทำการตรวจสอบอักขระต่อๆไปว่าเป็นตัวอักษรหรือตัวเลข โดยการที่จะตรวจสอบว่าอักขระเป็นตัวอักษรหรือตัวเลขนั้นก็สามารถตรวจสอบได้จากการประกาศค่านิพจน์ปกติของตัวอักษรและตัวเลขที่กำหนดไว้

การกำหนดกฎในการแปล เป็นการกำหนดการทำงานเมื่อมีการตรวจพบรูปแบบเหมือนที่ตรงกับนิพจน์ปกติที่กำหนดในโปรแกรม กล่าวคือ เมื่อตรวจสอบโทเค็นที่อยู่ในภาษาก็จะมีการทำงานตามกฎของการทำงานที่กำหนดไว้ ดังนั้นโทเค็นแต่ละตัวจะมีกฎในการทำงานที่แตกต่างกัน ทั้งนี้ขึ้นอยู่กับการกำหนดและออกแบบ ซึ่งสามารถเขียนเป็นกฎในการแปลได้ดังนี้

$$Digit \rightarrow [0-9]$$
$$NaturalNumber \rightarrow Digit+$$
$$SingNaturalNumber \rightarrow (+|-)?\, NaturalNumber$$
$$Number \rightarrow SingNaturalNumber("."\, NaturalNumber)?$$

3.4 ข้อกำหนดในการตั้งกฎของไวยากรณ์

การวิเคราะห์วากยสัมพันธ์เป็นขั้นตอนของการทำงานการตรวจสอบความสัมพันธ์ระหว่างโทเค็นที่ประกอบกันเป็นประโยค หรือวากยสัมพันธ์ หรือโครงสร้างของโปรแกรม โดยมีโครงสร้างของประโยคตามรูปแบบของกฎแห่งไวยากรณ์ที่กำหนดไว้ โดยจะรับข้อมูลเข้าซึ่งอยู่ในรูปของโทเค็นจากขั้นตอนการวิเคราะห์ศัพท์ (ในทางปฏิบัติแล้วโทเค็นไม่เชิงเป็นข้อมูลเข้าของตัววิเคราะห์กระจาย)

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

เพื่อนำมาตรวจสอบว่าลำดับของโทเค็นที่รับเข้ามานั้นมีการจัดกลุ่ม และเรียงลำดับได้ถูกต้องตามกฎของไวยากรณ์หรือไม่

3.4.1 ไวยากรณ์

ไวยากรณ์เป็นเครื่องมือในการกำหนดโครงสร้างของประโยค ซึ่งมีแนวคิดในการทำงานคล้ายกับการกำหนดโครงสร้างของคำโดยอาศัยนิพจน์ แต่ไวยากรณ์จะมีการทำงานที่ซับซ้อนกว่าการกำหนดนิพจน์ทางคณิตศาสตร์

$$< \exp r > \rightarrow < \exp r >< op >< \exp r > | (< \exp r >) | < number >$$
$$< op > \rightarrow + | - | * | / | \hat{} | \#$$

ไวยากรณ์ข้างต้นจะเห็นว่าในขณะที่กำหนดกฎไวยากรณ์ของ <expr> นั้น จะยังไม่ทราบโครงสร้างของกฎได้อย่างชัดเจน เนื่องจากยังมีสัญลักษณ์ไม่สิ้นสุดเป็นองค์ประกอบของกฎอยู่ นอกจากนี้องค์ประกอบของกฎยังมี <expr> เป็นองค์ประกอบด้วย ซึ่งโครงสร้างและข้อกำหนดของไวยากรณ์ลักษณะนี้เรียกว่า ไวยากรณ์แบบเวียนเกิด

3.4.2 ไวยากรณ์กับนิพจน์ปกติ

เมื่อพิจารณาตัวอย่างการกำหนดนิพจน์ทางคณิตศาสตร์ข้างต้น โดยเปรียบเทียบกับนิพจน์ปกติจะพบว่ามีการกำหนดให้ <number> มีคุณสมบัติดังนี้

$$< number >=< digit >< digit > *$$
$$< digit >= 0|1|2|3|4|5|6|7|8|9$$

โดยในนิพจน์มีการดำเนินการพื้นฐานอยู่สามอย่างคือ การเลือกเอาอย่างใดอย่างหนึ่งโดยใช้เครื่องหมาย "|" ในการแทนทางเลือก การหลอมรวม โดยการเขียนสิ่งที่ต้องการหลอมรวมให้อยู่ติดกัน และการใช้เครื่องหมาย "=" ในการแทนการอธิบายชื่อนิพจน์ปกติ ในการกำหนดไวยากรณ์ก็มีแนวคิดในการกำหนดที่คล้ายคลึงกัน แต่จะไม่มีสัญลักษณ์ที่ใช้แทนการทำซ้ำตั้งแต่ 0 ครั้งขึ้นไปคือ "*" (แต่จะมีการทำซ้ำโดยอาศัยการเวียนเกิดแทน) และการใช้เครื่องหมายลูกศรแทนเครื่องหมายเท่ากับ เนื่องจากชื่อของไวยากรณ์นั้นไม่สามารถแทนได้โดยการอธิบายด้วยตัวมันเอง แต่จะเกี่ยวพันกับไวยากรณ์อื่น ๆ

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

จากไวยากรณ์ข้างต้นจะพบว่า มีการใช้งานนิพจน์ปกติเป็นองค์ประกอบในการกำหนดไวยากรณ์ด้วย ซึ่งประกอบด้วยนิพจน์ปกติที่ใช้ในการแทนโทเค็นที่เป็นโทเค็นที่อักขระเดี่ยว 6 โทเค็น คือ + - * / (และ) และชื่อ number ซึ่งเป็นโทเค็นที่เป็นลำดับของตัวเลข ลำดับการแปลงในการสร้างสายอักขระ เช่น

$$(number - number) * number$$

สายแรกสร้างสายอักขระได้ดังนี้

$$\begin{aligned} <exrp> &\rightarrow <\exp r><op><\exp r> \\ &\rightarrow (<\exp r>)<op><\exp r> \\ &\rightarrow (<\exp r><op><\exp r>)<op><\exp r> \\ &\rightarrow (number <op><\exp r>)<op><\exp r> \\ &\rightarrow (number- <\exp r>)<op><\exp r> \\ &\rightarrow (number - number)<op><\exp r> \\ &\rightarrow (number - number)* <\exp r> \\ &\rightarrow (number - number)* number \end{aligned}$$

โปรแกรมวิเคราะห์ศัพท์เพื่อตรวจสอบรูปแบบของโทเค็นในไวยากรณ์

กำหนดกฎไวยากรณ์

expression : *NUMBER*
| *IDENTIFIER*
| expression'+'expression
| expression'–'expression
| expression'*'expression
| expression'/'expression
| expression'^'expression
| '('expression')
sin(expression)
cos(expression)
tan(expression)
arcsin(expression)
arccos(expression)
arctan(expression)

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved