

003/200

新수능 국어 최적화 기출 분석

시간을 재며 문제를 풀어보았다면 느끼겠지만
 도무지 정보처리가 깔끔하지 않습니다.
 정보 완결성이 없고, 결국 몇 번을 읽어도 처리가 안 됩니다.
우연히 선지를 골라낼 수는 있으나,
필연적 접근은 아닙니다. 배운 내용 그대로 갑니다.

T+I.A (Topic + Information Architecture)	
1 P 정보량 폭발 S 구조+처리 가능 정보에 집중 디지털 통신 시스템 구성 1.송신기, 2.채널, 3수신기 부호화 과정	1단락 세부 내용 간략 정리 원리 : 기호의 발생확률↑ 정보량↓ 기호의 발생확률↓ 정보량↑ 세부 내용 - 데이터 : 기호 집합의 기호 조합 정보량 : 정보의 크기 평균정보량:기호집합의 엔트로피 평균정보량:각 기호의 발생확률과 정보량을 서로 곱 모두 더함
2 P 정보량 폭발 S 구조+처리 가능 정보에 집중 송신기 ① 소스 부호화 ② 채널 부호화 ③ 선 부호화 ① 소스 부호화 : 데이터 압축 기호 → 0/1 부호로 변환 세부내용 또 정보량 터짐... 마지막 '허프만 부호화' 기호의 발생확률↑ 비트수↓ 기호의 발생확률↓ 비트수↑	1단락 끝. C 기호들 동일 발생확률 → E 기호 집합의 엔트로피 최댓값 정보량이 터지면 구조 잡고 주제 잡고 처리 가능한 정보에 집중
3 수치 발췌독 ② 채널 부호화 : 정보 추가 오류 검출 및 정정 '삼중 반복 부호화' 0과 1을 각각 000과 111로 과반으로 판단 부호율 0.33 수치 발췌독	2단락 ① 소스 부호화 : C데이터 압축 E기호→0/1부호로 변환 3단락 ② 채널 부호화 : 정보 추가 오류 검출 및 정정 '삼중 반복 부호화' 0과 1을 각각 000과 111로 과반으로 판단
4 정보처리 도구 : 표그리기 ③ 선 부호화 : 전압 비트 0 : 유지 비트 1 : 변환	4단락 ③ 선 부호화 : 전압 비트 0 : 유지 비트 1 : 변환

4~5단락 표를 우선 정리하고 접근하겠습니다. (사실, 중2때 배운 조건부 확률임)

38. 윗글에서 알 수 있는 내용으로 적절한 것은?

① 영상 데이터는 채널 부호화 과정에서 압축된다.

정보가치 높은 내용. 채널 부호화는 압축이 아니라 정보 추가!

② 수신기에는 부호를 기호로 복원하는 기능이 있다.

상식. 배경지식. 하향식으로 처리합시다.

당연히 있어야죠. 정답!

이와 관련하여 뒤에 내용을 덧붙이겠습니다.

③ 잉여 정보는 데이터를 압축하기 위해 추가한 정보이다.

상식. 배경지식. 하향식 처리.

잉여 정보가 있는데 어떻게 압축이 되나요

④ 영상을 전송할 때는 잡음으로 인한 오류가 발생하지 않는다.

상식. 배경지식. 하향식 처리.

당연히 오류 있겠죠.

⑤ 소스 부호화는 전송할 기호에 정보를 추가하여 오류에 대비하는 과정이다.

채널부호화.

이렇게 주제. 정보가치 높은 내용 + 하향식으로 처리를 하면 바로 보이는데, 이것을 예전 접근법으로 발췌독 한다?

난리 납니다.

여러분

다시 문제 풀었던 것으로 돌아가

1번 선지 풀기 위한 근거를 찾고, 시간 재고 다시 와주십시오.

대신. 매우 정확히요.

논리적으로 완벽하게!

디지털 통신 시스템은 송신기, 채널, 수신기로 구성되며, ㉔ 전송할 데이터를 빠르고 정확하게 전달하기 위해 부호화 과정을 거쳐 전송한다. 영상, 문자 등인 데이터는 ㉕ 기호 집합에 있는 기호들의 조합이다. 예를 들어 기호 집합 {a, b, c, d, e, f}에서 기호들을 조합한 add, cab, beef 등이 데이터이다. 정보량은 어떤 기호가 발생했다는 것을 알았을 때 얻는 정보의 크기이다. 어떤 기호 집합에서 특정 기호의 발생 확률이 높으면 그 기호의 정보량은 적고, 발생 확률이 낮으면 그 기호의 정보량은 많다. 기호 집합의 평균 정보량을 기호 집합의 엔트로피라고 하는데 모든 기호들이 동일한 발생 확률을 가질 때 그 기호 집합의 엔트로피는 최댓값을 갖는다.

우선 1단락에서 위의 내용이 필요합니다.

- ① 영상 데이터는 채널 부호화 과정에서 압축된다.

이 내용을 처리하기 위해서는

영상 데이터는

영상이 데이터라는 것을 처리해야 하니

영상, 문자 등인 데이터

이 내용이 있어야 하고

다시 또,

데이터를 빠르고 정확하게 전달하기 위해 부호화 과정

데이터는 ㉕ 기호 집합에 있는 기호들의 조합이다.

이렇게 데이터 전송에는 부호화 과정과

데이터를 기호로 바꾼다는 내용을 처리 해야 해요.

송신기에서는 소스 부호화, 채널 부호화, 선 부호화를 거쳐 기호를 ㉖ 부호로 변환한다. 소스 부호화는 데이터를 압축하기 위해 기호를 0과 1로 이루어진 부호로 변환하는 과정이다. 어떤 기호가 110과 같은 부호로 변환되었을 때 0 또는 1을 비트라고 하며 이 부호의 비트 수는 3이다. 이때 기호 집합의 엔트로피는 기호 집합에 있는 기호를 부호로 표현하는 데 필요한 평균 비트 수의 최솟값이다. 전송된 부호를 수신기에서 원래의 기호로 ㉗ 복원하려면 부호들의 평균 비트 수가 기호 집합의 엔트로피보다 크거나 같아야 한다. 기호 집합을 엔트로피에 최대한 가까운 평균 비트 수를 갖는 부호들로 변환하는 것을 엔트로피 부호화라 한다. 그중 하나인 '허프만 부호화'에서는 발생 확률이 높은 기호에는 비트 수가 적은 부호를, 발생 확률이 낮은 기호에는 비트 수가 많은 부호를 할당한다.

이어서

압축된다.

는 내용을 처리하려면

소스 부호화는 데이터를 압축하기 위해 기호를 0과 1로 이루어진 부호로 변환

위의 내용이 있어야

압축할 때는 소스부호화를 쓰는게 보입니다.

문제는 채널부호화가 이 내용이 아니라는 것을 확인하기 위해 채널 부호화 까지 처리해야하죠?

채널 부호화는 오류를 검출하고 정정하기 위하여 부호에 잉여 정보를 추가하는 과정이다. 송신기에서 부호를 전송하면 채널의 잡음으로 인해 오류가 발생하는데 이 문제를 해결하기 위해 잉여 정보를 덧붙여 전송한다. 채널 부호화 중 하나인 '삼중 반복 부호화'는 0과 1을 각각 000과 111로 부호화한다. 이때 수신기에서는 수신한 부호에 0이 과반수인 경우에는 0으로 판단하고, 1이 과반수인 경우에는 1로 판단한다. 즉 수신기에서 수신된 부호가 000, 001, 010, 100 중 하나라면 0으로 판단하고, 그 이외에는 1로 판단한다. 이렇게 하면 000을 전송했을 때 하나의 비트에서 오류가 생겨 001을 수신해도 0으로 판단하므로 오류는 정정된다. 채널 부호화를 하기 전 부호의 비트 수를, 채널 부호화를 한 후 부호의 비트 수로 나눈 것을 부호율이라 한다. 삼중 반복 부호화의 부호율은 약 0.33이다.

네.

이 내용 까지 있어야 합니다..

채널 부호화는 오류를 검출하고 정정하기 위하여 부호에 잉여 정보를 추가하는 과정이다.

- ① 영상 데이터는 채널 부호화 과정에서 압축된다.

이 선지를 논리적으로 완벽히 처리하기 위해서 걸리는 시간이란.

네.

현실적으로 가능한가요?

제발 안 되는 것 고집 맙시다.

이 선지만 그렇지도 않아요.

- ④ 영상을 전송할 때는 잡음으로 인한 오류가 발생하지 않는다.

이 4번 선지도 다시 풀던 때로 돌아가서 완벽한 논리적 근거 잡아보죠?

④ 영상을 전송할 때는 잡음으로 인한 오류가 발생하지 않는다.

다시 또, '영상'이 데이터고, 그 데이터는 기호조합이고...

디지털 통신 시스템은 송신기, 채널, 수신기로 구성되며, ① 전송할 데이터를 빠르고 정확하게 전달하기 위해 부호화 과정을 거쳐 전송한다. 영상, 문자 등인 데이터는 ② 기호 조합에 있는 기호들의 조합이다. 예를 들어 기호 집합 {a, b, c, d, e, f}에서 기호들을 조합한 add, cab, beef 등이 데이터이다. 정보량은 어떤 기호가 발생했다는 것을 알았을 때 얻는 정보의 크기이다. 어떤 기호 조합에서 특정 기호의 발생 확률이 높으면 그 기호의 정보량은 적고, 발생 확률이 낮으면 그 기호의 정보량은 많다. 기호 조합의 평균 정보량을 기호 조합의 엔트로피라고 하는데 모든 기호들이 동일한 발생 확률을 가질 때 그 기호 조합의 엔트로피는 최댓값을 갖는다.

1단락에서 위의 내용 그대로 필요합니다.

채널 부호화는 오류를 검출하고 정정하기 위하여 잉여 정보 추가하는 과정이다. 송신기에서 부호를 전송하면 채널의 잡음으로 인해 오류가 발생하는데 이 문제를 해결하기 위해 잉여 정보를 덧붙여 전송한다. 채널 부호화 중 하나인 '삼중 반복 부호화'는 0과 1을 각각 000과 111로 부호화한다. 이 때 수신기에서는 수신한 부호에 0이 과반수인 경우에는 0으로 판단하고, 1이 과반수인 경우에는 1로 판단한다. 즉 수신기에서 수신된 부호가 000, 001, 010, 100 중 하나라면 0으로 판단하고, 그 이외에는 1로 판단한다. 이렇게 하면 000을 전송했을 때 하나의 비트에서 오류가 생겨 001을 수신해도 0으로 판단하므로 오류는 정정된다. 채널 부호화를 하기 전 부호의 비트 수를, 채널 부호화를 한 후 부호의 비트 수로 나눈 것을 부호율이라 한다. 삼중 반복 부호화의 부호율은 약 0.33이다.

3단락 위의 내용이 또 필요합니다.

그래야 잡음으로 인한 오류 발생을 확인할 수 있어요.

절대로 불가능한 이런 풀이.

이제 그만...

이어지는 39번 문항.

1~2단락 정보군집으로 처리 해야 하는데, 일단 처리가 안 되면 넘겨도 좋습니다. 추후 다시 보겠습니다.

39. 윗글을 바탕으로, 2가지 기호로 이루어진 기호 집합에 대해 이해한 내용으로 적절하지 않은 것은?

- ① 기호들의 발생 확률이 모두 1/2인 경우, 각 기호의 정보량은 동일하다.
- ② 기호들의 발생 확률이 각각 1/4, 3/4인 경우의 평균 정보량이 최댓값이다.
- ③ 기호들의 발생 확률이 각각 1/4, 3/4인 경우, 기호의 정보량이 더 많은 것은 발생 확률이 1/4인 기호이다.
- ④ 기호들의 발생 확률이 모두 1/2인 경우, 기호를 부호화하는 데 필요한 평균 비트 수의 최솟값이 최대가 된다.
- ⑤ 기호들의 발생 확률이 각각 1/4, 3/4인 기호 조합의 엔트로피는 발생 확률이 각각 3/4, 1/4인 기호 조합의 엔트로피와 같다.

글의 주제인 부호화를 묻습니다.

40. 윗글의 '부호화'에 대한 내용으로 적절한 것은?

1단락
디지털 통신 시스템 구성 1.송신기. 2.채널. 3수신기
송신기-부호화 과정

2단락 ① 소스 부호화 : C데이터 압축 E기호→0/1부호로 변환

3단락 ② 채널 부호화 : 정보 추가 오류 검출 및 정정 '삼중 반복 부호화' 0과 1을 각각 000과 111로 과반으로 판단

4단락 ③ 선 부호화 : 전압 비트 0 : 유지 비트 1 : 변환

- ① 선 부호화에서는 수신기에서 부호를 전기 신호로 변환한다.

선 부호화 소주제 맞습니다. 다만 부호화는 송신기에서 일어납니다! 틀린 내용.

- ② 허프만 부호화에서는 정보량이 많은 기호에 상대적으로 비트 수가 적은 부호를 할당한다.

?? 1~2단락 처리 안 됩니다. 답 없으면, 필요시 발췌독 갑니다.

- ③ 채널 부호화를 거친 부호들은 채널로 전송하기 전에 잉여 정보를 제거한 후 선 부호화한다.

채널 부호화는 잉여정보 추가하는 것 아닌가요?

역시 소주제! 틀린 내용.

- ④ 채널 부호화 과정에서 부호에 일정 수준 이상의 잉여 정보를 추가하면 부호율은 1보다 커진다.

수치 발췌독! 부호율 0.33. 틀린 내용!

- ⑤ 삼중 반복 부호화를 이용하여 0을 부호화한 경우, 수신된 부호에서 두 개의 비트에 오류가 있으면 오류는 정정되지 않는다.

삼중 반복 부호화의 원리 이해 적용입니다.

3 수치 발췌독
 ㉔ 채널 부호화 : 정보 추가 오류 검출 및 정정
 '삼중 반복 부호화' 0과 1을 각각 000과 111로
 과반으로 판단

옳은 내용! 정답!

즉

0을 부호화 하면 000으로 보내는데

1비트 오류가 생겨 001 010 100 등으로 나오면

과반 0으로 판단하니 오류 검출 해결 되지만

2비트 오류 생기면 011 10 110

과반 1로 판단하니 오류 검출 안 되고 오류 그대로 들어갑니다.

41. 윗글을 바탕으로 <보기>를 이해한 내용으로 적절한 것은?

<보기>

날씨 데이터를 전송하려고 한다. 날씨는 '맑음', '흐림', '비', '눈'으로만 분류하며, 각 날씨의 발생 확률은 모두 같다. 엔트로피 부호화를 통해 '맑음', '흐림', '비', '눈'을 각각 00, 01, 10, 11의 부호로 바꾼다.

- ① 기호 집합 {맑음, 흐림, 비, 눈}의 엔트로피는 2보다 크겠군.

1단락 처리 못 했으니 넘어가고, 필요시 발췌독하여 처리합니다.

원래도 그렇게 풀었쥬?

- ② 엔트로피 부호화를 통해 4일 동안의 날씨 데이터 '흐림비 맑음흐림'은 '01001001'로 바뀌겠군.

보기 바탕. 0110인데 바로 틀렸네요?

'맑음', '흐림', '비', '눈'을 각각 00, 01, 10, 11의 부호로 바꾼다.

- ③ 삼중 반복 부호화를 이용하여 전송한 특정 날씨의 부호를 '110001'과 '101100'으로 각각 수신하였다면 서로 다른 날씨로 판단하겠군.

삼중 반복 부호화

110 001

1 0

101 100

1 0

결과는 둘다 10 = 10으로 동일!

- ④ 날씨 '비'를 삼중 반복 부호화와 차등 부호화를 이용하여 부호화하는 경우, 기준 신호가 양(+)의 전압이면 '음, 양, 음, 음, 음, 음'의 전압을 갖는 전기 신호로 변환되겠군.

비 <보기>	삼중 반복 부호화	차등 부호화						
		기준	1	1	1	0	0	0
10	111000	+	-	+	-	-	-	-
			바뀜	바뀜	바뀜	유지	유지	유지

바로 정답!

정보 완결성 X + 정보량 ↑

주제와 구조

그리고

처리 가능한 정보에 집중.

그 힘을 봐주세요.

39. 윗글을 바탕으로, 2가지 기호로 이루어진 기호 집합에 대해 이해한 내용으로 적절하지 않은 것은?

- ① 기호들의 발생 확률이 모두 1/2인 경우, 각 기호의 정보량은 동일하다.
- ② 기호들의 발생 확률이 각각 1/4, 3/4인 경우의 평균 정보량이 최댓값이다.
- ③ 기호들의 발생 확률이 각각 1/4, 3/4인 경우, 기호의 정보량이 더 많은 것은 발생 확률이 1/4인 기호이다.
- ④ 기호들의 발생 확률이 모두 1/2인 경우, 기호를 부호화하는 데 필요한 평균 비트 수의 최솟값이 최대가 된다.
- ⑤ 기호들의 발생 확률이 각각 1/4, 3/4인 기호 집합의 엔트로피는 발생 확률이 각각 3/4, 1/4인 기호 집합의 엔트로피와 같다.

아무리 읽어도 정보처리가 되지 않습니다.

이유를 살펴보면...

정보 완결성 X + 정보량 ↑

1단락 세부 내용 간략 정리
 원리 :
 기호의 발생확률↑ 정보량↓
 기호의 발생확률↓ 정보량↑

세부 내용 -
 데이터 : 기호 집합의 기호 조합
 정보량 : 정보의 크기
 평균정보량: 기호집합의 엔트로피
 평균정보량: 각 기호의 발생확률과 정보량을 서로 곱 모두 더함

1단락 끝.
 C 기호들 동일 발생확률 →
 E 기호 집합의 엔트로피 최댓값

먼저, 기호 집합에서 기호를 조합하여 그것을 데이터로 놓는 것까지는 알겠는데,

정보량... 부분이

add 라는 정보가 있다면 정보량은 3자리로 처리를 하는 것인지 아니면 어떻게 처리를 하는 것인지 불명확 하며,

이는

2단락에도 이어져

0과 1 1비트로 처리하는데

그러면 a를 0

b를 1

그러면 나머지는 어떻게??

도무지 처리가 안 되기 시작하죠.

확신이 없습니다.

당연하죠.

정보 완결성 X

정보가 없어요.

그게 끝이 아닙니다.

평균 정보량을

기호 집합의 엔트로피라 하는데

엔트로피라는 단어도 생소하며

무엇보다.

평균 정보량의 정의를

각 기호의 발생확률과 정보량을 서로 곱하여 더한다

물론 통계를 공부한 학생들이면

이것이 확률 통계의 평균 즉, 기댓값의 정의이니

$$E(x) = \sum_{i=1}^n x_i p_i$$

위와 같은 공식을 생각할 수는 있는데

문제는, 정보량에 대한 정의도, 그 확률에 대한 정의도 없는데 계산은 도대체 어떻게 해야하는지...

물론 시간을 충분히 가지고 꾸역 꾸역 하면 안 될 것은 없지만.

그 시간이 수십분이 될거예요.

실전에서 절대 불가.

출제 의도도 아닙니다.

그렇다면??

정보 완결성 X + 정보량 ↑

주제와 구조

그리고 처리 가능한 정보에 집중.

다시 문항을 보니. 그리고 가장 정보가치가 높은 것은

1단락 끝 + 2단락 초반...?

기호 집합의 평균 정보량*을 기호 집합의 엔트로피라고 하는데 모든 기호들이 동일한 발생 확률을 가질 때 그 기호 집합의 엔트로피는 최댓값을 갖는다.

C : 기호들이 동일한 발생 확률

E : 기호 집합의 엔트로피는 최댓값

- ① 기호들의 발생 확률이 모두 1/2인 경우, 각 기호의 정보량은 동일하다.

C : 기호들이 동일한 발생 확률

E : 기호 집합의 엔트로피는 최댓값

기호 발생확률 같으면 엔트로피 최댓값인데 정보량이 같다??

판단 불가...

- ② 기호들의 발생 확률이 각각 1/4, 3/4인 경우의 평균 정보량이 최댓값이다.

C : 기호들이 동일한 발생 확률

E : 기호 집합의 엔트로피는 최댓값

기호 발생 확률 같으면 엔트로피 최댓값인데

발생확률이 다르네?

아 이거 틀렸네! 정답!

- ③ 기호들의 발생 확률이 각각 1/4, 3/4인 경우, 기호의 정보량이 더 많은 것은 발생 확률이 1/4인 기호이다.
- ④ 기호들의 발생 확률이 모두 1/2인 경우, 기호를 부호화하는 데 필요한 평균 비트 수의 최솟값이 최대가 된다.
- ⑤ 기호들의 발생 확률이 각각 1/4, 3/4인 기호 집합의 엔트로피는 발생 확률이 각각 3/4, 1/4인 기호 집합의 엔트로피와 같다.

다시 한 번 말씀드립니다.

4차 산업혁명시대

사회가 요구하는 인재 육성

국어과의 사명은

국어교과로서의 과목도 있지만

도구교과

즉 정보처리 역량을 길러주는 교과입니다.

4차 산업 혁명 시대의 정보 특성은

Big data

이것을 처리하는 방법은?

특징	해법
Big Data + Data Crawling	정보가치 판단 가치 높은 정보에 집중 Data Structure + I.A

Data Mining	(Information Architecture) 자료 구조와 정보 설계 분석
Integrity (무결성 또는 완결성)	처리 가능한 정보에 집중
결론	
시간 안에 유의미한 정보처리 가능	

연습할만한 ebs 지문 중 특히 작년 수특에 어려웠던 지문을 비롯, 학평 지문을 엮어 추후 커리에서 연습하겠습니다.

등단 근처도 가지 못하는 배작가의 하참은 작품 4.

<엄마의 가방>

문득, 수십 년 넘도록 버리지 못했던 엄마의 가방이 떠오른다.

그 시절, 그리고 지금 우리 엄마는 행복했을까 생각해 본다.

평범한, 때로는 가난했던 그 집에서 태어난 우리 엄마도

꿈이 있었고, 젊음이 있었고, 사랑이 있었고, 삶이 있었다.

그 엄마의 꿈과 젊음과 삶의 대가로 나는 태어났고 자랐다.

단 한 번도 내색은 하지 않았지만,

때로는 엄마는 삶의 무거움에 나를 원망했을지도 모른다.

어린 시절 아무 생각 없이 들어보았던 엄마의 가방은,

그런 엄마의 삶의 무게가 그대로 느껴졌었다.

어느새 나는 우리 엄마가 나를 낳은 나이도 훌쩍 넘겨,

가만히 앉아있는 것만으로도 건강이 좋지 않은 나이가 되었다.

아무리 살아도 삶은 여전히 무겁고, 아프다.

하지만, 엄마는 더 아팠겠지. 그리고 더 무거웠겠지.

항상, 우리 엄마는 그랬다. 자신의 아픔과 무게보다.

내 아픔에 더 아파했고, 내 무게에 더 무거워했다.

평범한 직업을 구해, 평범한 결혼을 해서, 평범한 가정을 꾸리고 살기 바란 우리 엄마의 기대가 맞았는지도 모르겠다.

나는 이런 삶을 견디기에는 너무 약한지도 모르겠다.

하지만, 엄마. 정말 쉽지는 않지만, 누구보다 의미있는 삶, 행복하게 살게요. 그게 엄마가 나를 해준 것들에 대한 유일한 보답 같아요.

- 2020년 2월 14일 아침 9시 21분.

- 아픈 몸을 견디며, 교재를, 작품을 쓰느라 날을 새고, 문득 우리 엄마도 이렇게 힘들고 아팠을까 싶어서.

2018학년도 수능

[38~42] 다음 글을 읽고 물음에 답하시오.

글의 전개 방식

글의 전개 방식

디지털 통신 시스템은 송신기, 채널, 수신기로 구성되며, ① 전송할 데이터를 빠르고 정확하게 전달하기 위해 부호화 과정을 거쳐 전송한다. 영상, 문자 등인 데이터는 ② 기호 집합에 있는 기호들의 조합이다. 예를 들어 기호 집합 {a, b, c, d, e, f}에서 기호들을 조합한 add, cab, beef 등이 데이터이다. 정보량은 어떤 기호가 발생했다는 것을 알았을 때 얻는 정보의 크기이다. 어떤 기호 집합에서 특정 기호의 발생 확률이 높으면 그 기호의 정보량은 적고, 발생 확률이 낮으면 그 기호의 정보량은 많다. 기호 집합의 평균 정보량*을 기호 집합의 엔트로피라고 하는데 모든 기호들이 동일한 발생 확률을 가질 때 그 기호 집합의 엔트로피는 최대값을 갖는다.

첫괄 두침

원리 인과

글의 전개 방식: 과정

송신기에서는 ③ 소스 부호화, 채널 부호화, 선 부호화를 거쳐 기호를 ④ 부호로 변환한다. ① 소스 부호화는 데이터를 압축하기 위해 기호를 0과 1로 이루어진 부호로 변환하는 과정이다. 어떤 기호가 110과 같은 부호로 변환되었을 때 0 또는 1을 비트라고 하며 이 부호의 비트 수는 3이다. 이때 기호 집합의 엔트로피는 기호 집합에 있는 기호를 부호로 표현하는 데 필요한 평균 비트 수의 최솟값이다. 전송된 부호를 수신기에서 원래의 기호로 ⑤ 복원하려면 부호들의 평균 비트 수가 기호 집합의 엔트로피보다 크거나 같아야 한다. 기호 집합을 엔트로피에 최대한 가까운 평균 비트 수를 갖는 부호들로 변환하는 것을 엔트로피 부호화라 한다. 그 중 하나인 '허프만 부호화'에서는 발생 확률이 높은 기호에는 비트 수가 적은 부호를, 발생 확률이 낮은 기호에는 비트 수가 많은 부호를 할당한다.

비트 18.9 EBS 연계

그림 이해

원리 인과

원리 이해 하향식

채널 부호화는 오류를 검출하고 정정하기 위하여 부호에 잉여 정보를 추가하는 과정이다. 송신기에서 부호를 전송하면 채널의 잡음으로 인해 오류가 발생하는데 이 문제를 해결하기 위해 잉여 정보를 덧붙여 전송한다. 채널 부호화 중 하나인 '삼중 반복 부호화'는 0과 1을 각각 000과 111로 부호화한다. 이때 수신기에서는 수신한 부호에 0이 과반수인 경우에는 0으로 판단하고, 1이 과반수인 경우에는 1로 판단한다. 즉 수신기에서 수신된 부호가 000, 001, 010, 100 중 하나라면 0으로 판단하고, 그 이외에는 1로 판단한다. 이렇게 하면 000을 전송했을 때 하나의 비트에서 오류가 생겨 001을 수신해도 0으로 판단하므로 오류는 정정된다. 채널 부호화를 하기 전 부호의 비트 수를, 채널 부호화를 한 후 부호의 비트 수로 나눈 것을 부호율이라 한다. 삼중 반복 부호화의 부호율은 약 0.33이다. 수치 발췌

원리 이해 하향식

채널 부호화를 거친 부호들을 채널을 통해 전송하려면 부호들을 전기 신호로 변환해야 한다. 0 또는 1에 해당하는 전기 신호의 전압을 결정하는 과정이 선 부호화이다. 전압의 결정 방법은 선 부호화 방식에 따라 다르다. 선 부호화

중 하나인 '차동 부호화'는 부호의 비트가 0이면 전압을 유지하고 1이면 전압을 변화시킨다. 차동 부호화를 시작할 때는 기준 신호가 필요하다. 예를 들어 차동 부호화 직전의 기준 신호가 양(+)의 전압이라면 부호 0110은 '양, 음, 양, 양'의 전압을 갖는 전기 신호로 변환된다. 수신기에서는 송신기와 동일한 기준 신호를 사용하여, 전압의 변화가 있으면 1로 판단하고 변화가 없으면 0으로 판단한다.

* 평균 정보량 : 각 기호의 발생 확률과 정보량을 서로 곱하여 모두 더한 것.

정보가치+하향식+원리이해

38. 뒷글에서 알 수 있는 내용으로 적절한 것은?

발췌가능 but 하향식+정보가치

- ① 영상 데이터는 채널 부호화 과정에서 압축된다. 추가
- ② 수신기에는 부호를 기호로 복원하는 기능이 있다. 상식
- ③ 잉여 정보는 데이터를 압축하기 위해 추가한 정보이다. 상식
- ④ 영상을 전송할 때는 잡음으로 인한 오류가 발생하지 않는다. 상식
- ⑤ 소스 부호화는 전송할 기호에 정보를 추가하여 오류에 대비하는 과정이다. 채널 압축

주제+구조(첫괄 두침)

39. 뒷글을 바탕으로, 2가지 기호로 이루어진 기호 집합에 대해 이해한 내용으로 적절하지 않은 것은?

- ① 기호들의 발생 확률이 모두 1/2인 경우, 각 기호의 정보량은 동일하다. 동일
- ② 기호들의 발생 확률이 각각 1/4, 3/4인 경우의 평균 정보량이 최대값이다. 다르다
- ③ 기호들의 발생 확률이 각각 1/4, 3/4인 경우, 기호의 정보량이 더 많은 것은 발생 확률이 1/4인 기호이다.
- ④ 기호들의 발생 확률이 모두 1/2인 경우, 기호를 부호화하는 데 필요한 평균 비트 수의 최솟값이 최대가 된다.
- ⑤ 기호들의 발생 확률이 각각 1/4, 3/4인 기호 집합의 엔트로피는 발생 확률이 각각 3/4, 1/4인 기호 집합의 엔트로피와 같다.

일단 skip

처리 가능한 정보에 집중

40. 윗글의 '부호화'에 대한 내용으로 적절한 것은?

- ① 선 부호화에서는 수신기에서 부호를 전기 신호로 변환한다. ? 구조, 주제
- ② 허프만 부호화에서는 정보량이 많은 기호에 상대적으로 비트 수가 적은 부호를 할당한다. ? 정보 처리 못함
- ③ 채널 부호화를 거친 부호들은 채널로 전송하기 전에 잉여 정보를 제거한 후 선 부호화한다. 주제, 원리
- ④ 채널 부호화 과정에서 부호에 일정 수준 이상의 잉여 정보를 추가하면 부호율은 1보다 커진다. 수치 발취득
- ⑤ 삼중 반복 부호화를 이용하여 0을 부호화한 경우, 수신된 부호에서 두 개의 비트에 오류가 있으면 오류는 정정되지 않는다. 원리, 주제 이해

처리 가능한 정보에 집중

41. 윗글을 바탕으로 <보기>를 이해한 내용으로 적절한 것은?

<보기>

날씨 데이터를 전송하려고 한다. 날씨의 '맑음', '흐림', '비', '눈'으로만 분류하며, 각 날씨의 발생 확률은 모두 같다. 엔트로피 부호화를 통해 '맑음', '흐림', '비', '눈'을 각각 00, 01, 10, 11의 부호로 바꾼다.

- ① 기호 집합 {맑음, 흐림, 비, 눈}의 엔트로피는 2보다 크겠군. ?
- ② 엔트로피 부호화를 통해 4일 동안의 날씨 데이터 '흐림비 맑음흐림'은 '01001001'로 바뀌겠군. 원리 이해
- ③ 삼중 반복 부호화를 이용하여 전송한 특정 날씨의 부호를 '110001'과 '101100'으로 각각 수신하였다면 서로 다른 날씨로 판단하겠군. 10 101 => 1 100 => 0
- ④ 날씨 '비'를 삼중 반복 부호화와 차동 부호화를 이용하여 부호화하는 경우, 기준 신호가 양(+)의 전압이면 '음, 양, 음, 음, 음, 음'의 전압을 갖는 전기 신호로 변환되겠군.
- ⑤ 삼중 반복 부호화와 차동 부호화를 이용하여 특정 날씨의 부호를 전송할 경우, 수신기에서 '음, 음, 음, 양, 양, 양'을 수신했다면 기준 신호가 양(+)의 전압일 때 '흐림'으로 판단하겠군.

110 => 1
001 => 0

42. 문맥을 고려할 때, 밑줄 친 말이 ㉠~㉥의 동음이의어가 아닌 것은?

- ① ㉠ : 공항에서 해외로 떠나는 친구를 전송(餞送)할 계획이다.
- ② ㉡ : 대중의 기호(嗜好)에 맞추어 상품을 개발한다.
- ③ ㉢ : 나는 가난하지만 귀족이나 부호(富豪)가 부럽지 않다.
- ④ ㉣ : 한번 금이 간 인간관계를 복원(復原)하기는 어렵다.
- ⑤ ㉤ : 이 작품은 그 화가의 오랜 노력의 결정(結晶)이다.

10=> 111000
+ - + - - - -