

全国高技术重点图书 • 通信技术领域

计算机中的纠错码技术

王新梅 张焕国 马建峰 覃中平 编著



人民邮电出版社

全国高技术重点图书·通信技术领域

计算机中的纠错码技术

王新梅 张焕国 编著
马建峰 覃中平

15.11.10

人民邮电出版社

图书在版编目(CIP)数据

计算机中的纠错码技术/王新梅等编著.-北京:人民邮电出版社,1999.10

ISBN 7-115-06761-9

I.计… II.王… III.计算机系统-纠错码 IV.TP301

中国版本图书馆 CIP 数据核字(98)第 09367 号

内 容 提 要

本书介绍了计算机系统编码技术的基础知识,论述了计算机存储系统和逻辑设计中应用的编码理论和技术。全书共分7章,分别介绍了基础知识、线性分组码与循环码、计算机内存储器中应用的纠错码、计算机外存储器中应用的纠错码、非对称错与单向错纠错码、逻辑设计中的编码技术和容错逻辑设计的实现等内容。

本书可作为大专院校计算机专业、通信工程专业、电子工程专业和数学专业高年级学生和研究生的教材和参考书,也可作为从事数字存储系统和逻辑电路设计的研究人员和工程技术人员的参考书。

全国高技术重点图书·通信技术领域

计算机中的纠错码技术

◆ 编 著 王新梅 张焕国 马建峰 覃中平

责任编辑 徐修存 王亚明

◆ 人民邮电出版社出版发行 北京崇文区夕照寺街14号

北京朝阳隆昌印刷厂印刷

新华书店总店北京发行所经销

◆ 开本:850×1168 1/32

印张:11.625

字数:302千字

1999年10月第1版

印数:1-3000册

1999年10月北京第1次印刷

ISBN 7-115-06761-9/TN·1262

定价:30.00元

《全国高技术重点图书》 出版指导委员会

主任：朱丽兰

副主任：刘 杲 卢鸣谷

委员：(以姓氏笔划为序)

王大中	王为珍	王守武	牛田佳	卢鸣谷
叶培大	刘 仁	刘 杲	朱丽兰	孙宝寅
师昌绪	任新民	杨牧之	杨嘉墀	陈芳允
陈能宽	张钰珍	张效详	罗见龙	周炳琨
欧阳莲	赵忠贤	顾孝诚	谈德颜	龚 刚
梁祥丰				

总干事：罗见龙 梁祥丰

《全国高技术重点图书·通信技术领域》 编审委员会

主任：叶培大

委员：陈俊亮 徐大雄 姚 彦

程时昕 陈芳烈 李树岭

前 言

随着信息技术的发展,存入和读出存储器的数据量越来越大,但与此同时,数据的出错概率也大大增加。保持数据的正确性仅依赖于器件和设备的可靠运算是不可行的。为了纠正存储、传输或处理过程中出现的错误,信息应具有某种冗余形式。目前,差错控制编码已成为一个比较完整的理论体系,该理论一直侧重于可靠通信系统的设计;但编码理论的应用并不仅限于通信系统,这种理论在计算机系统的设计方面也有广泛的应用。实际上,编码技术已成为计算机系统设计的一个组成部分。如果没有这种技术,许多系统是不可能实现的。本书主要讨论计算机存储系统和逻辑设计中的编码技术。

全书共 7 章,除第一章和第二章外,内容分为两部分:存储系统中的编码技术和逻辑设计中的编码技术。

第一章是全书的基础。在这一章引入了各种信道模型,给出了计算机系统中的错误类型,讨论了线性分组码的一些基本概念。第二章首先论述了编码理论中的数学基础,其中包括群、环、域和线性空间等概念,然后比较系统地介绍了线性码的基本理论体系,对线性分组码、汉明码、循环码、BCH 码和 RS 码的性质和实现作了全面的概括。

第三章论述了计算机内存储器中的纠错码,详细讨论了汉明码、存储器芯片码以及字节纠错码的结构和实现,并结合银河-I 型计算机内存储器编码方案的设计,说明了编码技术在实际计算机系统中的应用。本章的最后还介绍了陪集码、多字节纠错码和 SbEC-DbED 码等编码方法。

第四章以银河-I 型计算机的磁盘存储系统中的编码方案为例,

从理论到实际详细讨论了弗尔码的结构和编译码实现。最后详细论述了目前理论和应用中颇受重视的磁盘阵列编码技术和光盘编码技术。

第五章给出了几类纠非对称和单向错误码,讨论了构造非对称错误和单向错误检错纠错码的一些条件,并以伯杰码、博斯-林码等具体的编码方案说明了码的构造过程和应具有的结构。

第六章和第七章讨论逻辑设计中的编码理论和技术。在第六章给出了故障安全、自测试、全自检、强故障安全等自检电路的基本概念和设计方法。第七章讨论了自检组合电路的实现,最后概述了利用纠错码技术提高芯片成品率的错误屏蔽技术。

本书第一章和第二章由王新梅编写,第三章和第四章由张焕国编写,第五章由覃中平编写,第六章和第七章由马建峰编写,并由王新梅同志负责全书审阅。

由于篇幅所限,本书并没有涉及到目前日益受到重视的、以提高磁和光存储系统存储密度和可靠性为目的的、部份响应信道的编译码理论和技术。

本书既包含计算机系统编码技术的基本内容,也包含作者和国内外其它学者在这一方面最新研究成果。因此,本书既可作为大专院校计算机专业、通信工程专业、电子工程专业和数学专业高年级学生和研究生的教材和参考书,也可作为从事数字存储系统和逻辑电路设计的研究人员和工程技术人员的参考书。

鉴于作者水平有限,书中错误、遗漏在所难免,恳请读者批评指正。本书得到国家自然科学基金会的部分资助。

作 者

目 录

第一章 引 论	1
1.1 错误图样、信道模型和错误分类.....	2
1.1.1 错误图样.....	3
1.1.2 信道模型.....	3
1.1.3 错误分类.....	7
1.2 分组码的基本概念.....	7
1.2.1 分组码的基本概念.....	7
1.2.2 分组码的码率.....	8
1.3 分组码的距离度量与纠错能力的关系.....	9
1.3.1 汉明(Hamming)重量与汉明距离.....	9
1.3.2 汉明距离与纠错能力的关系.....	12
参考文献.....	14
第二章 线性分组码与循环码	15
2.1 数学基础.....	15
2.1.1 群(G).....	15
2.1.2 环(Ring).....	19
2.1.3 域(Field).....	22
2.1.4 线性空间.....	24
2.2 线性分组码.....	27
2.2.1 基本概念.....	27
2.2.2 码的一致校验矩阵与生成矩阵.....	29
2.2.3 伴随式与标准阵列译码.....	37
2.3 汉明码及其扩展.....	42

2.3.1	汉明码的构造	42
2.3.2	汉明码的扩展与扩展码	44
2.3.3	汉明限与完备码	46
2.4	循环码	47
2.4.1	循环码的定义	47
2.4.2	循环码的多项式描述	48
2.4.3	循环码的矩阵描述	52
2.4.4	系统码的生成矩阵与校验矩阵	55
2.4.5	缩短循环码	57
2.4.6	循环码的伴随式计算和错误检测	58
2.5	循环码的编码电路	61
2.5.1	多项式乘除法电路	62
2.5.2	循环码编码器	70
2.6	BCH 码与 RS 码	74
2.6.1	最小多项式与本原多项式	75
2.6.2	BCH 码	79
2.6.3	里德—索洛蒙(Reed-Solomon, RS)码	98
	参考文献	104
第三章	计算机内存储器中应用的纠错码	106
3.1	汉明码及其变形码	106
3.1.1	截短汉明码及扩展汉明码(SEC-DED 码)	106
3.1.2	最佳奇权码	108
3.2	最佳奇权码在银河-I 型计算机中的应用	110
3.2.1	H 矩阵的选择	110
3.2.2	编译码器的设计	113
3.2.3	编译码器的检测	120
3.2.4	性能评估	120
3.3	存储器芯片码	121
3.3.1	备份行(或列)方案	122

3.3.2	纠错编码方案	122
3.4	$GF(2^b)$ 上的汉明码	125
3.5	$GF(2^b)$ 上汉明码的最佳设计	128
3.6	$GF(2^b)$ 上的汉明码在计算机内存存储器中的应用	133
3.6.1	按字节组织内存存储器结构	133
3.6.2	编码方案	134
3.6.3	多字节错的检测概率	141
3.7	藤原英二码	143
3.8	陪集码	148
3.9	多字节错误纠错码(tbEC码)	158
3.10	RS码的二元阵映射及SbEC-DbED码	162
	参考文献	165
第四章	计算机外存储器中应用的纠错码	168
4.1	循环码的纠检错能力	168
4.2	弗尔(Fire)码	174
4.2.1	弗尔码	175
4.2.2	推广弗尔码	176
4.3	弗尔码的编码和译码	177
4.3.1	弗尔码的编码	177
4.3.2	弗尔码的译码	178
4.4	交错技术	191
4.5	磁盘存储器中的纠错码	192
4.5.1	弗尔码在银河-I型计算机磁盘系统中的应用	193
4.5.2	$GF(2^b)$ 上的汉明码在巨型计算机磁盘系统中的应用	200
4.6	磁盘阵列中的纠错码	205
4.6.1	编码	205
4.6.2	译码	206
4.6.3	纠双错算法	207
4.7	光盘存储器中的纠错码	209

4.7.1 乘积码	210
4.7.2 光盘系统中的 CIRC	211
4.8 磁带存储器中的纠错码	212
参考文献	227
第五章 非对称错与单向错纠错码	228
5.1 对称错、非对称错与单向错	228
5.2 非对称错纠检错码的充要条件	229
5.3 构造非对称错纠检错码	231
5.4 单向错纠检错码的充分条件	235
5.5 单向错纠检错码的构造	239
5.5.1 可检所有单向错(AUED)的伯杰(Berger)码	239
5.5.2 可检 t 个单向错(t -UED)的博斯—林(Bose-Lin)码	242
5.5.3 博斯(Bose)与雷欧(Rao)的等重 1-EC AUED 码	247
5.5.4 布莱姆(Blaum)与蒂伯尔格(Tiborg)所构造的系统 t -EC AUED 码	249
5.6 一些评注	253
参考文献	254
第六章 逻辑设计中的编码技术	257
6.1 基本的检错码	258
6.2 自检的概念	261
6.2.1 一般概念	261
6.2.2 检验器的概念	272
6.3 自检组合电路	277
6.3.1 自检算术逻辑部件	278
6.3.2 自检可编程逻辑阵列	299
6.3.3 自检 ECC 电路	307
6.4 自检时序电路	312
参考文献	313
第七章 容错逻辑设计的实现	315

7.1 自检检验器	315
7.1.1 奇偶校验码检验器	315
7.1.2 两轨码检验器	317
7.1.3 n 中取 m 码检验器	322
7.1.4 可分码检验器	333
7.1.5 广义预测检验器	337
7.2 自检处理器和微处理器	346
7.2.1 自检处理器	346
7.2.2 自检微处理器	349
7.3 错误屏蔽技术	350
参考文献	356

第一章 引 论

在现代信息传输、存储和交换系统以及各种通信网中,计算机起着核心作用。如何保证计算机的可靠运行,已成为人们日益关注的焦点。纠错码、各种容错技术和容错算法,正是提高计算机运行可靠性的关键技术之一。

在计算机系统中,各个子系统以及部件之间进行着大量的数据交换和处理。例如计算机内的 CPU 和主存储器之间,以及大型计算机与外存储器之间,数据交换率可达每秒几百 Mbit 至几千 Mbit;在数字通信系统如光纤通信中,数据传输速率也已高达每秒几千 Mbit,甚至更高。因此,即使系统的可靠性非常高,但某些不可避免的因素,如大气干扰、电子元器件热噪声、某些元器件(如磁盘、光盘等存储介质)在制造过程中的微小缺陷,以及逻辑设计和程序设计中的某些疏忽和遗漏、部件参数的变化等,都可能引起交换、存储和传输过程中数据的错误,从而引起严重的后果,甚至使整个系统不能正常运行,直至瘫痪。如何提高计算机系统中数据传输、交换、存储、逻辑计算、程序运行等方面的可靠性和容错性,已日益引起人们的广泛注意,并已成为研究的热点问题之一。所谓容错,是指当系统中某些部件出现故障,或者逻辑设计和程序设计中出现错误后,系统仍能按预定要求运行,完成预定任务。

使计算机具有容错能力的方法有很多。其中之一是在系统设计中利用纠错码,使得在产生错误的数字没有引起严重后果之前,就把错误检测出来并自动纠正。另一方法是在计算机的电路设计、逻辑设计和算术运算中,采用具有容错能力的算法和程序,以及具有容错能力的逻辑电路和运算电路,并应用自动故障诊断系统,使得一旦产

生错误或故障以后,能自动纠错或诊断出来,或者能转移到正确的程序中运行。

本章和二、三、四章将研究如何利用纠错码纠正或发现已传数据组中的错误,以及计算机内、外存储器中的错误。其它各章将讨论容错电路设计和容错算法等。为此我们首先讨论传输数据信道的差错情况及其模型。

1.1 错误图样、信道模型和错误分类

无论何种类型的计算机系统、数据存储和数据通信系统中,各部件或子系统、通信收发之间的数据交换、处理,存储器中数据的读写,以及逻辑运算等,都可以用如图 1.1 所示的通信系统模型表示。其中信源是发出数据的发源地,在计算机系统中一般是以长为 $k\text{bit}$ 一组的二进制信息组 $m = (m_0, m_1, \dots, m_{k-1})$, 其中 m_i 是二进制码元(以后如无特别说明,所指的数据均指由二进制的 0、1 码元组成),如在计算机系统中经常以 8bit 为一组组成一个字节(也称为一个 byte)。信息组 m 经过纠错编码器后加上多余度码元(或校验元),得到一组长为 $n \geq k$ 的码组 $C = (c_0, c_1, \dots, c_{n-1})$, C 通过信道后传给收端。

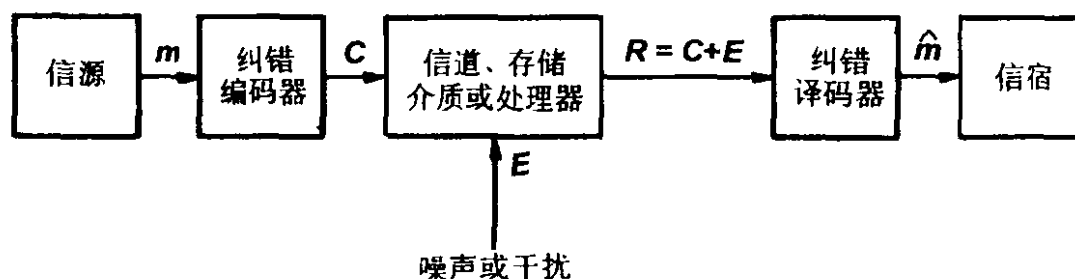


图 1.1 通信系统模型

图中的信道(或存储介质、数据处理器等)是传输数据的物理媒质,包括各种无线信道(如通信中的短波、超短波、微波、卫星等信道)、有线信道(如光纤、电缆等)、磁盘和光盘等存储介质和其相应的

读写电路,以及逻辑运算电路和数据处理器等。在数据通信系统中这一部分还包括调制解调器等。由于部件缺陷或系统故障等各种原因所引起的数据错误,我们都归结或等效为由于信道中的噪声或干扰所造成的错误序列 $E = (e_0, e_1, \dots, e_{n-1})$ 对所传码组 C 的作用结果。

因此, C 通过信道后到达收端时已成为 $R = C + E$, $R = (r_0, r_1, \dots, r_{n-1})$, $r_i = c_i + e_i$ 。纠错译码器,利用所加的多余度码元(校验元)和信息元之间的约束关系,对 R 中的数据进行纠错,恢复出原来的 k 个信息元 \hat{m} (\hat{m} 一般等于 m ;但若 R 中错误很多,已超过了纠错译码器的纠错能力,则 $\hat{m} \neq m$)。 \hat{m} 送至信宿。信宿就是信源所发出的信息组最后要到达的目的地,可以是计算机内部的某一子系统和部件。在数据通信网中也可以是计算机或各种终端设备或人。

1.1.1 错误图样

由前面讨论可知,由于信道干扰或噪声所造成的错误序列 $E = R - C$,也是一个由 0、1 组成的长为 n 的二进制序列 $(e_0, e_1, \dots, e_{n-1})$ 。在模 2 运算下, $E = R - C = R + C$ 。 E 中的每个码元 $e_i = r_i - c_i = r_i + c_i$,若 $r_i = c_i$,则 $e_i = 0$,否则 $e_i = 1$ 。可知, E 中的某一位为 1 说明 R 中相应的这一位产生了错误。例如若传输的码组 $C = (c_0, c_1, \dots, c_7) = (11\ 11\ 11\ 11)$,收到的 $R = (r_0, r_1, \dots, r_7) = (1101\ 1101)$,则

$$E = R - C = (r_0 - c_0, r_1 - c_1, \dots, r_7 - c_7) = (00100010) = (e_0, e_1, \dots, e_7)$$

E 中的 $e_2 = 1, e_6 = 1$,可知 R 中的第 2 位和第 6 位发生了错误, E 完全反映了接收码组 R 中的错误情况,因此也称 E 为错误图样。而纠错编译码器的任务就是从 R 中确定出错误图样 E ,然后由 $R - E = \hat{C}$ 中可恢复出传送的信息组 m 。

1.1.2 信道模型

由于下面几章仅讨论错误序列 E 对所传码组 C 的作用(所引起

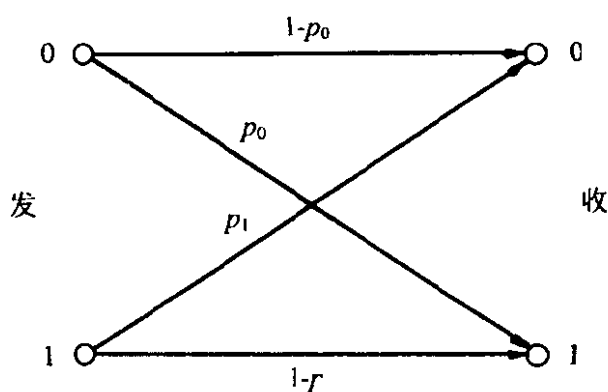


图 1.2 二进制信道模型

的错误), 以及如何消除接收码组 R 中的 E 的影响, 恢复出正确的信息组, 因此可以把图 1.1 的通信系统模型用图 1.2 所示的二进制信道模型表示。设发送端以相等的概率发送信息 0 和 1。经信道传输, 由于信道噪声或干扰的影响, 分别以概率 $1-p_0$ 和 $1-p_1$ 接收 0 和 1, 因此 p_0

和 p_1 分别是 0 错成 1(称为 0 错误)和 1 错成 0(称为 1 错误)的概率, 它们也称为信道的转移概率。

一、二进制对称信道(BSC)

如果 $p_0 = p_1 = p_e$, 则这种信道称为二进制对称信道(BSC), BSC 也称为随机信道或无记忆信道。在这种信道中的错误, 是由于随机噪声(如白高斯噪声)或随机干扰所引起的, 错误码元之间没有相关性。这是一种比较典型的和常见的信道模型。我们下一章所要讨论的纠错码主要是针对纠正这一类信道所产生的随机错误。

二、二进制非对称信道(BAC)和 Z 信道

在图 1.2 所示的二进制信道模型中, 若 $p_0 \neq p_1$, 则这种信道称为二进制非对称信道(BAC)。存储介质中由于制造工艺缺陷所引起的错误往往是非对称的, 如磁盘或光盘中某一微小区间或某一轨迹中某一段磁粉脱落, 则相应的这一段所存数据所引起的错误是非对称的。而且往往是 $p_1 \gg p_0$, 甚至 $p_0 = 0$, 此时, 图 1.2 所示的信道模型成为图 1.3(a)所示的 Z 信道模型。图 1.3(b)是 $p_1 = 0$ 的 Z 信道模型。通常, 我们认为在 Z 信道中所产生的错误往往是 1 错误或 0 错误, 而不能在同一数据组中同时并存这两种错误, 因此这种信道所产生的错误称为单向错误, 纠正这种单向错误比纠正随机错误要容易。

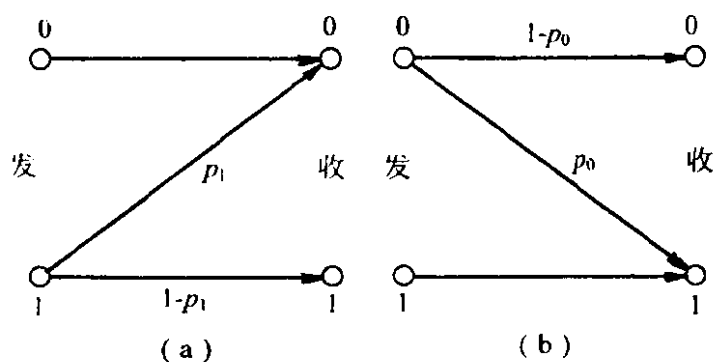


图 1.3 Z 信道模型

三、删除信道(BEC)

在数据通信系统中,解调器或存储系统中的读检测器对输入信号的判决,在某些情况下不能给出是 0 还是 1 的判决,而给出一个未知值 x 。这种信道称为二进制删除信道(BEC),如图 1.4 所示。在这种信道中,一般认为要么判决 0 或 1 的结果完全正确,要么以概率 p (删除概率)不能进行确定的判决而输出 x 。对于这种信道,纠错译码器的任务仅仅是确定 x 的值是 0 还是 1,而删除符号位 x 的位置是已知的,因而译码器较为容易实现。

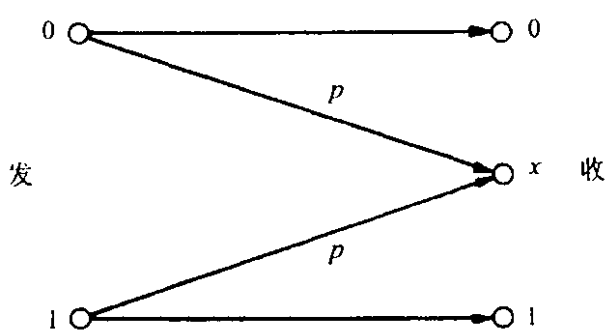


图 1.4 BEC 模型

四、突发信道

在数据通信系统中的某些信道,如无线信道或某些有线信道,以及计算机存储系统,由于各种干扰或读写头接触不良,所引起的错误往往不是随机错误,而是成群成串地出现,表现为错误图样 E 中 1 的个数较多且相对密集。如发的码组 $C = (11\ 11\ 11\ 11)$,收到的码组 $R = (11010011)$,相应的错误图样 $E = (00101100)$ 。这种产生比较密集错误的信道称为突发信道或有记忆信道,产生的错误之间有相

关性。错误图样 E 中的第一个 1 与最后一个 1 之间的长度称为突发长度 b , 该例中 $b = 4$ 。

描述这种突发信道的模型有很多^[2], 其中一个比较常用的突发信道模型如图 1.5 所示, 称为双状态马氏链模型或 Gilbert 模型 (Gi 模型), 相应的信道称为 Gi 信道。该模型说明, Gi 信道的状态始终处于好 (G) 或坏 (B) 两个状态之一。当处在 G 状态时, 信道不产生错误; 当处在 B 状态时, 以 $(1 - p)$ 的概率产生错误。B 和 G 两个状态之间分别以概率 P 和 p 转移, 显然 $P = 1 - Q, p = 1 - q, Q$ 和 q 分别是信道停留在 G 状态和 B 状态的概率。

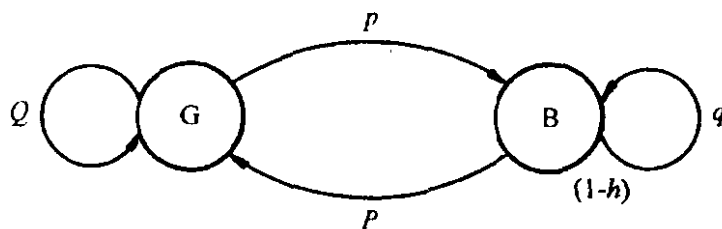


图 1.5 Gilbert 信道模型

比较图 1.2 和图 1.5 可知, 在 BSC 中只有一个参数 $p_e = p_0 = p_1$; 而在 G 信道中有 3 个参数, 因此, 要估计突发信道中使用各种纠错措施后达到的性能, 要比 BSC 困难得多。

五、算术信道

在计算机系统中还有一个与一般数据通信不一样的重要的信道, 即算术信道。

设移位寄存器 A 中存储了代表数据 N 的二进制数组 $(a_0, a_1, \dots, a_{n-1})$:

$$N = (A)_2 = \sum_{i=0}^{n-1} a_i 2^i \quad a_i \in \{0, 1\}$$

由于运算器的故障或其它原因, 使数据 N 产生了错误, 成为

$$N' = (A')_2 = \sum_{i=0}^{n-1} a'_i 2^i \quad a'_i \in \{0, 1\}$$

在这种情况下错误图样 $E = N - N'$ 。可知此时每个 a'_i 不一定是 a_i