

บทที่ 8

ภาษาโปรแกรม

- 8.1 หลักของภาษาโปรแกรม
- 8.2 รูปแบบของภาษาโปรแกรม
- 8.3 ประวัติความเป็นมา
- 8.4 เอกสารอ้างอิงและเว็บไซต์ที่ควรรู้

วัตถุประสงค์

- แนะนำภาษาโปรแกรม
- แนวคิดและที่มาของภาษาโปรแกรม
- ภาษาโปรแกรมในรูปแบบต่างๆ

ว

วิธีการหรือเครื่องมือที่ใช้ในการติดต่อสื่อสารระหว่างมนุษย์ก็คือภาษามนุษย์ คอมพิวเตอร์ก็มีลักษณะคล้ายกัน คือ นอกจากจะใช้ระบบเลขฐานสองหรือภาษาเครื่องในการติดต่อสื่อสารระหว่างและภายในคอมพิวเตอร์แล้ว ยังมีภาษาโปรแกรม (Programming language) เป็นสื่อกลางในการสื่อสารกันระหว่างโปรแกรมเมอร์ ภาษาโปรแกรมมีข้อแตกต่างจากภาษามนุษย์ที่สำคัญอยู่ 2 ประการ ประการแรก ภาษาโปรแกรมสามารถใช้ในการสื่อสารระหว่างมนุษย์กับเครื่องคอมพิวเตอร์ได้ ประการที่สอง ภาษาโปรแกรมมีขอบเขตที่ใช้ในการแสดงความหมายแคบกว่าภาษามนุษย์ เพราะเพียงต้องการช่วยให้สื่อสารแนวความคิดทางคอมพิวเตอร์ทำได้ง่าย ดังนั้น ภาษาโปรแกรมจึงมีข้อกำหนดที่แตกต่างจากภาษามนุษย์ ซึ่งเราจะศึกษาเกี่ยวกับข้อกำหนดเหล่านี้ในรายวิชานี้เอง

นักศึกษาสาขาเทคโนโลยีสารสนเทศจำเป็นต้องศึกษาเกี่ยวกับภาษาโปรแกรม ในฐานะเป็นผู้ใช้ภาษาของวันนี้และภาษาโปรแกรมใหม่ๆ ที่จะเกิดขึ้นในอนาคต สำหรับนักศึกษาสาขาวิทยาการคอมพิวเตอร์นั้น จำเป็นต้องรู้ลึกซึ้งในแง่ของการออกแบบภาษา ดังนั้น จึงจำเป็นต้องมีความรู้ความเข้าใจภาษาโปรแกรมอย่างกระจ่าง รู้ถึงลักษณะของภาษา ข้อดีและข้อเสียของภาษาโปรแกรมหลากหลายรูปแบบ รวมถึงการนำไปประยุกต์ใช้ให้เหมาะสมกับงานได้ อีกทั้งยังเป็นพื้นฐานในการเรียนรู้ภาษาโปรแกรมใหม่ๆ ได้ง่ายขึ้น และเป็นพื้นฐานที่สำคัญในการสร้างภาษาโปรแกรมภาษาใหม่ได้ การศึกษาภาษาโปรแกรมและการนำไปใช้เพียงภาษาเดียวจึงไม่สามารถก่อให้เกิดความรู้ความเข้าใจได้อย่างถ่องแท้ ในรายวิชานี้จะกล่าวถึงหลักการทั่วไปของภาษาโปรแกรม และกล่าวถึงภาษาโปรแกรมหลายๆ รูปแบบ

8.1 หลักของภาษาโปรแกรม (Principles)

ภาษาโปรแกรมเป็นเครื่องมือที่สั่งให้คอมพิวเตอร์ทำงานได้ตามที่ต้องการ และเป็นเครื่องมือหรือวิธีการที่โปรแกรมเมอร์ใช้ในการสื่อสารกันด้วย นอกจากนี้ยังอาจมองได้ว่าภาษาโปรแกรมเป็นสัญลักษณ์ที่ใช้สำหรับขั้นตอนการแก้ปัญหา (Notation for algorithms) ซึ่งเป็นวิธีการแสดงความสัมพันธ์ระหว่างแนวคิด เพื่อใช้ในการควบคุมอุปกรณ์คอมพิวเตอร์

ภาษาโปรแกรมแต่ละภาษาสามารถพิจารณาว่า เป็นเซตของข้อกำหนดอย่างเป็นทางการของ Syntax (วากยสัมพันธ์) ซึ่งเป็นส่วนของคำศัพท์พื้นฐานที่ใช้ในโครงสร้างของภาษา และ Semantic ซึ่งก็คือความหมายของคำศัพท์เหล่านั้นทั้งในแง่ของคำและในแง่ของการทำงานในทางปฏิบัติ ซึ่งข้อกำหนดเหล่านี้จะช่วยในเข้าใจการทำงานของภาษาได้เป็นอย่างดี ข้อกำหนดเหล่านี้มักรวมถึง

- ข้อมูลและโครงสร้างข้อมูล
- คำสั่งและลำดับการทำงาน
- ปรวิญญาในการออกแบบ
- สถาปัตยกรรมของภาษา

หลักการของการออกแบบภาษาโปรแกรม อาจแบ่งออกเป็น 3 ประเด็นหลัก ได้แก่

- วากยสัมพันธ์ (syntax)
- ชื่อและชนิด (names and types)
- ความหมาย (semantics)

แนวคิดของหลายๆ อย่างของหลักการออกแบบนี้ถูกหยิบยกมาจากภาษาศาสตร์และคณิตศาสตร์ ซึ่งจะกล่าวโดยย่อในบทนี้ และกล่าวโดยละเอียดในบทถัดๆ ไป

วากยสัมพันธ์

วากยสัมพันธ์ของภาษาเป็นสิ่งที่ใช้อธิบายส่วนประกอบของโครงสร้างโปรแกรมที่ถูกต้อง ซึ่งสามารถตอบคำถามได้ในหลายๆ เรื่อง เช่น รูปแบบของไวยากรณ์ที่ใช้ในการเขียนโปรแกรมในภาษานั้นๆ เป็นอย่างไร กลุ่มคำและสัญลักษณ์พื้นฐานที่โปรแกรมเมอร์ใช้ในการเขียนโปรแกรมที่โครงสร้างถูกต้องมีอะไรบ้าง

โครงสร้างที่เกี่ยวข้องกับวากยสัมพันธ์ของภาษาในยุคปัจจุบันจะกำหนดโดยใช้รูปแบบทางภาษาศาสตร์ที่เรียกว่า *context-free grammar* ซึ่งกล่าวถึงในบทที่ 9

ชื่อและชนิด

คำศัพท์ของภาษาโปรแกรมจะรวมถึงกฎของการตั้งชื่อต่างๆ ด้วย เช่น ตัวแปร ฟังก์ชัน คลาส พารามิเตอร์ ชื่อต่างๆ เหล่านี้ยังเกี่ยวเนื่องถึงคุณสมบัติต่างๆ ตลอดช่วงชีวิตของโปรแกรมอีกด้วย เช่น ขอบเขต (scope) การมองเห็น (visibility) และการเชื่อมโยง (binding)

ชนิดในที่นี้หมายถึงชนิดของข้อมูลหรือค่าที่โปรแกรมสามารถจัดการได้ เช่น ชนิดข้อมูลพื้นฐาน ชนิดข้อมูลแบบมีโครงสร้าง และชนิดข้อมูลที่มีโครงสร้างซับซ้อน ชนิดข้อมูลพื้นฐานได้แก่ จำนวนเต็ม จำนวนจริง ตัวอักษร ค่าบูลีน ชนิดข้อมูลแบบมีโครงสร้างประกอบด้วยสตริง ลิสต์ ทรี และตารางแฮช โครงสร้างข้อมูลที่ซับซ้อนประกอบด้วยฟังก์ชันและคลาส ซึ่งเราจะศึกษาเกี่ยวกับชื่อและชนิดข้อมูลนี้ในบทที่ 10

ความหมาย

ความหมายของโปรแกรมจะกำหนดโดย semantics นั่นคือ เมื่อรันโปรแกรม ผลกระทบของแต่ละคำสั่งที่เกิดกับค่าของตัวแปรในโปรแกรมจะถูกกำหนดโดยความหมายของภาษา ดังนั้น เมื่อเราเขียนโปรแกรม เราจะต้องเข้าใจแนวคิดพื้นฐานก่อน เช่นว่าเมื่อเราใช้คำสั่งในการกำหนดค่าจะเกิดผลอะไรขึ้นกับตัวแปร ถ้าเรามี semantic model ที่ไม่ขึ้นกับแพลตฟอร์มใดๆ เราก็จะสามารถนำ model นี้ไปใช้ได้กับเครื่องที่แตกต่างกันได้ การศึกษาในหัวข้อนี้จะทำให้ทราบถึงการทำงานของตัวแปลภาษา การเชื่อมโยงระหว่างตัวแปลภาษากับข้อกำหนดทาง semantic ของภาษา ซึ่งการศึกษาในหัวข้อนี้โดยในหัวข้อนี้เราจะไม่ลงลึกถึงรายละเอียด จึงจะไม่ขอกล่าวไว้ในเอกสารนี้

8.2 รูปแบบของภาษาโปรแกรม (Paradigms)

คำว่า Paradigm โดยทั่วไปจะหมายถึงรูปแบบของความคิดที่ลักษณะเป็นไปในแนวทางเดียวกัน ดังนั้น paradigm ในทางการโปรแกรมนั้นจึงหมายถึง รูปแบบของวิธีการคิดในการแก้ปัญหาภายใต้ลักษณะเฉพาะของโปรแกรมและภาษา รูปแบบของภาษาโปรแกรมสามารถแบ่งออกได้เป็น 4 รูปแบบหลัก คือ

- ภาษาเชิงคำสั่ง (Imperative programming)
- ภาษาเชิงวัตถุ (Object-oriented programming - OOP)

- ภาษาเชิงฟังก์ชัน (Functional programming)
- ภาษาเชิงตรรกะ (Logic programming)

ภาษาโปรแกรมบางภาษาก่อแบบมาเพื่อรองรับมากกว่าหนึ่งรูปแบบ ตัวอย่างเช่น ภาษา C++ ซึ่งเป็นทั้งภาษาเชิงคำสั่งและภาษาเชิงวัตถุ หรือภาษาเชิงทดลองอย่างเช่น ภาษา Leda ก็ออกแบบมาให้รองรับทุกรูปแบบ ในขณะที่ภาษาโปรแกรมในยุคแรกๆ เช่น PL/I หรือ Algol 68 หรือ Ada มักออกแบบโดยรองรับเพียงรูปแบบเดียวเพื่อการใช้งานทั่วไป

ภาษาเชิงคำสั่ง

ภาษาเชิงคำสั่งเป็นรูปแบบที่เก่าแก่ที่สุด เนื่องจากมีรากฐานมาจากคอมพิวเตอร์ที่มีสถาปัตยกรรมแบบ Von Neumann ที่ออกแบบโดยมีลักษณะที่ต้องเก็บโปรแกรมและตัวแปรไว้ในหน่วยความจำ และโปรแกรมจะประกอบด้วยชุดคำสั่งที่ทำการคำนวณ กำหนดค่าให้กับตัวแปร รับข้อมูลเข้า แสดงผล รวมทั้งการควบคุมลำดับการทำงานของชุดคำสั่ง

การสร้างกลุ่มคำสั่ง (block) โดยใช้ Procedural abstraction เป็นสิ่งจำเป็นสำหรับภาษาเชิงคำสั่ง ซึ่งประกอบด้วย

- คำสั่งกำหนดค่า (Assignment)
- คำสั่งวนซ้ำ (Loop)
- คำสั่งทำงานตามลำดับ (Sequence)
- คำสั่งเงื่อนไข (Conditional statement) และ
- คำสั่งในการจัดการความผิดปกติของโปรแกรม (Exception handling หรือสิ่งที่ไม่คาดว่าจะเกิดขึ้นในโปรแกรม)

ตัวอย่างภาษาเชิงคำสั่งที่เด่นๆ ได้แก่ ภาษาโคบอล (Cobol) ฟอรัทเรน (Fortran) ซี (C) એડા (Ada) และเพิร์ล (Perl)

ภาษาเชิงวัตถุ

ภาษาเชิงวัตถุจะกำหนดรูปแบบของโปรแกรมในลักษณะของกลุ่มของวัตถุที่มีการติดต่อสื่อสารกันโดยการส่ง message เพื่อเปลี่ยนสถานะ การส่ง message นี้ทำให้ข้อมูลของวัตถุมีลักษณะเป็นแบบ active แทนที่จะเป็นแบบ passive ลักษณะเช่นนี้เป็นสิ่งที่ทำให้ภาษาเชิงวัตถุแตกต่างจากภาษาเชิงคำสั่งอย่างชัดเจน การสร้างกลุ่มคำสั่งจะอาศัยหลักการพื้นฐานดังต่อไปนี้

- การแบ่งแยกวัตถุ (Object classification)
- การสืบทอดคุณสมบัติ (Inheritance)
- การส่งผ่านข้อความ (Message passing)

ตัวอย่างภาษาเชิงวัตถุ ได้แก่ ภาษาสโมลทอล์ก (Smalltalk) ซีพลัสพลัส (C++) จาวา (Java) และซีชาร์ป (C#)

ภาษาเชิงฟังก์ชัน

เป็นภาษาที่จำลองปัญหาทางคอมพิวเตอร์ในรูปของฟังก์ชันทางคณิตศาสตร์ ซึ่งประกอบด้วยข้อมูลเข้า (หรือ domain) และผลลัพธ์ (หรือ range) ภาษาโปรแกรมในกลุ่มนี้แตกต่างจากภาษาที่ใช้คำสั่งกำหนดค่า ตัวอย่างเช่น คำสั่งกำหนดค่า $x = x+1$ เป็นคำสั่งที่ไม่สมเหตุสมผลทางคณิตศาสตร์และในภาษาเชิงฟังก์ชันด้วย ฟังก์ชันจะมีการสัมพันธ์และรวมกับฟังก์ชันอื่นโดยการใช้

- การประกอบฟังก์ชัน (functional composition)
- การใช้เงื่อนไข (conditional)
- การวนซ้ำ (recursion)

ตัวอย่างภาษาเชิงฟังก์ชันที่สำคัญ ได้แก่ ภาษาลิสป์ (Lisp) สคีม (Scheme) แฮสเคิล (Haskell) และภาษาเอ็มแอล (ML)

ภาษาเชิงตรรกะ

เป็นภาษาเชิงประกาศ (declarative language) ที่แก้ปัญหาโดยประกาศผลของโปรแกรมที่น่าจะได้รับแทนที่จะระบุว่าจะทำอะไรจึงจะได้ผลลัพธ์ บางครั้งภาษากลุ่มนี้อาจเรียกว่าเป็นภาษา rule-based language เนื่องจากโปรแกรมมีลักษณะคล้ายกับเซตของกฎ หรือข้อบังคับของปัญหาแทนที่จะเป็นลำดับการทำงานของคำสั่ง

การแปลความการประกาศของโปรแกรมเชิงตรรกะ จะสร้างเซตของการแก้ปัญหาที่เป็นไปได้ทั้งหมดตามที่โปรแกรมระบุไว้ ภาษาเชิงตรรกะเหมาะสำหรับการแก้ปัญหาที่มีรายละเอียดไม่สมบูรณ์

ตัวอย่างภาษาเชิงตรรกะที่สำคัญ ได้แก่ ภาษาโพรล็อก (Prolog)

8.3 ประวัติความเป็นมา

คอมพิวเตอร์ในยุคแรกควบคุมแบบ Hardwired ซึ่งเป็นการควบคุมการทำงานของฮาร์ดแวร์โดยฮาร์ดแวร์ ทำให้ไม่มีความยืดหยุ่นในการใช้งาน ดังนั้นจึงมีแนวคิดในการสร้างคอมพิวเตอร์โดยใช้ชุดของคำสั่งในการควบคุมการทำงาน จึงเกิดการพัฒนาภาษาโปรแกรมในยุคแรก คือ ภาษาเครื่องและภาษาแอสเซมบลีที่เริ่มต้นใช้เมื่อประมาณทศวรรษที่ 1940 นับตั้งแต่นั้นเป็นต้นมา ก็มีการพัฒนาภาษาโปรแกรมหลายร้อยภาษารวมถึงภาษาย่อย (dialect) ต่างๆ ด้วย ภาษาโดยมากจะมีช่วงอายุการใช้งานที่จำกัด แต่ก็ยังมีบางภาษาที่ยังคงได้รับความนิยมแพร่หลายมาจนทุกวันนี้ หลายภาษาเป็นส่วนสำคัญที่มีอิทธิพลต่อการออกแบบภาษาใหม่ในอนาคต สิ่งที่ยังกระตุ้นให้เกิดการพัฒนาภาษาโปรแกรมในช่วงหลายทศวรรษที่ผ่านมา ก็คือวิวัฒนาการของคอมพิวเตอร์ที่มีประสิทธิภาพดีขึ้น และการประยุกต์ใช้โดยผู้ใช้งานในหลากหลายสาขา ซึ่งสาขาที่มีอิทธิพลอย่างมากต่อภาษาโปรแกรมมีดังนี้

- | | |
|--------------------------------|---------------------|
| • ปัญญาประดิษฐ์ | • เทคโนโลยีสารสนเทศ |
| • การศึกษา | • ระบบและเครือข่าย |
| • วิทยาศาสตร์และวิศวกรรมศาสตร์ | • เวิร์ลไวด์เว็บ |

ภาษาโปรแกรมอาจจำแนกออกตามลักษณะหรือยุคของการพัฒนาภาษาได้เป็น 3 ยุคหลัก ดังต่อไปนี้

ภาษาเครื่อง (Machine language)

จัดเป็นภาษาในยุคที่ 1 ซึ่งถือกำเนิดมาพร้อมกับคอมพิวเตอร์ เป็นภาษาที่เครื่องสามารถเข้าใจได้โดยตรง ภาษาเครื่องนี้ยังเป็นที่มาของคำว่า *โค้ด (code)* ในปัจจุบันอีกด้วย ซึ่งหมายถึงข้อความที่ประกอบขึ้นเป็นโปรแกรมในภาษาใดๆ ภาษาเครื่องจัดเป็นภาษาระดับต่ำ (Low level language) เนื่องจากทำความเข้าใจได้ยาก จึงไม่เป็นที่นิยมใช้ในปัจจุบัน แต่มีข้อดีคือทำงานได้เร็ว

ภาษาสัญลักษณ์ (Symbolic language)

เนื่องจากภาษาเครื่องเข้าใจได้ยาก และภาษาอังกฤษใช้คำมากเกินไปจนกว่าจะที่ใช้ในการแสดงความหมาย ภาษาแอสเซมบลีจึงได้ถือกำเนิดขึ้น โดยใช้สัญลักษณ์และชื่อแทนตัวเลข ทำให้เข้าใจได้ง่ายขึ้น ภาษาเครื่องจัดเป็นภาษาระดับต่ำ แต่ยังมีข้อเสียคือเป็นภาษาที่ขึ้นอยู่กับเครื่อง การใช้งานโปรแกรมจะต้องทำการแปลจากภาษาสัญลักษณ์ให้เป็นภาษาเครื่องก่อน โดยใช้ตัวแปลภาษา คือ แอสเซมเบอ

ภาษาระดับสูง (High level language)

เป็นภาษาที่ใช้คำและสัญลักษณ์ที่คุ้นเคย อ่านแล้วเข้าใจได้ง่าย มีความใกล้เคียงกับภาษาอังกฤษมากขึ้น มีจุดเด่นคือ ไม่ขึ้นอยู่กับเครื่อง นั่นคือเราสามารถนำโปรแกรมไปรันบนเครื่องที่ต่างกันได้โดยทำการแก้ไขโปรแกรมเพียงเล็กน้อยหรือไม่ต้องแก้ไขเลย ผู้เขียนโปรแกรมไม่จำเป็นต้องเข้าการทำงานภายในคอมพิวเตอร์มากนัก โดยมากภาษาระดับสูงนี้จะมีชุดคำสั่งพื้นฐานให้ใช้งาน (Program libraries) ทำให้สะดวกในการเขียนโปรแกรม การใช้งานโปรแกรมจะต้องทำการแปลให้เป็นภาษาเครื่องก่อน โดยใช้ตัวแปลภาษา คือ คอมไพเลอร์ (Compiler) หรืออินเทอร์พรีเตอร์ (Interpreter)

ภาษาระดับสูงนี้อาจเป็นออกได้เป็น 3 ยุคย่อยๆ ได้แก่

- ภาษายุคที่ 3 (3GL – Third General Language) เป็นภาษาระดับสูงที่มีลักษณะที่ต้องระบุขั้นตอนการทำงานให้แก่คอมพิวเตอร์ ภาษาในกลุ่มนี้ได้แก่ ภาษาโปรแกรมส่วนใหญ่ในปัจจุบัน
- ภาษายุคที่ 4 (4GL – Fourth General Language) เป็นภาษาที่ไม่จำเป็นต้องระบุวิธีการ ขั้นตอนการทำงานในโปรแกรม แต่บอกว่าต้องการอะไร ทำให้พัฒนาได้เร็ว ใช้งานและบำรุงรักษาง่าย ภาษาในกลุ่มนี้มักจะเป็นโปรแกรมที่รวบรวมการจัดการฐานข้อมูลด้วย
- ภาษายุคที่ 5 (5GL – Fifth General Language) หรือภาษาธรรมชาติ เป็นภาษาที่ใช้ภาษาพูดหรือเขียนของมนุษย์ในการสั่งงาน ใช้ความรู้ทางปัญญาประดิษฐ์ในการสร้างภาษา ภาษาในกลุ่มนี้กำลังได้รับความสนใจและอยู่ในระหว่างการศึกษาวิจัยและพัฒนา

8.4 เอกสารอ้างอิงและเว็บไซต์ที่ควรรู้

Allen B. Tucker and Robert E. Noonan. Programming Languages – Principles and Paradigms.

สุจิตรา อุดลย์เกษม. เอกสารประกอบการสอนองค์ประกอบคอมพิวเตอร์และภาษา.

<http://dariosantarelli.wordpress.com/2008/06/22/programming-languages-history/>

