



자료와 정보

대단원 개관

이 단위에서는 정보 기술을 활용하여 정보를 효율적으로 관리하고 생산하는 능력과 태도를 함양하는 데 중점을 두고 있다. 자료와 정보를 효율적으로 처리하기 위해 디지털 정보의 특성을 이해하고 실생활에 존재하는 다양한 형태의 자료와 정보를 디지털로 표현해 보도록 한다. 또한 문제 해결에 필요한 자료를 컴퓨팅 도구를 활용하여 수집, 관리하고 정보를 효과적으로 전달하기 위해 구조화하는 데 중점을 둔다.

대단원 학습 목표

- 자료와 정보의 차이를 이해하고 정보 처리 과정 및 다양한 자료의 유형을 설명할 수 있다.
- 컴퓨터의 자료 표현 원리를 이해하고, 정보의 이진 표현을 설명할 수 있다.
- 정보 기술을 활용하여 정보를 효율적으로 관리하고 생산하는 능력과 태도를 기른다.
- 정보 기술을 활용하여 문제 해결에 필요한 자료와 정보를 수집하고 효율적으로 구조화하는 능력과 태도를 기른다.

핵심 역량

- 1 정보 기술 활용 능력 - 아날로그 및 디지털 자료 수집, 자료 분류, 자료 관리 능력
- 2 추상화 - 자료를 표현하기 위한 표, 다이어그램, 그래프 표현 능력
- 3 공유와 협업 능력 - 자료 공유, 정보의 구조화 능력

대단원 지도 계획

중단원	소단원	차시	교과서 쪽수	학습 주제
Ⅱ- 1장 자료와 정보의 표현	1-1 자료의 유형	6	p. 38~p. 41	<ul style="list-style-type: none"> • 자료와 정보의 차이 • 정보 처리 • 자료의 유형
	1-2 정보의 디지털 표현	7~8	p. 42~p. 54	<ul style="list-style-type: none"> • 아날로그와 디지털 • 컴퓨터에서의 자료 표현 • 정보의 이진 표현
Ⅱ- 2장 자료와 정보의 분석	2-1 자료의 수집	9~10	p. 58~p. 63	<ul style="list-style-type: none"> • 자료 수집의 필요성 • 자료의 수집과 가공 • 자료를 이용한 문제 해결
	2-2 정보의 구조화	11~12	p. 64~p. 70	<ul style="list-style-type: none"> • 정보 구조의 개념 • 정보 구조화의 방법

6차시

소단원 지도 계획

1-1 자료의 유형

대단원 ▶ II 자료와 정보

중단원 ▶ 1 자료와 정보의 표현

핵심 역량 ▶ 정보 기술 활용 능력, 추상화, 공유와 협업 능력

수업 시간 ▶ 1시간

교과서 쪽수 ▶ 38~41쪽

학습 목표 ▶ 자료와 정보의 개념을 이해하고 차이점을 구분할 수 있다.

수업 방법 ▶ 강의, 조사, 토론, 발표

기능 ▶ 분석하기, 표현하기, 수집하기, 관리하기

단계	교수·학습 활동	학습 자료	지도상의 유의점
도입	<ul style="list-style-type: none"> 교과서 37쪽 동기 유발 자료를 통해 자료와 정보의 의미와 관계에 대해 이야기를 나누며 일상생활에서의 다양한 예시 자료를 찾아본다. 학습 문제와 학습 활동을 확인한다. 	교과서 37쪽 그림 자료	<ul style="list-style-type: none"> 자료가 모여 정보가 되었을 때 사람들에게 제 공해 줄 수 있는 유익한 점에 대해 이야기를 나누어 볼 수 있도록 한다.
전개	1 자료와 정보의 차이 <ul style="list-style-type: none"> 정보 사회에서 중요하게 다루는 자료와 정보의 의미와 차이점에 대해 알아본다. 교과서 38쪽 그림 자료를 보며 자료와 정보의 차이점에 대해 살펴본다. 일상생활에서 자료를 유용한 정보로 활용하는 사례와 문제 해결에 필요한 자료에 대해 함께 생각하고 발표 해 본다. [함께 해 보기]에서 제시된 문항을 바탕으로 생각해 본다. 	교과서 38쪽 그림 자료	<ul style="list-style-type: none"> [함께 해 보기] 활동은 모둠별 과제로 제시해 주 어 조사 활동을 통해 발표해 보도록 하는 것도 좋다.
	2 정보 처리 <ul style="list-style-type: none"> 컴퓨터를 활용한 정보 처리 과정에 대해 알아본다. 교과서 40쪽에 제시된 그림을 보며 컴퓨터를 이용한 정보 처리 과정의 흐름을 이해한다. 소셜 네트워크 서비스에서 나타나는 자료를 모아 유익한 정보로 활용한 사례를 살펴본다. [함께 해 보기] 활동을 활용한다. 	교과서 40쪽 그림 자료	
	3 자료의 유형 <ul style="list-style-type: none"> 정보를 표현하기 위한 다양한 자료의 형태에 대해 이 해한다. 교과서 41쪽에 있는 그림을 보며 일상생활에서 사용 하고 있는 자료의 형태에 대해 생각해 본다. 	교과서 41쪽 그림 자료	
정리	<ul style="list-style-type: none"> 자료와 정보의 의미와 차이, 컴퓨터를 활용한 정보 처리 과정에 대해 다시 한 번 확인하며 정리한다. 	정리 및 평가	
평가	<ul style="list-style-type: none"> 자료와 정보는 어떤 차이점을 갖고 있는가? 컴퓨터를 활용한 정보 처리 과정은 어떤 특징을 갖고 있는가? 정보를 표현하기 위한 자료의 유형에는 어떤 것들이 있는가? 	자기 평가지 동료 평가지 관찰 평가지	

7~8차시

소단원 지도 계획

1-2 정보의 디지털 표현

대단원 ▶ II 자료와 정보

중단원 ▶ 1 자료와 정보의 표현

핵심 역량 ▶ 정보 기술 활용 능력, 추상화, 공유와 협업 능력

수업 시간 ▶ 2시간

교과서 쪽수 ▶ 42~54쪽

학습 목표 ▶ 아날로그와 디지털의 차이를 알고, 컴퓨터의 자료 표현 원리를 설명할 수 있다.
정보의 이진 표현이 어떻게 이루어지는지 설명할 수 있다.

수업 방법 ▶ 강의, 조사, 토론, 발표, 실습

기능 ▶ 분석하기, 표현하기, 수집하기, 관리하기

단계	교수·학습 활동	학습 자료	지도상의 유의점
도입	<ul style="list-style-type: none"> 아날로그 정보와 디지털 정보의 차이점을 알아보고 컴퓨터에 입력되는 다양한 정보의 표현 방법에 대해 알아보자. 학습 문제와 학습 활동을 확인한다. 	교과서 42쪽 그림 자료	
전개	<ul style="list-style-type: none"> 아날로그와 디지털 아날로그와 디지털의 의미와 차이점에 대해 알아본다. 	교과서 42쪽 그림 자료	<ul style="list-style-type: none"> 아날로그 정보와 디지털 정보의 차이를 분석 하게 한다.
	<ul style="list-style-type: none"> 컴퓨터에서의 자료 표현 숫자, 문자, 그림, 소리 등 아날로그 자료가 디지털로 변환되어 컴퓨터 내부에서 처리되는 과정을 이해한다. 	교과서 43쪽 그림 자료	<ul style="list-style-type: none"> 컴퓨터 기기의 특징에 대한 이해를 바탕으로 디지털 과정을 이해할 수 있도록 한다.
	<ul style="list-style-type: none"> 정보의 이진 표현 십진법과 이진법에 대해 이해하고 십진수를 이진수 로, 이진수를 십진수로 변환하는 과정을 알아본다. 문자 정보, 그림 정보, 소리 정보, 동영상 정보의 표현 과정과 압축의 필요성과 방법에 대해 알아본다. 	지도서 81쪽	지도서 [추가 선택 활동] <ul style="list-style-type: none"> 이진수 카드놀이를 통해 십진수를 이진수로 변환하는 과정을 자연스럽게 이해할 수 있다. [함께 해 보기] 오류 카드 찾기 놀이하기, 그림 표현과 압축, 언플러그드 활동을 통해 다양한 정보의 표현 과정에 대해 이해해 본다.
	[실습 활동] – 개별 실습 <ul style="list-style-type: none"> 이진수 나타내기 그림 편집 프로그램을 이용하여 BMP 형식의 파일을 JPG 파일로 변환하기 목소리를 녹음하여 저장하고 MP3 파일로 변환하기 동영상 편집 프로그램으로 동영상 파일 만들기 애니메이션 제작 앱을 활용하여 동영상 만들기 [읽기 자료] <ul style="list-style-type: none"> 컴퓨터 과학의 아버지인 앨런 튜링 이야기 [토론하기] <ul style="list-style-type: none"> ‘정보의 압축이 반드시 좋다고 할 수 있을까?’에 대해 생각하고 토론해 본다. 	그림 편집 프로그램, 동영상 편집 프로그램, 애니메이션 제작 모바일 앱	<ul style="list-style-type: none"> 제시된 실습 과제를 학생별로 선택하여 해결 해 볼 수 있도록 한다.
정리	<ul style="list-style-type: none"> 아날로그와 디지털의 의미와 변환 과정에 대해 다시 한 번 확인하며 정리한다. 		
평가	<ul style="list-style-type: none"> 아날로그와 디지털에 대해 이해하고 있는가? 아날로그 자료가 디지털로 표현되는 과정을 이해하고 있는가? 각 정보의 이진 표현에 대해 알고 있는가? 	자기 평가지 동료 평가지 관찰 평가지	

2-1 자료의 수집

대단원 ▶ II 자료와 정보

중단원 ▶ 2 자료와 정보의 분석

핵심 역량 ▶ 정보 기술 활용 능력, 추상화, 공유와 협업 능력

수업 시간 ▶ 2시간

교과서 쪽수 ▶ 58~63쪽

학습 목표 ▶ 자료 수집의 필요성을 알고 문제 해결에 활용할 수 있다.

수업 방법 ▶ 강의, 조사, 토론, 발표, 실습

기능 ▶ 분석하기, 표현하기, 수집하기, 관리하기

단계	교수·학습 활동	학습 자료	지도상의 유의점
도입	<ul style="list-style-type: none"> 교과서 58쪽에 나타나 있는 날씨 자료를 활용하는 실생활의 예를 살펴보고 동기를 유발한다. 학습 문제와 학습 활동을 확인한다. 	교과서 58쪽 그림 자료	<ul style="list-style-type: none"> 학생들이 직접 날씨 앱을 활용하여 관련 자료를 찾아보도록 한다.
전개	1 자료 수집의 필요성 <ul style="list-style-type: none"> 정보를 얻기 위한 자료 수집의 필요성에 대해 이해한다. 실세계의 자료를 컴퓨터를 활용하여 문제 해결에 적용한 사례를 살펴본다. 	교과서 59쪽 그림 자료	
	2 자료의 수집과 가공 <ul style="list-style-type: none"> 실세계와 디지털 세계에서 자료를 수집하고 가공하는 과정에 대해 이해하고, 빅 데이터에 대해 알아본다. 	교과서 60쪽 그림 자료	<ul style="list-style-type: none"> 자료가 넘쳐나는 정보 사회에서 빅 데이터의 중요성과 가치에 대해 함께 알아본다.
	3 자료를 이용한 문제 해결 <ul style="list-style-type: none"> 수집된 자료를 이용하여 문제를 해결하는 과정에 대해 알아본다. 	교과서 62쪽 그림 자료	<ul style="list-style-type: none"> 실생활의 다양한 문제 해결 예시 자료를 통해 이해할 수 있도록 한다.
	[실습 활동] - 모둠별 실습 <ul style="list-style-type: none"> 온라인 설문을 활용하여 자료를 수집하고, 수집한 자료를 바탕으로 문제를 해결해 본다. 다양한 주제를 선정할 수 있으며, 수집한 자료를 가공, 분석한 후 반드시 발견된 문제점과 개선 방안에 대해 토론해 볼 수 있도록 한다. 보고서로 작성하여 발표하도록 한다. 	스마트폰 컴퓨터 온라인 설문 도구 보고서	<ul style="list-style-type: none"> 수집된 자료를 표, 다이어그램 등 다양한 시각적 형태로 구조화하여 표현하도록 하여 정보 활용 목적에 효과적이지 토론해 보도록 한다.
정리	<ul style="list-style-type: none"> 자료 수집의 필요성을 알고 문제 해결에 활용할 수 있음을 확인한다. 	교과서 62쪽 그림 자료	<ul style="list-style-type: none"> 수집된 자료를 표, 다이어그램 등 다양한 시각적 형태로 구조화하여 표현하도록 하여 정보 활용 목적에 효과적이지 토론해 보도록 한다.
평가	<ul style="list-style-type: none"> 정보를 얻기 위한 자료 수집의 필요성에 대해 이해하는가? 자료의 수집과 가공에 대해 알고 있는가? 자료를 이용한 문제 해결 과정에 대해 이해하고 실생활 문제 해결에 적용할 수 있는가? 		

2-2 정보의 구조화

대단원 ▶ II 자료와 정보

중단원 ▶ 2 자료와 정보의 분석

핵심 역량 ▶ 정보 기술 활용 능력, 추상화, 공유와 협업 능력

수업 시간 ▶ 2시간

교과서 쪽수 ▶ 64~70쪽

학습 목표 ▶ 정보 구조의 개념을 설명할 수 있다.

정보를 구조화하는 방법의 종류를 사례를 들어 설명할 수 있다.

수업 방법 ▶ 강의, 조사, 토론, 발표, 실습

기능 ▶ 분석하기, 표현하기, 수집하기, 관리하기

단계	교수·학습 활동	학습 자료	지도상의 유의점
도입	<ul style="list-style-type: none"> 교과서 57쪽의 도입 만화를 보며 정보 구조의 필요성에 대해 함께 이야기해 본다. 학습 문제와 학습 활동을 확인한다. 	교과서 57쪽 그림 자료	<ul style="list-style-type: none"> 정보 구조의 필요성에 대해서는 실생활에서 찾아보도록 한다.
전개	1 정보 구조의 개념 <ul style="list-style-type: none"> 정보 구조의 개념과 필요성에 대해 알아본다. 선형 구조와 비선형 구조에 대해 알아본다. 	교과서 64쪽 그림 자료	
	2 정보 구조화 방법 및 사례 <ul style="list-style-type: none"> 리스트, 스택, 큐, 배열, 계층, 그래프의 개념에 관해 일상생활에서의 예와 컴퓨터에서 활용되는 예를 통해 알아본다. 	교과서 66쪽 그림 자료	<ul style="list-style-type: none"> 정보 구조화의 각 방법에 대해 학생들이 이해할 수 있는 수준에서의 일상생활 예와 컴퓨터에서 활용되는 예를 설명해준다.
	[실습 활동] - 모둠별 실습 <ul style="list-style-type: none"> 다양한 응용 소프트웨어(검색엔진, 워드프로세서, 엑셀, 파워포인트 등)를 활용하여 주제 하나를 선택 후 정보를 구조화해 본다. 필요한 자료를 확인하고 다양한 시각적 형태로 구조화하여 표현하도록 한다. 보고서로 작성하여 발표하도록 한다. [토론하기] <ul style="list-style-type: none"> 6단계 이론(Small World)이 맞는지 '그래프' 구조를 이용하여 토론해 본다. [단원 평가]	다양한 응용 소프트웨어	<ul style="list-style-type: none"> 보고서를 작성하는 데 일정 기간이 필요하므로 단원이 끝나기 전에 미리 공지하여 준다. 인터넷 검색을 활용한 자료의 수집, 응용 소프트웨어를 활용한 자료의 분류, 관리, 공유를 수행하도록 하고 이러한 과정을 보고서로 작성해 보도록 한다.
정리	<ul style="list-style-type: none"> 자료 수집의 필요성을 알고 문제 해결에 활용할 수 있음을 확인한다. 		
평가	<ul style="list-style-type: none"> 정보 구조의 개념과 필요성에 대해 이해하는가? 다양한 정보 구조화 방법에 대해 알고 있는가? 정보의 구조화 과정을 이해하여 정보를 효과적으로 전달할 수 있는 형태로 구조화하였는가? 	자기 평가지 동료 평가지 관찰 평가지	

- 생활 속에서 두 가지 형태로 정보를 표현하고 전달하는 예를 찾아보며 컴퓨터에 대해 이해해 본다.

‘생활 속의 이진수와 디지털 정보’는 이진수로 나타낼 수 있는 일상생활 속의 예를 찾아봄으로써 0과 1 두 가지 수로 정보를 표현하고 전달하는 디지털 기기인 컴퓨터를 이해하기 위한 활동이라고 할 수 있다. 두 가지 형태의 깜빡임, 두 가지 모음, 두 가지 다른 것들은 적절히 조합되어 모든 형태의 정보를 전달할 수 있다. 이러한 표현 방법을 통하여 디지털의 속성에 익숙해지도록 하는 것이 컴퓨터 하드웨어와 소프트웨어에 숨겨진 언어와 내부 구조를 깊이 이해하는 데 기초가 될 수 있다.

컴퓨터에서 숫자, 소리, 음악, 그림, 영화 등을 저장하고 통신하기 위하여 다양한 부호가 사용된다. 컴퓨터는 인간이 지닌 눈, 귀, 입, 손가락과 같은 기관을 정확히 복제하여 가지고 있을 수 없기

때문에 인간이 사용하는 부호를 직접 사용하는 것은 불가능하다. 그럼에도 기술의 발전에 따라 개인용 데스크톱 컴퓨터에서 점차 사람이 의사소통에 사용하는 모든 형태의 정보를 취득하고, 저장하고, 가공하고, 표현하는 것이 가능해지고 있으며, 시각적인 표현(글자나 그림), 청각적 표현(말, 소리, 음악) 혹은 그 조합(애니메이션과 영화)이 이용되고 있다. 말을 사용할 때 입과 귀라는 인간의 기관이 필요하고 글을 사용하기 위하여 손과 눈이라는 다른 기관이 필요한 것처럼, 모든 형태의 정보에는 각각 적합한 고유의 부호가 요구된다.

출처: 'CODE', 찰스 페츨드, 인사이트

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.

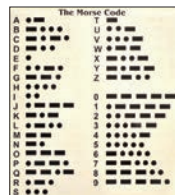
생활 속의 이진수와 디지털 정보

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)



집시 외출
(급방 들어온다는 뜻)



아웃마을에 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.

• 자료(資料, data, 데이터)

- ▶ 자료는 수, 영상, 단어 등의 형태로 된 의미 단위이다. 보통 연구나 조사 등의 바탕이 되는 재료를 말하며, 자료를 의미 있게 정리하면 정보가 된다.
- ▶ 질적 자료(quantitative data, 정성적 자료)는 수치로 측정이 불가능한 자료이다. 분류 자료 또는 범주형 자료(categorical data)라고도 한다.
 - 예 전화번호, 성별, 혈액형, 계급, 순위, 등급, 종교 분류 등
- ▶ 양적 자료(quantitative data, 정량적 자료)는 수치로 측정이 가능한 자료이다. 수치적 자료(numerical data)라고도 한다.

예 온도, 지능 지수, 절대 온도, 가격, 주가 지수, 실업률, 매출액, 기업 내 과장의 수 등.

• 정보(情報, information, info.)

특정 목적을 위하여 광(光) 또는 전자적 방식으로 처리되어 부호, 문자, 음성, 음향 및 영상 등으로 표현된 모든 종류의 자료 또는 지식을 말한다. 정보는 일상용어에서 전문 용어까지 다양한 뜻으로 사용된다. 이를테면 언어, 화폐, 법률, 자연환경 속의 빛이나 소리, 신경, 호르몬 등의 생체 신호부터 비롯한 모든 것을 정보라고 할 수 있다.

출처: 위키백과

자료의 유형

1-1

핵심질문

정보 사회에서의 자료와 정보가 중요한 이유는 무엇일까? 그 둘의 차이점은 무엇일까?

- 자료에 의미와 가치를 부여하면 정보가 되고, 이 정보를 활용하고 통합하여 의미를 부여하여 자신의 것으로 만들면 지식이 되는 과정을 더하여 설명한다.
- 자료가 모여 정보가 되었을 때 사람들에게 제공해 줄 수 있는 유익한 점에 대해 이야기 나누어 볼 수 있도록 한다.

참고자료

• 자료와 정보의 차이



자료가 정보가 되기도 하지만 정보가 또다시 다른 정보를 만들기 위한 자료가 되기도 한다. 이러한 관계를 위의 그림과 같은 예시 자료를 통해 이해할 수 있도록 한다.

• 용어 정리

- **정보 사회**: 정보를 가공, 처리, 유통하는 활동이 활발하여 사회 및 경제의 중심이 되는 사회를 말한다.



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.

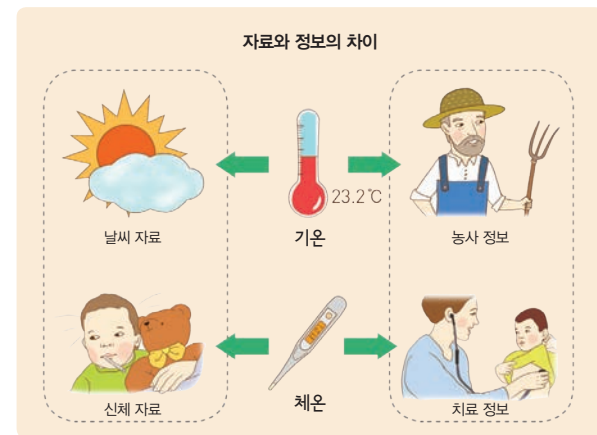
• 정보와 지식의 차이점

지식이란 정보를 활용하는 사람이 가지고 있는 기존의 정보를 통합하여 정보를 해석하고 의미를 부여하여 자신의 것으로 만드는 것을 의미한다.

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.



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

38

II 장 자료와 정보

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.



II-1장 자료와 정보의 표현

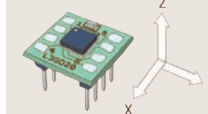
39

• 불쾌지수

불쾌지수란 기온과 습도의 조합으로 사람이 느끼는 온도를 표현한 것이다.

• 가속도 센서

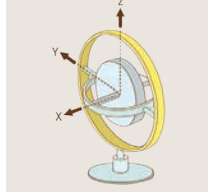
이동하는 물체의 가속도나 충격의 세기를 측정하는 센서이며, x축, y축, z축 세 방향의 단위 시간당 속도를 측정한다.



▲ 가속도 센서

• 자이로 센서

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



▲ 자이로 센서

함께 해 보기 예시 답안

1 시간, 위도, 경도, 속도로 내비게이션 활용. 길을 찾고, 목적지까지 걸리는 시간을 예측할 수 있다.

- 2
- 기온, 습도
 - 위도, 경도
 - X축, Y축, Z축 값

• 이 외에도 일상생활에서의 다양한 자료와 정보를 활용하고 있는 예시를 학생들이 찾아보도록 한다.

• **함께 해 보기** 활동은 모둠별 과제로 제시해 주어 조사 활동을 통해 발표 해 보도록 하는 것도 좋다.

핵심질문

컴퓨터를 이용한 정보 처리 과정은 어떻게 이루어지는가?

• 자료 처리와 정보 처리

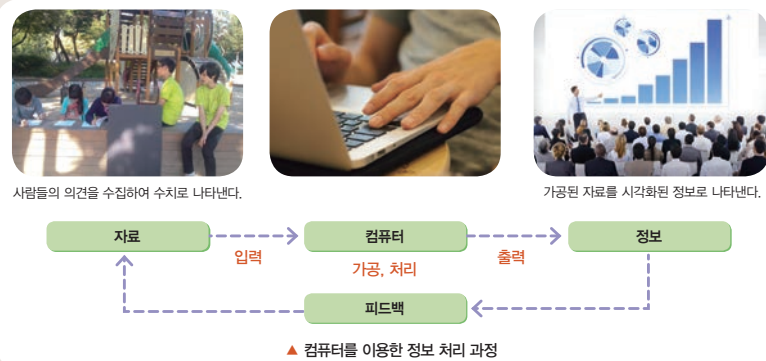
자료 처리는 자료를 정보나 지식으로 바꾸는 컴퓨터 처리를 일컫는 말이며, 정보 처리는 자료 처리 과정을 포함하는 좀 더 넓은 의미이다.

• 컴퓨터를 이용한 자료 처리의 특성

- 신속성: 자료 처리 속도가 빠름
- 신뢰성(정확성): 자료의 처리 결과가 정확함
- 공유성: 시간과 공간의 제약 없이 많은 사람과 통신망을 통해 공유
- 대용량성: 대용량의 자료를 처리
- 자동성: 컴퓨터 프로그램에 의해 자동으로 처리
- 전달성: 시간과 공간의 제약 없이 먼 거리까지 전달

함께 해 보기 예시 답안

- 구글과 MIT, 하버드가 만나 만든 전 세계 질병 지도 'Health Map(<http://www.healthmap.org/en>)' (구글 지도에 구글 뉴스의 정보와 의학 저널, 지역 정보 등을 매칭시키고 WHO 정보까지 가미해서 전 세계 질병 현황을 한눈에 볼 수 있음)
- 소셜 네트워크를 통해 영화 평점, 교우 관계, 최신 트렌드 등을 비교하고 예측하여 정보로 활용할 수 있다.



구글 검색어로 독감을 예측한 사례

지난 2008년 구글은 '구글 플루 트렌드'(Google Flu 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(LI) 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.

40

II 자료와 정보



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.



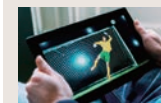
▲ 문자 자료의 예



▲ 그림 자료의 예



▲ 소리 자료의 예



▲ 동영상 자료의 예

II-1장 자료와 정보의 표현

41

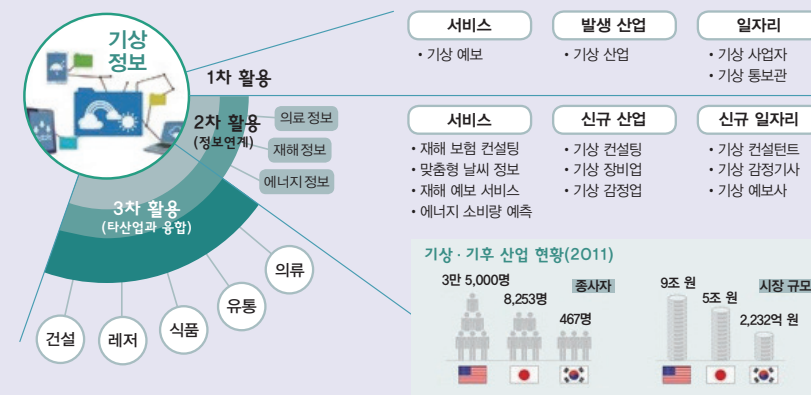
참고자료

구글은 자사 검색 데이터를 활용해 독감 유행 수준을 예측한다. 독감이 유행하면 관련 검색어가 많아지는 등 독감 관련 주제를 검색하는 사람 수와 실제로 독감 증상이 있는 사람 수 사이에 밀접한 관계가 있다고 보았기 때문이다. 실제로 미국 질병통제예방센터 데이터와 비교해 본 결과 구글의 예측은 신뢰성이 높게 나왔다. 이 같은 관계를 바탕으로 구글은 2008년부터 '독감 트렌드'라는 독감 예측 서비스를 선보였다. 그러나 구글도 틀릴 때가 있다. 구글은 순전히 사용자의 검색 데이터를 바탕으로 독감 유행 정도를 파악한다. 이 중엔 정말 독감과 비슷한 증상 때문에 독감 관련 검색어를 입력하는 사용자도 있겠지만, 단순한 호기심 때문에 독감

관련 검색어를 입력하는 사람도 있기 때문이다. 대용량 데이터를 처리하고 분석하는 데 있어서 무조건적인 맹신은 조심하여야 한다.

출처: <http://www.biater.net/archives/143856>

공공 데이터 이렇게 활용한다



공공 데이터 포털 사이트 (www.data.go.kr)

정부가 보유한 다양한 공공 데이터를 개방하여 누구나 편리하고 손쉽게 활용할 수 있도록 하고 있다. 왼쪽 그림은 기상 정보의 활용 예시이다. 이처럼 공공 데이터를 활용하여 해결할 수 있는 문제나 새로이 창출되는 일자리를 알아볼 수 있다.

정보의 디지털 유형

1-2

핵심질문

아날로그와 디지털은 무엇이며, 그 둘은 어떤 차이점이 있을까?

- 아날로그와 디지털을 비교하며 서로의 차이를 분석하게 한다.
- 디지로그(digilog): 디지털(digital)과 아날로그(analog)를 결합한 말이다. 본격적인 디지털 시대가 오면 모든 것이 디지털 방식으로 변화할 것 같지만, 실제로는 그렇지 않다는 것이다. 인간적이고 부드러운 아날로그 방식이 그대로 유지되면서 디지털과 아날로그 방식 간의 융합이 이루어지는 것을 말한다.

용어 정리

• **아날로그(analogue)**: 어떤 자료를 '길이', '각도' 또는 '전류'와 같이 외부적인 원인에 의해 연속적으로 변하는 것들을 물리량으로 나타내는 일. 자동차의 속도를 바늘의 각도로 표시해 주는 속도 측정계, 수은주의 길이로 온도를 나타내는 온도계, 상대적으로 얇게 패이거나 깊게 패인 여러 홈과 바늘의 마찰로 인해 녹음된 소리가 나오는 음반(LP)이 아날로그의 예이다. 디지털에 대비되어 쓰인다. 우리가 거시적인 자연에서 얻는 신호는 대개 아날로그이다. 이를테면 빛의 밝기, 소리의 높낮이나 크기, 바람의 세기 등이 있다.

• **디지털(digital)**: 아날로그의 반대되는 개념으로 자료를 연속적인 실수가 아닌, 특정한 최소 단위를 갖는 이산적(離散的)인 수치를 이용하여 처리하는 방법을 말한다. 이 용어는 손가락을 뜻하는 라틴어 낱말 digit에서 나온 것으로, 숫자를 세는 데 쓰인다.

출처: <https://ko.wikipedia.org/wiki>



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.

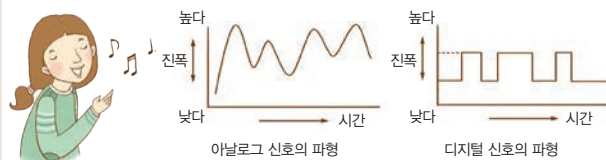
1 아날로그와 디지털

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.

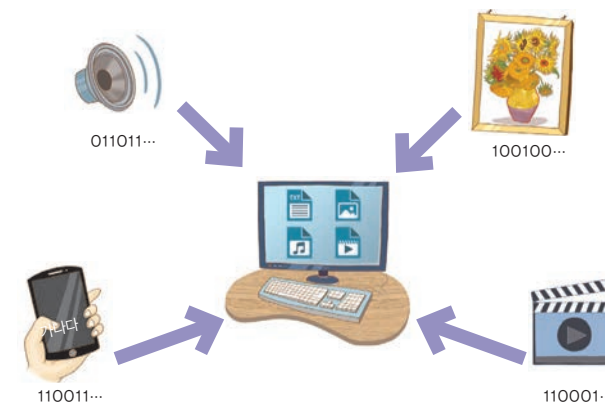


▲ 아날로그 신호와 디지털 신호

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.

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

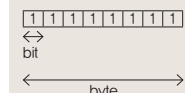
2비트	상태	2진수
⬜	⬜⬜	00
	⬜⬜	01
	⬜⬜	10
	⬜⬜	11

• **유니코드(Unicode)**
바이트 2개(16비트)로 구성되어 총 65,536(2¹⁶)개의 문자를 표현할 수 있다.

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



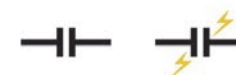
• **바이트(byte)**
비트 8개를 모은 것으로 하나의 문자를 표현하는 단위로 사용된다.



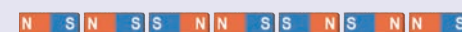
핵심질문

컴퓨터에서는 문자, 숫자, 소리, 그림, 동영상 등의 정보를 어떤 형태로 표현하는가? 그리고 왜 그런 형태로 표현하는가?

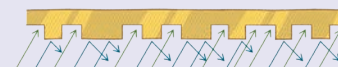
- 컴퓨터 기기의 특징에 대한 이해를 바탕으로 디지털 과정을 이해할 수 있도록 한다.
- 비트가 많아질수록 표현할 수 있는 정보의 양도 많아진다는 것을 교과서에 있는 그림처럼 경우의 수를 구해 봄으로써 확인해 볼 수 있다.
- 이진수를 사용하는 컴퓨터: 컴퓨터는 전기를 사용하는 전자 기기이다. 전류의 흐름, 즉 저전압은 0으로, 고전압은 1로 나타낼 수 있기 때문에 0과 1이라는 두 개의 값만을 사용하며 이를 이진수라고 한다. 컴퓨터의 모든 데이터는 이진수로 표현된다. 컴퓨터에서 처리하는 정보는 0과 1의 연속된 수로 표현된다.



• 정보가 전화선이나 무선으로 전송될 때에는 고음의 소리와 저음의 소리로 1과 0을 표시한다. 자기 디스크(하드 디스크나 플로피 디스크)와 테이프에 비트(bit)를 기록할 때에는 코팅된 표면의 자기장 방향으로 나타낸다(N-S극 혹은 S-N극).



오디오 CD, CD-ROM과 DVD 저장 장치는 빛을 반사하느냐 반사하지 않느냐로 비트(bit)를 나타낸다.



컴퓨터가 2개의 값을 사용하는 이유는 이러한 장치가 만들기 쉽기 때문이다. 10개의 반사 단계를 지닌 CD(0~9까지의 수를 표현 가능)를 만들 수도 있었지만 그렇게 만들 경우 복잡하여 아주 비싸고 정밀한 장치가 필요해진다. 또 다른 이유는 컴퓨터가 고전압과 저전압, N극과 S극 등을 통해 0과 1만을 저장할 수 있기 때문이다. '빛난다'와 '빛나지 않는다'라고 쓰는 것보다 '1', '0'으로 쓰는 것이 더 빠르다. 컴퓨터로 나타내는 모든 정보(문서, 그림, 노래, 영상, 숫자, 프로그램과 애플리케이션 등)는 이진수로 구성된다.

출처: CSUnplugged v3.1, computing.or.kr

- 십진수를 이진수로, 이진수를 십진수로 바꾸어 표현할 수 있는가?
- 문자, 그림, 소리, 동영상 정보의 표현 과정과 압축의 필요성, 방법에 대해 설명할 수 있는가?

- 십진수를 이진수로 바꾸어 보는 활동은 교과서에 제시된 계산식으로도 할 수 있지만 '이진수 카드 놀이 활동'으로도 쉽게 이해할 수 있다. 지도서 81쪽에 있는 카드 자료를 출력하여 학생들에게 나누어 주어 활동해 볼 수도 있으며, 여의치 않을 경우 학생들이 작은 종이를 5장으로 나누고 그 안에 동그라미 수를 직접 칠하며 늘어나는 수의 규칙을 발견하면서 이진수 활동을 위한 준비를 할 수도 있다.



참고자료

• 이진수를 팔진수와 십육진수로 변환하기

컴퓨터는 내부적으로 모든 자료를 이진수로 표현하여 처리한다. 컴퓨터가 처리하는 2진수를 10진수로 변환하거나 2진수 체계를 직접 사용하는 것은 매우 불편하다. 2진수를 8진수 또는 16진수를 사용하여 보다 간단히 표현할 수 있다. 2진수가 [0, 1], 10진수가 [0, 1, 2, 3, 4, 5, 6, 7, 8, 9]들의 숫자로 표현되는 것처럼 8진수는 [0, 1, 2, 3, 4, 5, 6, 7], 16진수는 [0, 1, 2, 3, 4, 5, 6, 7, 8, 9, a, b, c, d, e, f]를 사용하여 표현된다. 기호 [a, b, c, d, e, f] (또는 [A, B, C, D, E, F])는 10진수의 [10, 11, 12, 13, 14, 15]에 해당한다.

이진수는 오른쪽 끝 숫자부터 시작하여 3자리씩 묶어 8진수로, 4자리씩 묶어 16진수로 간단히 변환할 수 있다.

- 8진수는 0에서 7까지의 8개 숫자로 나타낼 수 있다.

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.

$$\begin{array}{r} 1976 \\ + 345 \\ \hline 2321 \end{array} \quad \begin{array}{r} 1976 \\ + 345 \\ \hline 2321 \end{array} \quad \begin{array}{r} 1976 \\ + 345 \\ \hline 2321 \end{array} \quad \begin{array}{r} 1976 \\ + 345 \\ \hline 2321 \end{array}$$

▲ 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 $2_{(2)}$ 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.

$$\begin{array}{r} 1011 \\ + 0011 \\ \hline 1001 \end{array} \quad \begin{array}{l} 2^3 \text{의 자리} \rightarrow 1 \\ 2^2 \text{의 자리} \rightarrow 0 \\ 2^1 \text{의 자리} \rightarrow 1 \\ 2^0 \text{의 자리} \rightarrow 1 \end{array}$$

▲ 이진수의 자리 올림

자릿수	자릿값	자릿수 × 자릿값
2 ³ 의 자리	1	(1 × 1)
2 ² 의 자리	0	(0 × 4)
2 ¹ 의 자리	1	(1 × 2)
2 ⁰ 의 자리	1	(1 × 1)

▲ 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

$$\begin{array}{r} 2 \overline{) 109} \\ 2 \overline{) 54} \dots 1 \\ 2 \overline{) 27} \dots 0 \\ 2 \overline{) 13} \dots 1 \\ 2 \overline{) 6} \dots 1 \\ 2 \overline{) 3} \dots 0 \\ 2 \overline{) 1} \dots 1 \end{array}$$

▲ 이진수를 십진수로 바꾸기

$$\begin{array}{r} 109 \\ = 1 \times 2^6 + 0 \times 2^5 + 0 \times 2^4 + 1 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0 \\ = 1 \times 64 + 0 \times 32 + 0 \times 16 + 1 \times 8 + 1 \times 4 + 0 \times 2 + 1 \times 1 \\ = 64 + 8 + 4 + 1 \\ = 77 \end{array}$$

- 16진수는 0에서 9까지의 10개 숫자와 A, B, C, D, E, F를 사용하여 나타낼 수 있다.

10진수	2진수	8진수	16진수
0	0000	0	0
1	0001	1	1
2	0010	2	2
3	0011	3	3
4	0100	4	4
5	0101	5	5
6	0110	6	6
7	0111	7	7

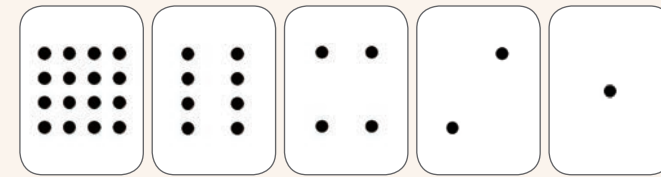
- 진법이 다양한 이유는 프로그래머가 쓰임새에 따라 자료를 편하게 표현할 수 있기 때문이다.

10진수	2진수	8진수	16진수
8	1000	10	8
9	1001	11	9
10	1010	12	A
11	1011	13	B
12	1100	14	C
13	1101	15	D
14	1110	16	E
15	1111	17	F

이진수 카드 놀이 하기

준비물

이진수 카드

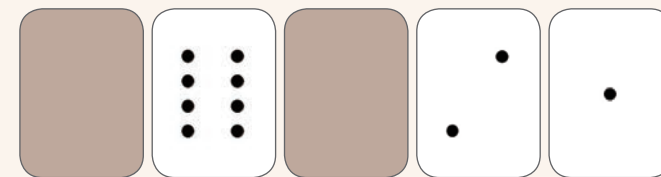


- 1 카드를 오른쪽에서 왼쪽 순으로 갈수록 수의 크기가 바뀌는 규칙을 찾아보자. 점의 개수가 두 배씩 커진다.

- 2 카드에 그려진 점의 개수를 더하여 여러 가지 수를 만들어 보자.

(필요한 점이 있는 카드는 앞면 그대로, 필요하지 않은 카드는 뒤집어 놓는다.)

- 11이라는 수를 나타내기 위해 아래 그림처럼 카드를 놓으면 된다. $[8+2+1=11]$

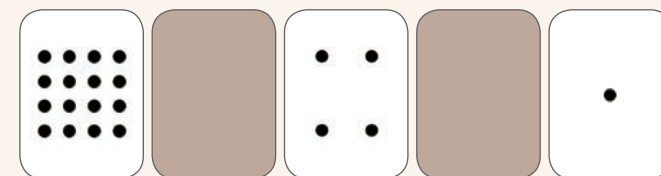


- 위의 카드로 만들 수 있는 가장 작은 수와 가장 큰 수를 생각해 보자.
가장 작은 수 0, 가장 큰 수 31
- 카드를 활용하여 가장 작은 수에서부터 가장 큰 수까지 만들어 보자.
- 카드를 뒤집을 때 나타내는 규칙을 찾아 보자.

[처음에는 오른쪽 카드가 뒤집히고, 두 번째는 바로 왼쪽 옆의 카드가 뒤집히고, 세 번째는 뒤집혔던 카드가 동시에 뒤집힌다. 이것이 자리가 커지면서 반복된다.]

- 3 카드의 앞면을 숫자 1, 뒷면을 0으로 표현해 보자.

- 위의 그림의 카드를 0과 1로 나타내면 '01011'로 나타낼 수 있다.
- 아래 그림을 0과 1로 나타내 보자. $[10101]$



- 친구와 함께 카드를 보여 주며 0과 1로 바꾸는 놀이를 해 보자.

- 4 오른손 다섯 손가락으로 활동을 해 보자.

- 손바닥을 보며 엄지손가락부터 1, 2, 4, 8, 16의 자릿값을 갖는다고 손가락을 펴 가며 십진수를 만들어 보자.
- 만약 왼손으로 자리가 더 늘어난다면 얼마나 더 많은 수를 표현할 수 있을까?

[32, 64, 128, 256, 512, 1024... 두 배씩 커지면서 표현할 수 있는 수가 많아진다.]

• 문자 정보를 표현할 수 있는 아스키코드, 유니코드에서 사용하는 문자의 코드 값은 아래 사이트를 통해 확인해 볼 수 있다. 교과서에 제시된 코드 값을 학생들이 직접 찾아보게 한다.

• <https://ko.wikipedia.org/wiki/미국정보교환표준부호>

• <http://graphemica.com/unicode/characters>



▲ 아스키코드 ▲ 유니코드

- 패리티 비트(parity bit)
 - ▶ 전송하고자 하는 자료의 각 문자(7비트)에 1비트를 더하여(8비트) 전송하는데 2가지 종류의 패리티 비트(홀수, 짝수)를 사용한다.
 - ▶ 짝수(even) 패리티는 전체 비트에서 1의 개수가 짝수가 되도록 패리티 비트를 정하는 것인데, 이를테면 데이터 비트에서 1의 개수가 홀수이면 패리티 비트를 1로 정한다.
 - ▶ 홀수(odd) 패리티는 전체 비트에서 1의 개수가 홀수가 되도록 패리티 비트를 정하는 방법이다.

7비트 데이터	패리티 포함 8비트	
	짝수	홀수
0000000 (0)	00000000	10000000
1010001 (3)	11010001	01010001
1101001 (4)	01101001	11101001
1111111 (7)	11111111	01111111

참/고/자/료

• 문자 이진 코드

서로 다른 언어를 쓰는 사람에게 의사 전달을 하기 위해서는 자신의 언어를 번역하여 상대방이 이해할 수 있는 형태로 바꾸어 전달해야 한다. 마찬가지로 인간이 사용하는 한글, 영문자, 특수 문자 등을 컴퓨터에서 사용하려면, 입력되는 문자를 컴퓨터가 이해할 수 있는 이진수 형태로 바꾸어 주어야 한다.

이때 각 문자를 서로 구분하기 위해 문자별로 약속된 이진수 값을 정해 두었는데, 이를 문자 이진 코드(character binary code)라고 한다.

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 구성되어 총 128(2⁷)개의 서로 다른 문자를 표현할 수 있다. 정보를 전송할 때는 오류 검사를 할 수 있는 1비트를 추가하여 총 8비트로 구성되어 사용한다.

7비트 아스키코드 표					
문자	아스키코드	문자	아스키코드	문자	아스키코드
A	1000001	a	1100001	@	1000000
B	1000010	b	1100010	=	0111101



오류 카드 찾기

1 게임 순서

두 명의 친구(A와 B)와 팀을 만든다.

- A: 양면을 무작위로 섞어 5×5 배열로 배치한다.
- 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.
- A: 배치된 카드 중 하나를 선택하여 뒤집는다(이때 B 학생은 보지 않게 뒤돌아 있다).
- B: A가 뒤집어 놓은 카드를 맞춘다.

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.



▲ 오류 카드 찾기 놀이를 하는 모습



▲ 인간과 컴퓨터의 문자 인식 과정 비교

• 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.

• 8비트 아스키코드

검사 비트	존 비트			디짓 비트			
C	7	6	5	4	3	2	1

100: 영문자 A~Q(0001~1111)
101: 영문자 P~Z(0000~1010)
011: 숫자 0~9(0000~1001)

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(2¹⁶)개의 서로 다른 문자를 표현할 수 있다.

유니코드의 예

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

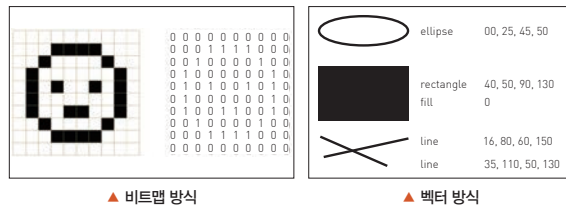
한글 코드는 컴퓨터에서 한글을 사용하기 위하여 개발된 코드로 조합형과 완성형이 있다. 두 가지의 글자가 구현되는 방식에 차이가 있는데, 조합형은 한글의 초성, 중성, 종성에 각각 코드를 부여하여 한글을 표현하며, 완성형은 이미 만들어진 한글 글자에 코드를 부여하여 불러오는 방식이다. 한글을 표현하기 위해 16비트를 부여하며, 우리나라는 유니코드에 규정된 한글 코드를 사용한다.

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

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.



▲ 비트맵 방식

▲ 벡터 방식

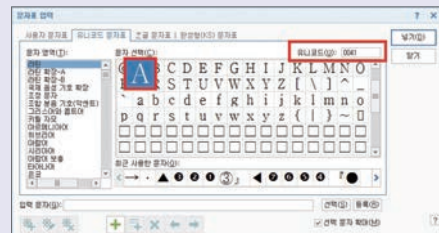
• 아스키코드(ASCII)의 경우 암호 찾기가 어려울 수 있으므로 아스키 85(85진법)를 활용해도 좋다. 만약 이진수로 활동할 경우에는 영어 단어 하나 정도를 제시해 줄 수 있도록 조건을 준다.

• 아스키코드 표

아스키코드(ASCII)는 문자 코드의 국제 표준이다. 0~127개 번호에 각각 문자가 할당되어 있다.
(아래 표에서 No 상단은 10진수, 하단 팔호 안은 16진수로 표기한 것)

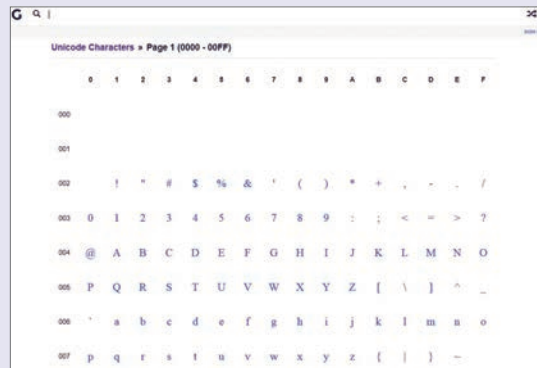
No	문자	No	문자	No	문자	No	문자	No	문자
0 (00)	NUL (null)	26 (1A)	SUB (substitute)	52 (34)	4	78 (4E)	N	104 (68)	h
1 (01)	SOH (start of heading)	27 (1B)	ESC (escape)	53 (35)	5	79 (4F)	O	105 (69)	i
2 (02)	STX (start of text)	28 (1C)	FS (file separator)	54 (36)	6	80 (50)	P	106 (6A)	j
3 (03)	ETX (end of text)	29 (1D)	GS (group separator)	55 (37)	7	81 (51)	Q	107 (6B)	k
4 (04)	EOT (end of transmission)	30 (1E)	RS (record separator)	56 (38)	8	82 (52)	R	108 (6C)	l
5 (05)	ENQ (enquiry)	31 (1F)	US (unit separator)	57 (39)	9	83 (53)	S	109 (6D)	m
6 (06)	ACK (acknowledge)	32 (20)	(Space)	58 (3A)	:	84 (54)	T	110 (6E)	n
7 (07)	BEL (bell)	33 (21)	!	59 (3B)	;	85 (55)	U	111 (6F)	o
8 (08)	BS (backspace)	34 (22)	"	60 (3C)	<	86 (56)	V	112 (70)	p
9 (09)	HT (horizontal tab)	35 (23)	#	61 (3D)	=	87 (57)	W	113 (71)	q
10 (0A)	LF (line feed)	36 (24)	\$	62 (3E)	>	88 (58)	X	114 (72)	r
11 (0B)	VT (vertical tab)	37 (25)	%	63 (3F)	?	89 (59)	Y	115 (73)	s
12 (0C)	FF (form feed)	38 (26)	&	64 (40)	@	90 (5A)	Z	116 (74)	t
13 (0D)	CR (carriage return)	39 (27)	.	65 (41)	A	91 (5B)	[117 (75)	u
14 (0E)	SO (shift out)	40 (28)	(66 (42)	B	92 (5C)	\	118 (76)	v
15 (0F)	SI (shift in)	41 (29))	67 (43)	C	93 (5D)]	119 (77)	w
16 (10)	DLE (data link escape)	42 (2A)	*	68 (44)	D	94 (5E)	^	120 (78)	x
17 (11)	DC1 (device control 1)	43 (2B)	+	69 (45)	E	95 (5F)	-	121 (79)	y
18 (12)	DC2 (device control 2)	44 (2C)	,	70 (46)	F	96 (60)	.	122 (7A)	z
19 (13)	DC3 (device control 3)	45 (2D)	-	71 (47)	G	97 (61)	a	123 (7B)	{
20 (14)	DC4 (device control 4)	46 (2E)	.	72 (48)	H	98 (62)	b	124 (7C)	
21 (15)	NAK (negative acknowledge)	47 (2F)	/	73 (49)	I	99 (63)	c	125 (7D)	}
22 (16)	SYN (synchronous idle)	48 (30)	0	74 (4A)	J	100 (64)	d	126 (7E)	~
23 (17)	ETB (end of trans. block)	49 (31)	1	75 (4B)	K	101 (65)	e	127 (7F)	DEL (delete)
24 (18)	CAN (cancel)	50 (32)	2	76 (4C)	L	102 (66)	f		
25 (19)	EM (end of medium)	51 (33)	3	77 (4D)	M	103 (67)	g		

▲ 유니코드를 사이트(<http://graphemica.com/unicode>)에서 찾기



▲ 16진수 0041이 A이다(아래아한글 문자표에서 찾기).

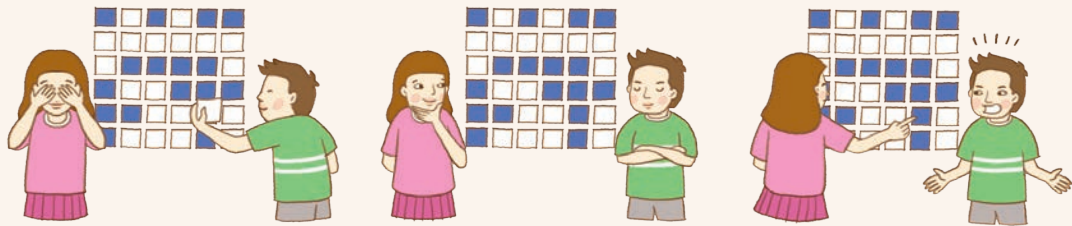
• 유니코드의 경우 코드 값을 제시해 준 사이트 검색 창에서 입력하면 해당 문자가 나오므로 좀 더 긴 문장을 암호로 만들어도 무난하다.



오류 카드 찾기 놀이하기

1 카드 마술하기

- 양면이 무작위로 섞여 있는 5×5의 사각 카드를 배치할 학생 뽑기
- 배치된 카드를 조금 더 복잡하게 만들기 위해 “좀 더 어렵게 만들어 보겠다”며 다른 행이나 열 추가하기
(이때 새로 추가되는 카드들은 속임수이므로 각각의 행과 열에서 색 카드가 짝수가 되도록 맞추기 위해 카드 선택하여 배열하기)



- 마술사인 나는 눈감고, 다른 친구는 하나의 카드 뒤집기
- 뒤집힌 카드 찾기!

2 마술의 속임수 알아보기

- 짝 활동으로 자신의 카드를 배열하기
- 각각의 행과 열에 있는 카드의 수를 짝수로 맞추기 위해 카드 추가 (추가된 카드를 ‘홀짝 맞춤 카드’라고 부름)
- 짝과 서로 순서를 바꿔 가며 마술 놀이하기

3 다양한 활동하기

- 카드 대신 두 가지 상태를 나타낼 수 있는 다른 물건으로 시도하기
(예 0과 1의 숫자 카드, 동전의 앞면 또는 뒷면 등)
- 훨씬 더 큰 카드의 배열로 놀이해 보기

준비물

학생별 양쪽 면이 다른 카드 36장
▶ 학종이 접기를 이용하면 자르는 번거로움 없이 활동할 수 있다.



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.

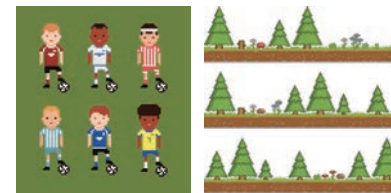
2 짝과 함께 아스키코드 또는 유니코드를 선택하여 코드값만 제시해 주고 과일 이름을 찾는 ‘코드 맞추기’ 놀이를 해 보자.

내가 만든 코드

친구 코드 해석하기

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.



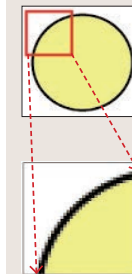
▲ 픽셀 그래픽 이미지

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 수치 정보

로 표현한다. 벡터 방식은 그래픽 프로그램을 이용하여 제품을 디자인하거나 3D 작품을 제작할 때 자주 사용한다.

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



확대된 비트맵 이미지

II-1장 자료와 정보의 표현

47

한글 해 보기 예시 답안

2 과일 이름을 영어(apple)로 할 경우

- 1 아스키코드 활용
(1100001, 1110000, 1110000, 1101100, 1100101)
- 2 과일 이름을 한글(사과)로 할 경우
 - 유니코드 활용
 - 호환용 한글자모(초성, 중성, 종성으로 구분) - (3145, 314F, 3131, 3157, 314F)
 - 한글 글자 마디 - (COAC, ACFC)

• 비트맵 이미지의 다양한 형식과 특징

확장자	압축	용도
JPEG [.jpg]	손실 압축	웹 페이지 큰 사진
GIF [.gif]	비손실 압축	애니메이션, 로고
PNG [.png]	비손실 압축	웹 페이지 작은 사진
BMP [.bmp]		컴퓨터에 저장되는 큰 사진

- 초기 휴대 전화의 화면은 84×48 픽셀로 구성되었지만 오늘날 아이폰 6S 화면은 1,920×1,080 픽셀이다. 학생들이 쓰고 있는 휴대 전화의 화면 해상도와 픽셀 밀도를 확인해 보도록 한다.

• 압축

컴퓨터의 저장 공간은 무한하지 않기 때문에 정보를 효율적으로 표현할 필요가 있다. 이러한 방법을 ‘압축’이라고 한다. 자료를 저장할 때는 인코딩을 하고, 보여 주거나 검색할 때에는 디코딩(압축 해제)을 함으로써 컴퓨터는 더 많은 정보를 저장하고 인터넷에 저장할 수 있다.

• 해상도

화면 해상도는 한 화면에 화소가 몇 개 포함되어 있는지를 의미한다. 대개 가로와 세로의 화소 수를 곱하는 형태로 표현한다. 컴퓨터 화면의 크기와 비디오 메모리의 용량 한계를 기준으로 여러 가지의 화면 해상도를 설정할 수 있다. 예를 들어 800×600의 화면 해상도는 가로 800개와 세로 600개의 화소를 가진 모니터를 나타내며 한 화면의 총 화소 수는 480,000개가 된다.

• 이미지 저장 방식

비트맵 이미지는 다양한 형식이 있다. 가장 자주 사용되는 것으로는 JPEG, BMP, GIF, PNG이 있으며, 이들 형식은 각각 장점과 한계점이 다르다.

확장자	심색도 지원	투명도	애니메이션	압축	파일 크기	용도
JPEG	픽셀당 24비트까지	아니요	아니요	예(손실)	매우 작음 1.19MB	웹 페이지상의 큰 사진
GIF	픽셀당 8비트까지	예	예	예(무손실)	작음 7.86MB	애니메이션 혹은 로고
PNG	픽셀당 48비트까지	예	아니요	예(무손실)	작음 9.94MB	웹 페이지상의 작은 사진
BMP	픽셀당 64비트까지	아니요	아니요	아니요	큼 35.4MB	컴퓨터에 저장되는 사진

▲ 이미지 저장 방식

'무손실'은 파일이 압축되었을 때 이미지 품질이 손상되지 않는 것을 의미한다.

● 압축의 필요성

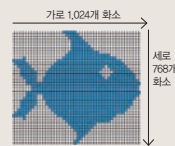
데이터 양이 많으면 저장 장치에 효율적으로 저장하기가 어려우며 통신 네트워크로 전달하기 어렵다.

● 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

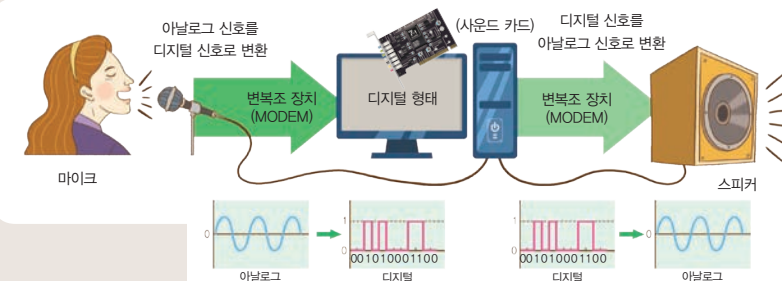
비트맵 방식의 그림은 저장된 파일이 크기 때문에 압축 방법을 사용하여 크기를 줄인다. 확장자가 jpg, gif, png 등의 파일이 대표적인 그림 압축 파일이다. jpg 파일의 경우 압축률이 높아 저장 공간을 줄이고 웹 페이지에서 사용하기 좋다. 하지만 심하게 압축하거나 저장을 반복할 경우 그림의 품질이 떨어지는 단점이 있다.

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.



'손실'은 파일이 압축되었을 때 이미지 품질이 일부 손상되는 것을 의미한다.

출처: 『중등 소프트웨어 교육 컴퓨팅 사고력 키우기』, 제임스 아벨라 외, 한국과학창의재단



그림의 표현과 압축

1 이미지를 이진수로 표현하고 그 수를 간단하게 표현하는 활동을 해보자.

- 비어 있는 흰색 칸은 0으로, 검은색 칸은 1로 나타낸다.
- 홀수 번째는 흰색 픽셀의 개수, 짝수 번째는 검은 픽셀의 개수로 나타내면 간단히 표현할 수 있다.
- 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 0 0 0 1 0 1 1 1 0 1 0 0 0 1 0 1 1 1 0	1(흰), 3(검) 0, 1, 3, 1 1, 3 0, 1, 3, 1 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
	→	
	→	
	→	
	→	
	→	
	→	
	→	
	→	
	→	
	→	

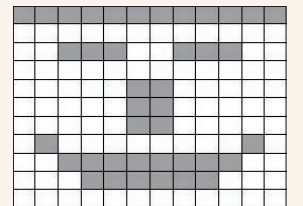
• 아날로그와 디지털 소리

소리는 아날로그 소리와 디지털 소리로 구분된다. 아날로그 소리는 악기나 음성과 같이 자연에서 만들어진 소리이고, 디지털 소리는 컴퓨터에 의해 0과 1로 변환되어 저장된 소리이다. 아날로그 소리는 풍부하고 미묘한 소리까지 표현하지만 연속적인 신호이기 때문에 전달할 때 약간의 잡음에도 신호가 변경되어 손상되는 경우가 많다. 반면 디지털 소리는 아날로그 소리에 비해 음질이 깨끗하고 선명하나 변환 과정에서 원음의 손실이 일부 발생할 수도 있다.

함께 해 보기 예시 답안

2

이미지 나타내기



0과 1로 표현하기

1	1	1	1	1	1	1	1	1	1
0	0	0	0	0	0	0	0	0	0
0	0	1	1	0	0	1	1	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	1	1	0	0	0
0	0	0	0	0	1	1	0	0	0
0	0	0	0	0	1	1	0	0	0
0	1	0	0	0	0	0	0	1	0
0	0	1	1	1	1	1	1	0	0
0	0	0	1	1	1	1	0	0	0

간단히 표현하기

0, 12
12
2, 3, 2, 3
12
5, 2
5, 2
5, 2
1, 1, 8, 1
2, 8
3, 6
12

• 두 번째 함께 해 보기 활동은 짝과 함께 활동할 수 있도록 한다.

- **MP3(MPEG-1 Audio Layer-3)**
가장 잘 알려진 압축 방식이다. 사람의 청각 기관의 특성을 이용하고 있는데 사람의 청각 기관에서 인식할 수 없는 미세한 부분은 버리는 방식을 이용한다. 청각 기관의 특성 중에 '시간적 차폐 효과'라 불리는 것이 있는데 이것은 사람의 청각 기관에서는 큰 소리를 들은 후 짧은 시간 동안 원래는 들려야 할 작은 소리를 듣지 못하는 현상을 말한다. 또 다른 특성인 '주파수 차폐 효과'는 특정 주파수의 소리가 그 주변 주파수의 부드러운 소리들을 가리는 현상을 말한다. MP3는 이러한 특성들을 이용함으로써 소리 품질을 유지하면서 압축을 한다.

출처: 『컴퓨터 과학총론』

● **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


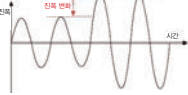
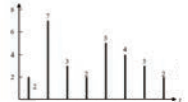


● **프레임(frame)**
프레임은 비디오나 영화, TV 등이 영상 매체를 전달할 때 화면에 보여주는 한 장의 이미지를 말한다.

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

파일 형식	특징
AVI (Audio Visual Interleave(d))	• 원도 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.

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.

값을 양자 값이라고 한다. 아날로그 신호를 4비트로 양자화를 하면 신호의 진폭을 16단계로 표현할 수 있다. 이와 같은 방법으로 16비트로 양자화를 한다면 보다 세밀한 65,536단계까지 값을 표현할 수 있다.

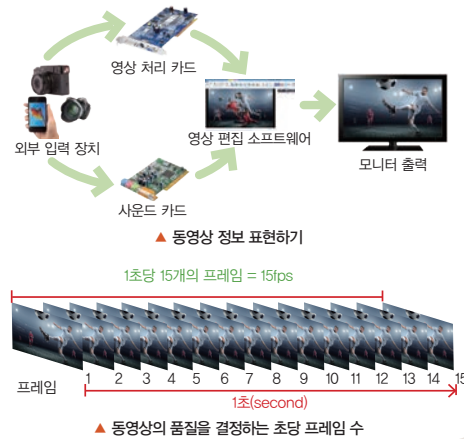
부호화

양자화된 신호들을 0과 1로 이루어진 2진수로 표현하는 것이다. 한 개의 비트로는 0과 1 두 가지 상태를 표현할 수 있으므로 신호의 크기가 4일 경우 이를 4비트로 부호화하면 '0100'으로 표현된다.

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.

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.



▲ 동영상 정보 표현하기

1초당 15개의 프레임 = 15fps

▲ 동영상의 품질을 결정하는 초당 프레임 수

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

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

참고자료

표본화

표본화란 아날로그 신호를 일정한 시간 간격으로 잘라내어 각 시간 간격마다 그 시간에 해당되는 진폭의 크기를 구하는 것이다. 이때 1초 동안 표본화한 개수를 표본화율이라 하며, 단위로 헤르츠(Hz)를 사용한다. 표본화율이 클수록 원음에 가깝게 표현할 수 있으나 처리할 자료의 양이 많아진다.

양자화

표본화에 의해 얻은 표본 값을 디지털화할 수 있는 값으로 처리하는 과정을 말한다. 즉 표본화를 통해 얻어진 표본 값을 숫자화할 때 근접한 정수의 값(반올림)으로 변화시키는 것으로 이때 얻어진

비디오의 압축

일반적으로 비디오 압축 기법은 영화를 필름에 녹화할 때와 같이 비디오를 일련의 그림들로 구성된 것으로 간주하여 처리한다. 일련의 그림을 압축할 때 화상들 중에 I-프레임이라고 불리는 일부 그림에 대해서만 그 전체를 인코딩한다. I-프레임들 사이의 그림들은 상대적 인코딩을 사용한다. 즉 이들에 대해서는 전체 그림 대신 이전 그림과의 차이만이 기록된다. I-프레임들은 대개 JPEG와 유사한 기법을 사용하여 압축된다.

실습 활동

- 1번에서 주어진 비트로 만들 수 있는 이진수를 만들 때 가능한 가짓수만큼의 수를 조합한다고 생각하면 쉽게 만들 수 있다. 예를 들면 8비트일 때 0에서부터 7까지의 수를 만든다고 생각하며 0은 000, 1은 001, 2는 010... 이런 식으로 십진수를 이진수로 표현하는 과정을 적용해 보도록 한다.
- 2번에서 사용하는 사진 또는 그림은 크기가 작아(저품질) 확대했을 때 그 차이를 확실하게 알아볼 수 있다.

실습 활동

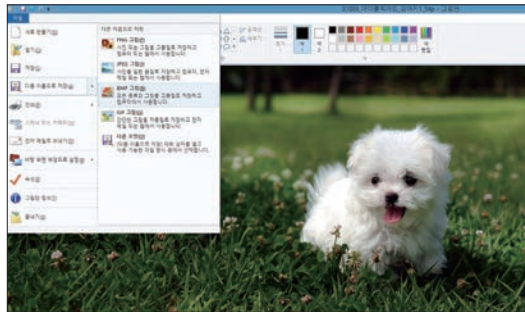
- 1 빈칸에 비트 조합으로 표현 가능한 숫자들을 모두 쓰고 그 개수를 적어 보자. 그리고 아래 물음에 답해 보자.

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

- 숫자 7을 이진수로 나타내어 보자.
- 4비트로 만들 수 있는 가장 큰 숫자는 무엇인가?
- 16가지의 기호를 분류하여 나타내고 싶으면 최소 몇 개의 비트가 필요한가?

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

- JPG 그림 파일을 불러오고, 그림을 선택한 뒤 크게 확장시켜 본다. 그림을 확장할수록 그림이 어떻게 보이는가?
- JPG 그림 파일의 확장자를 BMP로 바꾸어 저장하여 보자. 파일의 크기가 어떻게 달라지는가?



* 참고

- BMP(BitMap)
- JPEG(Joint Photographic Coding Experts Group)

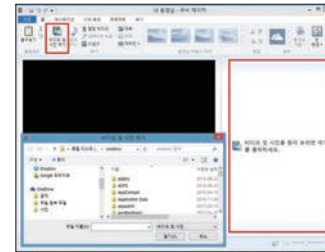
52

II 자료와 정보

- 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.



1 무비 메이커를 실행시킨다.

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

3 메뉴 상단에서 [음악 추가]를 클릭하여 음악을 불러온다.



4 재생 단추를 클릭해서 사진과 음악을 감상한다.

자기 평가

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.			

II-1장 자료와 정보의 표현

53

- 실습 활동 중 동영상을 만들어 보는 활동은 수행 평가 자료로 활용하여도 좋다.

- 3번 소리 파일의 변환은 프로그램이 없을 경우 온라인 변환기 프로그램을 사용하도록 한다.

- 4번 무비 메이커를 활용한 동영상을 만들 때는 주제를 정하여 하나의 완성된 작품을 만들어 볼 수 있도록 한다. 시간이 부족할 경우 사용법을 수업에서 익히고 난 뒤 과제로 제시해 줄 수 있다.

- 자기 평가를 통해 자신의 학습을 스스로 평가하고 진단해 볼 수 있도록 한다.

참고자료

1비트	2비트	3비트	4비트
0, 1	00, 01, (10), (11)	000, 001, 010, 011, 100, 101, 110, 111	0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111
2가지	4가지	(8)가지	(16)가지
1비트 조합으로 표현 가능한 숫자의 총 개수	2비트 조합으로 표현 가능한 숫자의 총 개수	3비트 조합으로 표현 가능한 숫자의 총 개수	4비트 조합으로 표현 가능한 숫자의 총 개수

- 1 ① 0111
② 15 [3비트는 7, 4비트는 15]
③ 4비트

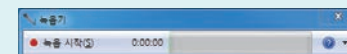
- 2 ① 그림이 흐릿해지면서 픽셀이 드러나며 계단 현상이 나타난다.
② 파일의 크기가 커진다.

- 3 • 파일 크기의 차이: 파일 크기가 줄어든다.

- 음질의 차이: 소리의 선명함이 떨어진다.

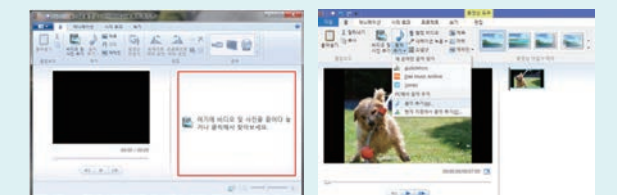
- 1 녹음기에 자신의 목소리를 녹음한다.

- 2 저장된 파일을 소리 변환 프로그램에서 MP3 파일로 변환하여 저장한다. 온라인 오디오 변환기 <http://media.io/> 사이트에서 파일을 업로드하여 변환 후 저장이 가능하다.



- 4 1 무비 메이커를 실행한다. → 2 사진을 아래의 빨간 네모 영역에 마우스로 끌어다 놓거나 [비디오 및 사진 추가] 기능을 활용하여 추가한다. → 3 메뉴 상단에서 [음악 추가]를 클릭하여 음악을

- 불러온다. → 4 재생 단추를 클릭해서 사진과 음악을 감상한다.



- 5 • 스마트폰이나 패드를 활용하여 '애니메이션 제작 앱'을 다운로드한다.

- 초당 프레임 속도(보통 6)를 설정하고 저장하기.
- 동작의 변화를 주면서 그림 그려 넣기.
- 동영상 제작하기.

읽기 자료 & 토론허기

- 앨런 튜링은 컴퓨터 과학의 아버지로 불리며 컴퓨터 과학에 대한 토대를 마련하였다. 교과서에 제시된 읽기 자료 외에 관련된 도서를 찾아 읽어 볼 수 있도록 한다.
- 토론허기에서는 자료의 압축이 무조건 좋다는 생각에서 벗어나 압축의 필요성에 대해 다시 한번 고민해 볼 수 있는 주제이다. 제시된 문제점 외에 압축이 필요하지 않은 예는 무엇이 있는지 찾아보도록 한다.

읽기 자료

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



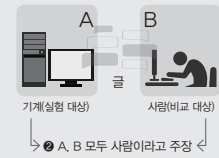
▲ 튜링 머신을 고안한 앨런 튜링

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

- 컴퓨터 화면을 통해 문자로만 대화



기계(실험 대상) 사람(비교 대상)
A, B 모두 사람이라고 주장
심사위원(제3자)
● 어느 쪽이 사람인지 구분할 수 없을 경우 A, B 둘 다 인간 수준의 사고 능력을 가진 것으로 판정



▲ 튜링 머신 1280px 모델

토론허기

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 토론허해 보자.

54

II 자료와 정보

참고자료

• 튜링 테스트(Turing test)

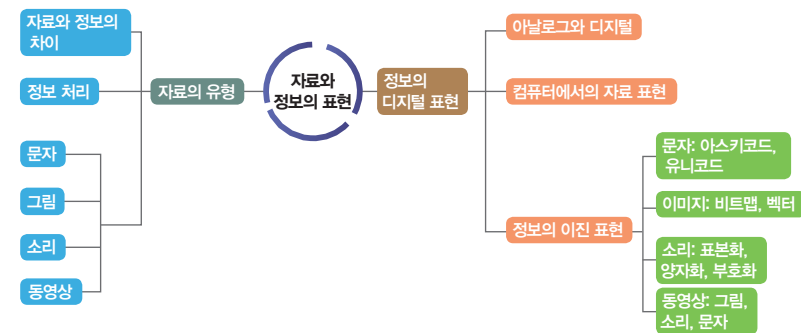
2014년 영국 레딩대는 '유진 구스트만'이 튜링 테스트를 통과하였다고 발표하였다. 인공지능이란 사람처럼 생각하는 컴퓨터 프로그램이다. 인공지능이 튜링 테스트를 통과했다는 말은 더 이상 컴퓨터 프로그램과 사람을 구분할 수 없다는 뜻이다. 하지만 유진 구스트만은 우크라이나에 사는 13세 소년으로 설정하여 대화가 어색하더라도 그냥 넘어갈 수 있도록 한 점이 이미 실험 과정에 있어 한계를 드러냈다는 비판을 받기도 하였다.

사람처럼 생각할 줄 아는 진짜 인공지능이라면 종합적인 판단력이 있어야 한다는 게 전문가의 평가다. 인터넷 서점이 내 구매

목록을 바탕으로 내 취향에 꼭 맞는 신간을 추천해 주거나, 애플 시리가 내 목소리를 알아듣고 친구에게 문자 메시지를 적어 보내주는 것, 운전하는 사람 없이 혼자 도로를 달리는 구글 무인 자동차 등은 얼핏 인공지능처럼 보이지만 진짜 인공지능은 아니다. 모두 특정한 상황에서 주어진 일만 해낼 수 있을 뿐이기 때문이다.

뉴욕대 인지과학자 게리 마르쿠스는 21세기에 걸맞게 튜링 테스트를 판올림해야 한다고 주장했다. "진짜 인공지능은 TV 프로그램이나 유튜브 동영상을 보고 그 내용에 관한 질문에 답할 수 있어야 합니다." 사람처럼 정보를 모아 종합적으로 판단을 내리는 컴퓨터 프로그램이야말로 진정한 인공지능이라는 것이 게리 마르쿠스의

배운 내용 정리하기



평가하기

- 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?

III-1장 자료와 정보의 표현

55

참고자료

배운 내용 정리 & 평가하기

- 배운 내용 정리하기를 통해 학습한 내용을 다시 한 번 되돌아본다.

평가하기 답안

- 1 자료, 정보
- 2 0, 1
- 3 바이트(byte)

자료와 정보의 분석

- 자료를 의미 있는 정보로 만드는 '데이터 과학자'에 대해 함께 생각해 보며 사회의 발달로 새로운 직업이 탄생하는 다른 예들도 함께 생각해 본다.

참/고/자/료

• 지능 정보 사회와 함께 성장하는 새로운 직업

지능 정보 기술(인공지능; AI)과 사물 인터넷·클라우드 컴퓨팅·빅 데이터·모바일(CBM 기술 등)은 온·오프라인에서 수집한 막대한 데이터를 통합·관리하고, 데이터를 분석하여 어떻게 행동할지 판단하며, 판단 결과를 바탕으로 인간이 사물과 손쉽게 소통·제어할 수 있도록 도와준다.

1 감지(sensing) 관련 기술을 개발하고 활용하는 직업이 필요하다.

■ 스마트 센서 개발자

자율 주행 차, 로봇, 드론 등을 악천후나 심야에도 사용할 수 있도록 인식률 높은 고감도 센서(카메라, 라이다, 레이더 등)를 개발.

■ 비전 인식 전문가

자율 주행 차, 로봇 등이 도심이나 가정에서 수신호, 차선 등 각종 영상 데이터를 인식하고 의미를 해석하기 위한 알고리즘 개발.

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.

빅 데이터 분석가?
나는 설문조사 결과 분석 같은 것을 좋아하니까
전선에 막을 것 같아~

데이터 사이언스 시대?
데이터 과학자는 누구일까?

56

II 자료와 정보

■ 사물·공간 스캐너

사물을 스캔하여 3D 프린팅에 필요한 디지털 파일을 만들거나 현실 공간을 스캔하여 가상 공간(VR)을 생성.

2 인간의 개입 없이 자율적으로 판단할 수 있는 지능화 관련 직업의 수요가 늘어날 것이다.

■ 생성적 디자이너

제품의 강도, 유연성, 사이즈 등 성능 목표를 설정하면 스스로 디자인을 만들어 내는 생성적 디자인(Generative Design), 소프트웨어를 이용하여 사용자의 취향이나 사용 환경에 맞춘 전자제품, 자동차, 운동용품 등 디자인.

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.

II-2장 자료와 정보의 분석

57

- 자료를 구조화하였을 때의 좋은 점에 대해 생각해 보며 동기 유발을 한다.

■ 오감 제어 전문가

오감을 통해 가상 현실을 경험하고, 가상 공간 내 사물을 이질감 없이 조작할 수 있는 기술을 개발.

■ 로봇 트레이너

직관적인 방식을 이용하여 로봇에게 주어진 업무를 교육·훈련하여 현장에 투입. 미국 리싱크로보틱스는 인간이 로봇 팔을 잡고 필요한 동작을 시연하는 것만으로

작업을 훈련시킬 수 있는 협업 로봇 백스터(Baxter)를 개발.

■ 인간·자동차 인터페이스 개발자

자율 주행 모드에서 운행하던 자동차가 돌발 상황에서 적시에 효과적으로 인간에게 경고를 하고 운전 권한을 넘기는 기술 개발.

이러한 기술을 기반으로 하는 직업 외에 문화적 지식과 감성을 기반으로 하는 콘텐츠 기획자, 예술가와 기술과 문화를 결합해 상품화하는 경영자, 광고 종사자, 웹 마스터 등도 유망한 직업이라고 할 수 있다.

출처: <http://www.msip.go.kr/webzine/posts.do?cateCd=C000000000002&webzineCd=wz-2016102815062>

자료의 수집

핵심질문

문제 해결에 있어 자료 수집은 왜 필요하며, 어떻게 자료를 수집할 수 있을까?

• 교과서 58쪽에 나타나있는 날씨 자료를 활용하는 실생활의 예를 살펴봄과 동기를 유발한다.

• 예전에 문제를 해결하기 위해 자료를 수집하였던 방법과 현재 시대가 바뀌면서 문제를 해결하기 위해 컴퓨터를 활용한 자료 수집 방법을 서로 비교하며 이야기를 나눈다.

자료 수집은 필요한 정보를 얻기 위한 수집 활동을 말한다. 예전에는 필요한 자료를 수집하기 위하여 직접 조사하여 글로 쓰거나 카메라로 사진을 찍는 등의 직접적인 조사 방법을 통하여 자료를 수집하였으나 디지털 기기가 발달함에 따라 많은 자료를 스마트폰이나 PC를 통해 수집할 수 있게 되었다.

자료를 이용한 문제 해결의 장점

새로운 관점과 문제 정의

▶ 직관, 경험, 주관에 의지하지 않고, 관찰 가능한 현상과 누구나 동의할 수 있는 수치화된 지표를 제시한다.

세밀하고 객관적인 분석

▶ 인간의 감각과 직관의 한계를 벗어나 정확도를 높이며, 편향과 오류의 가능성을 낮춘다.

분석 결과의 효율적이고 정확한 구현

▶ 모델과 알고리즘의 생성으로 반복적인 의사 결정의 효율성을 높인다. 데이터를 통해 얻은 결론은 비슷한 다른 문제에도 적용할 수 있다.

투명하고 강력한 의사소통

▶ 사실에 근거한 결론으로 설득의 힘을 지닌다.



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



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

◀ 날씨 정보의 예

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

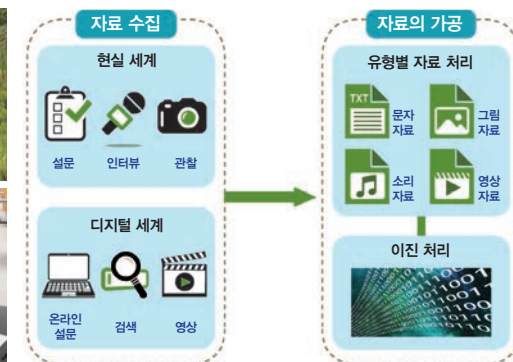
내비게이션에 나타나는 도로와 각 지역에 대해 흩어져 있는 자료들이 내가 길을 찾고자 할 때는 중요한 정보가 되어 문제 해결을 도와준다.

- 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.

자료 수집 유의 사항

- ① 일관성 - 동일한 수집 방법을 유지해야 함.
- ② 유연성 - 수집 목표 달성 여부에 따라 수집 시기를 유연하게 결정해야 함.
- ③ 무작위성 - 오류를 최소화하기 위해 표본 선정의 무작위성이 보장되어야 함.

• 빅 데이터의 긍정적인 측면을 교과서에서는 제시하였으나 사생활 침해와 보안 측면에서 우려될 수 있는 문제점들도 함께 생각해 볼 수 있도록 한다.

빅 데이터

기존 데이터베이스 관리 도구로 데이터를 수집·저장·관리·분석할 수 있는 역량을 넘어서는 대량의 정형 또는 비정형 데이터 집합 및 이러한 데이터로부터 가치를 추출하고 결과를 분석하는 기술을 의미한다. 수 테라바이트의 용량을 지닌 위키백과의 텍스트 및 이미지 자료는 빅 데이터의 고전적 사례에 속한다. 다양한 종류의 대규모 데이터에 대한 생성·수집·분석·표현을 그 특징으로 하는 빅 데이터 기술의 발전은 다변화된 현대 사회를 더욱 정확하게 예측하여 효율적으로 작동하게 하고, 개인화된 현대 사회 구성원마다 맞춤형 정보를 제공·관리·분석을 가능하게 하며 과거에는 불가능했던 기술을 실현시키기도 한다. 이같이 빅 데이터는 정치, 사회,

경제, 문화, 과학 기술 등 전 영역에 걸쳐서 사회와 인류에게 가치 있는 정보를 제공할 수 있는 가능성을 제시하며 그 중요성이 부각되고 있다.

하지만 빅 데이터의 문제점은 바로 사생활 침해와 보안 측면에 자리하고 있다. 빅 데이터는 수많은 개인들의 수많은 정보의 집합이다. 그렇기에 빅 데이터를 수집·분석할 때 개인들의 사적인 정보까지 수집하여 관리하는 빅브라더의 모습이 될 수도 있는 것이다. 그리고 그렇게 모은 데이터가 보안 문제로 유출된다면 이 역시 거의 모든 사람의 정보가 유출되는 것이기에 큰 문제가 될 수 있다.

출처: 위키백과(<https://ko.wikipedia.org/wiki/>)

핵심질문

실세계에서 수집한 아날로그 자료를 수집하고 가공하는 과정은 어떻게 되는가?

- 자료가 넘쳐나는 정보 사회에서 빅 데이터의 중요성과 가치에 대해 함께 알아본다.
- 자동 수집 - 편리하며, 누락이나 오류 없이 일관된 데이터를 얻을 수 있음.
→ 온라인 설문 이 이에 해당한다.
- 수동 수집 - 사람의 주관적인 감정, 만족도 등과 관련된 데이터를 정확하게 얻을 수 있음.
→ 직접 조사 방식으로 면 대 면 상황에서 자료 수집이 이루어진다.

함께 해 보기 예시 답안

- ① 사는 지역, 직업, 사이트를 알게 된 경로 등
- ② 구매 물품, 확인한 물품 내역 등

참고자료

직접 조사 방법과 인터넷 조사 방법의 차이

구분	직접 조사 방법	인터넷 조사 방법
비용	표본 수의 증가에 비례	상대적으로 적은 비용
조사 기간	2~5주	1~10일
구조화 설문	어려움	가능
설문 형식	문자 중심	멀티미디어 기반
실시간 분석	불가	가능
응답자 접근성	제한적	시간/공간적 제약 없음
조사의 연속성	어려움	가능

로그 파일(log file)

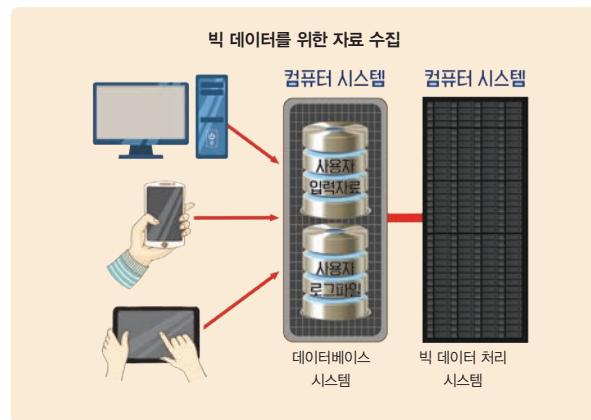
사용자가 컴퓨터를 실행하는 중에 발생하는 이벤트나 각기 다른 사용자의 통신 과정상에 나타난 메시지를 기록한 파일

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.



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	①
사용자의 사용 흔적을 컴퓨터가 자동 저장하여 자료 수집	로그인 시간, 사이트 방문 횟수 등	②

용어 정리

- **데이터 과학자**: 데이터(자료)에 담긴 패턴이나 예측을 통해 문제를 해결하는 사람

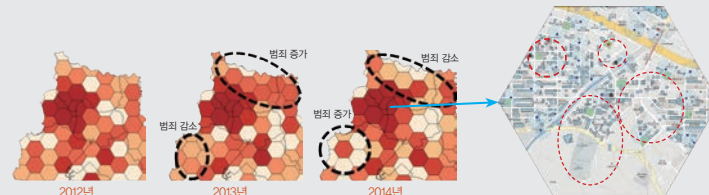
읽기자료

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 마니볼 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 예측은 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 구매 욕을 높이도록 하였다. 이러한 배치 방법은 판매되는 물건의 자료와 구매자들의 자료를 수집하여 분석한 중요한 정보에 따른 것이다.

핵심질문

수집된 자료를 이용하여 문제를 해결하는 과정은 어떻게 되는가?

- 실생활의 다양한 문제 해결 예시 자료를 통해 이해할 수 있도록 한다.

데이터 시각화(data visualization)

데이터 분석 결과를 쉽게 이해할 수 있도록 시각적으로 표현하고 전달하는 과정을 말한다. 데이터 시각화의 목적은 도표(graph)라는 수단을 통해 정보를 명확하고 효과적으로 전달하는 것이다. 디지털 시대의 학습자는 영상, 디지털 매체를 비롯해 소셜 네트워크 서비스(SNS), 클라우드 서비스 등을 통한 무한정의 데이터에 노출되어 있으며 이와 같은 환경의 변화에 따라 데이터 표현과 수용 방식의 변화가 요구되고 있다. 다시 말해 수많은 정보를 시각적으로 묘사하고 필요한 정보를 효율적이고 명확하게 제공하는 데이터 시각화의 필요성이 증가하고 있다. 최근 빅 데이터 기술 시장의

성장과 함께 시각화 기술도 중요한 기술 요소로 자리 잡고 있다. 또한 전통적인 시각화 기술은 주로 시스템 로그나 실험 분석 결과 등에 대한 통계 정보를 그래프로 보여 주는 방식인 데 반해, 빅 데이터의 시각화 측면에서는 모든 데이터를 살펴보는 것에 제약이 따르기 때문에 시각화의 기술적인 요소와 더불어 데이터를 요약하고, 한눈에 살펴볼 수 있도록 돕는 시각화 방법론적 요소의 중요성이 커지고 있다. 데이터 시각화에서는 미적 형태와 기능성 두 가지를 모두 가져가는 것으로 대개 데이터들의 연결과 그룹핑을 표현하는 데 초점을 둔다.

출처: 위키백과

- 이에 해당하는 내용을 학생들이 엑셀을 활용할 수 있는 실습 형태로 재구성하여 적용해 보아도 좋다.

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

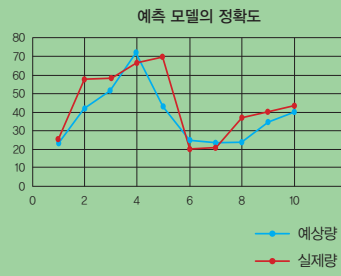
자료를 수집, 가공, 분석하여 실제 문제 해결에 활용하는 예를 살펴보자.

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				어린이날		43	69
6						24	20
7	흐림					23	21
8			어버이날			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.



문제 해결

- 법정 공휴일의 경우 판매량이 늘어난다.
- "생일 파티" 예약 시스템을 활용하여 재료를 준비한다.
- 주말에 판매량이 늘어난다.

참/고/자/료

• 데이터 시각화의 분야

- ▶ **정보 시각화**(information visualization)는 보통 대규모 데이터를 색채, 통계(도표, 그래프 등), 이미지 등을 활용해 요약적으로 표현하는 것을 의미한다.
- ▶ **과학적 시각화**(scientific visualization)는 실험 결과나 시뮬레이션 데이터 등 복잡한 데이터를 쉽게 탐색할 수 있도록 3차원 그래픽 기술 등을 활용하여 시각화하는 기술이다.

▶ **인포그래픽**(information graphics)은 인포메이션과 그래픽의 합성어로, 복잡한 수치나 글로 표현되어 있는 다량의 정보를 차트, 지도, 다이어그램, 로고, 일러스트레이션 등을 활용하여 한눈에 파악할 수 있도록 하는 디자인이다.

실 | 습 | 활 | 동

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

1 문제 선정하기

④ 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 형식

우리 반 친구들의 스마트폰 사용 실태 분석

우리 반 친구들의 스마트폰 사용 실태를 분석하여 문제점 진단과 그에 따른 해결 방법을 함께 알아보려고 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.

자 기 평 가

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 토론했을 수 있다.			

II-2장 자료와 정보의 분석

63

실습 활동

핵·심·질·문

온라인 설문을 활용하여 자료를 수집하고, 수집한 자료를 바탕으로 문제를 해결할 수 있는가?

- 다양한 주제를 선정할 수 있으며, 수집한 자료를 가공·분석한 후 반드시 발견된 문제점과 개선·방안에 대해 토론했을 수 있도록 한다.
- 보고서로 작성하여 발표하도록 한다.
- 수집된 자료를 표, 다이어그램 등 다양한 시각적 형태로 구조화하여 표현하도록 하여 정보 활용 목적에 효과적인지 토론했을 수 있도록 한다.

• 실습 예시 자료

- 1 나의 문제: 체중의 급격한 증가로 인한 건강 문제
- 2 자료 수집하기
 - 수집할 자료: 나의 식사 횟수, 식단, 섭취 칼로리, 운동 시간, 체중 증감 등
 - 자료 수집 방법: 수동 수집(식사 횟수, 식단, 칼로리), 자동 수집(건강 앱 활용해 걸음, 걷기, 달리기, 계단 오르기 자료 수집 등)
 - 자료 수집 기간: 2주일
 - 자료의 관리: 엑셀에 자료 기록
- 3 체중 증가 원인 분석하기

II-2장 자료와 정보의 분석 • 101

정보의 구조화

2-2

핵심질문

정보 구조란 무엇이며 왜 필요한가?
선형 구조와 비선형 구조는 무엇이며
그에 따른 종류에는 무엇이 있는가?

- 교과서 57쪽의 도입 만화를 보며 정보 구조의 필요성에 대해 함께 이야기해 본다.
- 정보 구조의 필요성에 대해서는 실생활에서 찾아보도록 한다.
- 구조화: 부분적 요소나 내용이 서로 관련되어 통일된 조직으로 만드는 것을 의미한다.

자료 구조(data structure)

자료를 효율적으로 이용할 수 있도록 컴퓨터에 저장하는 방법이 다. 신중히 선택한 자료 구조는 보다 효율적인 알고리즘을 사용할 수 있게 한다. 이러한 자료 구조의 선택 문제는 대개 추상적 자료 구조의 선택으로부터 시작하는 경우가 많다. 효과적으로 설계된 자료 구조는 실행 시간 혹은 메모리 용량과 같은 자원을 최소한으로 사용하면서 연산을 수행하도록 해 준다. 자료 구조에는 여러 종류가 있으며, 이러한 각각의 자료 구조는 컴퓨터에서의 연산 및 목적에 맞추어져 있다. 다양한 프로그램을 설계할 때 어떠한 자료 구조를 선택할지를 가장 우선적으로 고려해야 한다.

출처: <https://ko.wikipedia.org/wiki/>

정보와 자료의 차이

자료와 정보의 의미를 명확하게 구분 짓기는 어렵다. 자료에 의미가 부여되어 정보가 되고, 정보가 다시 자료가 되기도 하기 때문에 큰 차이를 두고 사용하지는 않는다. 다만 관점의 차이로 정보는 일상생활에서 주로 다루며, 자료는 컴퓨터 프로그래밍에서 처리하고 표현하는 부분에서 주로 다룬다는 것으로 그 의미를 구분 지을 수 있다.

구조화시켜 놓으면 관리와 제어가 쉬워진다

학교는 수백 명의 학생을 한정된 학교 시설 안에서 수업 계획, 시간표 관리, 성적 관리, 학생 관리, 입학 관리 등 많은 업무를 수행하여야 한다. 이럴 때 수백 명의 학생을 효과적으로 '구조화'하여 효율적이고 빠르게 관련 업무를 수행할 수 있다. 예를 들어 학생들에게 학년, 반, 번호를 부여하고, 성적에 따라 등급을 나누어 관리하기도 하고, 교실에서는 키 순서 또는 이름 순서대로 자리를 정해 앉는 것과 같은 기본적인 구조 체계를 사용할 수 있다.

출처: 『데이터로 표현하는 세상』, 김현철



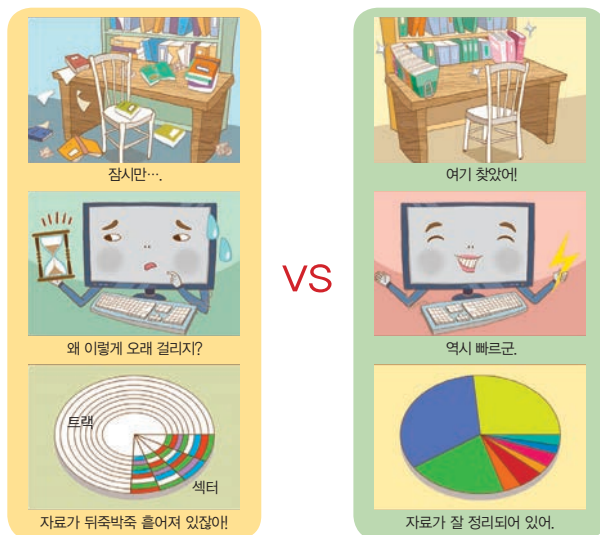
2-2 정보의 구조화

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

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.

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

● 생활 속 대표적인 구조화 방법은 다음과 같다.

- 목록
- 계층
- 표
- 다이어그램



▲ 정보 구조의 유형

- 신중히 선택한 자료 구조는 보다 효율적인 알고리즘을 사용할 수 있게 한다는 필요성의 관점에서 설명한다.
- 정보를 구조화하여 관리와 통제를 쉽게 할 수 있는 일상생활의 예 - 출석부, 회사 조직도, 재고 관리, 공항 시스템, 항만 시스템, 설문 조사 등

핵심질문

정보 구조화의 방법 및 사례를 들어 설명할 수 있는가?

- 리스트, 스택, 큐, 배열, 계층, 그래프의 개념에 관해 일상생활에서의 예와 컴퓨터에서 활용되는 예를 알아본다.
- 정보 구조화의 각 방법에 대해 학생들이 이해할 수 있는 수준에서의 일상생활 예와 컴퓨터에서 활용되는 예를 비교하면서 이해할 수 있도록 한다.

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

● 리스트의 연결 구조
· 단순 연결 리스트



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

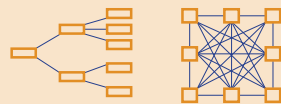
한 원소 뒤에 하나의 원소가 존재한다.
원소 간의 순서를 고려한다. 전후(인접) 원소 간에 1:1 관계로 나열된다.



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

비선형 구조는 자료를 일렬이 아니라 특별한 형태로 저장하는 구조로 계층적 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.

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.

생활 속의 리스트 구조
☞ 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

☞ 연탄 쌓기, 컵 쌓기



▲ 스택의 원리

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

☞ 순번 대기표, 줄서기

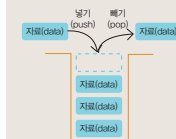


▲ 큐의 원리

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.



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



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

- 자료의 특성에 따라 정보 구조화의 방법이 달라질 수 있음을 이해한다.
- 리스트의 항목들에 접근하는 방식을 통해 스택과 큐라는 유형의 리스트 두 가지를 얻을 수 있다.
- 스택은 종종 백 트래킹 활동의 기초 구조로 사용된다(백 트래킹은 시스템에 진입하는 것과 반대 방향으로 시스템에서 빠져나오는 과정을 말하는데 흔히 숲에서 빠져나가는 길을 찾기 위해 지나갔던 길을 되돌아오는 과정에 비유한다).
- 큐는 데이터를 한 장소에서 다른 장소로 옮길 때 임시 저장 영역으로 사용되는 버퍼의 기초 구조로 종종 사용된다. 데이터 항목들은 버퍼에 도착하는 대로 큐의 테일 쪽에 저장되며, 최종 목적으로 보낼 시점이 되면 데이터 항목들은 큐의 헤드 쪽에 나타나는 순서대로 전달된다. 따라서 항목들은 도착한 순서대로 전달된다.

출처: 『컴퓨터 과학 총론』

• 리스트

리스트는 자료를 일정 순서에 맞춰 일렬로 저장하는 구조이며, 데이터와 링크의 조합으로 만들어진 노드들의 선형 집합이다. 자료들이 위치 정보를 알고 있는 포인터 변수에 의해 서로 이어져 있는 것처럼 보이게 만든 형태이다. 가장 단순하면서도 널리 쓰인다. 배열을 이용해서 리스트를 구현할 수 있다. 자료의 양이 많지 않고, 자료의 삽입/삭제가 빈번하게 이루어질 경우에는 연결 리스트가 효율적이다.

어떤 데이터 집합이나 리스트로 간주될 수 있다. 예를 들면 텍스트는 상징들의 리스트, 2차원 배열은 행들의 리스트, CD에 기록된 음악은 음의 리스트로 볼 수 있다. 보다 일상적인 리스트의 예로는 초청자 명단, 쇼핑 목록, 수강생 명단, 재고 목록 등을 들 수 있다.

출처: 『컴퓨터 과학 총론』

• 배열

배열은 번호(인덱스)와 번호에 대응하는 데이터들로 이루어지며, 같은 종류의 자료를 '직사각형'처럼 생긴 연속적인 공간에 저장하는 구조이다. 하나의 행으로 이루어지는 1차원 배열, 행들과

열들로 이루어지는 2차원 배열, 이를 더 확장한 3차원 배열이 있다. 배열은 자료의 양이 많지만 자료의 삽입/삭제가 거의 없고, 자료를 관리하거나 반복된 작업을 처리하는 데 쉽게 활용할 수 있다.



int Array[5] = { 3, 2, 7, 1, 9 }

▲ 깃발이 번호(인덱스)이고, 통 속에 저장된 값이 저장된 데이터

- 트리 구조의 주된 목적은 자료의 탐색을 하기 위해서이다.

배열과 연결 리스트의 차이점

배열의 특성은 메모리 공간에서 연속이라는 점이다. 그래서 처리 속도가 빠르지만, 메모리가 연속이어야 하기 때문에 삽입과 삭제가 번거롭다. 하지만 연결 리스트는 메모리상의 연속이지 않아도 된다. 왜냐하면 다음 데이터에 대한 링크 정보를 지니고 있기 때문이다.

배열



리스트



CPU의 처리 속도에서 배열이 리스트보다 훨씬 빠르다. 번호가 3인 자료에 접근하려면 배열의 경우 직접적으로 접근이 가능하지만 리스트의 경우 0번에서부터 3개의 화살표를 따라 이동해야만 접근이 가능하다.

참/고/자/료

배열에서의 자료 삽입/삭제

자료의 삽입/삭제가 빈번하게 이루어질 경우 삽입/삭제가 이루어진 위치의 다음부터 모든 자료의 위치를 변경해야 하므로 매우 취약하다.



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

시각표

배열의 예

시각표

엑셀을 이용한 배열표 만들기

엑셀

2 비선형 구조

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.

생활 속의 계층 구조

가계도, 회사 조직도

생활 속의 그래프 구조

전철 노선도, 인맥 지도

2 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.

68

II 자료와 정보

테이블(표)

테이블(표)형 구조는 두 개 이상의 항목을 하나의 표 형태로 구성한 구조이다. 일주일 시간표의 경우 요일과 시간을 항목으로 구성한 테이블형으로 나타내는 것이 가장 일반적이며 보기에도 쉽게 파악이 되어 매우 효율적이다.

▲ 일상생활에서의 예-시간표

	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

▲ 엑셀을 이용한 표만들기

정점(vertex)

간선(edge)

방향 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.

○○ 중학교

1학년

2학년

3학년

▲ 계층형 구조화

II-2장 자료와 정보의 분석

69

- 그래프를 이용하여 '나와 친구들과의 관계'를 나타내어 볼 수도 있다.
- 일상생활에서의 정보 구조화의 예를 찾아 보도록 하자.

합계 해 보기 예시 답안

각 정점은 위치를 나타내고 간선은 다리를 나타낸다. 그래프에 정점과 간선이 모두 포함되어 있는 사이클로서 오일러 사이클(Euler cycle)이라고 부른다.

그래프 이론(graph theory)

그래프 이론은 유한 개의 점으로 이루어진 집합과 점들 사이의 관계를 연구하는 학문이다. 수학과 컴퓨터 과학의 한 분야로, 특정 집단 내 대상들 간의 관계를 그래프로 나타내어 연구한다. 아래 그림은 아마존, 구글, 페이스북, 마이크로소프트, 애플, 오라클의 네

트위크를 그래프로 표현한 것이다. 이 그래프로 소통과 관계가 어떻게 이루어져 있는지 쉽게 파악할 수 있다.

출처: <http://guestbook.blog.naver.com/>

106 • II 부 자료와 정보

II-2장 자료와 정보의 분석 • 107

수식 계산에서의 스택 구조 활용 예

$$3 + 8 * 9 - 7 =$$

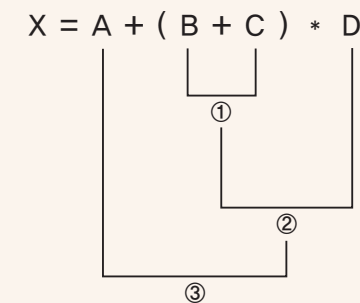
숫자 스택	연산자 스택	
1 3	+	2개의 레지스터에 수식의 순서대로 3은 숫자 레지스터에 입력(push)되고 +는 연산자 레지스터에 입력(push)된다.
2 8 3	* +	2개의 레지스터에 8과 *가 입력(push)된다. 연산자의 우선순위에 의해 *가 + 보다 먼저이므로 곱하기할 다음의 숫자를 받는다.
3 9 8 3	* +	9가 입력(push)되면 연산 순위가 높은 곱하기를 하기 위해 숫자 레지스터에서 9와 8을 꺼낸다(pop).
4 72 3	+	곱하기 연산자(*)를 꺼내(pop) 9*8을 하여 결과 값을 숫자 레지스터에 입력(push)한다.
5 72 3	- +	빼기(-)를 연산자 레지스터에 입력(push)하면 연산자 우선순위가 +와 -가 같기 때문에 빼기 연산자를 꺼내(pop) 72+3을 실행한다.
6 75		72+3의 결과 값 75를 숫자 레지스터에 입력(push)하고 더하기 연산자는 꺼낸다(pop).
7 7 75	-	숫자 레지스터에 7을 입력(push)하고 연산자 레지스터에 빼기 연산자를 입력(push)한다.
8 68	-	75-7의 최종 결과 값을 숫자 레지스터에 입력(push)한다.

* 숫자와 연산자를 넣는 공간을 '레지스터'라고 한다.

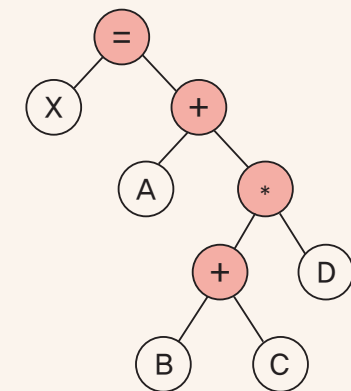
계층 구조를 이용한 수식 표기법

예제 1 $X = A + (B + C) * D$

연산자의 우선순위 표기

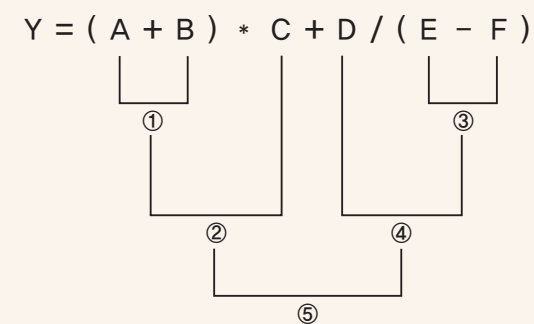


계층 구조 표현

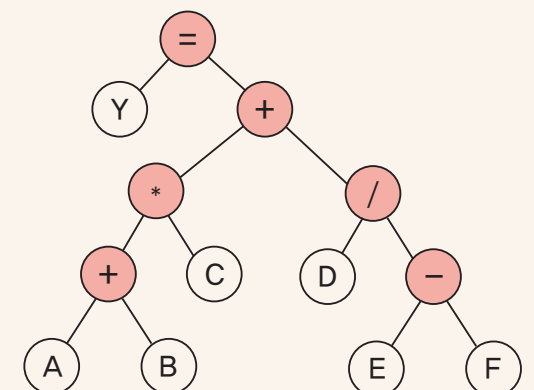


예제 2 $Y = (A + B) * C + D / (E - F)$

연산자의 우선순위 표기



계층 구조 표현



실습 활동 & 토론하기

- 실습 과제를 수행하기에는 자료를 수집해야 하는 기간이 필요하므로 사전에 과제를 공지하여 주도록 한다.
- 교과서에 제시된 토론 주제 중 하나를 선정하여 토론하여도 좋고, 학생들이 직접 주제를 선정하여도 좋다.
- 직접 조사, 인터넷 검색을 활용한 자료의 수집, 응용 소프트웨어를 활용한 자료의 분류·관리·공유를 수행하도록 하고 이러한 과정을 보고서로 작성해 보도록 한다.
- 모둠별 프로젝트이므로 계획 단계부터 역할을 분명히 하여 무임승차하는 학생이 없도록 구성하도록 한다.
- 자기 평가를 통해 자신의 학습을 스스로 평가하고 진단해 볼 수 있도록 한다.

실 습 활 동

❶ 친구들이 사 먹는 간식에 대한 실태를 조사하고 분석하여 보고서로 작성하여 보자.

활동 순서	활동 내용	활용할 수 있는 응용 소프트웨어
① 실태 조사	친구들이 사 먹는 간식에 대해 직접 조사하여 목록 작성하기 (☞ 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
② 유형 분류	직접 조사한 자료를 유형별로 분류하기 (☞ 장소별, 횟수별, 먹는 양, 먹는 시간대, 간식 유형별 등)	
③ 디지털 자료화	엑셀을 이용하여 표 구조화하고 정렬하기	
④ 실태조사를 바탕으로 필요한 정보 수집하기	검색 엔진을 활용하여 정보 수집하기 (☞ 칼로리, 식품 첨가물, 위생 상태 등)	
⑤ 보고서 작성하기	자료와 정보를 분석하여 보고서로 작성하기 (☞ 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.			

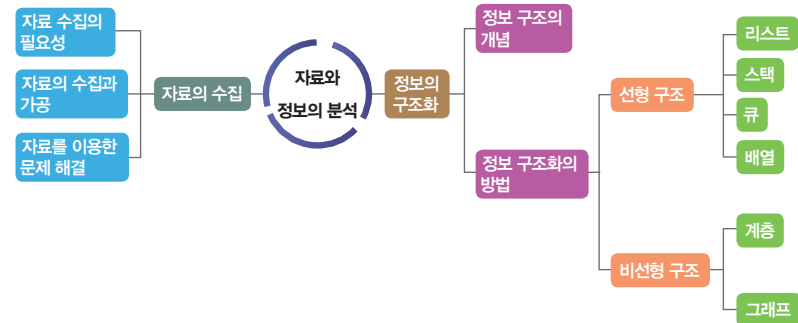
토 론 하 기

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.

70

Ⅱ 자료와 정보

배운 내용 정리하기

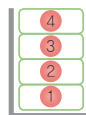


평 가 하 기

❶ 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.

❷ 아래 자료가 스택 구조라면 가장 오래된 자료는 몇 번일까?



❸ 정보를 구조화하였을 때 좋은 점 한 가지를 써 보자.

배운 내용 정리 & 평가하기

- 배운 내용 정리하기를 통해 학습한 내용을 다시 한 번 되돌아본다.

평 가 하 기 예시 답안

- ① ① 계층 구조 ② 그래프
- ② ①
- ③
 - 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 변경 가능하다.
 - 자료 검색을 빠르게 할 수 있다.

참 고 자 료

• 세상은 네트워크로 연결되어 있다 '6단계 분리 이론'

1967년 미국의 하버드 대학교 교수였던 스탠리 밀그램(Stanley Milgram)은 '6단계 분리'라고 불리는 실험을 통해 세상 사람들은 생각보다 짧은 단계, 즉 6단계만 거치면 세상 모든 사람이 연결된다는 것을 제시하였다.

밀그램은 네브래스카 오마하에 사는 사람 가운데 무작위로 160명을 선택하여 최종 목적인 매사추세츠 보스턴의 증권 중개인과 더 가까울 것 같은 지인에게 소포를 전달해 달라고 요청하였다. 이 실험에 참여한 실험자들은 편지의 수신인을 알지 못하는 상태였기 때문에 여러 가지 추측을 통해 자신이 아는 사람 중에서 편지의 수취인을 알 만한 사람에게 편지를 전해 주었다. 그리고 편지를

다른 사람에게 전해 주기 전, 반드시 편지 봉투에 이를 받은 사람의 이름을 적어 넣게 하였다. 열핏 보면 실패로 끝날 것 같았던 이 실험에서 상당수의 편지가 주인에게 제대로 전달되었고, 편지 봉투에 쓰인 이름이 생각보다 적었다. 어떤 편지는 3번 만에 전달되기도 하였고, 어떤 편지는 열 사람을 거쳐 전달되기도 했지만 평균적으로 약 6명의 사람을 통해 전달되었다는 사실이 밝혀졌다. 이 실험의 결론은 미국에 살고 있는 사람들의 경우 전혀 모르는 사람이라도 여섯 단계만 거치면 서로 연결이 된다는 것인데 이를 '작은 세상 효과(small-world effect)'라고 부른다.

이렇게 6단계의 분리 현상이 일어나는 것은 세상이 무작위적이 지 않기 때문이고 또한 현실 세상에서 모두가 연결되어야만 이루

어지는 것이 아니기 때문이다. 두 개의 클러스터 사이에 단 하나의 링크만 있어도 두 개의 클러스터는 하나로 이어질 수 있다. 나의 친구들은 서로 알지 못하는 두 개의 클러스터라고 할 수 있지만 이 사이에 '나'라는 링크가 있음으로 해서 두 그룹은 하나로 이어지고, 각 그룹에서 알고 있던 정보가 쉽게 다른 그룹으로 넘어갈 수도 있다.

이처럼 클러스터와 클러스터를 잇는 단 하나의 링크가 세상 사람들을 여섯 단계로 이어질 수 있게 만드는 요소가 될 수 있다. 네트워크로 이루어진 세상은 아무리 크다 할지라도 놀랄 만큼 가까운 거리를 유지하고 있다는 것을 알 수 있다.

출처: 사이언스타임즈(<https://goo.gl/XMrzR3>)

Ⅱ-2장 자료와 정보의 분석

71

참 고 자 료

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 |

[해설] 일상생활에서 쓰이는 다양한 자료의 유형에 대한 사례이다.

2 십진수를 이진수로, 이진수를 십진수로 변환해 보자.

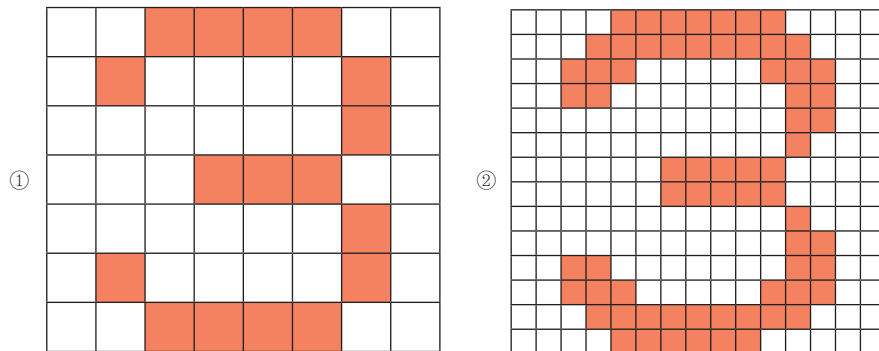
- | | |
|-----------------------------------|-----------------------------------|
| ① $18 \rightarrow (10010_{(2)})$ | ② $1011_{(2)} \rightarrow (11)$ |
| ③ $53 \rightarrow (110101_{(2)})$ | ④ $110111_{(2)} \rightarrow (55)$ |

[해설] 18과 53은 2로 차례로 나누면서 남는 나머지를 적어 이진수로 바꾸거나 아니면 이진수 각 자리의 값을 활용하여 이진수로 나타낼 수도 있다.

$$1011_{(2)} = (1 \times 2^3) + (0 \times 2^2) + (1 \times 2^1) + (1 \times 2^0) = 11,$$

$$110111_{(2)} = (1 \times 2^5) + (1 \times 2^4) + (0 \times 2^3) + (1 \times 2^2) + (1 \times 2^1) + (1 \times 2^0) = 55$$

3 아래 두 가지 모눈종이에 숫자 3을 색칠하여 나타내 보자.



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

- 1) 두 개의 그림 중 어느 그림이 자세하게 보이는가? ②번
- 2) 모눈종이의 흰 픽셀은 0, 검은 픽셀은 1로 나타낸다면 어떤 것이 1의 값을 더 많이 포함하는가? ②번, 위의 모눈 종이에서는 거의 두 배의 1값을 포함한다.
- 3) 그림을 더 자세하게 표현하는 것과 픽셀은 어떤 관계가 있는가? 같은 크기의 그림에서 픽셀의 수가 많아질수록 더 자세히 표현된다.

4 수집된 자료를 이용하여 문제를 해결하는 과정을 차례로 나열해 보시오.

- | | | | |
|---------|---------|---------|---------|
| ① 문제 해결 | ② 자료 가공 | ③ 자료 수집 | ④ 자료 분석 |
|---------|---------|---------|---------|

(③) → (②) → (④) → (①)

[해설] 자료를 이용한 문제 해결 과정은 자료 수집→자료 가공→자료 분석→문제 해결의 과정을 거친다.

5 () 안에 들어갈 말을 찾아 쓰시오.

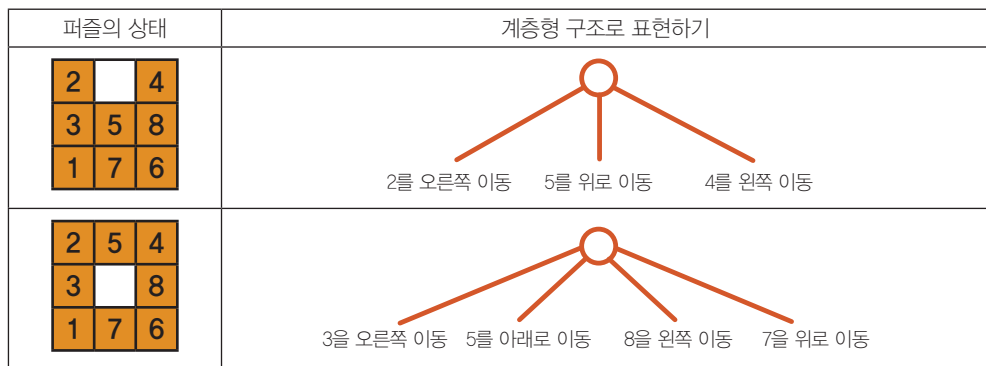
자료, 정보, 아날로그, 디지털

- ① 단순한 관찰이나 측정을 통해 수집한 사실이나 어떤 값을 말한다. (자료)
- ② 단순히 나열되어 있는 자료를 더욱 유용하고 의미 있도록 가치를 부여한 것을 말한다. (정보)
- ③ 연속적으로 변화하는 값을 일정한 간격으로 끊어 불연속적인 값으로 양을 표현하는 것을 말한다. (디지털)
- ④ 연속적으로 변화하는 값으로 양을 표현하는 것을 말한다. (아날로그)

[해설] 자료와 정보, 아날로그와 디지털에 관한 개념이다.

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

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



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

