- 作业讲解
  - TC第31.1节练习12、13
  - TC第31.2节练习4、5、6、9
  - TC第31.3节练习5
  - TC第31.4节练习2、3
  - TC第31.5节练习2、3
  - TC第31.6节练习2、3

TC第31.2节练习4

EUCLID (a,b)

```
1. while b ≠ 0
2.      t = b
3.      b = a % b
4.      a = t
5. return a
```

算法比较简单，但要注意最后一步应返回a。

TC第31.2节练习5

$$F_{k+1} \approx {\varnothing^{k+1}}\big/{\sqrt{5}}$$

由Theorem 31.11得： $b < {\varnothing^{k+1}}\big/{\sqrt{5}}$

可得：$k \le 1 + \log_{\varnothing} b$

所以调用次数至多为$1 + \log_{\varnothing} b$

该情况是最坏情况下的复杂度，即b|a =1

当算法运行到结果gcd(a,b)时，立刻返回，即相当于求EUCLID(a/gcd(a,b),b/gcd(a,b))的复杂度，由上可知：$k \le 1 + \log_{\varnothing} b/gcd(a,b)$。

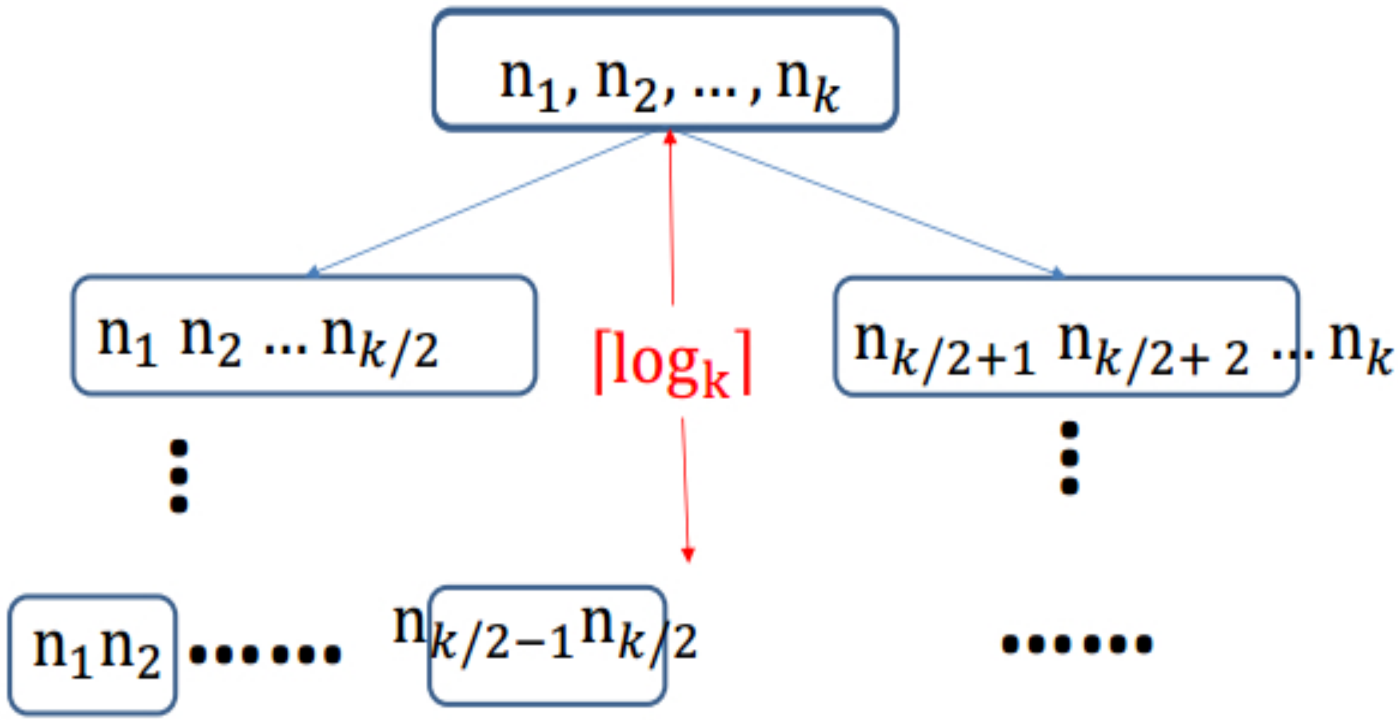
TC第31.2节练习9

将$n_1, n_2, \ldots, n_k$分成两个部分A,B，其中n_i只会在其中一部分中出现。如:

$$A = n_1 n_2 \ldots n_{k/2}$$

$$B = n_{\frac{k}{2}+1} n_{\frac{k}{2}+2} \ldots n_k$$

由gcd(A, B) = 1可知，A中的所有数与B中的所有数都互质，下面只需说明A、B内的所有数两两互质即可，利用递归思想，构建树：

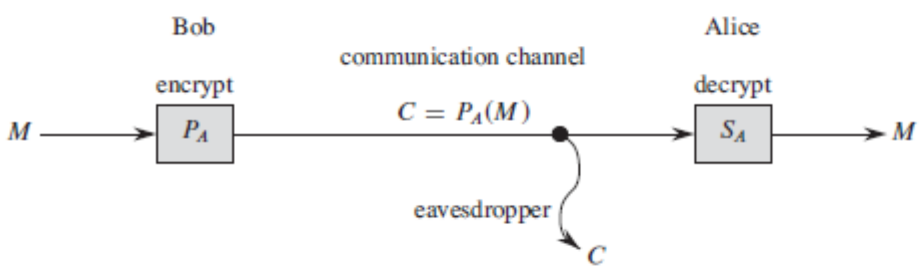


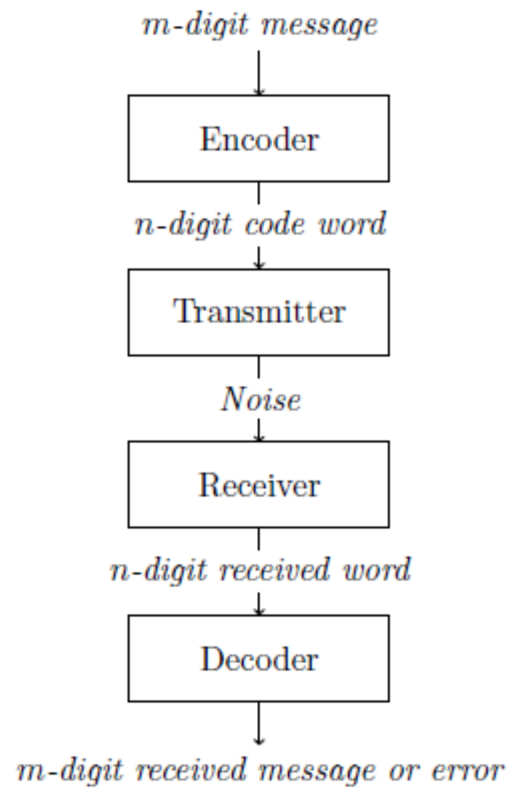
树高为$\lceil log_k \rceil$，每一层中n_i都出现只一次

- 教材讨论
  - TJ第8章

问题1：编码

- 同样是“编码→信道→解码”，你认为这两周讨论的问题有哪些区别？
- 你能结合这两个公式解释编码、查错、解码的具体步骤吗？
  - $Gx=y$
  - $Hy=0$





问题2：奇偶校验

- 上周我们提到过简单的奇偶校验码(m+1)，现在你对它有什么新的认识？你能用这周所学内容来解释它吗？

$$H = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{x} = (x_1, x_2, x_3, x_4, x_5, x_6)^{\mathsf{T}}$$

$$\mathbf{0} = H\mathbf{x} = \begin{pmatrix} x_2 + x_3 + x_4 \\ x_1 + x_2 + x_5 \\ x_1 + x_3 + x_6 \end{pmatrix}$$

$$H = (1\ 1\ 1\ 1)$$

$$x = (x_1, x_2, x_3, x_4)^T$$

$$0 = Hx = x_1 + x_2 + x_3 + x_4$$

Theorem 8.7 *Let $H \in \mathbb{M}_{m\times n}(\mathbb{Z}_2)$ be a canonical parity-check matrix. Then* $\mathrm{Null}(H)$ *consists of all $\mathbf{x} \in \mathbb{Z}_2^n$ whose first $n-m$ bits are arbitrary but whose last m bits are determined by $H\mathbf{x} = \mathbf{0}$. Each of the last m bits serves as an even parity check bit for some of the first $n-m$ bits. Hence, H gives rise to an $(n, n-m)$-block code.*

问题2：奇偶校验 (续)

- 现在，你学习Hamming code是不是更容易了？

The following general algorithm generates a single-error correcting (SEC) code for any number of bits.

1. Number the bits starting from 1: bit 1, 2, 3, 4, 5, etc.
2. Write the bit numbers in binary: 1, 10, 11, 100, 101, etc.
3. All bit positions that are powers of two (have only one 1 bit in the binary form of their position) are parity bits: 1, 2, 4, 8, etc. (1, 10, 100, 1000)
4. All other bit positions, with two or more 1 bits in the binary form of their position, are data bits.
5. Each data bit is included in a unique set of 2 or more parity bits, as determined by the binary form of its bit position.

Bit position		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	
Encoded data bits		**p1**	**p2**	**d1**	**p4**	**d2**	**d3**	**d4**	**p8**	**d5**	**d6**	**d7**	**d8**	**d9**	**d10**	**d11**	**p16**	**d12**	**d13**	**d14**	**d15**	
Parity bit coverage	**p1**	X		X		X		X		X		X		X		X		X		X		
	p2		X	X			X	X			X	X			X	X			X	X		...
	p4				X	X	X	X					X	X	X	X					X	
	p8								X	X	X	X	X	X	X	X						
	p16																X	X	X	X	X	

- Hamming code怎么编码？怎么解码？怎么查错？怎么纠错？
- 同样是奇偶校验码，m+1和Hamming code各有什么优缺点？

问题2：奇偶校验 (续)

- 如果我们用Hamming code将4位数据编码为7位，你能根据G和H在编码、查错、解码中的用法，直接写出Hamming code对应的G和H吗？

Bit position		1	2	3	4	5	6	7
Encoded data bits		p1	p2	d1	p4	d2	d3	d4
Parity bit	p1	X		X		X		X
	p2		X	X			X	X
	p4				X	X	X	X

$\mathbf{x}=(d1,d2,d3,d4)^T$
$\mathbf{y}=(p1,p2,d1,p4,d2,d3,d4)^T$

G

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

H

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

- 你的结果和教材中的形式相符吗？如果不，你能解释吗？

$$H = (A \mid I_m)$$

$$G = \begin{pmatrix} I_{n-m} \\ \hline A \end{pmatrix}$$

问题3：linear code

- 实际上我们只是要找一种奇偶校验码，为什么要刻意选择 linear code，它的特殊性质能给我们带来什么好处？

问题3：linear code (续)

A code is a *linear code* if it is determined by the null space of some matrix $H \in \mathbb{M}_{m\times n}(\mathbb{Z}_2)$.

- 你觉得“linear”在这里是什么意思？
  - codeword的linear combination仍是codeword
  - 即：所有codeword构成了一个linear subspace
- linear subspace和null space of matrix之间是什么关系？
  - 每个linear subspace都可以表示为某个矩阵的null space
- 现在你感觉到linear code的第一个好处了吗？
  - 查错很方便：Hy=0

问题3：linear code (续)

- “linear”这个性质，在这个定理证明的哪一步中被用上了？
  你能解释每一步推导的理由吗？

Theorem 8.5 *Let $d_{\min}$ be the minimum distance for a group code C. Then $d_{\min}$ is the minimum of all the nonzero weights of the nonzero codewords in C. That is,*

$$d_{\min} = \min\{w(\mathbf{x}) : \mathbf{x} \neq \mathbf{0}\}.$$

PROOF. Observe that

$$\begin{aligned} d_{\min} &= \min\{d(\mathbf{x}, \mathbf{y}) : \mathbf{x} \neq \mathbf{y}\} \\ &= \min\{d(\mathbf{x}, \mathbf{y}) : \mathbf{x} + \mathbf{y} \neq \mathbf{0}\} \\ &= \min\{w(\mathbf{x} + \mathbf{y}) : \mathbf{x} + \mathbf{y} \neq \mathbf{0}\} \\ &= \min\{w(\mathbf{z}) : \mathbf{z} \neq \mathbf{0}\}. \end{aligned}$$

- 你感受到这个定理的重大意义了吗？
  这就是linear code的第二个好处！

问题4：查错和纠错

- 从查错和纠错的角度
  - d_{min}=1意味着什么？
  - d_{min}=2呢？
  - d_{min}=3呢？
- 如果要求能查出所有n位错误，d_{min}=？
- 如果要求能纠正所有n位错误，d_{min}=？

- 在纠错时，你其实做了一个什么假设？
  - We will assume that transmission errors are rare, and, that when they do occur, they occur independently in each bit; that is, if p is the probability of an error in one bit and q is the probability of an error in a different bit, then the probability of errors occurring in both of these bits at the same time is pq. <u>We will also assume that a received n-tuple is decoded into a codeword that is closest to it; that is, we assume that the receiver uses maximum-likelihood decoding.</u>

问题4：查错和纠错 (续)

- H要满足什么条件才能实现d_{min}=2？为什么？ d_{min}=3呢？

$$
\begin{aligned}
d_{\min} &= \min\{d(\mathbf{x},\mathbf{y}) : \mathbf{x} \neq \mathbf{y}\} \\
&= \min\{d(\mathbf{x},\mathbf{y}) : \mathbf{x}+\mathbf{y} \neq \mathbf{0}\} \\
&= \min\{w(\mathbf{x}+\mathbf{y}) : \mathbf{x}+\mathbf{y} \neq \mathbf{0}\} \\
&= \min\{w(\mathbf{z}) : \mathbf{z} \neq \mathbf{0}\}.
\end{aligned}
$$

Theorem 8.12 *Let H be an $m \times n$ binary matrix. Then the null space of H is a single error-detecting code if and only if no column of H consists entirely of zeros.*

Theorem 8.13 *Let H be a binary matrix. The null space of H is a single error-correcting code if and only if H does not contain any zero columns and no two columns of H are identical.*

问题4：查错和纠错 (续)

Theorem 8.13 *Let H be a binary matrix. The null space of H is a single error-correcting code if and only if H does not contain any zero columns and no two columns of H are identical.*

- 因此，在满足这个条件的前提下，$H=(A|I_m)$最多有几列？
- 我们为什么希望列越多越好？
- 这个方法的最大编码率是多少？　　$(2^m-(1+m)) / (2^m-1)$
- Hamming code的最大编码率又是多少？

Bit position		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
Encoded data bits		p1	p2	d1	p4	d2	d3	d4	p8	d5	d6	d7	d8	d9	d10	d11	p16	d12	d13	d14	d15	
Parity bit coverage	p1	X		X		X		X		X		X		X		X		X		X		
	p2		X	X			X	X			X	X			X	X			X	X		…
	p4				X	X	X	X					X	X	X	X					X	
	p8								X	X	X	X	X	X	X	X						
	p16																X	X	X	X	X	

Block length	$2^r - 1$ where $r \geq 2$
Message length	$2^r - r - 1$
Rate	$1 - r/(2^r - 1)$

- 你发现什么了吗？
- 你还能找到编码率更高的方法吗？

问题4：查错和纠错 (续)

- 如果Hy≠0，我们怎么纠错，或者说，哪一位错了？为什么？

Theorem 8.15 *Let $H \in \mathbb{M}_{m\times n}(\mathbb{Z}_2)$ and suppose that the linear code corresponding to H is single error-correcting. Let $\mathbf{r}$ be a received n-tuple that was transmitted with at most one error. If the syndrome of $\mathbf{r}$ is $\mathbf{0}$, then no error has occurred; otherwise, if the syndrome of $\mathbf{r}$ is equal to some column of H, say the ith column, then the error has occurred in the ith bit.*

- 我们今天讨论了这么多，“群”去哪儿了？