



MathStandard

เอกสารประกอบการเรียน

เทคโนโลยีและวิทยาการคำนวณ

เรื่อง ข้อมูลดิจิทัล

เตรียมสอบเข้า ม.4

โครงการห้องเรียนวิทยาศาสตร์ วมว. รอบ 2

www.mathstandard.com

คำชี้แจง

ข้อมูลดิจิทัล

จัดทำโดย MathStandard

ฉบับ 2.0

ข้อมูลดิจิทัลอาศัยสัญลักษณ์และข้อตกลงเพื่อแทนจำนวน ข้อความ ภาพ และสถานะของระบบ การแปลความหมายจึงต้องพิจารณาทั้งค่าของบิต ลำดับการอ่าน และรูปแบบของข้อมูล เนื้อหาครอบคลุมหน่วยข้อมูล เลขฐานสอง รหัส การบีบอัด ภาพดิจิทัล และการตรวจสอบความถูกต้อง รวมถึงการวิเคราะห์ขนาด จัดเก็บและการเปรียบเทียบวิธีภายใต้ข้อจำกัด

ความรู้พื้นฐานที่ใช้ ได้แก่ จำนวนเต็ม การหารเอาเศษ เลขยกกำลัง อัตราส่วน และการอ่านตาราง ในการคำนวณขนาดข้อมูลต้องแยกบิตกับไบต์ ข้อมูลหลักกับส่วนประกอบ และจำนวนที่คำนวณได้กับจำนวนหน่วยจัดเก็บเต็มที่ต้องใช้ เงื่อนไขการบรรจุและการเติมเต็มระบุในแต่ละสถานการณ์

สัญลักษณ์ 101_2 หมายถึงจำนวนฐานสอง ส่วนสายบิตที่ใช้เป็นรหัสอาจไม่ใช่ค่าจำนวนโดยตรง ดัชนีแถวคอลัมน์ และตำแหน่งบิตเริ่มที่ 1 นับแถวจากบนลงล่าง คอลัมน์และบิตจากซ้ายไปขวา เว้นแต่คำถามกำหนดต่างออกไป การติดกันของจุดภาพใช้กฎที่ระบุในคำถาม ไม่ถือว่าการแตะมุมเป็นการเชื่อมกันโดยอัตโนมัติ

การตรวจพาริตีและการแก้บิตต้องพิจารณาสมมติฐานจำนวนบิตที่ผิด ผลตรวจที่ผ่านไม่ได้ยืนยันว่าข้อมูลถูกต้องเสมอ ส่วนคำสั่งแปดบิตและการใช้ XOR เป็นแบบจำลองที่กำหนดกฎไว้เฉพาะ ไม่ใช่ข้อรับรองความปลอดภัยของระบบรักษาความลับจริง

คำตอบที่ใช้วิธีต่างจากเฉลยพิจารณาจากความถูกต้องของหลักการและความสอดคล้องกับข้อกำหนด หากถามค่าต่ำสุดหรือจำนวนมากที่สุด ต้องมีทั้งวิธีที่ได้จริงและเหตุผลยืนยันขอบเขต คะแนนแต่ละข้อระบุในเกณฑ์ตรวจท้ายเล่ม

สารบัญ

เนื้อหาและแบบฝึกหัด

ตอนที่ 1 ข้อมูลและการแทนข้อมูลในระบบดิจิทัล	4
แบบฝึกหัดตอนที่ 1	8
ตอนที่ 2 เลขฐานสองและขอบเขตการแทนจำนวน	10
แบบฝึกหัดตอนที่ 2	14
ตอนที่ 3 รหัสและการบีบอัดข้อมูล	16
แบบฝึกหัดตอนที่ 3	21
ตอนที่ 4 ภาพดิจิทัลและขนาดข้อมูล	23
แบบฝึกหัดตอนที่ 4	28
ตอนที่ 5 ความถูกต้องและการตรวจสอบข้อมูล	31
แบบฝึกหัดตอนที่ 5	36
สรุปสาระสำคัญ	39
แบบฝึกหัดบูรณาการ	41
เฉลยเฉพาะคำตอบ	43
เฉลยวิธีคิดและเกณฑ์ตรวจ	45
คำศัพท์และเอกสารอ้างอิง	58

ตอนที่ 1

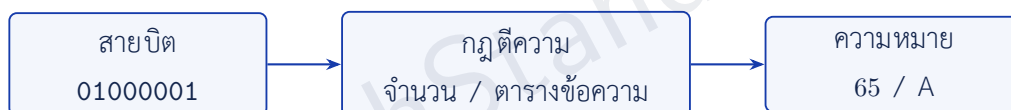
ข้อมูลและการแทนข้อมูลในระบบดิจิทัล

1.1 ข้อมูล สัญลักษณ์ และข้อตกลงในการแปลความหมาย

ข้อมูลอาจเป็นจำนวน ข้อความ ภาพ เสียง หรือสถานะของอุปกรณ์ การแทนข้อมูล คือ การกำหนดสัญลักษณ์ให้สอดคล้องกับความหมายที่ต้องการ เมื่อเก็บหรือส่งในรูปแบบดิจิทัล สัญลักษณ์จะถูกแปลงเป็นลำดับค่าที่แยกกันได้ โดยระบบฐานสองใช้ 0 และ 1 การกำหนดสองค่านี้ไม่บังคับว่าศูนย์ต้องหมายถึงปิดหรือสีขาวเสมอ ความหมายขึ้นกับข้อตกลงของระบบ

สายบิตเดียวกันอาจตีความได้หลายแบบ เช่น 01000001 มีค่าจำนวนไม่ติดลบเท่ากับ 65 เมื่อใช้ค่าน้ำหนักฐานสอง แต่อาจแทนอักษร A เมื่อใช้ข้อตกลงรหัสข้อความ หรือแทนสีหนึ่งสีในตารางสีที่กำหนด การเห็นบิตโดยไม่ทราบชนิดข้อมูลและกฎแปลความหมายจึงยังไม่เพียงพอสำหรับสรุปผล

ข้อมูลเชิงจำนวนที่ต้องนำไปคำนวณต่างจากรหัสที่ประกอบด้วยตัวเลข รหัสห้อง 007 อาจต้องคงศูนย์นำหน้า ขณะที่จำนวนเจ็ดเขียนเป็น 7 ได้ การแทนข้อมูลต้องรักษารายละเอียดที่มีความหมายต่อการใช้งาน รวมถึงขอบเขตของแต่ละรายการ เช่น ข้อความสิ้นสุดที่ใดและจุดใดเป็นบิตตรวจสอบ



ภาพที่ 1 สายบิตและกฎตีความร่วมกันกำหนดความหมายของข้อมูล

ข้อมูลที่จำเป็นต่อการตีความ

ระบุชนิดข้อมูล ตารางรหัสหรือค่าน้ำหนัก ลำดับการอ่าน และความยาวของแต่ละส่วน หากมีบิตเพิ่มเติมหรือบิตตรวจสอบ ต้องแยกออกจากข้อมูลหลักตามกติกา

1.2 บิต ไบต์ และขนาดของข้อมูล

บิตเป็นหน่วยข้อมูลที่มีสองค่า คือ 0 กับ 1 ส่วนไบต์ประกอบด้วยแปดบิต ตัวอย่างเช่น 00110110 มีแปดตำแหน่งจึงยาวหนึ่งไบต์ ไม่ว่ามีบิตหนึ่งอยู่ที่ตำแหน่ง การนับเฉพาะจำนวนบิตหนึ่งไม่ได้บอกความยาวของข้อมูลทั้งหมด

สัญลักษณ์ b หรือ bit ใช้กับบิต และ B ใช้กับไบต์ ตัวพิมพ์เล็กและใหญ่มีความหมายต่างกัน ถ้าขนาดข้อมูลเป็น s ไบต์ จะมี 8s บิต ส่วนข้อมูล d บิตที่ต้องใช้หน่วยจัดเก็บเต็มไบต์ต้องใช้จำนวนไบต์ที่เพียงพอรับบิตทั้งหมด หากหารแปดไม่ลงตัวต้องเพิ่มอีกหนึ่งไบต์ ไม่ถึงเศษบิตที่ยังเป็นข้อมูล

8 บิต = 1 ไบต์

0	1	0	0	0	0	0	1
1	2	3	4	5	6	7	8

ภาพที่ 2 หนึ่งไบต์มีแปดตำแหน่ง รวมทั้งตำแหน่งที่มีค่าเป็นศูนย์

การแปลงหน่วยและการปัดขึ้น

$$1 \text{ B} = 8 \text{ bit}, \quad S = \left\lceil \frac{d}{8} \right\rceil \text{ B}$$

d คือจำนวนบิตที่ต้องเก็บ S คือจำนวนไบต์เต็มที่ใช้ และ $\lceil x \rceil$ คือจำนวนเต็มที่น้อยที่สุดซึ่งไม่น้อยกว่า x บิตว่างในไบต์สุดท้ายมี $8S - d$ บิต

การปัดขึ้นใช้เมื่อหาจำนวนหน่วยจัดเก็บที่ต้องมี ส่วนการหาจำนวนแฟ้มเต็มที่ใส่ในพื้นที่ที่กำหนดใช้การปัดลง เพราะเศษพื้นที่ที่ยังไม่พอหนึ่งแฟ้มไม่สามารถนับเป็นแฟ้มเพิ่มได้ ทิศทางการปัดจึงขึ้นกับคำถาม ไม่ใช่การใช้การปัดแบบเดียวกันทุกกรณี

1.3 หน่วยฐานสิบและหน่วยฐานสอง

หน่วยฐานสิบใช้ตัวคูณกำลังของสิบ เช่น กิโลเท่ากับ 10^3 และเมกะเท่ากับ 10^6 ส่วนหน่วยฐานสองใช้ชื่อเฉพาะ เช่น คิบิเท่ากับ 2^{10} และเมบิเท่ากับ 2^{20} ทั้งสองระบบให้จำนวนไบต์ต่างกันแม้ชื่อใกล้เคียงกัน การเปรียบเทียบความจุต้องแปลงเป็นหน่วยเดียวกันก่อน

หน่วย	สัญลักษณ์	จำนวนไบต์
กิโลไบต์	kB	1,000
คิบิไบต์	KiB	1,024
เมกะไบต์	MB	1,000,000
เมบิไบต์	MiB	1,048,576

การเขียน KB ในบางบริบทอาจใช้ไม่ตรงกัน จึงต้องพิจารณานิยามที่ระบุ ไม่อนุমানว่าเป็น 1,000 หรือ 1,024 โดยไม่มีข้อมูลกำกับ ในการคำนวณต่อไปนี้ใช้ kB และ MB เป็นฐานสิบ ส่วน KiB และ MiB เป็นฐานสอง หากคำถามให้นิยามหน่วยไว้เฉพาะ ให้ใช้ค่านั้นอย่างสอดคล้องตลอดคำตอบ

ความจุกับอัตราส่งเป็นปริมาณต่างชนิด ขนาด 2 kB เป็นจำนวนข้อมูล แต่ 2 kbit/s เป็นจำนวนบิตที่ส่งได้ต่อวินาที การคำนวณเวลาส่งโดยหารขนาดด้วยอัตราต้องใช้หน่วยบิตหรือไบต์ให้ตรงกัน และระบุว่ายังไม่รวมเวลารอหรือข้อมูลประกอบหากไม่ได้กำหนดไว้

1.4 โครงสร้างแฟ้ม ข้อมูลหลัก และส่วนเติมเต็ม

แฟ้มประกอบด้วยลำดับไบนารีที่จัดตามรูปแบบหนึ่ง ส่วนหัวอาจบอกชนิดข้อมูล ขนาดภาพ จำนวนรายการ หรือวิธีบีบอัด ส่วนข้อมูลหลักเก็บค่าที่ต้องการ และอาจมีส่วนตรวจสอบหรือส่วนเติมเต็มเพื่อให้ครบหน่วยจัดเก็บ หน้าที่ของแต่ละส่วนขึ้นกับรูปแบบ ไม่ใช่ทุกแฟ้มมีโครงสร้างเดียวกัน

นามสกุล เช่น .txt หรือ .png เป็นข้อมูลบอกประเภทที่คาดหวัง แต่การเปลี่ยนชื่อนามสกุลเพียงอย่างเดียวไม่ได้แปลงโครงสร้างข้อมูลภายใน ข้อความตัวหนึ่งก็ไม่จำเป็นต้องใช้หนึ่งไบต์เสมอ ใน UTF-8 รหัสอักขระใช้ตั้งแต่หนึ่งถึงสี่ไบต์ เช่น A ใช้หนึ่งไบต์ ส่วน ก ใช้สามไบต์ จำนวนอักขระที่มองเห็นจึงไม่ใช่จำนวนไบต์โดยตรง

ส่วนหัว รูปแบบ / ขนาด	ข้อมูล หลัก จำนวนบิตตามเนื้อหา	ส่วนเติมเต็ม ตามกฎจัดเก็บ
--------------------------	-----------------------------------	------------------------------

ขนาดรวมต้องนับทุกส่วนตามข้อกำหนด

ภาพที่ 3 โครงสร้างแฟ้มจำลองแยกส่วนหัว ข้อมูลหลัก และส่วนเติมเต็ม

ขนาดตามแบบจำลองต้องมีหน่วยตรงกัน

ถ้าส่วนหัวมี h ไบต์ และข้อมูลหลัก d บิตบรรจุต่อกันจนจบแล้วเติมเต็มไบต์ ขนาดรวมคือ

$$S = h + \left\lceil \frac{d}{8} \right\rceil B$$

สูตรนี้ไม่รวมส่วนประกอบอื่นที่ไม่ได้ระบุ และไม่ใช้แทนขนาดแฟ้มจริงทุกชนิด

หากแต่ละรายการหรือแต่ละแถวต้องเริ่มในไบต์ใหม่ ต้องปัดขึ้นแยกแต่ละส่วนแล้วจึงรวม ไม่รวมบิตทั้งหมดก่อนปัดครั้งเดียว ส่วนพื้นที่ที่จัดสรรเป็นบล็อกอาจมากกว่าขนาดแฟ้มอีกชั้นหนึ่ง ต้องแยกขนาดข้อมูลกับพื้นที่จัดสรรตามกติกาของระบบ

ตัวอย่างที่ 1 สายบิตหนึ่งชุดกับสองความหมาย

กำหนดให้ตารางรหัสข้อความแทน A ด้วย 01000001 จงระบุจำนวนบิต จำนวนไบต์ และค่าฐานสิบถ้าอ่านสายเดียวกันเป็นจำนวนไม่ติดลบ

วิธีคิด มีแปดตำแหน่งจึงเป็นหนึ่งไบต์ น้ำหนักที่เป็นหนึ่งอยู่ที่ 64 และ 1 จึงได้ค่าจำนวน 65 ส่วนความหมายว่า A มาจากตารางรหัส ไม่ใช่จากจำนวนตำแหน่ง

ตอบ ยาว 8 บิต หรือ 1 ไบต์ มีค่าจำนวน 65 และแทน A ตามตารางที่กำหนด

ตัวอย่างที่ 2 แยกส่วนหัวก่อนปิดจำนวนไบต์

แฟ้มจำลองมีสัญลักษณ์สืบทอด ตัวละห้าบิต บรรจุต่อกันแล้วเต็มเต็มเฉพาะท้าย มีส่วนหัวสามไบต์ จงหาขนาดรวมและบิตเต็มเต็ม

วิธีคิด ข้อมูลหลักใช้ $10(5) = 50$ บิต ต้องใช้เจ็ดไบต์เพราะหกไบต์รับได้เพียงสี่สิบแปดบิต ไบต์ข้อมูลเจ็ดไบต์มีห้าสิบหกบิต จึงเติมหกบิต เมื่อรวมส่วนหัวสามไบต์ได้สิบไบต์

ตอบ รวม 10 ไบต์ และมีบิตเต็มเต็ม 6 บิต

ตัวอย่างที่ 3 แปลงหน่วยก่อนหาจำนวนแฟ้ม

พื้นที่ 2 KiB ใช้เก็บแฟ้มขนาดแฟ้มละ 250 B ได้กี่แฟ้มเต็ม และเหลือกี่ไบต์ ไม่มีพื้นที่จัดสรรเพิ่มเติม

วิธีคิด พื้นที่เท่ากับ $2(1024) = 2048$ B แปลงแฟ้มใช้ 2000 B เหลือสี่สิบแปด แต่เก็บแฟ้มใช้ 2250 B ซึ่งเกินพื้นที่ จึงนับได้เพียงแปด

ตอบ ได้ 8 แฟ้ม เหลือ 48 B

แบบฝึกหัดตอนที่ 1

ข้อมูลและการแทนข้อมูล

จงแสดงวิธีคิดและระบุหน่วยเมื่อมีการคำนวณ ข้อที่ให้ผู้สร้างตารางหรือภาพต้องแสดงผลให้ครบตามเงื่อนไข

ข้อ 1 สายบิต 00101100 มีความยาวกี่บิต คิดเป็นกี่ไบต์ และมีค่าเท่าใดเมื่ออ่านเป็นจำนวนฐานสองไม่ติดลบ หากไม่ทราบตารางรหัส จะระบุได้หรือไม่ว่าแทนตัวอักษรใด

ตอบ

ข้อ 2 อุปกรณ์บันทึกสัญญาณ 13 ตัว ตัวละ 6 บิต โดยบรรจุต่อกันและเต็มเต็มเฉพาะท้ายให้ครบไบต์ มีส่วนหัว 4 ไบต์ จงหาขนาดแฟ้มรวมและจำนวนบิตเต็มเต็ม

ตอบ

ข้อ 3 พื้นที่จัดเก็บสองแห่งมีขนาด 3 kB และ 3 KiB ตามลำดับ ต้องเก็บแฟ้มขนาด 1010 B ต่อแฟ้ม โดยไม่มีพื้นที่ส่วนอื่น จงหาจำนวนแฟ้มเต็มและพื้นที่เหลือของแต่ละแห่ง

ตอบ

ข้อ 4 ช่องสื่อสารส่งข้อมูลหลักได้คงที่ 96 kb/s ต้องส่งข้อมูล 48 kB ไม่มีส่วนหัวและไม่มีเวลารอ กำหนด k เป็น 1000 จงหาอัตราส่งเป็น B/s และเวลาที่ใช้

ตอบ

ข้อ 5 ข้อความ AnAn ใช้ UTF-8 โดยกำหนดให้ A ใช้หนึ่งไบต์ และ n ใช้สามไบต์ ไม่มีอักขระอื่น หากส่วนหัวใช้สองไบต์ จงหาขนาดรวมเป็นไบต์และบิต พร้อมอธิบายว่าเหตุใดการนับตัวอักษรที่เห็นแล้วถือว่าตัวละหนึ่งไบต์จึงไม่ถูกต้อง

ตอบ

ข้อ 6 พิจารณาข้อความสามข้อ ได้แก่ (ก) เปลี่ยนนามสกุลแล้วเนื้อหาภายในย่อเปลี่ยนรูปแบบ (ข) การบีบอัดทุกวิธีทำให้แฟ้มเล็กลงเสมอ และ (ค) การแทนข้อความด้วยเลขฐานสองเพียงอย่างเดียวรับรองความลับ จงระบุว่าข้อใดถูก พร้อมให้เหตุผลแต่ละข้อ

ตอบ

เฉลยวิธีคิดและเกณฑ์ตรวจ

แบบฝึกหัดทุกตอน

แต่ละข้อมีคะแนนเต็มสี่คะแนน เกณฑ์พิจารณาจากหลักฐานในคำตอบ ยอมรับวิธีอื่นที่ให้เหตุผลถูกต้อง และสอดคล้องกับเงื่อนไข

ข้อ 1

วิธีคิด นับได้แปดตำแหน่ง เท่ากับหนึ่งไบต์ ตำแหน่งที่เป็นหนึ่งมีน้ำหนัก 32, 8, 4 จึงมีค่า $32 + 8 + 4 = 44$ การตีความเป็นตัวอักษรต้องอาศัยตารางรหัส จึงยังระบุตัวอักษรไม่ได้

ตอบ 8 บิต 1 ไบต์ ค่าจำนวน 44 และยังระบุตัวอักษรไม่ได้

กลับไปโจทย์ข้อ 1

เกณฑ์ตรวจ (4 คะแนน): ความยาวและหน่วย 1 คำนำน้ำหนักและผลรวม 2 ข้อจำกัดการตีความ 1

ข้อ 2

วิธีคิด ข้อมูลหลักใช้ $13(6) = 78$ บิต แก้วไบต์รับได้เพียง 72 บิต จึงต้องใช้สลิปไบต์ซึ่งรับได้ 80 บิต เดิม $80 - 78 = 2$ บิต แล้วรวมส่วนหัวเป็น $10 + 4 = 14$ ไบต์

ตอบ 14 ไบต์ และเติม 2 บิต

กลับไปโจทย์ข้อ 2

เกณฑ์ตรวจ (4 คะแนน): ข้อมูลหลัก 1 บิตเป็นสลิปไบต์ 1 ส่วนเติมเต็ม 1 ขนาดรวม 1

ข้อ 3

วิธีคิด 3 kB เท่ากับ 3000 B สองแฟ้มใช้ 2020 B เหลือ 980 B แต่สามแฟ้มใช้ 3030 B จึงเกิน ส่วน 3 KiB เท่ากับ 3072 B สามแฟ้มใช้ 3030 B เหลือ 42 B สี่แฟ้มใช้ 4040 B จึงเกิน

ตอบ 3 kB เก็บได้ 2 แฟ้ม เหลือ 980 B ส่วน 3 KiB เก็บได้ 3 แฟ้ม เหลือ 42 B

กลับไปโจทย์ข้อ 3

เกณฑ์ตรวจ (4 คะแนน): แปลงสองหน่วย 1 จำนวนเต็มสองกรณี 2 พื้นที่เหลือสองกรณี 1

ข้อ 4

วิธีคิด อัตราส่งคือ 96000 บิตต่อวินาที หากด้วยแปดได้ 12000 B/s ข้อมูลมี 48000 B จึงใช้เวลา $48000/12000 = 4$ วินาที

ตอบ 12000 B/s และ 4 วินาที

กลับไปโจทย์ข้อ 4

เกณฑ์ตรวจ (4 คะแนน): แปลงบิตเป็นไบต์ 2 แปลงขนาด 1 คำนวณเวลาและหน่วย 1