Ⅱ-1장 자료와 정보의 표현 ····· 36
1-1 자료의 유형 38
1-2 정보의 디지털 표현 42

Ⅱ-2장 자료와 정보의 분석 ····· 56
2-1 자료의 수집 58
2-2 정보의 구조화 64

Ⅱ 자료와 정보

대단원을 시작하면서

인터넷과 정보 기술에 의해 데이터(자료)의 양이 많아짐에 따라 흩어져 있는 아날로그 상태의 자료를 디지털로 변환하여 처리하고 가공하여 의미 있는 정보로 만드는 일이 중요해지고 있다. 정보는 사용 목적에 따라 고유한 형식을 갖고 있으며 문서나 소리, 그림 및 동영상 등 다양한 형태로 표현할 수 있다.

Ⅱ부에서는 아날로그 자료의 특성을 이해하고 디지털 정보로 표현하는 방법을 알아본다. 그리고 문제 해결을 위해 자료를 수집하고 구조화하여 정보를 처리하는 방법에 대해 살펴보고 다가오는 데이터 사이언스 시대에 적응할 수 있는 능력을 기르도록 한다.

Ⅱ-1장

자료와 정보의 표현

학습 목표

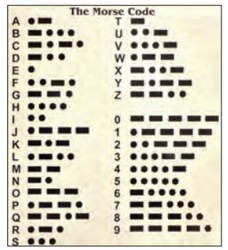
- 자료와 정보의 개념을 이해하고 그 차이점을 구분할 수 있다.
- 아날로그와 디지털의 차이를 알고 컴퓨터의 자료 표현 원리를 설명할 수 있다.
- 정보의 이진 표현은 어떻게 이루어지는지 설명할 수 있다.

생활 속의 이진수와 디지털 정보

우리는 보통 0~9의 십진수를 사용하지만 컴퓨터는 0과 1로 표현하는 이진수를 사용한다. 이진수로 표현되는 디지털 정보의 개념은 우리 생활 속에서 쉽게 찾아볼 수 있다.

'집 안의 전등을 켜고 끄는 스위치'
'원거리에 신호를 전달하기 위한 모스 부호'
'시각장애인의 문자 이해를 돕기 위한 점자'
'막대기 네 개로 즐기는 윷놀이'
'대문의 나무 막대로 집주인의 외출을 표시하는 제주도의 정낭'

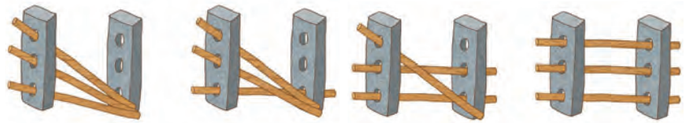
▲ off = 0, on = 1



▲ ●= 0, ▬ = 1

▲ 평평 = 0, 볼록 = 1

▲ 윗면 = 0, 아랫면 = 1



집에 사람이 있음 (집에 사람이 있으니 들어오라는 뜻)
잠시 외출 (금방 들어온다는 뜻)
이웃마을에 외출 중 (오늘 중으로 들어온다는 뜻)
집 멀리 외출 중 (며칠 있다가 들어온다는 뜻)

▲ 내리기 = 0, 올리기 = 1

다음은 자료를 활용한 정보 추천 기능을 소개한 것이다.

탐구 자료 부록 2(178p)

생·각·해·보·기

위의 만화를 보고 자료와 정보의 차이를 생각해 보자.

1-1 자료의 유형

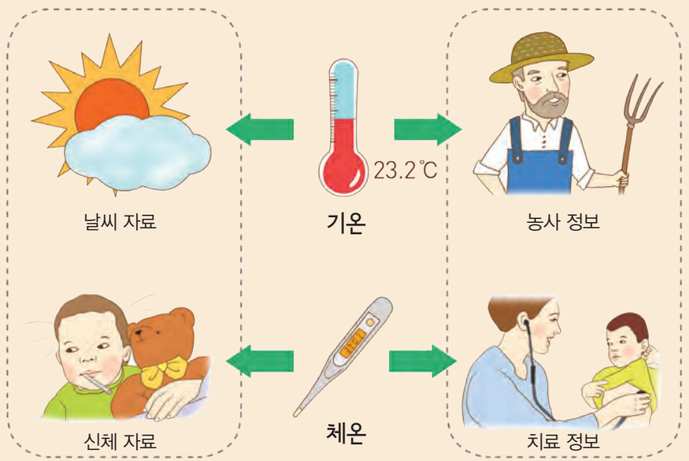
나의 키와 몸무게는 단지 숫자에 불과한 자료이지만 의자를 만드는 사람에게는 편한 의자를 만드는 데 중요한 정보이다. 성적, 날씨, 기차 시간 등도 단순한 수치 자료이지만 그것을 필요로 하는 사람에게는 유용한 정보가 된다. 우리 주변에 자료가 되는 사례를 찾아보고 자신에게 도움이 되는 정보로 나타내 보자.

1 자료와 정보의 차이

우리가 살아가는 세상은 수많은 자료로 넘쳐나고 있다. 정보 기술과 인터넷의 비약적인 발전으로 인하여 자료의 양이 급격하게 증가하였고 이를 우리의 삶에 유용하게 활용하기 위한 빅 데이터(big data) 세상으로 변화하고 있다.

정보 사회에서는 일상생활의 문제를 해결하는 데 정보를 효과적으로 빠르게 분석하고 활용하는 것이 매우 중요하다. 자료는 정보와 의미가 다르고 지식과도 차이가 있다. 다음 예를 통해 자료와 정보의 차이점에 대해 알아보자.

자료와 정보의 차이



자료(data)는 현실 세계에서 관찰이나 측정을 통해 얻은 단순한 값이다. 정보(information)는 자료의 처리와 분석을 통해 의미와 가치를 부여한 것을 말한다.

● **정보와 지식의 차이점**
지식이란 정보를 활용하는 사람이 가지고 있는 기존의 정보를 통합하여 정보를 해석하고 의미를 부여하여 자신의 것으로 만드는 것을 의미한다.

● **정보의 특성**
- 적시성: 이용하고자 하는 사람이 필요할 때 제공되어야 한다.
- 공공성: 많은 사람이 공유할수록 더 큰 가치를 가진다.
- 개별성: 같은 정보라도 사용자에 따라 가치가 달라진다.
- 활용성: 사용자가 활용할 경우에만 가치를 갖게 된다.
- 무형성: 정보는 형태가 없고, 사실에 대한 내용만 있다.
- 교환성: 정보가 다른 매체로 변환되어도 의미는 변하지 않는다.

기상대에서 측정한 온도와 습도는 단순히 숫자인 자료에 불과하지만 농사꾼에게는 1년 농사를 결정짓는 중요한 정보이다. 병원에서 환자의 열을 측정한 온도는 단순한 수치 자료에서 환자의 생명을 좌우하는 중요한 정보가 될 수 있다.

객관적인 자료에 의미를 부여하여 유용한 형태로 가공하고 활용했을 때 정보의 특성과 가치가 드러난다.

자료를 정보로 활용하기

1 다음의 자료를 우리에게 도움이 되는 정보로 활용하는 사례를 들어 보자.
시간, 날짜, 위도, 경도, 수압, 거리, 속도, 무게

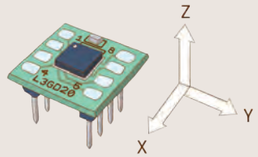
2 다음의 문제를 해결하기 위해 어떠한 자료가 필요한지 설명하여 보자.
- 불쾌지수를 측정하는 기계를 만들 때 필요한 자료는 무엇일까?
- 내비게이션의 도로 정보를 제공하기 위해 필요한 자료는 무엇일까?
- 스마트폰 화면을 회전시키는 가속도 센서에서 필요한 자료는 무엇일까?

2 정보 처리

정보 처리는 기존의 정보를 가공하여 보다 가치 있는 정보를 만들어 내는 것을 말한다. 이렇게 가공된 정보는 문제 해결에 유용하게 활용되며 컴퓨터가 중요한 도구로 이용된다. 컴퓨터를 이용한 정보 처리 과정은 실세계의 자료를 수집하여 입력하고 컴퓨터에 의해 가공·처리된 뒤 출력을 통해 정보로 활용하는 것이다. 컴퓨터에 의해 출력된 정보를 다시 입력한 뒤 재가공하여 새로운 정보를 얻는 피드백 과정을 거치기도 한다.

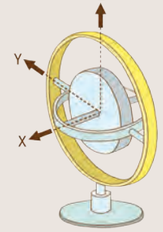
● **불쾌지수**
불쾌지수란 기온과 습도의 조합으로 사람이 느끼는 온도를 표현한 것이다.

● **가속도 센서**
이동하는 물체의 가속도나 충격의 세기를 측정하는 센서이며, x축, y축, z축 세 방향의 단위 시간당 속도를 측정한다.



▲ 가속도 센서

● **자이로 센서**
자이로 센서는 가속도를 측정하는 가속도 센서와 달리 각속도(시간당 회전하는 각도)를 측정한다.



▲ 자이로 센서

사람들의 의견을 수집하여 수치로 나타낸다.

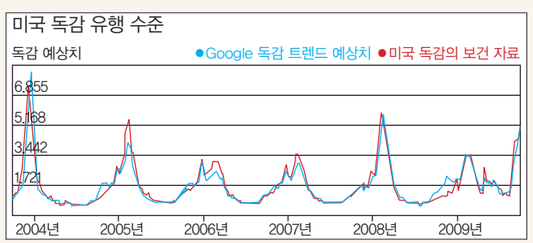
가공된 자료를 시각화된 정보로 나타낸다.

▲ 컴퓨터를 이용한 정보 처리 과정

함께 해 보기

구글 검색어로 독감을 예측한 사례

지난 2008년 구글은 '구글 플루 트렌드(Google Flue Trends)'라는 빅 데이터 기반의 독감 발생 감지 서비스를 공개했다. 사람들이 독감에 걸렸을 때 검색하는 단어를 바탕으로 언제 독감이 유행할지를 예측한 사례이다.



▲ 미국질병통제센터에서 공개한 인플루엔자 의사환자(ILI) 데이터

• 구글의 사례와 유사하게 페이스북과 트위터 등의 소셜 네트워크 서비스에서 나타나는 자료를 모아 우리에게 유익한 정보로 활용한 사례를 살펴보자.

3 자료의 유형

생일을 맞은 친구에게 우리는 스마트폰을 이용하여 축하 문자, 케이크 이모티콘, 생일 축하곡이나 동영상을 보내기도 한다.

이렇게 정보를 표현하기 위한 자료의 형태는 문자, 소리, 그림, 동영상 등이 있으며 각기 서로 다른 특징을 가지고 있다. 이러한 자료는 사용 목적과 형태, 정보 기기의 종류에 따라 다르게 표현된다.

▲ 정보를 생성하기 위한 자료의 형태

1 문자

문자 자료는 사람들이 사용하는 음성 언어를 한글, 영문자, 한자, 숫자, 특수 기호 등으로 표현한 것을 말한다. 신문, 책, 검색어 등에서 사용하는 형태이다.

컴퓨터에서 키보드로 문자 자료를 입력하면 이진 코드로 변환되며 다시 우리가 알아볼 수 있는 문자 형태로 변환되어 모니터, 프린터에 출력된다.

2 그림

그림 자료는 사진, 약도, 표지판, 포스터 등과 같이 사물의 특징을 상징적으로 나타낸 것을 말하며 점과 선, 색을 이용하여 표현된다.

컴퓨터에서는 디지털 카메라, 이미지 스캔 장치를 통해 입력되며, 프린터나 모니터 등으로 출력할 수 있다.

3 소리

소리 자료는 사람의 음성이나 음악 등과 같이 귀로 들을 수 있는 청각적인 형태로 표현된다. 컴퓨터에서는 마이크나 녹음기를 통해 입력받아 처리하며, 이어폰이나 스피커를 통해 출력된다.

4 동영상

동영상 자료는 영화나 텔레비전에서처럼 그림과 소리, 문자를 동시에 사용하며, 시간에 따라 움직이는 영상으로 표현한다.

컴퓨터에서는 휴대 전화나 캠코더 등으로 촬영한 것을 입력받아 처리되며, 모니터와 스피커를 통해 동시에 출력할 수 있다.

▲ 문자 자료의 예

▲ 그림 자료의 예

▲ 소리 자료의 예

▲ 동영상 자료의 예

1-2 정보의 디지털 표현

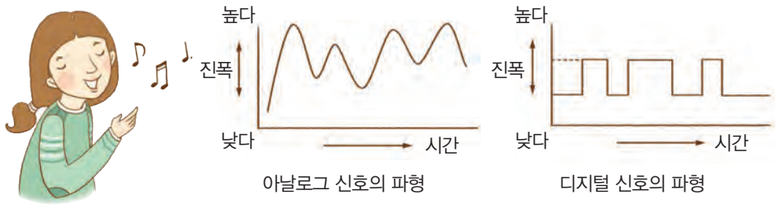
사람이 이해하는 아날로그 정보와 컴퓨터가 이해하는 디지털 정보는 표현하는 방법에 있어서 차이가 있다. 두 가지의 차이점을 알아보고 컴퓨터에 입력되는 다양한 정보의 표현 방법에 대해서도 알아보자.

1 아날로그와 디지털

일상의 자료와 정보는 그 형태에 따라 아날로그와 디지털로 구분할 수 있다. 아래 그림의 시계는 아날로그와 디지털 방식의 예이다.

▲ 아날로그 시계와 디지털 시계

아날로그는 연속적으로 변하는 것들을 물리량으로 나타내는 것으로 빛의 밝기, 소리의 높낮이나 크기, 바람의 세기 등 자연에서 얻는 신호는 대개 아날로그이다. 디지털은 연속적으로 변하는 값을 불연속적인 형태로 표현하는 방식이다.



아날로그 신호의 파형 디지털 신호의 파형

▲ 아날로그 신호와 디지털 신호

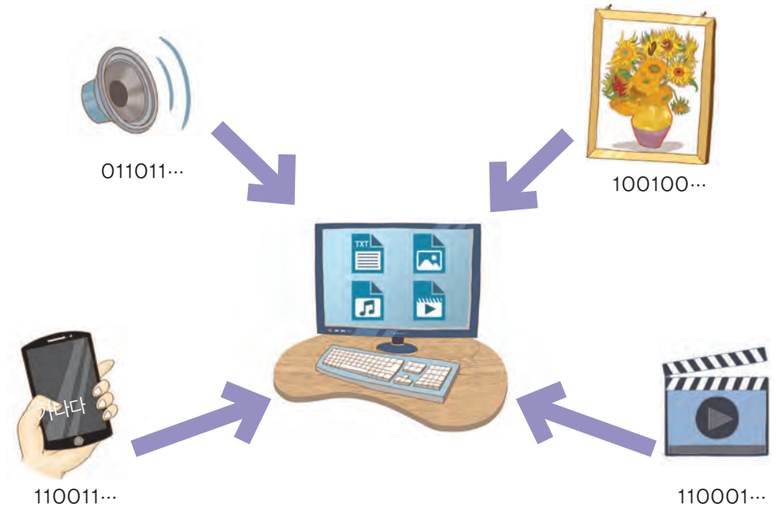
아날로그는 실제 세계에서 사용되는 신호이므로 인간은 감각 기관을 통해 자연스럽게 받아들이며 미세한 신호까지도 자세하게 느낄 수 있다. 아날로그 신호는 똑같이 여러 번 표현하거나 다른 신호로 변형하기 어려운 대신 디지털 신호는 복제나 변형이 쉬운 이점이 있다.

● **아날로그**(analog)
그리스어 'analogia'에서 유래했으며, '닮음, 유사'라는 뜻을 담고 있다.

● **디지털**(digital)
손가락을 뜻하는 라틴어 낱말 'digit'에서 나온 것이다.

2 컴퓨터에서의 자료 표현

컴퓨터는 전기가 흐르는 상태(on)와 흐르지 않는 상태(off)의 두 가지 신호를 사용한다. 두 가지 전기 신호를 0과 1의 숫자로 대체하여 디지털로 표현된 정보를 생성한다. 디지털화의 과정을 통하여 문자, 숫자, 소리, 그림, 동영상 등 어떠한 형태의 정보라도 0과 1이라는 이진 숫자의 형태로 표현할 수 있다.



▲ 컴퓨터에서 자료 표현

비트(bit)는 정보를 두 가지 상태를 나타내는 최소 단위로 0과 1, 참과 거짓, on과 off의 값으로 표현할 수 있다. 하나의 비트는 0 또는 1로 두 가지 상태를 표현할 수 있고 두 개의 비트를 사용하면 네 가지 정보를 표현할 수 있다. 비트의 개수를 추가할수록 더 많은 정보를 표현할 수 있다.

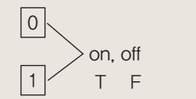
컴퓨터는 0과 1을 사용하여 정보를 처리하기 때문에 여러 개의 비트를 이용하면 문자, 그림, 소리, 동영상 등을 표현할 수 있다.

1비트	상태	2진수
		0
		1

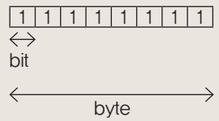
2비트	상태	2진수
		00
		01
		10
		11

● **유니코드**(Unicode)
바이트 2개(16비트)로 구성되어 총 65,536(2^{16})개의 문자를 표현할 수 있다.

● **비트**(bit)
비트는 'binary digit'의 줄임말로 정보를 표현하는 최소 단위이다. 0과 1로 나타낸다.



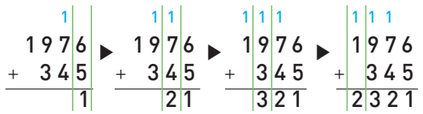
● **바이트**(byte)
비트 8개를 모은 것으로 하나의 문자를 표현하는 단위로 사용된다.



3 정보의 이진 표현

1 정보의 이진 처리

우리가 일상생활에서 쓰는 수는 0에서 9까지의 10개 숫자를 사용하는 십진수이다. 십진수는 수의 자리가 하나씩 올라감에 따라 자리의 값이 10배씩 커지는 수의 표시 방법인 십진법으로 표현한 수이다.

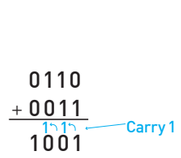


▲ 십진수의 자리 올림

$376 = 3 \times 100 + 7 \times 10 + 6 \times 1$
$= 3 \times 10^2 + 7 \times 10^1 + 6 \times 10^0$

▲ 십진수 376의 전개식

이진법은 수를 표현하는 숫자가 0과 1이므로 $1_{(2)}$이 2개이면 자리 올림이 발생하여 $10_{(2)}$으로 표현되고, $10_{(2)}$이 두 개이면 다시 자리 올림이 발생하여 $100_{(2)}$으로 표현된다. 즉, 이진법은 수의 자리가 하나씩 올라감에 따라 자리의 값이 2배씩 커지는 수의 표시 방법이다. 이진수 $1011_{(2)}$이 나타내는 자릿값은 다음과 같다.



▲ 이진수의 자리 올림

1011

	값	자릿수 × 자릿값
2^0의 자리 ▶	1	(1 × 1)
2^1의 자리 ▶	2	(1 × 2)
2^2의 자리 ▶	0	(0 × 4)
2^3의 자리 ▶	8	(1 × 8)

▲ 이진수 $1011_{(2)}$의 자릿값

컴퓨터는 정보를 처리하기 위해 이진수를 사용한다. 따라서 인간이 사용하는 자료를 컴퓨터로 처리하기 위해서는 십진수를 이진수로 또는 이진수를 십진수로 변환하는 과정이 필요하다.

십진수를 이진수로, 이진수를 십진수로 변환하는 과정은 다음과 같다.

▲ 십진수를 이진수로 바꾸기 ▲ 이진수를 십진수로 바꾸기

● **정보의 이진표현**

❶ 카드 나열하기

❷ 보이는 카드의 점의 개수 세기

= 9

❸ 점이 보이는 카드는 1, 뒤집힌 카드는 0으로 나타내기

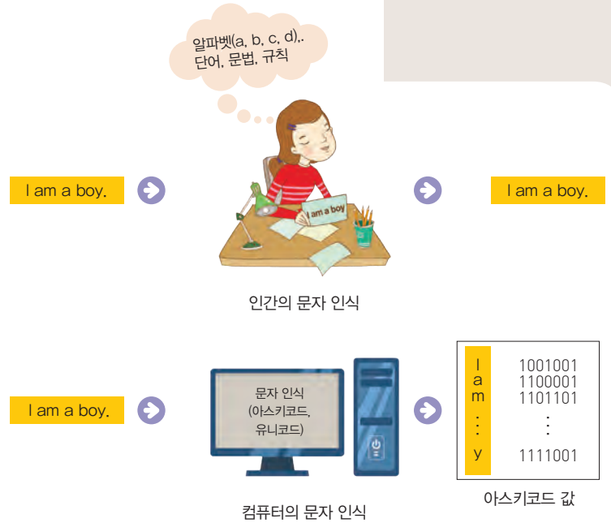
= 21 → 십진수

1 0 1 0 1 → 이진수

2 문자 정보의 표현

컴퓨터에서는 문자를 이진수로 변환하여 처리한다. 각각의 문자는 고유한 비트로 구성된 이진 코드로 표현된다. 문자 표현을 위한 이진 코드는 개발 국가나 제작 회사 등에 따라 다양하게 개발되어 있으며 아스키코드, 유니코드, 한글 코드 등이 있다.

아스키코드(ASCII)는 영문 알파벳을 사용하는 대표적인 문자 코드이다. 알파벳 대·소문자, 0에서 9까지의 숫자, 특수 기호 등을 표현하기 위해 7비트로 구성하여 총 128(2^7)개의 서로 다른 문자를 표현할 수 있다. 정보를 전송할 때는 오류 검사를 할 수 있는 1비트를 추가하여 총 8비트로 구성하여 사용한다.



▲ 인간과 컴퓨터의 문자 인식 과정 비교

7비트 아스키코드 표

문자	아스키코드	문자	아스키코드	문자	아스키코드
A	1000001	a	1100001	@	1000000
B	1000010	b	1100010	=	0111101

● **아스키코드(ASCII)**

'아스키(ASCII)'란 'American Standard Code for Information Interchange'의 약자로 미국 정보 교환 표준 부호이다. 1940년대와 1950년대에 다양한 종류의 코드가 개발되어 사용되면서 통신 문제들이 발생하자 이를 해결하기 위해 채택한 것이다.

함께 해 보기

오류 카드 찾기

1 게임 순서

두 명의 친구(A와 B)와 팀을 만든다.

❶ A: 양면을 무작위로 섞어 5 × 5 배열로 배치한다.

❷ B: 배치된 카드의 아랫쪽에 카드 한 줄(양면이 무작위로 섞인 카드 5개), 오른쪽에 카드 한 줄(양면이 무작위로 섞인 카드 6개)을 추가한다. 이 때, 검은 색 면의 카드 또는 흰색 면의 카드의 수가 짝수가 되도록 추가한다.

❸ A: 배치된 카드 중 하나를 선택하여 뒤집는다(이때 B 학생은 보지 않게 뒤돌아 있는다).

❹ B: A가 뒤집어 놓은 카드를 맞힌다.

▲ 오류 카드 찾기 놀이를 하는 모습

2 생각해 보기

- B 친구는 어떻게 뒤집어 놓은 카드를 맞힐 수 있었을까 생각해 보자.
- 짝과 서로 순서를 바꾸어 가며 오류 카드 찾기 놀이를 해 보자.

유니코드(Unicode)는 전 세계의 모든 문자를 컴퓨터에서 일관되게 표현하고 다룰 수 있도록 만든 국제 문자 부호 체계이다.

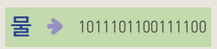
아스키코드로는 영어 외의 다른 문자로 표현하는 데 한계가 있어 이를 보완하기 위하여 유니코드가 개발되었다. 16비트(2바이트)로 구성되어 총 65,536(2^{16})개의 서로 다른 문자를 표현할 수 있다.

유니코드의 예

문자	유니코드
가	AC00
♥	2665
1	0031
A	0041
+	002B

- **한글 코드 구현 방식**

조합형

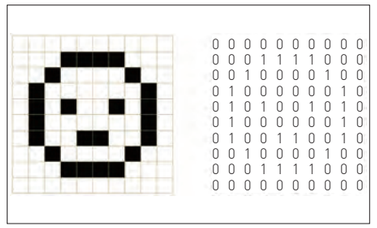


완성형

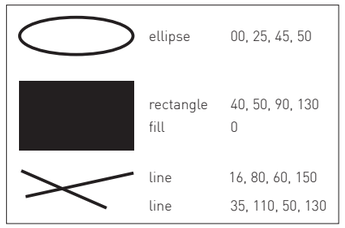
한글 코드는 컴퓨터에서 한글을 사용하기 위하여 개발된 코드로 조합형과 완성형이 있다. 두 가지는 글자가 구현되는 방식에 차이가 있는데, 조합형은 한글의 초성, 중성, 초성에 각각 코드를 부여하여 한글을 표현하며, 완성형은 이미 만들어진 한글 글자에 코드를 부여하여 불러오는 방식이다. 한글을 표현하기 위해 16비트를 부여하며, 우리나라는 유니코드에 규정된 한글 코드를 사용한다.

3 그림 정보의 표현

문자 정보와 마찬가지로 그림 정보도 컴퓨터로 처리하기 위해서 이진수로 표현한다. 그림 정보는 문자 정보보다 더 많은 비트를 갖는다. 컴퓨터는 이러한 그림 정보를 비트맵(bitmap) 방식과 벡터(vector) 방식으로 저장한다.



▲ 비트맵 방식



▲ 벡터 방식

비트맵 방식

컴퓨터 화면을 가로와 세로로 일정하게 나누어 보면 모니터 화면이 더 이상 나눌 수 없는 하나의 작은 점들로 구성된 것을 알 수 있다. 이 작은 점 하나를 화소 또는 '픽셀(pixel)'이라고 한다.

초기의 컴퓨터 게임에서는 픽셀화된 그래픽을 사용하곤 했다. 화소를 이용하여 그림을 표현하는 방식을 비트맵 방식이라고 한다. 비트맵 방식은 그림을 스캔 또는 화면을 캡처하거나 디지털 카메라로 촬영된 사진을 저장할 때 주로 사용한다.

비트맵 방식은 화려한 색상과 복잡한 그림을 쉽게 표현할 수 있는 장점이 있는 반면, 저장된 그림 파일의 크기가 크고 그림을 확대, 축소할 경우에 일그러지거나 깨지는 등의 변형이 나타나는 단점이 있다.

▲ 픽셀 그래픽 이미지

벡터 방식

벡터 방식은 그림을 선과 도형의 집합으로 표현하는 방식이다. 예를 들어 원을 나타낼 경우 원의 중심 좌표, 원의 반지름, 원의 색상 등과 같은 수학적인 수식 정보들로 표현한다. 벡터 방식은 그래픽 프로그램을 이용하여 제품을 디자인하거나 3D 작품을 제작할 때 자주 사용한다.

함께 해 보기

아스키코드와 유니코드 값 변환

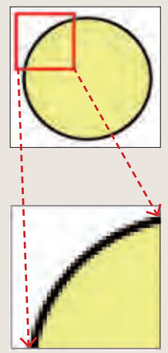
1 자신의 이름을 아스키코드와 유니코드의 값으로 바꾸어 보자.

- 아스키코드: 인터넷에서 '아스키코드 표'를 검색하여 활용한다.
- 유니코드: 한글 문서에서 'Ctrl+F10' 버튼을 누르면 코드 표가 나타나며, 해당 문자를 찾은 후 오른쪽 상단에 보면 유니코드를 확인할 수 있다. 인터넷에서 '유니코드 표'를 검색해도 된다.

2 짝과 함께 아스키코드 또는 유니코드를 선택하여 코드값만 제시해 주고 과일 이름을 찾는 '코드 맞히기' 놀이를 해 보자.

내가 만든 코드	
친구 코드 해석하기	

- **앨리어싱(aliasing) 현상**
  비트맵 이미지를 확대하면 각 점이 그대로 커져 경계선 부분에 계단처럼 보이는 현상이 나타난다.

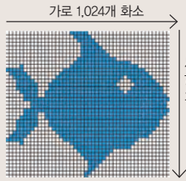


확대된 비트맵 이미지

● **압축의 필요성**
데이터 양이 많으면 저장 장치에 효율적으로 저장하기가 어려우며 통신 네트워크로 전달하기 어렵다.

● **손실 압축**
자료의 일부를 버리고 압축하는 방법으로 압축 성능이 우수하나 원본과 같은 형태로 복원하기 어렵다.

● **무손실 압축**
중복되는 부분을 대체 또는 제거하여 자료를 압축하는 방법으로 압축 효율은 떨어지나 원본과 같은 형태로 복원이 가능하다.



▲ 화면상에 1,024×768 = 786,432개의 화소

벡터 방식의 그림은 비트맵 방식과 달리 수학적 수식을 사용하기 때문에 그림을 확대하거나 축소하여도 그림의 변형이 일어나지 않고 파일의 용량에도 변함이 없는 장점이 있다. 하지만 실세계의 장면을 실감나게 표현하는 데 한계가 있다.

그림 파일의 압축

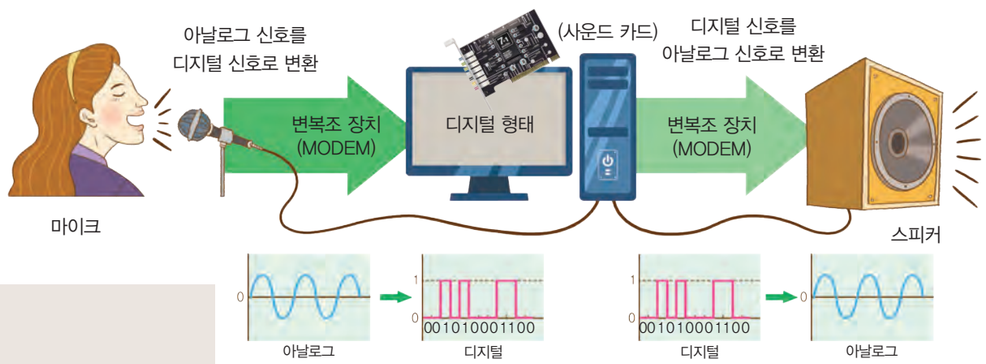
비트맵 방식의 그림은 저장된 파일이 크기 때문에 압축 방법을 사용하여 크기를 줄인다. 확장자가 jpg, gif, png 등의 파일이 대표적인 그림 압축 파일이다. jpg 파일의 경우 압축률이 높아 저장 공간을 줄이고 웹 페이지에서 사용하기 좋다. 하지만 심하게 압축하거나 저장을 반복할 경우 그림의 품질이 떨어지는 단점이 있다.

해상도

컴퓨터에서 사용하는 그림은 같은 모양이라도 품질에 따라 시각적으로 느끼는 정도가 다르다. 이처럼 그림의 품질을 결정하는 중요한 요소가 해상도이다. 해상도란 그림을 몇 개의 화소를 사용해 표현했는지 또는 하나의 화소를 얼마나 자세히 표현했는지에 대한 그림의 세밀한 정도를 말한다. 해상도가 높으면 그림이 선명해지지만 파일의 용량은 커진다.

4 소리 정보의 표현

현실 세계의 소리를 컴퓨터에서 이용하기 위해서는 0과 1의 값을 가진 디지털 신호로 변환하는 과정이 필요하다. 이를 위해 컴퓨터는 마이크와 본체 내부의 사운드 카드와 같은 신호 변환 장치를 이용한다.

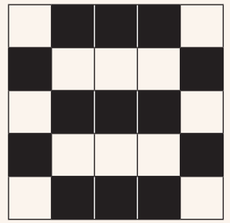


함께 해 보기

그림의 표현과 압축

1 이미지를 이진수로 표현하고 그 수를 간단하게 표현하는 활동을 해보자.

❶ 비어 있는 흰색 칸은 0으로, 검은색 칸은 1로 나타낸다.
❷ 홀수 번째는 흰색 픽셀의 개수, 짝수 번째는 검정 픽셀의 개수로 나타내면 간단히 표현할 수 있다.
❸ 첫 칸이 검은색으로 시작될 경우: 처음에는 흰색이 없다는 신호로 '0'을 써 준다.

이미지	0과 1로 표현하기	간단히 표현하기
	0 1 1 1 0	1(흰), 3(검)
	1 0 0 0 1	0, 1, 3, 1
	0 1 1 1 0	1, 3
	1 0 0 0 1	0, 1, 3, 1
	0 1 1 1 0	1, 3

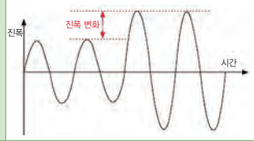
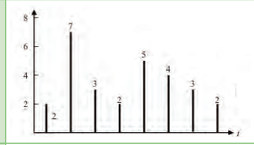
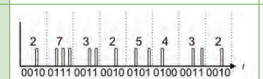
2 위와 같은 방법으로 이미지를 그리고 수로 표현해 보자.

이미지 나타내기	0과 1로 표현하기	간단히 표현하기

3 위에서 간단히 표현한 숫자를 짝의 활동지에 서로 써 주고 난 뒤 숫자만 보고 이미지를 그려 보자.

간단히 표현하기		이미지 나타내기
	→	
	→	
	→	
	→	
	→	
	→	
	→	

컴퓨터를 이용하여 아날로그 소리를 디지털 소리로 변환하기 위해서는 표본화, 양자화, 부호화 과정을 수행한다.

입력		실세계의 소리를 마이크와 같은 장치를 이용하여 입력한다.
표본화		아날로그 신호를 일정한 시간 간격으로 잘라 내어 각 시간 간격마다 그 시간에 해당되는 진폭의 크기를 구한다.
양자화		표본화에 의해 얻은 표본 값을 디지털화한 숫자 값으로 처리한다.
부호화		숫자로 양자화된 값들을 0과 1로 이루어진 2진수로 표현한다.
출력		2진수로 변환된 값은 컴퓨터로 처리하여 저장하거나 스피커를 통해 소리를 출력한다.

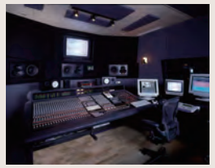
소리 파일의 압축

소리 정보는 음의 파형을 이진수로 표현하기 때문에 문자나 그림 정보보다 많은 정보를 포함한다. WAV 파일은 압축 없이 저장된 소리 파일로 음질은 좋으나 파일의 용량이 큰 단점이 있다. MP3 형식의 소리 파일은 CD 음질을 유지하면서도 소리 파일의 용량을 1/10 정도로 압축한 소리 파일이다. MP3 압축 방식은 우리 귀에 들리지 않는 영역의 소리와 잡음은 제거하고 귀에 들리는 부분만 저장하여 파일의 크기를 줄이는 방법을 사용한다.

미디

악기는 일정한 규칙과 패턴을 담은 악보를 소리의 형태로 표현하는 도구이다. 악기 소리를 컴퓨터로 처리하기 위해 미디(MIDI)라는 코드 규약을 사용한다. 미디 코드는 음표, 속도, 악기, 박자 등의 약속된 규칙을 이진수로 표현한다. 표준화된 미디 코드를 이용하여 피아노 연주와 같은 음악을 컴퓨터로 출력할 수 있다.

● **멀티미디어**
일반적으로 컴퓨터는 문자, 소리, 숫자, 이미지, 동영상 정보 중에서 두 가지 이상의 요소를 합쳐 표현이 가능한데 이와 같이 두 가지 이상의 정보 유형을 합쳐서 디지털 형태로 표현하는 것을 '멀티미디어(multimedia)'라고 한다.



▲ 디지털 레코딩 스튜디오

● **프레임(frame)**
프레임은 비디오나 영화, TV 등이 영상 매체를 전달할 때 화면에 보여주는 한 장의 이미지를 말한다.

● **프레임 율(frame rate)**
1초당 보여 주는 프레임의 개수로 fps(frames per second)로 표현한다.

파일 형식	특징
AVI (Audio Visual Interleave(d))	• 윈도 기본 재생 동영상 • 재생할 때 압축 해제 가능한 코덱이 필요함.
MPEG (Moving Picture Experts Group)	• 압축률이 높아 CD와 방송용으로 많이 사용함.

● **미디(MIDI)**
'Musical Instrument Digital Interface'의 약어로 여러 가지 전자적인 음악 장치들을 연결하여 서로 제어할 수 있도록 하는 표준 인터페이스를 말한다.

5 동영상 정보의 표현

동영상 정보는 정지된 여러 장의 프레임을 보여주는 것으로 TV나 영화 등에서 사용된다. 동영상은 프레임 단위의 그림을 일정한 시간 동안 보여주며 잔상 효과를 이용하여 움직이는 것처럼 표현한다. 실세계의 장면을 디지털 형식의 동영상 정보로 표현하기 위해서는 오른쪽 그림과 같은 과정을 거치게 된다. 자연스러운 동영상을 만들기 위해서는 일반적으로 1초당 프레임이 15장 정도는 되어야 한다.

▲ 동영상 정보 표현하기

▲ 동영상의 품질을 결정하는 초당 프레임 수

동영상의 품질은 1초당 프레임 수와 프레임의 한 화소당 표현하는 색상의 비트 수에 의해 결정된다. 1초당 프레임의 수가 많을수록, 한 화소당 표현하는 색상의 비트 수가 많을수록 보다 선명한 동영상을 표현할 수 있다. 그러나 동영상을 선명하게 표현할수록 동영상 파일의 크기는 커진다.

동영상의 압축

용량이 큰 동영상을 컴퓨터에 저장하고, 네트워크를 통해 빠르게 전송하기 위해서는 효율적인 압축 방법이 필요하다. 대표적인 압축 방법이 MPEG이다. 운동장을 힘차게 뛰어가는 육상 선수의 동영상 장면을 압축하는 사례를 살펴보자. 움직이는 육상 선수를 제외하고 멈춰 있는 나무는 매 프레임에 고정되어 나타나므로 중복된 같은 자료 값을 갖는다. 이렇게 중복된 값을 제거하여 저장함으로써 압축이 가능하다.

▲ 동영상 압축하기

● **코덱**
영상, 소리 등의 미디어 정보를 압축하거나 해제하는 하드웨어나 소프트웨어를 말한다.

실 습 활 동

1 빈칸에 비트 조합으로 표현 가능한 숫자들을 모두 쓰고 그 개수를 적어 보자. 그리고 아래 물음에 답해 보자.

비트	1비트	2비트	3비트	4비트
비트 조합으로 표현 가능한 숫자	0, 1	00, 01, ( ), ( )	000, _____ _____	0000, _____ _____
총 가짓수	2가지	4가지	( )가지	( )가지

① 숫자 7을 이진수로 나타내어 보자.

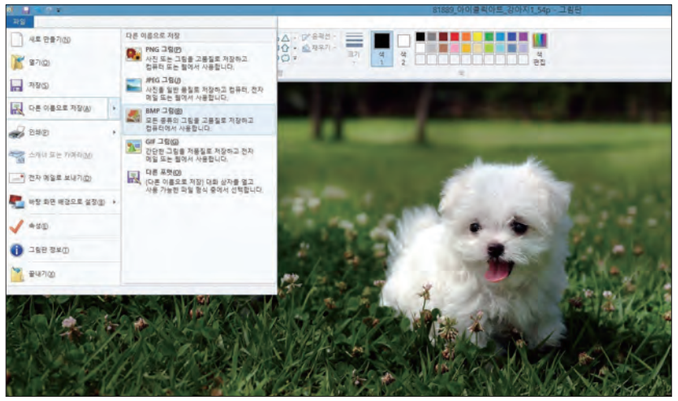
② 4비트로 만들 수 있는 가장 큰 숫자는 무엇인가?

③ 16가지의 기호를 분류하여 나타내고 싶으면 최소 몇 개의 비트가 필요한가?

2 그림판을 이용하여 BMP 형식의 파일을 JPG 형식의 파일로 변환해 보자.

① JPG 그림 파일을 불러오고, 그림을 선택한 뒤 크게 확장시켜 본다. 그림을 확장할수록 그림이 어떻게 보이는가?

② JPG 그림 파일의 확장자를 BMP로 바꾸어 저장하여 보자. 파일의 크기가 어떻게 달라지는가?

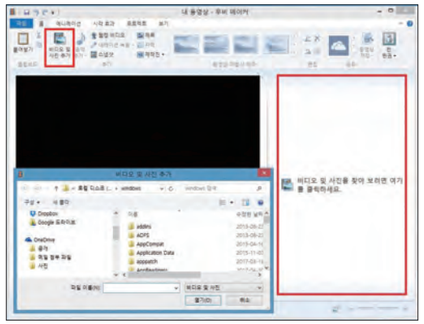


＊ 참고
- BMP(BitMaP)
- JPEG(Joint Photographic Coding Experts Group)

3 컴퓨터로 음악이나 자신의 목소리를 녹음하고 WAV 파일로 저장한 파일과 MP3로 변환한 파일의 크기와 음질 차이를 비교하여 보자.

	WAV	MP3
파일 크기		
음질 차이		

4 윈도 무비 메이커로 5장의 사진과 음악 파일을 이용하여 동영상 파일을 만들어 보자.

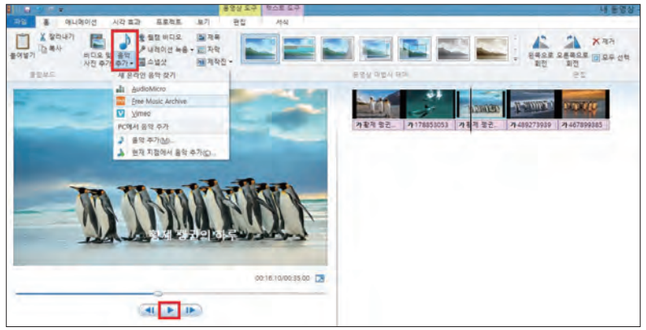


❶ 무비 메이커를 실행시킨다.

❷ 사진을 아래의 빨간 네모 영역에 마우스를 클릭하여 불러오거나 [비디오 및 사진 추가] 기능을 활용하여 추가한다.

❸ 메뉴 상단에서 [음악 추가]를 클릭하여 음악을 불러온다.

❹ 재생 단추를 클릭해서 사진과 음악을 감상한다.



자 기 평 가

평가 내용	잘함	보통	노력 요함
1. 십진수와 이진수의 변환을 할 수 있다.			
2. 그림 파일을 압축하여 변환할 수 있다.			
3. 이미지와 소리 파일을 이용하여 동영상을 만들 수 있다.			

읽기자료

앨런 매티슨 튜링(Alan Mathison Turing, 1912~1954)

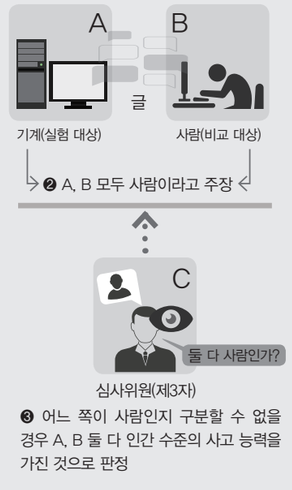
▲ 튜링 머신을 고안한 앨런 튜링

영국의 수학자, 암호학자, 논리학자이자 컴퓨터 과학의 선구적 인물이다. 인간의 생각을 계산과 알고리즘으로 표현할 수 있는 모델을 제안하였다. 인간의 사고 과정을 추상적인 처리 과정으로 표현하여 실제 기계에서 구동 가능한 모델을 만들었다. 이러한 기계를 튜링 기계라 하는데 지금 우리가 사용하는 컴퓨터 구성 장치와 기능의 토대가 되었다. 또한 기계가 지능을 가질 수 있는지 확인할 수 있는 튜링 테스트를 제안하기도 하였다. 튜링 테스트는 사람처럼 대화하는 방법을 사용하여 컴퓨터가 지능이 있는지 확인하는 방법이다.

노벨상처럼 컴퓨터 과학에 중요한 업적을 남긴 사람들에게 매년 시상하는 튜링상은 그의 이름을 따서 제정한 것이다. 튜링은 컴퓨터 과학 이론과 인공지능 분야에 지대한 공헌을 했기 때문에 '컴퓨터 과학의 아버지'라고 불린다.

인공지능 판별 기준

튜링 테스트 개요



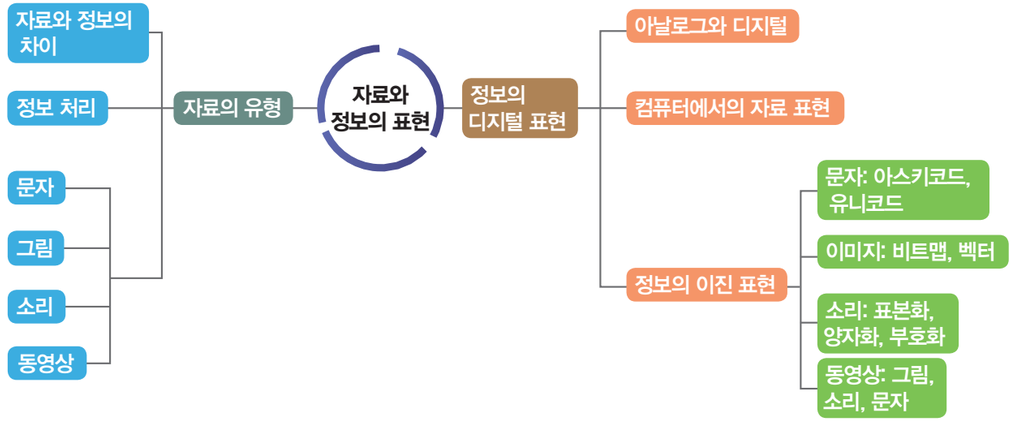
❸ 어느 쪽이 사람인지 구분할 수 없을 경우 A, B 둘 다 인간 수준의 사고 능력을 가진 것으로 판정

▲ 튜링 머신 1280px 모델

토·론·하·기

자료의 크기를 줄여 관리하거나 저장, 전송을 하는 압축 알고리즘이 언제나 좋다고 볼 수 있을까? 은행의 CCTV 영상 또는 병원 의료 사진을 많이 압축할 경우 생기는 문제를 사례로 들어 토론해 보자.

배운 내용 정/리/하/기



평/가/하/기

1 ( ) 안에 알맞은 말을 넣어 보자.

단순한 관찰이나 측정을 통해 얻어지는 값이나 사실을 ( )라 하고, 이를 실제 생활에 도움이 될 수 있는 형태로 바꾸거나 정리한 결과를 ( )라고 한다.

2 디지털 정보를 표현할 수 있는 두 개의 숫자는 무엇인가?

( )과 ( )

3 문자를 표현하는 최소 단위로, 8개의 비트로 표현하는 단위는 무엇인가? ( )

Ⅱ-2장

자료와 정보의 분석

학습 목표
- 자료 수집의 필요성을 알고 문제 해결에 활용할 수 있다.
- 정보 구조의 개념을 설명할 수 있다.
- 정보를 구조화하는 방법의 종류를 사례를 들어 설명할 수 있다.

자료의 바다에서 가치 있는 정보를 만드는 '데이터 과학자'

데이터 과학자는 자료를 가공한 뒤 자료가 갖고 있는 의미와 가치를 분석하고 처리하여 의사 결정에 도움을 줄 수 있도록 하는 전문가를 말한다.

자료의 가공에는 자료의 수집, 통합, 패턴을 분석하여 처리하는 일 등이 포함된다.

사용자가 원하는 검색 결과를 정확하게 보여 주고, 소셜 네트워크 서비스(SNS)에서 친구 관계를 자동으로 알려 주며, 온라인 쇼핑몰에서 구입한 자료를 바탕으로 필요한 물품을 추천하는 편리한 기능들은 데이터 과학자의 역할로 탄생하였다.

미래에는 자료의 양이 기하급수적으로 증가하므로 데이터 과학자의 중요성은 더욱 높아질 것이다.

다음은 일상생활에서 볼 수 있는 자료 구조의 예를 소개한 것이다.

생·각·해·보·기

위의 만화를 보고 자료를 구조화했을 때와 그렇지 않았을 때의 차이를 생각해 보자.

2-1 자료의 수집

문제 해결에 필요한 정보를 얻기 위해서는 자료를 수집하는 것에서부터 시작한다. 객관적이고 체계적인 자료 수집을 통해 문제 해결을 위한 방법을 예측할 수 있기 때문이다. 자료 수집의 필요성과 방법에 대해 알아보자.

1 자료 수집의 필요성

자료 수집은 필요한 정보를 얻기 위한 수집 활동을 말한다. 실세계의 현상을 수집한 자료는 컴퓨터로 처리하기 위해 수치화한다. 이렇게 수치화된 자료를 바탕으로 정보를 생성하여 일상생활의 다양한 문제를 해결할 수 있다. 예를 들어 캠핑 여행 계획을 세우기 위해서는 다양한 자료를 검색하게 된다. 여행에서 중요한 것은 날씨이므로 우선 여행지의 날씨 정보를 살펴보자.

날씨 정보

◀ 날씨 정보의 예

① 여행지의 주간 날씨 예보 자료를 검색한다.
② 강수 확률, 기온, 풍속 등의 자료를 확인한다.
③ 원하는 날짜의 날씨 정보를 확인한다.

기상청의 날씨 자료를 바탕으로 캠핑 준비에 대한 예측이 가능하다. 위에 제시된 날씨 정보에서 일요일과 같이 강수 확률이 높고 뇌우가 발생할 경우 계곡에서의 캠핑 대신 안전한 장소를 선택할 수 있다. 컴퓨터를 활용하기 이전에는 사람들이 문제를 해결하기 위해 알맞은 양의 자료를 수집하여 분석을 하였다. 하지만 정보 통신 기술의 발달로 순식간에 대량의 자료가 만들어지면서 사람이 직접 자료를 수집하고 분석하는 데 한계가 발생하였다. 컴퓨터를 활용하면서 그 한계를 극복할 수 있게 되었다. 이제 우리는 수집한 자료를 바탕으로 사람들의 생각이나 의견, 트렌드 등을 분석하고 예측할 수 있으며 이를 통해 실세계의 문제를 해결할 수 있게 되었다.

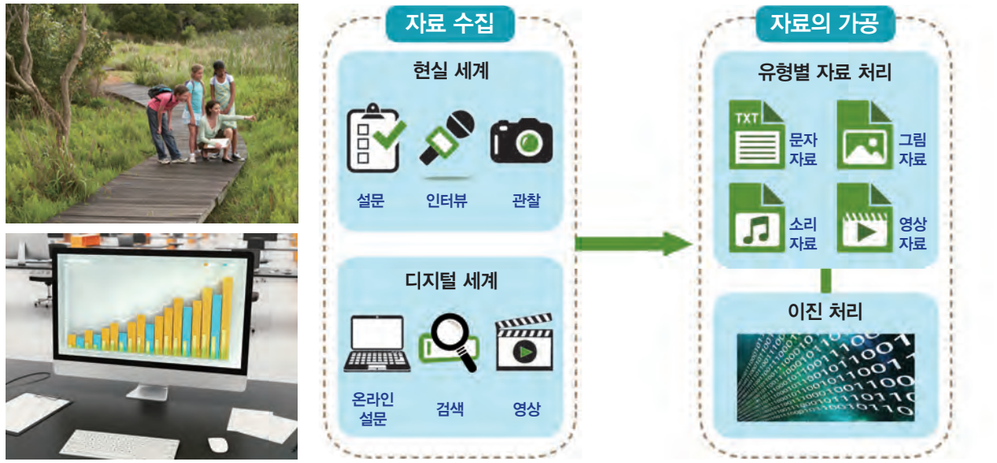
길 찾기 정보

내비게이션에 나타나는 도로와 각 지역에 대해 흩어져 있는 자료들이 내가 길을 찾고자 할 때는 중요한 정보가 되어 문제 해결을 도와준다.

- 수집 자료: 출발 시간, 교통수단, 비용, 도로 상황, 버스의 배차 시간, 이동 거리 등
- 정보 활용: 목적지로 가는 효율적인 경로

2 자료의 수집과 가공

자료 수집에는 직접 조사하거나 디지털 기기를 이용하는 방법이 있다. 직접적인 조사 방법을 통한 자료 수집은 시간과 비용이 많이 소요되므로 현재는 빠른 시간에 적은 비용으로 많은 자료를 수집하기 위해 인터넷을 주로 활용한다.



▲ 자료의 수집과 가공 절차

현실 세계에서 수집된 아날로그 자료는 컴퓨터로 처리하기 위해 이진 처리되어 저장되고 문자, 그림, 소리, 영상 등의 정보로 표현된다.

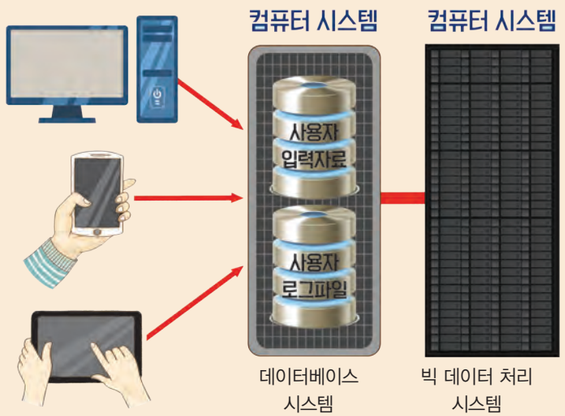
● **로그 파일**(log file)
사용자가 컴퓨터를 실행하는 중에 발생하는 이벤트나 각기 다른 사용자의 통신 과정상에 나타난 메시지를 기록한 파일

● **빅 데이터**(big data)
일반적으로 사용되는 데이터의 수집, 관리 및 처리할 수 있는 소프트웨어의 수용 한계를 넘어서는 크기의 데이터를 말한다.

인터넷에서 자료가 생성되는 경우는 사용자들이 직접 자료를 입력하여 생성하기도 하지만 컴퓨터가 사용자의 사용 흔적을 토대로 자료를 생성하기도 한다. 수많은 사용자가 생성한 자료와 컴퓨터가 생성한 자료가 무수히 수집되면서 빅 데이터의 시대가 열렸다.

빅 데이터는 다양한 종류의 대규모 자료에서 가치를 추출하고 결과를 분석하는 기술을 말한다. 빅 데이터 기술의 발전으로 인하여 사회를 정확하게 예측하여 효율적으로 작동하게 하고, 사회 구성원에게 맞춤형 정보를 제공·관리·분석 가능하게 되었다.

빅 데이터를 위한 자료 수집



함께 해 보기

온라인 상품 판매와 홍보를 위한 자료 수집

신발을 판매하는 온라인 쇼핑몰을 창업하고자 한다. 인터넷 사용자들에게 상품을 효과적으로 판매하거나 홍보하기 위해 홈페이지를 통해 수집할 수 있는 자료를 작성해 보자.

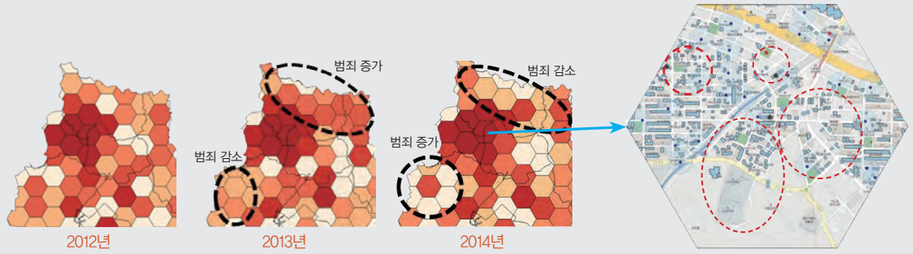
수집 방법	수집 자료	수집할 자료
사용자가 입력한 자료 수집	성별, 나이 등	
사용자의 사용 흔적을 컴퓨터가 자동 저장하여 자료 수집	로그인 시간, 사이트 방문 횟수 등	

읽기 자료

실세계의 자료를 컴퓨터로 분석하여 문제 해결에 적용한 사례

1. **선거 결과 예측**: 소셜 네트워크 서비스(SNS) 여론 분석을 통해 투표 결과를 예측한다.
2. **추천 상품 표시**: 온라인 서점이나 온라인 쇼핑몰에서 고객들의 구매 내역을 분석해 소비자의 소비 취향과 관심사를 파악하여 추천 상품을 표시한다.
3. **야구의 경기 데이터 분석**: 메이저리그 베이스볼(MLB)의 머니볼 이론은 경기 데이터를 철저하게 분석해 오직 데이터를 기반으로 적재적소에 선수들을 배치해 승률을 높인다는 게임 이론이다. 실제 25위에 해당하는 야구팀이 이를 통해 메이저리그 최초로 20연승을 올리는 신기록을 세운 바 있다.
4. **'범죄 예상 지도'와 '안전 지도'**: 빅 데이터를 활용해 사용자의 위치와 관련된 위험 지역과 안전한 지역의 데이터를 분석해 '범죄 예상 지도'와 '안전 지도'를 만들 수 있으며, 안심 귀가를 할 수 있도록 사용자에게 실시간 정보를 제공하고 사고 발생 시 주변 지인과 정보를 교환할 수 있는 시스템도 고안할 수 있다. 과학적인 데이터를 토대로 범죄의 발생 빈도 예상은 물론 이를 통해 CCTV의 취약점까지 파악할 수 있다.

▲ 사용자의 위치와 관련된 위험 지역 표시



▲ 범죄 발생 추이 변화 지역별 분석

▲ 다차원 분석 CCTV 최적지 선정

5. **상품의 배치 방법 연구**: 미국의 월마트에서는 고객들이 이동하는 경로와 성별, 시간대에 따라 물건이 판매되는 것이 다르다는 것을 파악하였고, 혼자 온 고객의 구매 내역과 자녀와 함께 온 부모들의 구매가 다르다는 것을 컴퓨터로 분석하였다. 예를 들면 아이들의 눈높이에 놓인 선반에는 사탕, 초콜릿, 과자 등을 배치하고, 주부들의 눈높이와 동선에 맞춰 주방 용품과 목욕 용품 등을 배치하여 구매욕을 높이도록 하였다. 이러한 배치 방법은 판매되는 물건의 자료와 구매자들의 자료를 수집하여 분석한 중요한 정보에 따른 것이다.

3 자료를 이용한 문제 해결

수집한 자료를 이용하여 문제를 해결하는 과정은 다음과 같다.

자료 수집 → **자료 가공** → **자료 분석** → **문제 해결**

햄버거 가게를 운영하는 부모님 돕기 프로젝트

자료를 수집, 가공, 분석하여 실세계 문제 해결에 활용하는 예를 살펴보자.

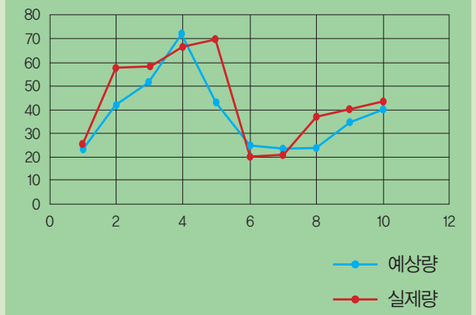
자료 수집 날씨, 요일, 행사 등에 따라 판매된 햄버거의 수치 자료를 수집한다.

자료 가공 수집한 자료를 스프레드시트에 입력하여 표의 형태로 가공한다.

5월	날씨	요일	행사1	행사2	행사3	예상량	실제량
1						23	25
2	비				야구	42	57
3		주말				51	58
4	맑음	주말				72	66
5				어린 이날		43	69
6						24	20
7	흐림					23	21
8			어버 이날			23	37
9					생일 파티	34	40
10		주말				40	43

자료 분석 그래프를 이용하여 예측 가능한 햄버거의 수요를 분석하고 재료의 구매량을 결정하여 비용을 줄인다.

예측 모델의 정확도



문제 해결
- 법정 공휴일의 경우 판매량이 늘어난다.
- '생일 파티' 예약 시스템을 활용하여 재료를 준비한다.
- 주말에 판매량이 늘어난다.

실습활동

온라인 설문을 활용하여 자료를 수집하여 다음의 자료를 바탕으로 문제를 해결하여 보자.

1 문제 선정하기

예 친구들의 스마트폰 사용 실태 분석을 통한 문제점 진단과 해결 방법 찾아보기

2 온라인 설문 도구를 사용하여 설문지 만들기

① 온라인 설문지를 개발할 도구를 선택한다.

② 설문 내용을 작성한다(설문 제목, 설문 문항, 답변 유형 등).

설문지 양식

우리 반 친구들의 스마트폰 사용 실태 분석

우리 반 친구들의 스마트폰 사용 실태를 분석하여 문제점 진단과 그에 따른 해결 방법을 함께 알아보고자 합니다. 솔직한 답변 부탁드립니다.

1. 하루에 스마트폰을 몇 시간 정도 사용하나요?

○ 1시간 미만 ○ 1시간~2시간 미만 ○ 2시간~3시간 미만
○ 3시간~4시간 미만 ○ 4시간 이상

2. 스마트폰으로 주로 이용하는 것은?(중복 답변 가능)

□ 소셜 네트워크 서비스(SNS)(카카오톡, 페이스북) □ 게임 □ 유튜브
□ 인터넷 강의 □ 카페나 블로그 □ 웹툰

3 자료 수집하기

① 작성된 설문지의 링크 주소를 친구들에게 알려 준다.

② 친구들은 스마트폰으로 접속하여 설문에 답변한다.

4 자료 가공 및 분석하기

① 응답 내용을 그래프 또는 표의 형태로 가공한다.

② 응답 요약을 토대로 시각화된 자료를 보며 분석한다.

5 문제 해결 및 공유하기

① 우리 반 친구들의 스마트폰 사용 실태를 파악한다.

② 분석된 결과를 통해 발견한 문제점과 개선 방안을 함께 토론한다.

자기평가

평가 내용	잘함	보통	노력 요함
1. 모둠원과 협력하여 문제를 해결할 수 있다.			
2. 온라인 설문 도구를 활용하여 설문지를 만들 수 있다.			
3. 일정 기간 동안 필요한 자료를 수집, 관리할 수 있다.			
4. 수집된 자료를 가공하여 분석할 수 있다.			
5. 친구들과 공유하여 함께 토론할 수 있다.			

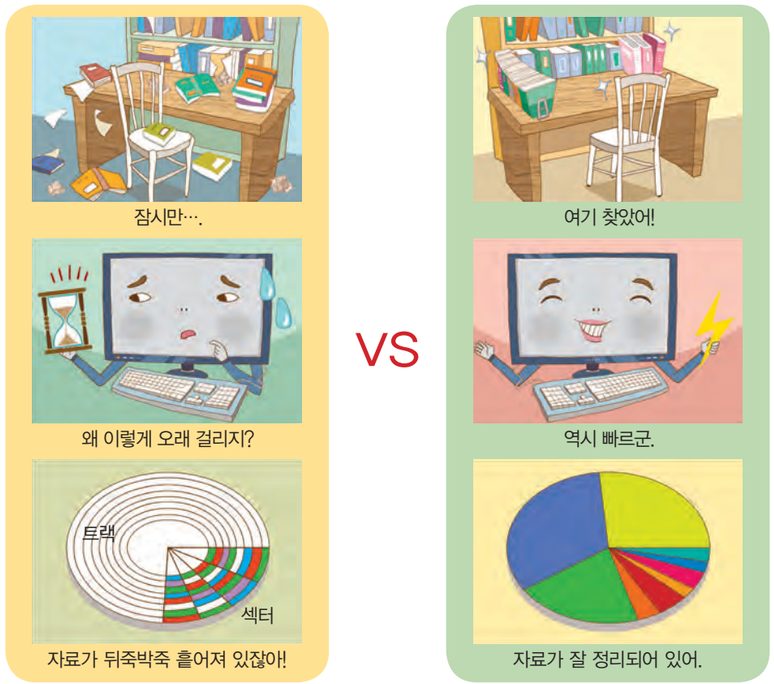
2-2 정보의 구조화

일상생활에서 우리는 물건들을 정리하기 위해 다양한 방법으로 구조화한다. 마구잡이로 흩어 놓는 것보다 차곡차곡 쌓아 정리하면 물건을 작은 공간에 많이 보관할 수 있기 때문이다. 컴퓨터에서는 입력된 자료를 효율적으로 처리하기 위해 어떻게 구조화하는지 알아보자.

1 정보 구조의 개념

우리는 일상생활에서 줄을 서서 차례를 기다리거나, 물건을 분류하여 배치하는 경우가 많다. 이렇게 정리를 하면 물건을 쉽게 찾거나 일을 빠르게 처리할 수 있다. 컴퓨터 역시 일을 효율적으로 처리하기 위해 정보를 구조화하는 방법을 사용한다.

'내가 찾는 것이 어디 있지?'



▲ 정보를 구조화했을 때의 장점

● **정보 구조화의 장점**
- 정보를 체계적으로 정리하여 관리할 수 있다.
- 정보의 흐름이나 절차를 쉽게 파악할 수 있다.
- 효율적으로 일을 처리할 수 있다.
- 빠진 부분을 쉽게 발견할 수 있다.
- 비슷한 구조로 쉽게 변경이 가능하다.

컴퓨터에서 처리하는 자료의 양이 많으면 많을수록 이들을 종류별로 구분하여 저장해 두면 관리하기가 쉽다. 이렇게 컴퓨터에 자료를 효율적으로 저장하기 위해 자료들을 구조화하는 방법을 정보 구조라고 한다.

정보 구조는 컴퓨터가 수행하는 일의 효율성과 관계가 있으며 또한 그 자체로써 컴퓨터의 기억 장치를 사용하게 된다. 즉 효율적인 정보 구조는 적은 양의 기억 장소를 사용하여 원하는 정보를 빠르게 검색할 수 있도록 도와준다.

● 적절한 자료 구조는 보다 효율적인 알고리즘을 사용할 수 있게 한다.

생활 속의 구조화된 정보

아파트는 여러 명이 공동으로 주거하는 곳이다. 효율적으로 관리하기 위해 동과 호수 등과 같이 구조화된 형태로 배치되었다.

학교는 많은 학생과 교사들이 일련의 교육 과정을 통해 수업을 진행하는 곳이다. 시간 관리, 성적 관리, 학생 관리, 입학 관리 등 많은 업무를 수행하기 위해 구조화된 형식을 사용한다.

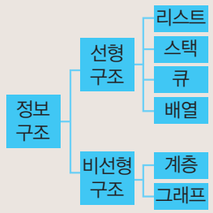
우체국은 우편물을 효율적으로 취급하기 위해 택배 업무, 우편배달, 주소 체계, 우편 규격 등의 업무를 구조적으로 설계하여 관리한다.

▲ 정보 관리의 효율성

이렇게 실생활의 정보를 구조화하면 업무를 효율적으로 처리할 수 있다. 정보를 구조화하여 관리와 통제를 쉽게 할 수 있는 일상생활의 또 다른 예로 무엇이 있을까?

정보 구조는 자료 간의 연결 형태와 모양에 따라 크게 선형 구조와 비선형 구조로 나눌 수 있다.

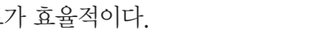
● 생활 속 대표적인 구조화 방법은 다음과 같다.
- 목록
- 계층
- 표
- 다이어그램



▲ 정보 구조의 유형

선형 구조는 자료를 순차적으로 직선 형태로 나열되어 있는 구조로 리스트, 스택, 큐, 배열이 있다.

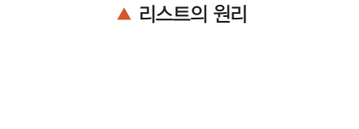
한 원소 뒤에 하나의 원소가 존재한다.
원소 간의 순서를 고려한다. 전후(인접) 원소 간에 1:1 관계로 나열된다.



▲ 선형 구조는 하나의 자료 뒤에 하나의 자료가 일 대 일 관계로 연결된다.

비선형 구조는 자료를 일렬이 아니라 특별한 형태로 저장하는 구조로 계층과 그래프가 있다. 계층적 그래프를 나타내기에 적합하다.

한 원소 뒤에 여러 개의 원소가 존재할 수 있다.
인접(전후) 원소 간에 일(一):다(多) 또는 다(多):다(多) 관계로 배치된다.



▲ 비선형 구조는 한 원소 뒤에 여러 개의 원소가 다 대 다 관계로 연결된다.

● 정보를 구조화할 때 정보의 특징, 정보 간의 관계, 구조화의 목적과 용도 등을 고려하여 가장 적절한 구조화 방법을 선택하여야 한다.

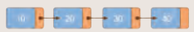
2 정보 구조화의 방법

1 선형 구조

① 리스트(List)

리스트는 자료의 순서를 매겨 일렬로 저장하는 구조이다. 가장 단순하면서도 널리 쓰인다. 자료의 양이 많지 않고, 자료의 삽입·삭제가 빈번하게 이루어질 경우에는 연결 리스트가 효율적이다.

● **리스트의 연결 구조**
• 단순 연결 리스트



생활 속의 리스트 구조
예 칸칸이 연결된 기차

사진 자료를 연결 리스트로 저장하여 스마트폰에서 사진을 넘겨 볼 수 있다.

▲ 리스트의 원리

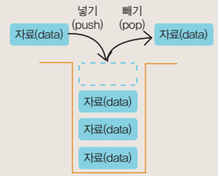
② 스택(Stack)

스택은 리스트의 한쪽 방향에서만 자료를 삽입하거나 삭제할 수 있는 구조이다. 자료를 삽입, 삭제하는 것이 한쪽에서만 이루어지기 때문에 가장 최근에 삽입된 자료가 가장 먼저 삭제된다. 웹 브라우저에서 뒤로 가기 기능이나 실행 취소(Ctrl+Z) 기능, 휴대폰 메시지, 인터넷 메일의 자료 정리, 산술식의 계산 등과 같은 처리에 이용된다.

생활 속의 스택 구조
예 연탄 쌓기, 컵 쌓기

'되돌아가기' 버튼을 이용하여 가장 마지막에 본 웹 페이지 순으로 돌아갈 수 있다.

▲ 스택의 원리

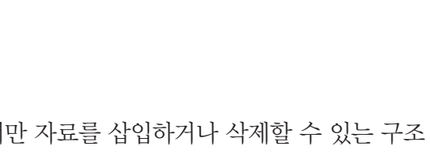


● 스택은 나중에 삽입한 자료가 먼저 삭제된다고 하여 후입선출(LIFO, Last In First Out) 구조라고도 한다.

③ 큐(Queue)

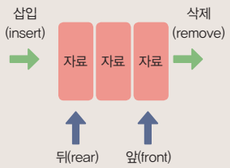
큐는 스택과 달리 한쪽 끝에서는 자료가 삭제되고 반대쪽 끝에서는 자료의 삽입만 가능한 구조이다. 가장 먼저 삽입된 자료가 가장 먼저 삭제된다. 프린터는 컴퓨터에서 인쇄 명령을 받은 순서대로 인쇄를 하는데, 이것은 바로 인쇄 명령이 먼저 받은 명령부터 실행되는 큐 구조로 되어 있기 때문이다.

생활 속의 큐 구조
예 순번 대기표, 줄서기



먼저 입력된 파일부터 출력

▲ 큐의 원리



● 큐는 가장 먼저 삽입한 자료가 가장 먼저 삭제된다고 하여 선입선출(FIFO, First-In, First-Out) 구조라고도 한다.

④ 배열(Array)

배열은 같은 유형의 자료를 연속적인 공간에 저장하는 구조이다. 번호와 그 번호에 대응하는 자료들로 이루어진다.

하나의 행으로 이루어지는 1차원 배열, 행들과 열들로 이루어지는 2차원 배열, 이를 더 확장한 3차원 배열 등이 있다.

0	1	2	3	4
3	2	7	1	9

● 1차원 배열의 예: 초록색 첫째 줄의 번호가 색인이고 둘째 줄의 값이 저장된 자료

● 2차원 배열은 행과 열을 가진 구조이며, 1차원 배열을 여러 개의 행으로 겹친 구조이다.

생활 속의 배열 구조

예 시간표

▲ 배열의 예

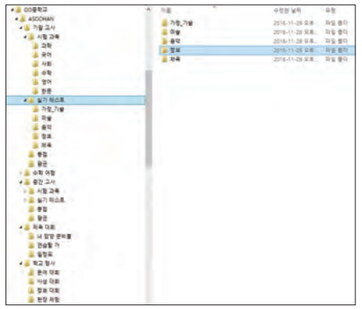
엑셀을 이용한 배열표 만들기

	A	B	C	D	E
1	번호	국어	수학	사회	정보
2	1	80	84	72	96
3	2	88	92	84	92
4	3	96	92	88	92
5	4	100	96	84	88
6	5	76	80	92	96
7	6	80	72	96	100
8	7	84	84	88	92
9	8	92	96	84	100
10	9	80	84	92	96
11	10	80	72	84	88

2 비선형 구조

❶ 계층

계층형 구조는 비선형 구조로서 한 지점에서 여러 갈래로 나누어져 연결되는 형태를 표현한 구조이다. 컴퓨터에서는 폴더 구조가 계층형 구조의 예라고 할 수 있다. 컴퓨터에서는 기억 장소의 할당, 자료의 정렬과 탐색 등에 이용된다.

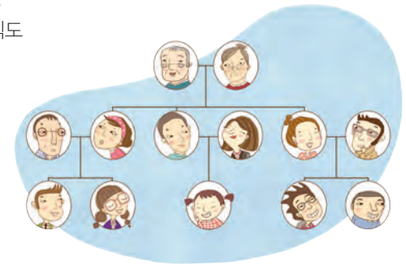


계층적 구조

● 계층형 구조는 나무를 닮은 자료 구조라 하여 '트리(tree)' 구조라고도 한다.

생활 속의 계층 구조

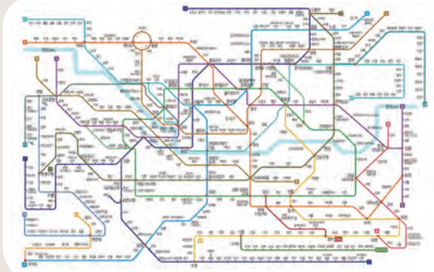
예 가계도, 회사 조직도



▲ 계층 구조의 예

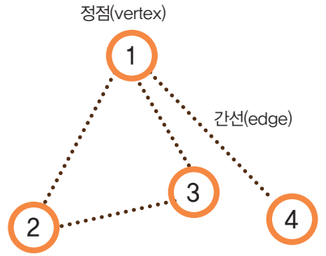
● 생활 속의 그래프 구조

예 전철 노선도, 인맥 지도

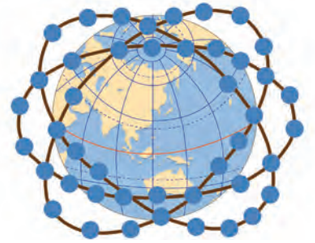


❷ 그래프

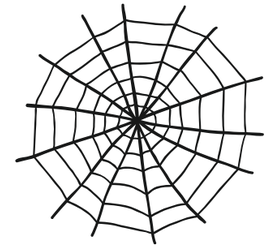
그래프형 구조는 연결되어 있는 개체들 간의 관계를 표현한 구조이다. 그래프를 이용하여 단순한 연결 관계뿐만 아니라 개체들이 연결된 연결선으로 방향을 나타내거나 선의 길이를 통해 개체들 간의 거리 정보도 나타낼 수 있다. 그래프는 모든 연결 구조를 표현할 수 있기 때문에 여러 분야에서 폭넓게 이용되고 있다.



방향이 없는 그래프



인터넷 망의 구조(대표적인 그래프 구조)



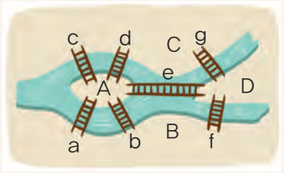
거미줄 구조

▲ 그래프 구조의 예

다리 건너기 문제

모든 다리를 한 번만 건너서 처음 출발했던 장소로 돌아오는 문제를 해결하기 위해 다음 그림을 구조화해 보시오.

모든 다리를 한 번만 건너 돌아오는 경로 문제 ▶



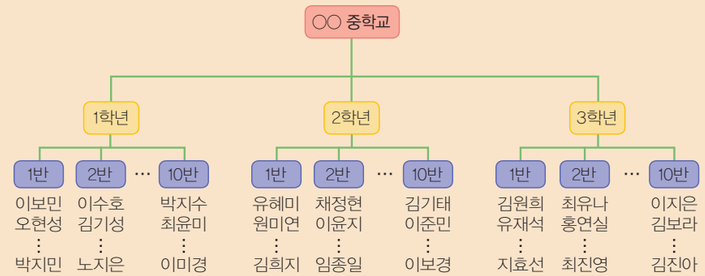
정보 구조화 방법의 선택은 이렇게!

일 년 동안 키가 얼마나 자랐는지 정리한 표를 살펴보자.

이름	김희지	박지수	이준민	이지은	최유나	박지민
키(cm)	4	2.3	6	3.7	3	6.6

이 표를 통해 키가 가장 많이 자란 학생이 '박지민'이라는 것을 알 수 있다. 하지만 우리 학교 전교생 1,000명 중에서 가장 많이 자란 학생을 찾으려면 어떻게 해야 할까? 표를 만든다 하더라도 모든 학생의 키를 일일이 비교해야 하므로 시간이 오래 걸리며, 그 과정에서 오류가 생길 수도 있을 것이다.

이럴 경우에는 계층형 구조를 이용하여 각 학급에서 가장 많이 자란 친구를 찾고, 같은 학년에서 가장 많이 자란 친구를 찾고, 마지막으로 전교생 중에서 찾는다면 빠르고 정확하게 처리할 수 있다. 빠른 탐색이 필요한 경우에는 리스트나 표(테이블) 형보다는 계층형이 적절하다고 할 수 있다.



▲ 계층형 구조화

실 | 습 | 활 | 동

1 친구들이 사 먹는 간식에 대해 실태를 조사하고 분석하여 보고서로 작성하여 보자.

활동 순서	활동 내용	활용할 수 있는 응용 소프트웨어
❶ 실태 조사	친구들이 사 먹는 간식에 대해 직접 조사하여 목록 작성하기 (예 장소, 횟수, 먹는 양, 먹는 시간, 간식의 종류 등)	• 정보 검색하기: 검색 엔진, 도서관 검색 • 워드프로세서로 문단 구조화하기, 정렬하기 • 엑셀로 표 구조화하기, 정렬하기, 계산하기, 그래프로 나타내기 • 파워포인트로 도형 구조화하기 • 앨범으로 그림 구조화하기
❷ 유형 분류	직접 조사한 자료를 유형별로 분류하기 (예 장소별, 횟수별, 먹는 양, 먹는 시간대, 간식 유형별 등)	
❸ 디지털 자료화	엑셀을 이용하여 표 구조화하고 정렬하기	
❹ 실태조사를 바탕으로 필요한 정보 수집하기	검색 엔진을 활용하여 정보 수집하기 (예 칼로리, 식품 첨가물, 위생 상태 등)	
❺ 보고서 작성하기	자료와 정보를 분석하여 보고서로 작성하기 (예 과소비, 영양 불균형, 위생 등)	

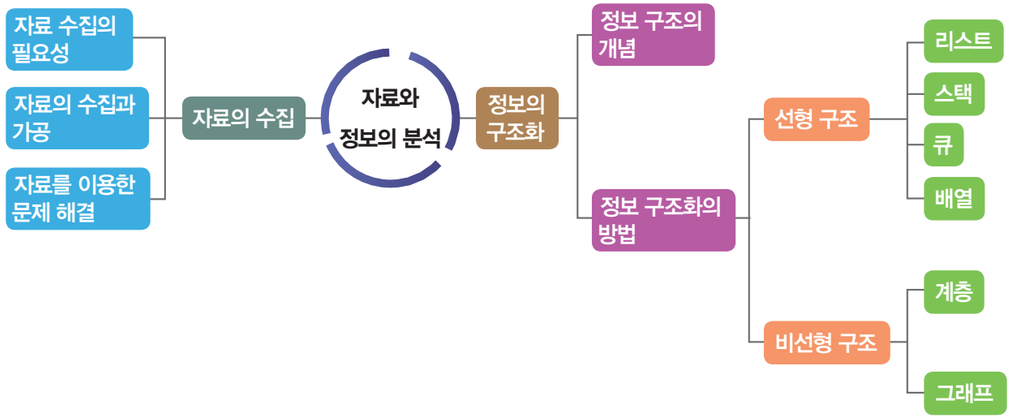
자 | 기 | 평 | 가

평가 내용	잘함	보통	노력 요함
1. 수집한 자료를 바탕으로 정보를 생성할 수 있다.			
2. 정보를 효과적으로 전달할 수 있는 형태로 구조화할 수 있다.			
3. 다양한 응용 소프트웨어를 활용할 수 있다.			
4. 보고서를 체계적으로 작성할 수 있다.			
5. 친구들과 공유하여 함께 토론할 수 있다.			

토 · 론 · 하 · 기

네트워크로 구성된 현대 사회는 6단계(또는 6명)만 거치면 모든 사람의 관계가 연결된다는 6단계 이론(Small World 이론)이 있다. 과연 이 6단계 이론이 맞는지 '그래프' 구조를 이용하여 친구들과 실험하고 토론해 보자.

배운 내용 정/리/하/기



평 | 가 | 하 | 기

1 다음의 내용이 설명하는 정보 구조화의 형태를 적어 보자.

① 한 지점에서 여러 갈래로 나누어져 연결되는 형태이다.

② 내용 요소 간의 상호 관계를 선으로 연결하는 구조이다.

2 아래 자료가 스택 구조라면 가장 오래된 자료는 몇 번일까?

3 정보를 구조화하였을 때 좋은 점 한 가지를 써 보자.

단원 평가

1 자료의 유형과 구체적인 사례를 연결해 보자.

① 문자 자료 •	• ㉠ 음향
② 소리 자료 •	• ㉡ 교통 표지판
③ 이미지 자료 •	• ㉢ 이름표
④ 동영상 자료 •	• ㉣ 체육대회 응원 영상

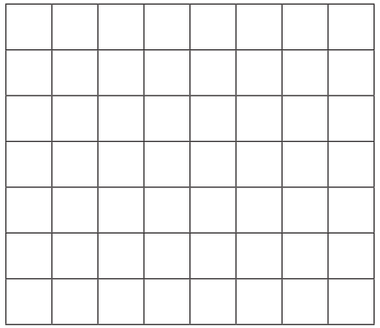
2 십진수를 이진수로, 이진수를 십진수로 변환해 보자.

① 18 → ( )　　② $1011_{(2)}$ → ( )

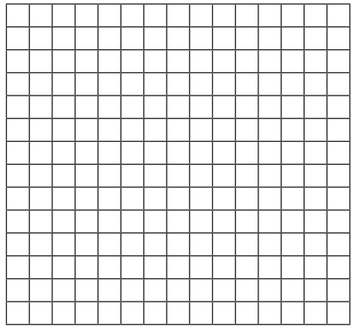
③ 53 → ( )　　④ $110111_{(2)}$ → ( )

3 아래 두 가지 모눈종이에 숫자 3을 색칠하여 나타내 보자.

①



②



※ 모눈종이에 숫자 3의 모양을 최대한 크게 그리고, 해당되는 칸은 꽉 채워 색칠한다.

1) 두 개의 그림 중 어느 그림이 자세하게 보이는가?

2) 모눈종이의 흰 픽셀은 0, 검은 픽셀은 1로 나타낸다면 어떤 것이 1의 값을 더 많이 포함하는가?

3) 그림을 더 자세하게 표현하는 것과 픽셀은 어떤 관계가 있는가?

4 수집된 자료를 이용하여 문제를 해결하는 과정을 차례로 나열해 보시오.

① 문제 해결	② 자료 가공	③ 자료 수집	④ 자료 분석

( ) → ( ) → ( ) → ( )

5 ( ) 안에 들어갈 말을 찾아 쓰시오.

자료, 정보, 아날로그, 디지털

① 단순한 관찰이나 측정을 통해 수집한 사실이나 어떤 값을 말한다. ( )

② 단순히 나열되어 있는 자료를 더욱 유용하고 의미 있도록 가치를 부여한 것을 말한다. ( )

③ 연속적으로 변화하는 값을 일정한 간격으로 끊어 불연속적인 값으로 양을 표현하는 것을 말한다. ( )

④ 연속적으로 변화하는 값으로 양을 표현하는 것을 말한다. ( )

6 조각퍼즐 맞추기 놀이를 계층형 구조로 표현하고자 한다. 다음 보기를 살펴보고, 계층형 구조로 나타내 보자.

이동 규칙 왼쪽의 가장 위쪽이 비어 있으므로 2를 왼쪽으로 이동하거나 3을 위로 이동할 수 있다.

[보기]

	2	4
3	5	8
1	7	6

퍼즐의 상태	계층형 구조로 표현하기	퍼즐의 상태	계층형 구조로 표현하기
2 ☐ 4 3 5 8 1 7 6		2 5 4 3 ☐ 8 1 7 6	

● 만약 위의 보기에 제시된 조각퍼즐 중 3을 위로 이동시키고 나면 그 다음의 계층형 구조는 어떻게 나타내야 할까?

퍼즐의 상태	계층형 구조로 표현하기
☐ 2 4 3 5 8 1 7 6	