Show All Answers Hide All Answers

6.02 Practice Problems: Error Correcting Codes

Problem 1. For each of the following sets of codewords, please give the appropriate (n,k,d) designation where n is number of bits in each codeword, k is the number of message bits transmitted by each code word and d is the minimum Hamming distance between codewords. Also give the code rate.

A. {111, 100, 001, 010}

Hide Answer

n=3, k=2 (there are 4 codewords), d = 2. The code rate is 2/3.

B. {00000, 01111, 10100, 11011}

Hide Answer

n=5, k=2 (there are 4 codewords), d = 2. The code rate is 2/5.

C. {00000}

Hide Answer

A bit of a trick question: n=5, k=0, d = undefined. The code rate is 0 -- since there's only one codeword the receiver can already predict what it will receive, so no useful information is transferred.

Problem 2. Suppose management has decided to use 20-bit data blocks in the company's new (n,20,3) error correcting code. What's the minimum value of n that will permit the code to be used for single bit error correction?

Hide Answer

`n` and `k=20` must satisfy the constraint that $n + 1 \le 2^{n-k}$. A little trial-and-error search finds `n=25`.

Problem 3. The Registrar has asked for an encoding of class year ("Freshman", "Sophomore", "Junior", "Senior") that will allow single error correction. Please give an appropriate 5-bit binary encoding for each of the four years.

Hide Answer

We want a (5,2,3) block code. For such a code, 00000 is a codeword by definition. Every other codeword must have weight at least 3, and 00111 is an obvious choice (or any permutation thereof).

We now need only two more codewords and each must have at least three ones, and must also have a Hamming distance of 3 from the second codeword above. A little bit of trial and error shows that 11011 and 11100 work.

So: 00000, 00111, 11011, 11100 should satisfy the Registrar

Problem 4. For any block code with minimum Hamming distance at least 2t + 1 between code words, show that:

$$2^{n-k} \geq 1 + \binom{n}{1} + \binom{n}{2} + \dots \binom{n}{t}.$$

Hide Answer

An (n, k) block code can represent in its parity bits at most 2^{n-k} patterns, and these must cover all the error cases we wish to correct, as well as the one case with no errors. When the minimum Hamming distance is 2t + 1, the code can correct up to t errors. The number of ways in which the transmission can experience 0,1,2,...,t errors is equal to 1 + (n choose 1) + (n choose 2) + ... + (n choose t), and clearly this number must not exceed 2^{n-k}.

Problem 5. Pairwise Communications has developed a block code with three data (D1, D2, D3) and three parity bits (P1, P2, P3):

```
P1 = D1 + D2
P2 = D2 + D3
P3 = D3 + D1
```

A. What is the (n,k,d) designation for this code.

   Hide Answer

   n = 6, k = 3. To determine d, list the 8 possible codewords (we'll use the order D1, D2, D3, P1, P2, P3):

   ```
   000000
   001011
   010110
   011101
   100101
   101110
   110011
   111000
   ```

   By inspection, the minimum Hamming distance d = 3.

B. The receiver computes three syndrome bits from the (possibly corrupted) received data and parity bits:

   ```
   E1 = D1 + D2 + P1
   E2 = D2 + D3 + P2
   E3 = D3 + D1 + P3.
   ```

   The receiver performs maximum likelihood decoding using the syndrome bits. For the combinations of syndrome bits listed below, state what the maximum-likelihood decoder believes has occured: no errors, a single error in a speciﬁc bit (state which one), or multiple errors.

   ```
   E1 E2 E3 = 000
   E1 E2 E3 = 010
   E1 E2 E3 = 101
   E1 E2 E3 = 111
   ```

   Hide Answer

   E1 E2 E3 = 000. No errors.
   E1 E2 E3 = 010. Error in P2.
   E1 E2 E3 = 101. Error in D1.
   E1 E2 E3 = 111. Multiple errors.

Problem 6. Dos Equis Encodings, Inc. specializes in codes that use 20-bit transmit blocks. They are trying to design a (20, 16) linear block code for single error correction. Explain whether they are likely to succeed or not.

Hide Answer

A single error correcting code must be able to uniquely identify n + 1 patterns (n error patterns and one without any errors). So 2^{n-k} must be $\geq$ n+1. That is not true when n=20 and k=16.

Problem 7. Consider the following (n,k,d) block code:

```
D0   D1   D2   D3   D4    | P0
D5   D6   D7   D8   D9    | P1
D10  D11  D12  D13  D14   | P2
-------------------------
P3   P4   P5   P6   P7    |
```

where D0-D14 are data bits, P0-P2 are row parity bits and P3-P7 are column parity bits. The transmitted code word will be:

D0 D1 D2 ... D13 D14 P0 P1 ... P6 P7

A. Please give the values for n, k, d for the code above.

   Hide Answer

   n=23, k=15, d=3.

   For a discussion of why d=3, see section 6.4.1 in the notes.

B. If `D0 D1 D2 ... D13 D14 = 0 1 0 1 0, 0 1 0 0 1, 1 0 0 0 1`, please compute P0 through P7.

   Hide Answer

   P0 through P7 = `0 0 0 1 0 0 1 0`

C. Now we receive the four following code words:

```
M1: 0 1 0 1 0, 0 1 0 0 1, 1 0 0 0 1, 0 0 0 1 1 0 1 0
M2: 0 1 0 1 0, 0 1 0 0 1, 1 0 0 0 1, 0 0 1 1 1 0 1 0
M3: 0 1 0 1 0, 0 1 0 0 1, 1 0 0 0 1, 1 1 0 1 1 0 1 0
M4: 0 1 0 1 0, 0 1 0 0 1, 1 0 0 0 1, 1 0 0 1 1 0 1 0
```

   For each of received code words, indicate the number of errors. If there are errors, indicate if they are correctable, and if they are, what the correction should be.

   Hide Answer

   M1:

```
0 1 0 1 0 | 0 | 0
0 1 0 0 1 | 0 | 0
1 0 0 0 1 | 0 | 0
---------------
1 1 0 1 0 |
---------------
0 1 0 0 0
```

   The syndrome bits are shown in red. Since there is only one error -- the parity bit P4 -- it must be that parity bit itself that had the error. So the correct data bits are:

```
M1: 0 1 0 1 0, 0 1 0 0 1, 1 0 0 0 1
```

M2:

```
0 1 0 1 0 | 0 | 0
0 1 0 0 1 | 0 | 0
1 0 0 0 1 | 1 | 1
----------------
1 1 0 1 0 |
----------------
0 1 0 0 0   0
```

There are parity errors detected for the third row and second column, so D11 must be in error. So the correct data bits are:

```
M2: 0 1 0 1 0, 0 1 0 0 1, 1 1 0 0 1
```

M3:

```
0 1 0 1 0 | 1 | 1
0 1 0 0 1 | 1 | 1
1 0 0 0 1 | 0 | 0
----------------
1 1 0 1 0 | 0
----------------
0 1 0 0 0   0
```

There are more than two row/column parity errors, indicating an uncorrectable multi-bit error.

M4:

```
0 1 0 1 0 | 1 | 1
0 1 0 0 1 | 0 | 0
1 0 0 0 1 | 0 | 0
----------------
1 1 0 1 0 |
----------------
0 1 0 0 0
```

The row and column parity bit indicate the data bit which got flipped (D1). Flipping it gives us the correct data bits:

```
M4: 0 0 0 1 0, 0 1 0 0 1, 1 0 0 0 1
```

Problem 8. The following matrix shows a rectangular single error correcting code consisting of 9 data bits, 3 row parity bits and 3 column parity bits. For each of the examples that follow, please indicate the correction the receiver must perform: give the position of the bit that needs correcting (e.g., D7, R1), or "no" if there are no errors, or "M" if there is a multi-bit uncorrectable error.

D1	D2	D3	R1
D4	D5	D6	R2
D7	D8	D9	R3
C1	C2	C3	—

1	1	1	1
1	1	0	0
0	1	1	0
0	1	1	–

0	0	1	0
0	1	1	1
0	1	1	0
0	0	1	–

1	1	0	0
1	0	0	1
0	0	1	0
0	0	1	–

1	0	1	0
1	0	0	1
1	1	0	0
1	1	1	–

1	1	1	0
1	1	1	1
1	0	0	0
0	0	1	–

Hide Answer

1) Error in column 3 parity bit.
2) M.
3) Error in D8.
4) No errors.
5) M.

Problem 9. Consider two convolutional coding schemes - I and II. The generator polynomials for the two schemes are

Scheme I: G0 = 1101, G1 = 1110
Scheme II: G0 = 110101, G1 = 111011

Notation is follows: if the generator polynomial is, say, 1101, then the corresponding parity bit for message bit `n` is

(x[n] + x[n-1] + x[n-3]) mod 2

where x[n] is the message sequence.

A. Indicate TRUE or FALSE
   a. Code rate of Scheme I is 1/4.
   b. Constraint length of Scheme II is 4.
   c. Code rate of Scheme II is equal to code rate of Scheme I.
   d. Constraint length of Scheme I is 4.

   Hide Answer

   The code rate (r) and constraint length (k) for the two schemes are

   I: r = 1/2, k = 4
   II: r = 1/2, k = 6

   So

   a. false
   b. false
   c. true
   d. true

B. How many states will there be in the state diagram for Scheme I? For Scheme II?

   Hide Answer

   Number of states is given by 2^{k-1} where k = constraint length. Following the convention of state machines as outlined in lecture, number of states in Scheme I is 8 and in Scheme II, 32.

C. Which code will lead to a lower bit error rate? Why?

   Hide Answer

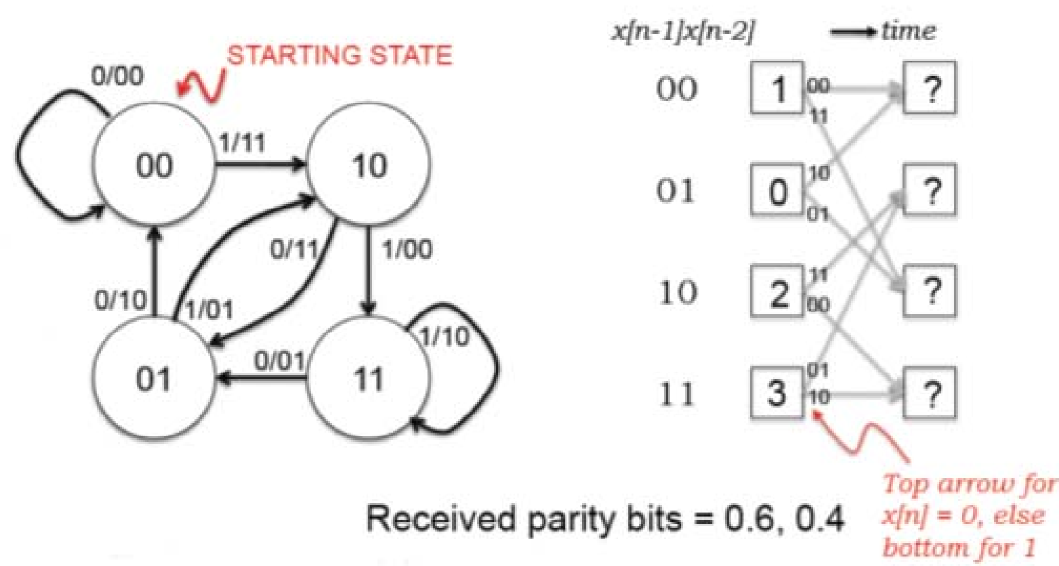
   Scheme II is likely to lead to a lower bit error rate. Both codes have the same code rate but different constraint lengths. So Scheme II encodes more history and since it is less likely that 6 trailing bits will be in error vs. 4 trailing bits, II is stronger.

D. Alyssa P. Hacker suggests a modification to Scheme I which involves adding a third generator polynomial G2 = 1001. What is the code rate `r` of Alyssa's coding scheme? What about constraint length `k`? Alyssa claims that her scheme is stronger than Scheme I. Based on your computations for `r` and `k`, is her statement true?

Hide Answer

For Alyssa's scheme `r` = 1/3, `k` = 4. Alyssa's code has a lower code rate (more redundancy), and given then she's sending *additional* information, the modified scheme I is stronger in the sense that more information leads to better error detection and correction.

Problem 10. Consider a convolution code that uses two generator polynomials: G0 = 111 and G1 = 110. You are given a particular snapshot of the decoding trellis used to determine the most likely sequence of states visited by the transmitter while transmitting a particular message:



A. Complete the Viterbi step, i.e., fill in the question marks in the matrix, assuming a hard branch metric based on the Hamming distance between expected an received parity where the received voltages are digitized using a 0.5V threshold.

Hide Answer

The digitized received parity bits are 1 and 0.

For state 0:

PM[0,n] = min(PM[0,n-1]+BM(00,10), PM[1,n-1]+BM(10,10)) = min(1+1,0+0) = 0
Predecessor[0,n] = 1

For state 1:

PM[1,n] = min(PM[2,n-1]+BM(11,10), PM[3,n-1]+BM(01,10)) = min(2+1,3+2) = 3
Predecessor[1,n] = 2

For state 2:

PM[2,n] = min(PM[0,n-1]+BM(11,10), PM[1,n-1]+BM(01,10)) = min(1+1,0+2) = 2
Predecessor[2,n] = 0 or 1

For state 3:

PM[1,n] = min(PM[2,n-1]+BM(00,10), PM[3,n-1]+BM(10,10)) = min(2+1,3+0) = 3
Predecessor[1,n] = 2 or 3

B. Complete the Viterbi step, i.e., fill in the question marks in the matrix, assuming a soft branch metric based on the square of the Euclidean distance between expected an received parity voltages. Note that your branch and path metrics will not necessarily be integers.

Hide Answer

For state 0:

PM[0,n] = min(PM[0,n-1]+BM([0,0],[0.6,0.4]), PM[1,n-1]+BM([1,0],[0.6,0.4])) = min(1+0.52,0+.32) = .32
Predecessor[0,n] = 1

For state 1:

PM[1,n] = min(PM[2,n-1]+BM([1,1],[0.6,0.4]), PM[3,n-1]+BM([0,1],[0.6,0.4])) = min(2+0.52,3+0.72) = 2.52
Predecessor[1,n] = 2

For state 2:

PM[2,n] = min(PM[0,n-1]+BM([1,1],[0.6,0.4]), PM[1,n-1]+BM([0,1],[0.6,0.4])) = min(1+0.52,0+0.72) = 0.72
Predecessor[2,n] = 1

For state 3:

PM[1,n] = min(PM[2,n-1]+BM([0,0],[0.6,0.4]), PM[3,n-1]+BM([1,0],[0.6,0.4])) = min(2+0.52,3+.32) = 2.52
Predecessor[1,n] = 2

C. Does the soft metric give a different answer than the hard metric? Base your response in terms of the relative ordering of the states in the second column and the survivor paths.

Hide Answer

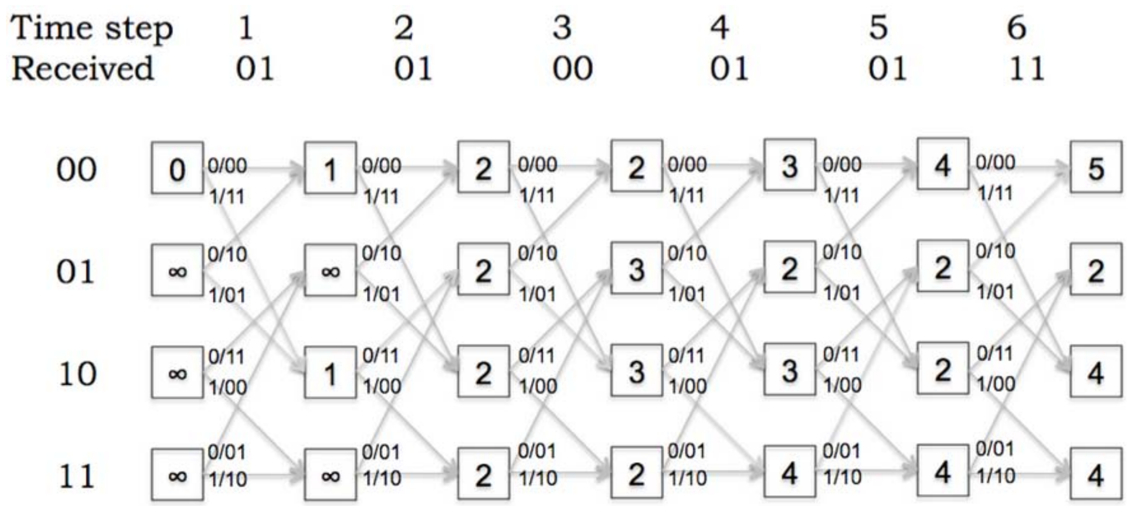
The soft metric certainly gives different path metrics, but the relative ordering of the likelihood of each state remains unchanged. Using the soft metric, the choice of survivor path leading to states 2 and 3 has firmed up (with the hard metric either of the survivor paths for each of states 2 and 3 could have been chosen).

D. If the transmitted message starts with the bits "01011", what is the sequence of bits produced by the convolutional encoder?

Hide Answer

sequence produced by encoder: 00 11 11 01 00.

The receiver determines the most-likely transmitted message by using the Viterbi algorithm to process the (possibly corrupted) received parity bits. The path metric trellis generated from a particular set of received parity bits is shown below. The boxes in the trellis contain the minimum path metric as computed by the Viterbi algorithm.



E. Referring to the trellis above, what is the receiver's estimate of the most-likely transmitter state after processing the bits received at time step 6?

Hide Answer

Most-likely transmitter state = state with smallest path metric = state 01.

F. Referring to the trellis above, show the most-likely path through the trellis by placing a circle around the appropriate state box at each time step and darkening the appropriate arcs. What is the receiver's estimate of the most-likely transmitted message?

Hide Answer

Tracing backwards through the trellis:

```
after state 6: state 01  (most-likely final state)
after state 5: state 10
after state 4: state 01
after state 3: state 11
after state 2: state 10
after state 1: state 00
```

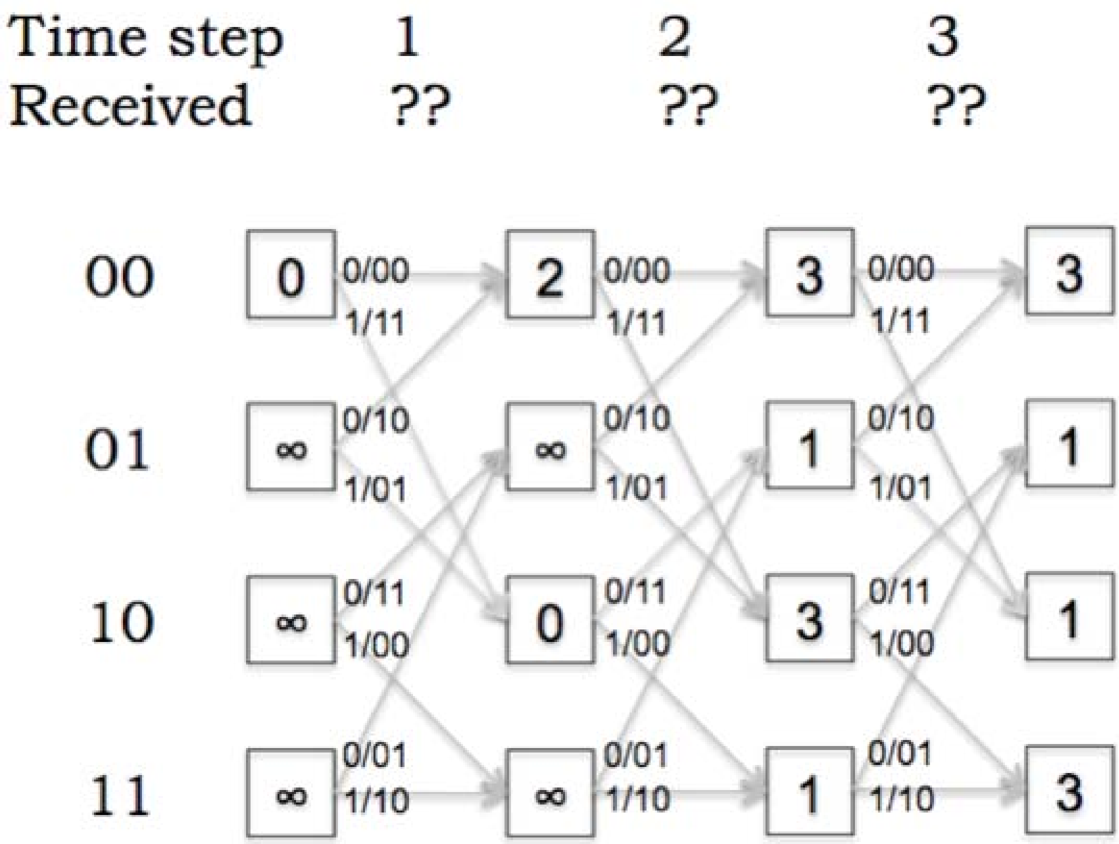
The message can be read off from the high-order bit of the transmitter state (now moving foward through the trellis): 011010.

G. Referring to the trellis above, and given the receiver's estimate of the most-likely transmitted message, at what time step(s) were errors detected by the receiver? Briefly explain your reason- ing.

Hide Answer

At time steps 1 and 2, where the path metric increments along the most-likely path.

H. Now consider the path metric trellis generated from a *different* set of received parity bits.



Referring to the trellis above, determine which pair(s) of parity bits could have been been received at time steps 1, 2 and 3. Briefly explain your reasoning.

Hide Answer

At time 1, the transition from state 00 to 10 has a branch metric BM(??,11)=0, so the parity bits must have been 11.

At time 2, looking at the transition from state 10 to states 01 and 11, we see that BM(??,11)=1 and BM(??,00)=1, so the possible pairs of parity bits are 01 and 10.

At time 3, looking at the transition from state 01 to state 10, we see that BM(??,01)=0, so the parity bits are 01.

Problem 11. Consider a binary convolutional code specified by the generators (1011, 1101, 1111).

A. What are the values of
    a. constraint length of the code
    b. rate of the code
    c. number of states at each time step of the trellis
    d. number of branches transitioning into each state
    e. number of branches transitioning out of each state
    f. number of expected parity bits on each branch

Hide Answer

a. 4
b. 1/3

c. $2^{4-1} = 8$
d. 2
e. 2
f. 3

A 10000-bit message is encoded with the above code and transmitted over a noisy channel. During Viterbi decoding at the receiver, the state 010 had the lowest path metric (a value of 621) in the final time step, and the survivor path from that state was traced back to recover the original message.

B. What is the likely number of bit errors that are corrected by the decoder? How many errors are likely left uncorrected in the decoded message?

Hide Answer

621 errors were likely corrected by the decoder to produce the final decoded message. We cannot infer the number of uncorrected errors still left in the message absent more information (like, say, the original transmitted bits).

C. If you are told that the decoded message had no uncorrected errors, can you guess the approximate number of bit errors that would have occured had the 10000 bit message been transmitted without any coding on the same channel?

Hide Answer

3*10000 bits were transmitted over the channel and the received message had 621 bit errors (all corrected by the convolutional code). Therefore, if the 10000-bit message would have been transmitted without coding, it would have had approximately 621/3 = 207 errors.

D. From knowing the final state of the trellis (010, as given above), can you infer what the last bit of the original message was? What about the last-but-one bit? The last 4 bits?

Hide Answer

The state gives the last 3 three bits of the original message. In general, for a convolutional code with a constraint length k, the state indicates the final k-1 bits of the original message. To determine more bits we would need to know the states along the most-likely path as we trace back through the trellis.

Consider a transition branch between two states on the trellis that has 000 as the expected set of parity bits. Assume that 0V and 1V are used as the signaling voltages to transmit a 0 and 1 respectively, and 0.5V is used as the digitization threshold.

E. Assuming hard decision decoding, which of the two set of received voltages will be considered more likely to correspond to the expected parity bits on the transition: (0V, 0.501V, 0.501V) or (0V, 0V, 0.9V)? What if one is using soft decision decoding?

Hide Answer

With hard decision decoding: (0V, 0.501V, 0.501V) -> 011 -> hamming distance of 2 from expected parity bits. (0V, 0V, 0.9V) -> 001 -> hamming distance of 1. Therefore, (0V, 0V, 0.9V) is considered more likely.

With soft decision decoding, (0V, 0.501V, 0.501V) will have a branch metric of approximately 0.5. (0V, 0V, 0.9V) will have a metric of approximmately 0.8. Therefore, (0V, 0.501V, 0.501V) will be considered more likely.

Problem 12. Indicate whether each of the statements below is true or false, and a brief reason why you think so.

A. If the number states in the trellis of a convolutional code is S, then the number of survivor paths at any point of time is S. Remember that if there is "tie" between to incoming branches (i.e., they both result in the same path metric), we arbitrarilly choose only one as the predecessor.

Hide Answer

True. There is one survivor per state.

The path metric of a state s1 in the trellis indicates the number of residual uncorrected errors left along the trellis path from the start state to s1.

Hide Answer

False. It indicates the number of likely corrected errors.

B. Among the survivor paths left at any point during the decoding, no two can be leaving the same state at any stage of the trellis.

Hide Answer

False. In fact, the survivor paths will likely merge at a certain stage in the past, at which point all of then will emerge from the same state.

C. Among the survivor paths left at any point during the decoding, no two can be entering the same state at any stage of the trellis. Remember that if there is "tie" between to incoming branches (i.e., they both result in the same path metric), we arbitrarilly choose only one as the predecessor.

Hide Answer

True. When two paths merge at any state, only one of them will ever be chosen as a survivor path.

D. For a given state machine of a convolutional code, a particular input message bit stream always produces the same output parity bits.

Hide Answer

False. The same input stream with different start states will produce different output parity bits.

Problem 13. Consider a convolution code with two generator polynomials: G0=101 and G1=110.

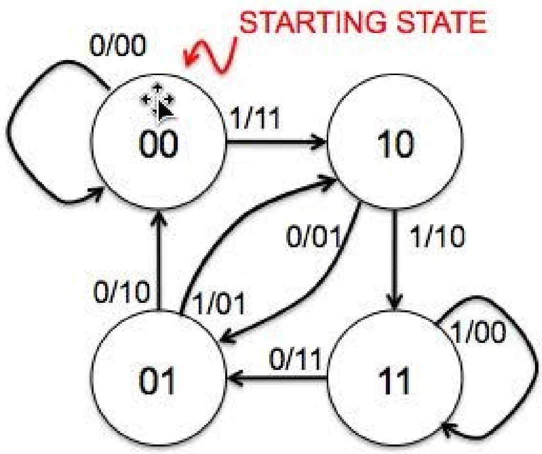
A. What is code rate r and constraint length k for this code?

Hide Answer

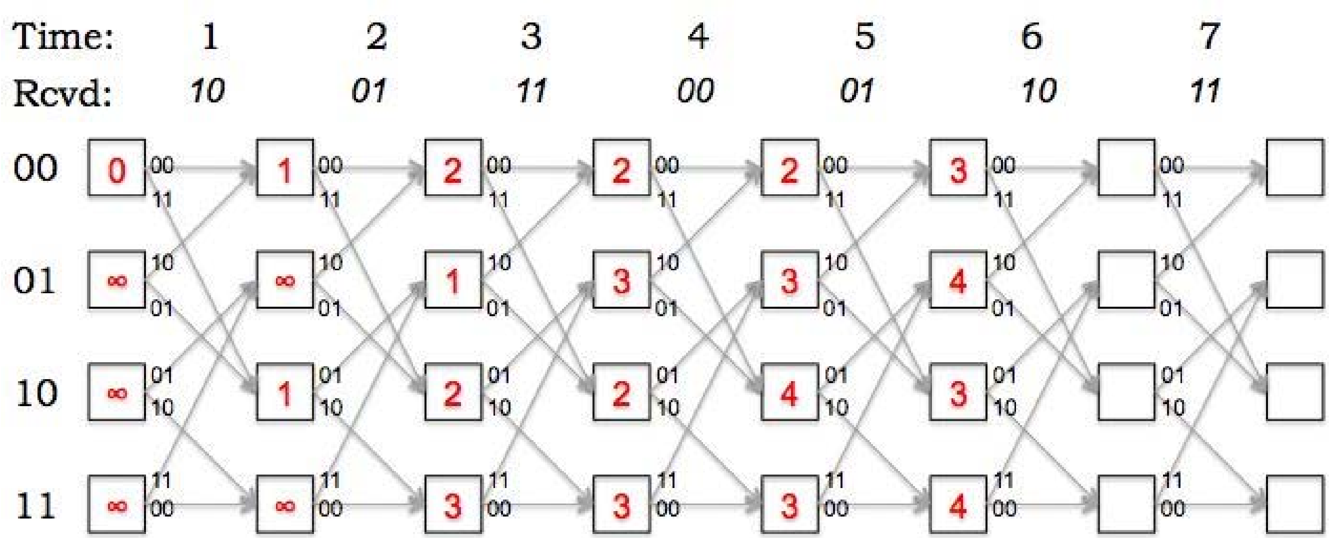
We send two parity bits for each message bit, so the code rate r is 1/2. Three message bits are involved in the computation of the parity bits, so the constraint length k is 3.

B. Draw the state transition diagram for a transmitter that uses this convolutional code. The states should be labeled with the binary string $x_{n-1} \ldots x_{n-k+1}$ and the arcs labeled with x_n/p_0p_1 where $x[n]$ is the next message bit and p_0 and p_1 are the two parity bits computed from G0 and G1 respectively.

Hide Answer

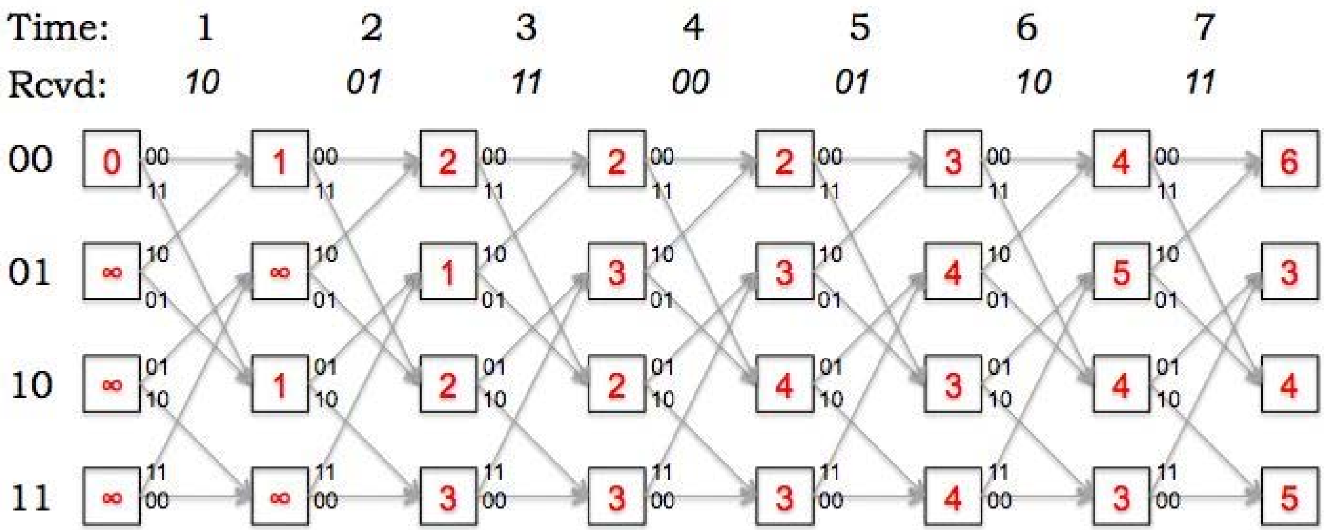


The figure below is a snapshot of the decoding trellis showing a particular state of a maximum likelihood decoder implemented using the Viterbi algorithm. The labels in the boxes show the path metrics computed for each state after receiving the incoming parity bits at time `t`. The labels on the arcs show the expected parity bits for each transition; the actual received bits at each time are shown above the trellis.



C. Fill in the path metrics in the empty boxes in the diagram above (corresponding to the Viterbi calculations for times 6 and 7).

Hide Answer



D. Based on the updated trellis, what is the most-likely final state of the transmitter? How many errors were detected along the most-likely path to the most-likely final state?

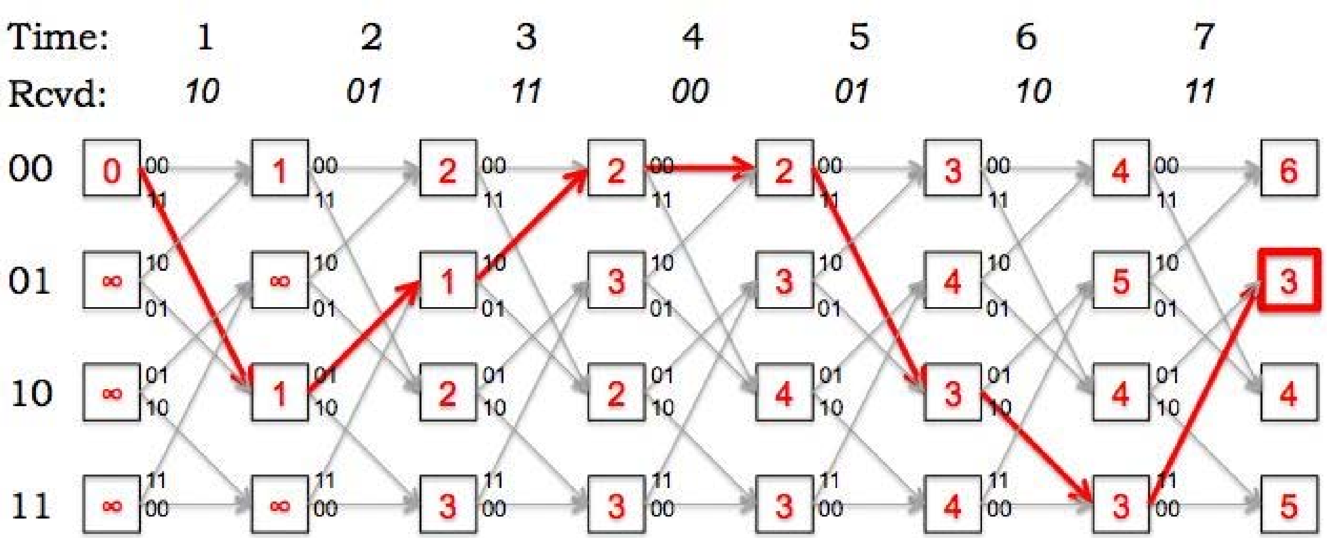
Hide Answer

The most-likely final state is 01, the state with the smallest path metric. The path metric tells us the total number of errors along the most-likely path leading to the state. In this example there were 3 errors altogether.

E. What's the most-likely path through the trellis (i.e., what's the most-likely sequence of states for the transmitter)? What's the decoded message?

Hide Answer

The most-likely path has been highlighted in red below. The decoded message can be read off from the state transitions along the most-likely path: 1000110.



F. Based on your choice of the most-likely path through the trellis, at what times did the errors occur?

Hide Answer

The path metric is incremented for each error along the path. Looking at the most-likely path we can see that there were single-bit errors at times 1, 3 and 5.