

기계 : 1. 기계적 동력장치 -> 불도저
2. 자동화 기계장치 -> 컴퓨터

단위 -> 손가락 한개가 가능한표현을 의미

단위 * 단위 -> n*n개의 표현이 가능해짐
손가락이 10개이면, 2^10개의 표현이 가능

그래서 조합은 굉장한 힘 -> 많은 표현이 가능하게 하는 힘이 있다.

수업에거 숙제가 있었다

정말 좋은 손가락 언어를 만들려면 많은 것들을 고려해야 한다.

100가지를 표현하고 싶은데, 단위가 2이면, 단위가 최소 7개가 필요하다

코드의 효율이 필요하다.
-> 매일 사용하는 것은 짧은 코드를 쓰고, 가끔 사용하는 것은 긴 코드를 사용해도 된다.

눈꺼풀 하나로만 의사 표현을 해야하는 상황이 온 것이다.

하나씩 spelling단위로 가리키면서 이게 맞는지의 여부를 물어보았다.

알파벳 순이 아닌, 빈도수 순으로 프랑스 알파벳을 불러줬다면 시간이 훨씬 더 절약할 수 있었을 것이다.

모스부호도 2개를 표현하는 단위의 조합으로 표현한다.

영어는 E, T가 가장 짧다.
한글은 ㅏ, ㅑ가 가장 짧다.

자주사용하는 코드에는 짧은 모스부호를 부여한다.
반대로 자주 사용하지 않는 코드에는 긴 모스부호를 부여한다.

모스부호에서 잠깐 쉬는 것을 통해 자간을 구분한다.

이중모음 -> 모음 2개를 붙여서 할 것 같다.

허프만 코딩 알고리즘도 비슷한 원리로 만들어져 있다.
사실 파일을 압축하는데 거의 대부분 허프만 코딩을 사용한다.

허프만 코딩에서는 하나의 코드가 다른 코드의 prefix가 절대로 되지 않도록 보장해 놓았다.
모스부호를 보낼 때의 실수를 수정하는 방법은 없었다.

정보이론 : 이 정보의 가치는?
정보의 entropy(혼란도, 엔트로피)
shannon이라는 사람이 bit라는 단위를 만든 사람이다.
디지털 선구자들 중 하나이다. (학교에서 오토바이 타고 저글링하는 걸 좋아했다)

이 사람은 정보의 가치를 정량적으로 측정하기를 원했다.
1940년대의 과학은 "물리량" 으로 세상을 이해하려는 시대였다.
-> 물질과 에너지로 세상을 이해하려고 했다.

shannon은 물질과 에너지 외의 제 3의 요소가 있다고 주장,
그 3요소는 정보(information)이라고 하는 것이다.
세상은 내면의 세상과 외부의 세상이 있다. -> 내부의 세상은 정보, 외부의 세상은 과학이라는 것이다.
내부의 정보가 외부의 과학적 세상으로 넘어오는 계기가 된다.

정보의 측정 모델을 만들고 싶어했다.
(이 사람은 근본적으로 물리학자였다.)

그렇다면 정보자체를 어떻게 측정하나?
-> 불확실성(uncertainty)으로 측정을 하자.
동전이 앞면이 나오겠는가? 아니면 뒷면이 나오겠는가? -> 최고의 불확실성이다. 50% vs 50% -> 불확실성 최고
그런데 누군가가 앞면에 특수 스티커가 붙어있다고 하면 앞면 20%, 뒷면 80%라고 하면 불확실성이 줄어든 것이다.
아예 앞면에 특수장치가 있어서 앞면이 나올 수가 없다는 정보를 얻으면, 앞면 0%, 뒷면 100%이 되는 것이다 -> 불확실성 최저
정보의 정량적 값은 그 정보를 얻고 난 뒤에 얼마나 불확실성이 줄었는가로 표현한다.

50% 50%일 때는 $\log_2 2 = 1$ 만큼의 불확실성이지만, -> 1bit -> 2가지 심볼이 나온다.
100% 0%일 때는 $\log_2 1 = 0$ 만큼의 불확실성이 되는 것이다. -> 0bit -> 1가지 심볼이 나온다
그리고 그 단위를 bit라고 하는 것이다.

심볼이 4개인 사면체 -> 2bit로 표현가능
심볼이 8개인 팔면체 -> 3bit로 표현가능.
나올 수 있는 심볼이 많아지면 -> uncertainty는 증가한다.

정리하면 $H = -$ 가 된다.

우리의 목적은 entropy를 줄이는 것이 좋은 코드를 설계하는 목적이다.

DNA의 ACGT를 생각해보자는 것이다.
만일 8번 관찰을 했는데, 그 DNA position의 uncertainty, entropy를 어떻게 계산하는가?
 $P_A = 4/8, P_C = 2/8, P_G = 1/8, P_T = 1/8$ 으로 H를 계산한다.

entropy가 낮으면 uncertainty가 낮다는 의미이다.

원래는 다 1/4의 확률로, 2bit가 나와야 하는데,
 $P_A = 4/8, P_C = 2/8, P_G = 1/8, P_T = 1/8$ 라는 정보를 얻어서
1.75bit로 엔트로피를 줄이면 그만큼 $P_A = 4/8, P_C = 2/8, P_G = 1/8, P_T = 1/8$ 이라는 정보가 가치(information gain)가 있는 것이다.

최대 엔트로피는 $\log_2 n$ 이다. -> 각각의 확률을 $1/n$ 으로 가정해야 하는 것이다.

엔트로피의 단위가 bit이고, 엔트로피간 차이도 bit단위이다.