

# SQLite Record Recovery

---



*zurum*

*heros86@korea.ac.kr*

*DFRC@CIST@KU*



1. SQLite 구조
2. 삭제된 영역
3. 레코드(Cell) 구조
4. 복원 기법

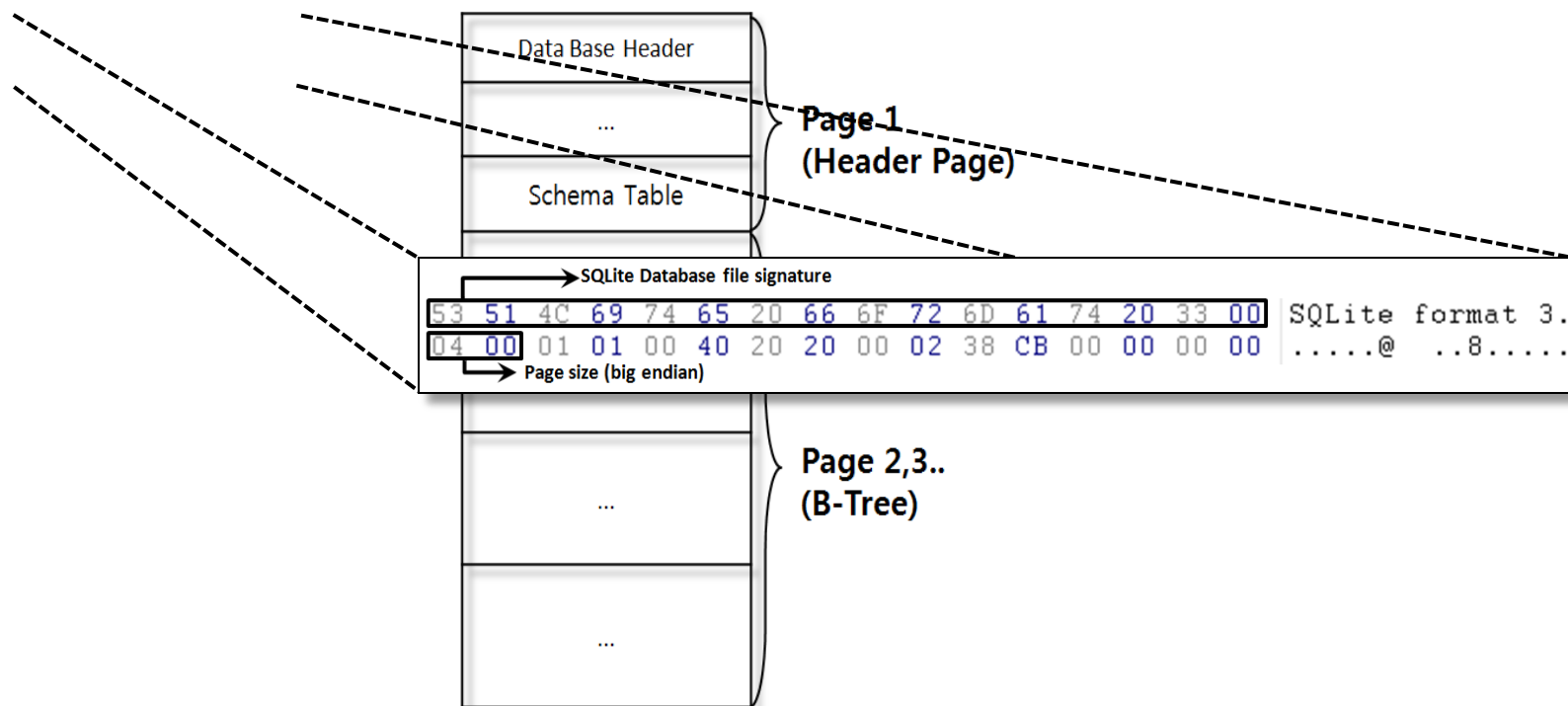
# Structure of SQLite DataBase

- File & Page Structure



## File Structure

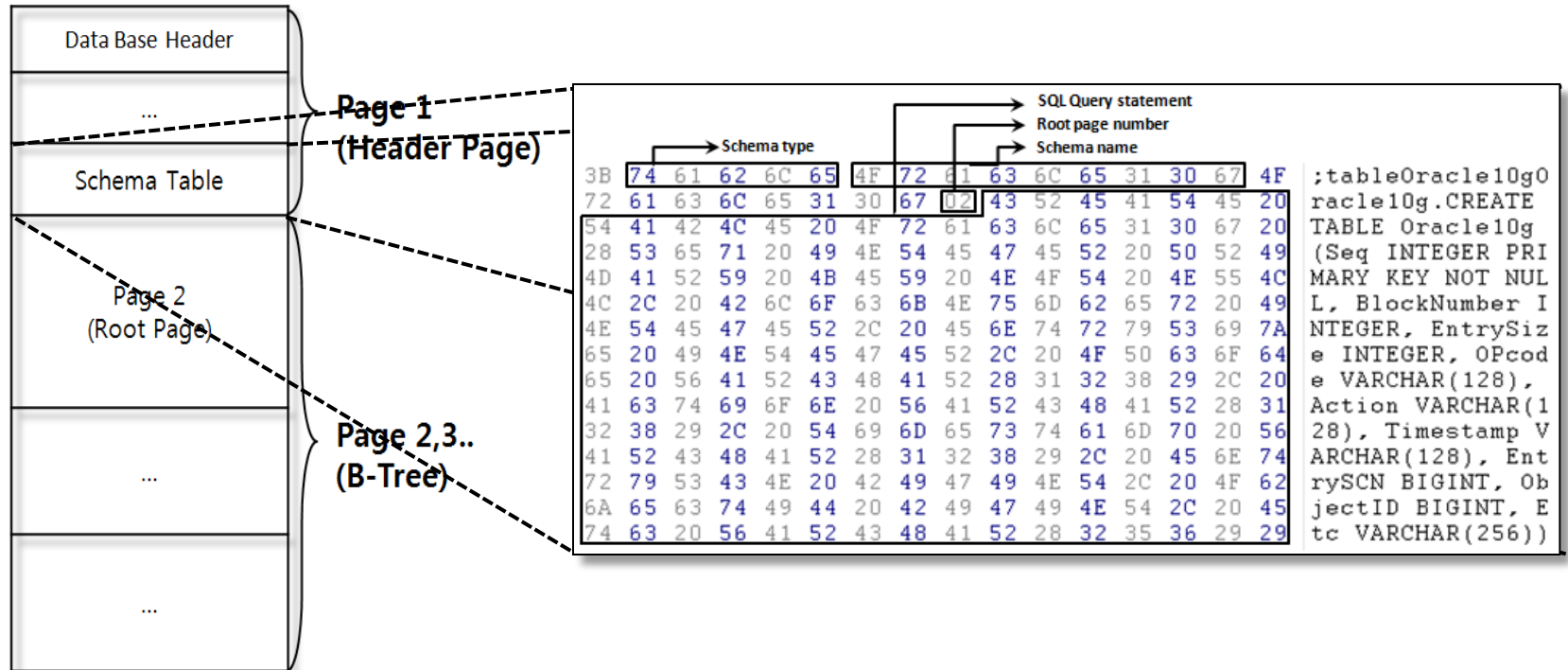
- 전체 구조 및 파일 헤더





## File Structure

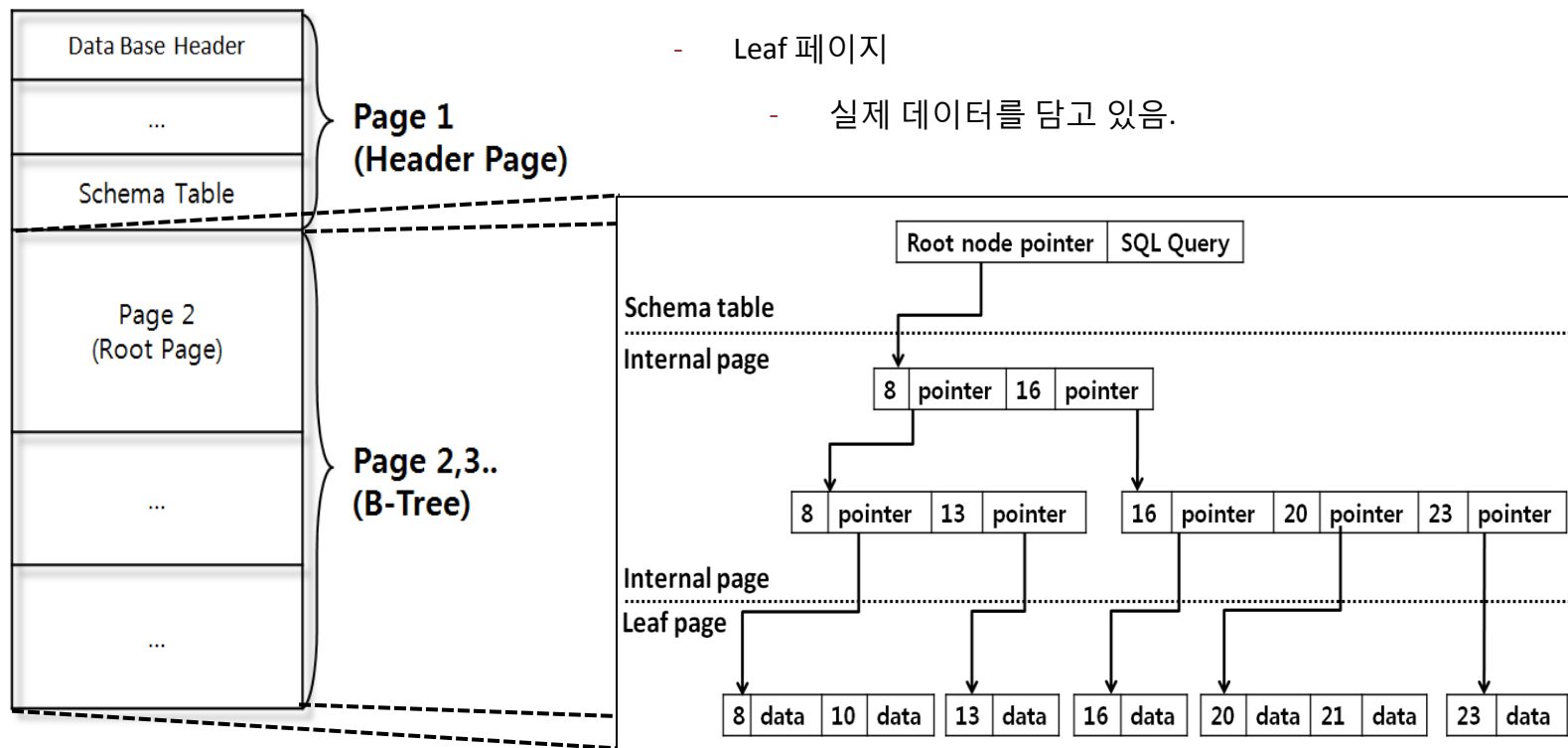
### 스키마 테이블





## File Structure

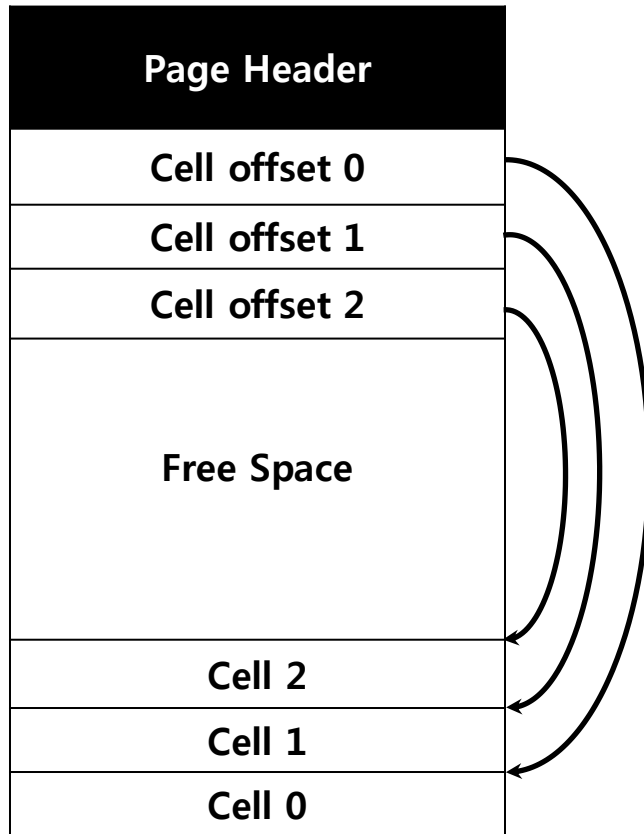
### Table B-tree



- Internal 페이지
  - 하위 페이지의 포인터를 가짐.
- Leaf 페이지
  - 실제 데이터를 담고 있음.



## Page Structure



- Page Header
  - Table – b-tree
- Offset 0
  - 0x05 - Internal Nodeleaf
  - 0x0D – Leaf Node
- Size
  - 12 Byte – Internal Node pages
  - 8 Byte - Leaf node pages
- Cell Offset
  - 2byte Big endian integer



## Page Structure

- Internal Page header

	Page flag	Number of record	Offset of the first bytes of the record		Page number of right most child-page
00000400	05	00 00	00 50	01 D6 00	00 00 1F C9
00000410	03 EF	03 E8	03 E2	03 DC	03 03 CE 03 C7 03 C0
00000420	offset of first block of free space			03	Num of fragmented free bytes
00000430	03 81	03 7A	03 73	03 6C	03 65 03 5E 03 57 03 50

- Leaf Page header

	Page flag	Number of record	Offset of the first bytes of the record	
00000800	0D	00 00	00 14	00 4B 00
00000810	02 F6	02 DB	02 C1	02 47 02 12 01 F8 01 DD
00000820	offset of first block of free space			00
00000830	Num of fragmented free bytes			00 00 00 00 00 00 00 00



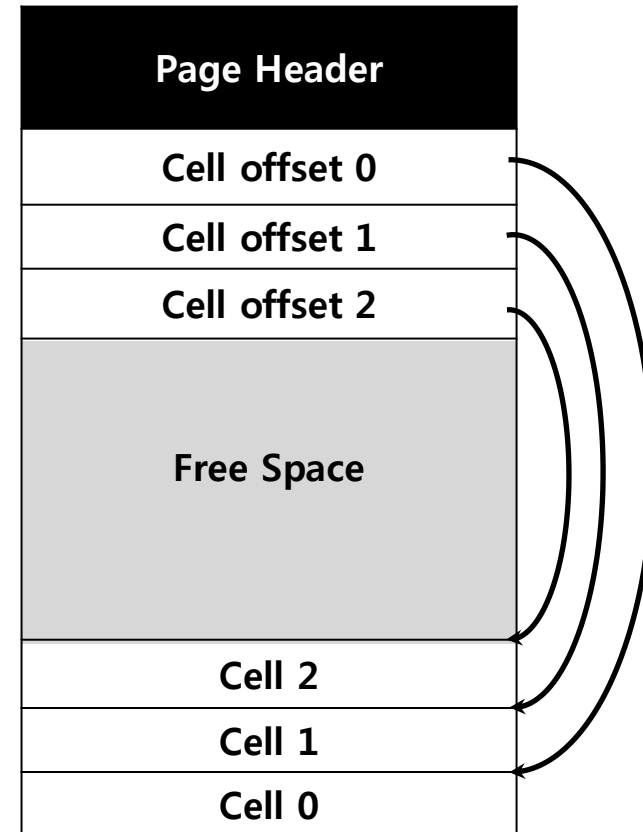
# Deleted area in database file

- Unallocated Area(Free space & Free block)



## Unallocated Area

- Free Space(비할당 공간)
  - 아직 셀(레코드)이 할당되지 않은 영역
  - 셀 오프셋과 마지막 셀 사이의 공간

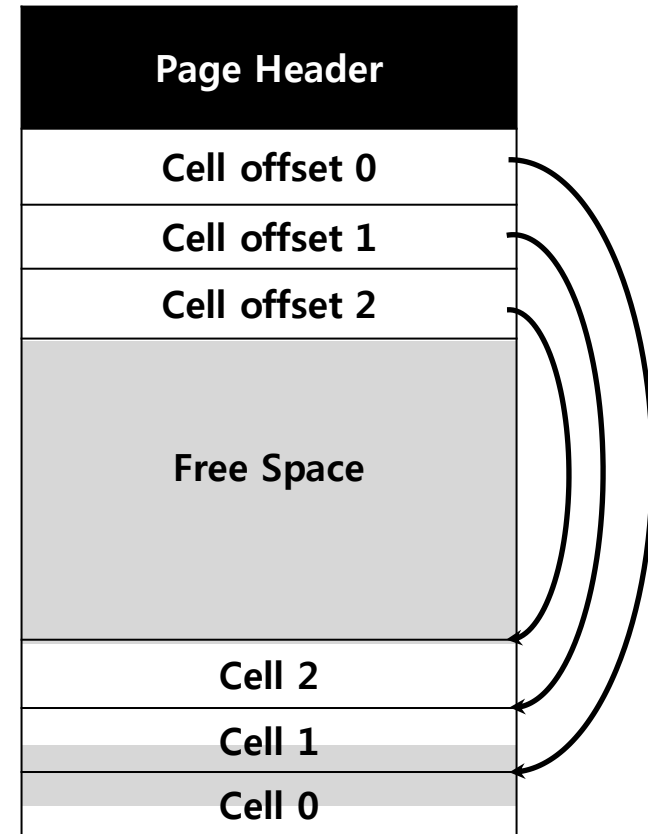




## Unallocated Space

### ▪ Free Block

- 셀이 할당되었다가 삭제된 영역
- 셀과 셀 사이에 존재





## Deleted area

### 삭제된 영역 수집

- 비할당 영역(Free Area)의 삭제 데이터

- ✓ 헤더와 오프셋 배열을 건너 뛴 후 페이지의 첫 번째 레코드가 나올 때까지 탐색

- 기본적으로 Free Area는 0으로 세팅되어 있음
- 따라서 0이 아닌 값이 존재하는 경우 삭제된 데이터로 판단 가능.

Number of record

Offset of the first byte of the record

00000000	0D	02	22	00	02	01	21	00	01	21	02	82	02	82	02	82	.."...!...!.,.,,	
00000010	02	82	02	82	03	83	03	E3	03	E3	03	E3	03	E3	03	E2	.,.,.f.ă.ă.ă.ă.ă	
00000020	03	E3	03	E3	03	E3	03	E3	03	E3	03	E3	03	E3	03	00	2C 81	.ă.ă.ă.ă.ă.ă.,,



## Deleted area

- 삭제된 영역 수집
  - 비할당 블록(Free Block)의 삭제 데이터
    - ✓ 비할당 블록의 데이터는 무조건 삭제된 데이터로 판단 가능
    - ✓ 비할당 블록 체인 (Chain of free block)

00000800	0D	03 10	00 12	00 4B	00 03	45 03	2B 02	F6 02	DB	.....K..E+.ö.Ü
00000810	02	C1 02	47 02	2D 02	12 01	F8 01	DD 01	C3 01	49	.Á.G.-...ø.Ý.Ă.I
00000820	01	27 01	14 00	FA 00	DF 00	C5 00	4B 00	4B 00	4B	./...ú.B.Ă.K.K.K
00000830	00	00 00	00 00	00 00	00 00	00 00	00 00	00 00	00	.....
00000840	00	00 00	00 00	00 00	00 00	00 00	00 00	78 14	0B 00 01	.....x....
⋮										
00000b10	03	D3	00 1B	01 02	15 0D	0D 03	01 15	02 00	DC 31	.Ó.....Ů1
00000b20	30	2F	34 12	63 D7	00 4E	4F 4E	45 18	03 0A	00 01	0.4.c*.NONE.....
⋮										
00000bd0	30	0D 0A	00 00	00 2D	01 02	13 0D	33 03	01 15	02	0.....-....3....
00000be0	01	7C 35	2E 31	31 34	2F 31	30 2F	32 30	30 39	20	. 5.114/10/2009
00000bf0	32	32 3A	30 31	3A 31	32 12	63 D7	00 4E	4F 4E	45	22:01:12.c*.NONE

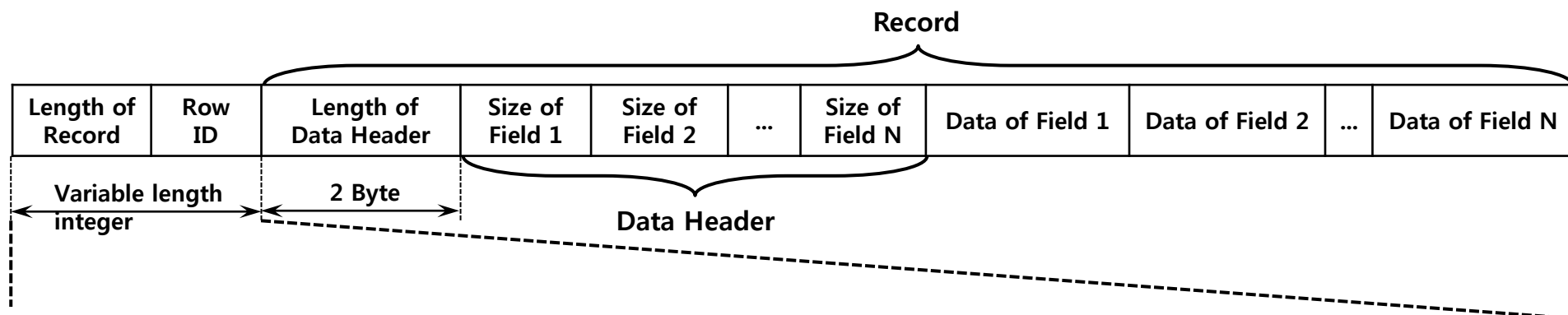
# Data Recovery

- Normal Cell & Deleted Cell
- Record Recovery



## Cell structure

- Variable length integer

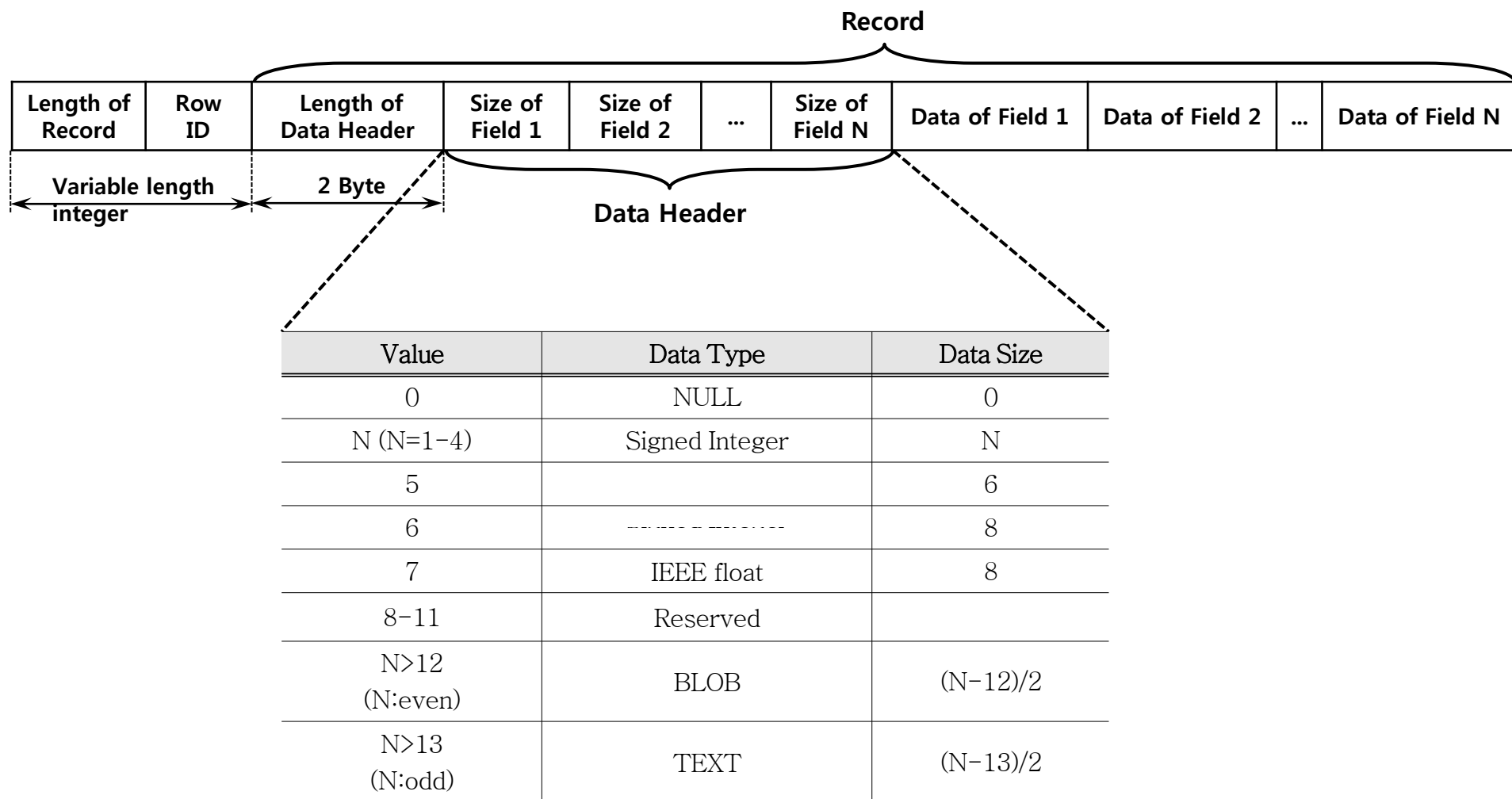


Byte	Value range	Bit pattern
1	7 bit	0XXXXXXXX
2	14 bit	1XXXXXXXX 0XXXXXXXX
3	21 bit	1XXXXXXXX 1XXXXXXXX 0XXXXXXXX
4	28 bit	1XXXXXXXX 1XXXXXXXX 1XXXXXXXX 0XXXXXXXX
5	35 bit	1XXXXXXXX 1XXXXXXXX 1XXXXXXXX 1XXXXXXXX 0XXXXXXXX
6	42 bit	1XXXXXXXX 1XXXXXXXX 1XXXXXXXX 1XXXXXXXX 1XXXXXXXX 0XXXXXXXX
7	49 bit	1XXXXXXXX 1XXXXXXXX 1XXXXXXXX 1XXXXXXXX 1XXXXXXXX 1XXXXXXXX 0XXXXXXXX
8	56 bit	1XXXXXXXX 1XXXXXXXX 1XXXXXXXX 1XXXXXXXX 1XXXXXXXX 1XXXXXXXX 1XXXXXXXX 0XXXXXXXX
9	64 bit	1XXXXXXXX 1XXXXXXXX 1XXXXXXXX 1XXXXXXXX 1XXXXXXXX 1XXXXXXXX 1XXXXXXXX 1XXXXXXXX XXXXXXXXXX



## Cell structure

- Data Header

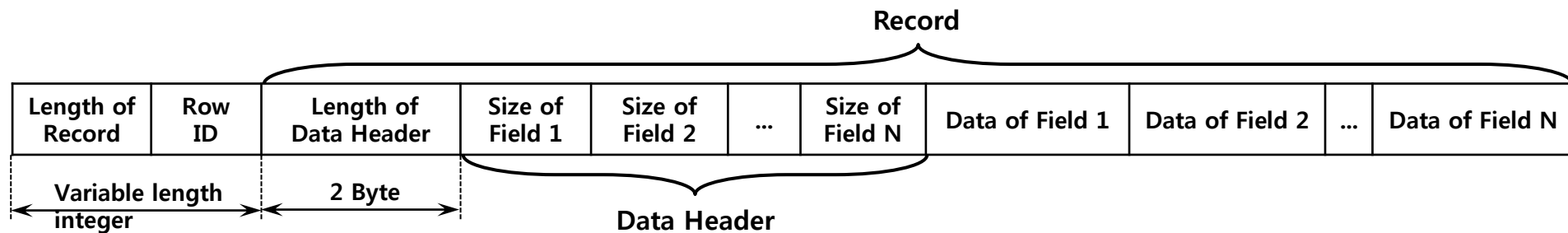




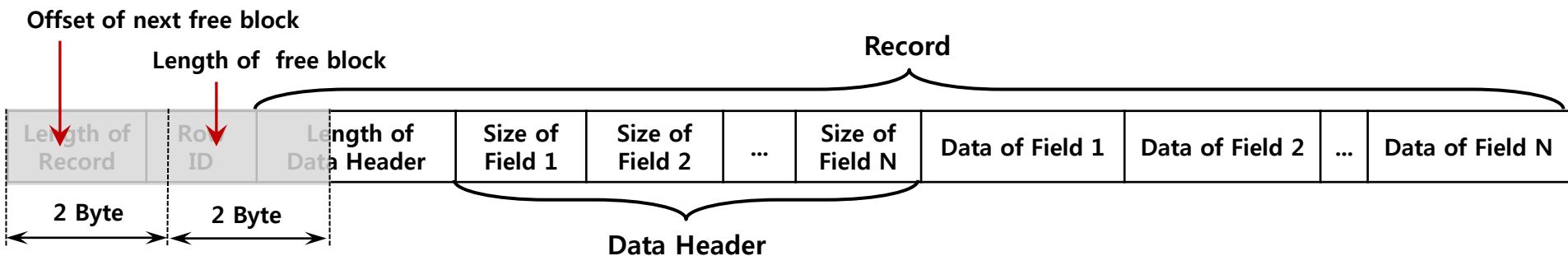


## Cell structure

- Normal Cell

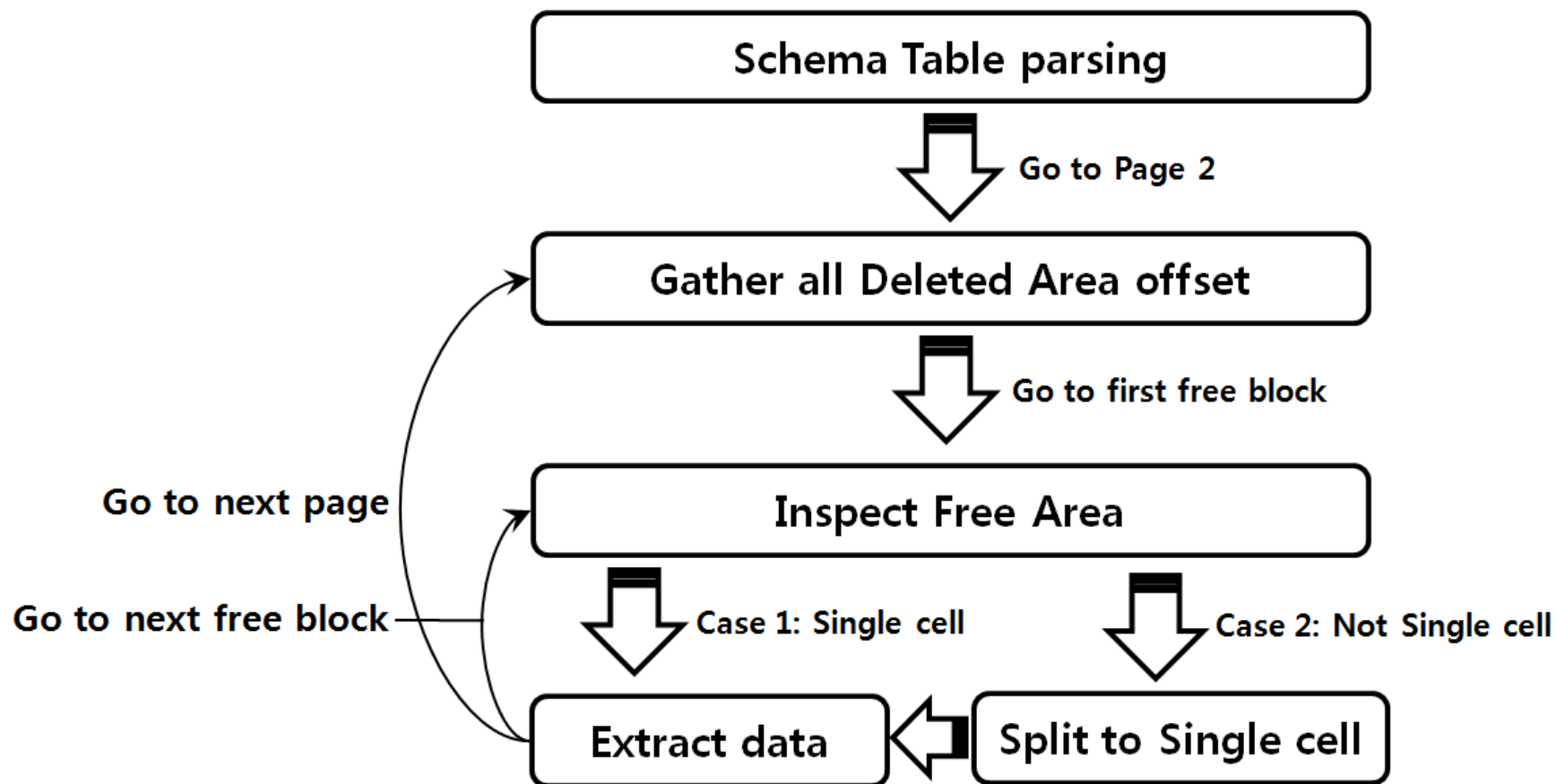


- Deleted Cell





## Record carving

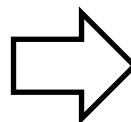




## Inspect Free Area

- simulation

A	B	C	D
Click here to define a filter			
1	10000000000	17	one
2	13245	34	two
3	9876854	51	three
4	99	51	four



A	B	C	D
Click here to define a filter			
2	13245	34	two
4	99	51	four

- Schema table

Index	Name	Declared Type
1	A	INTEGER
2	B	INTEGER
3	C	INTEGER
4	D	VARCHAR(256)

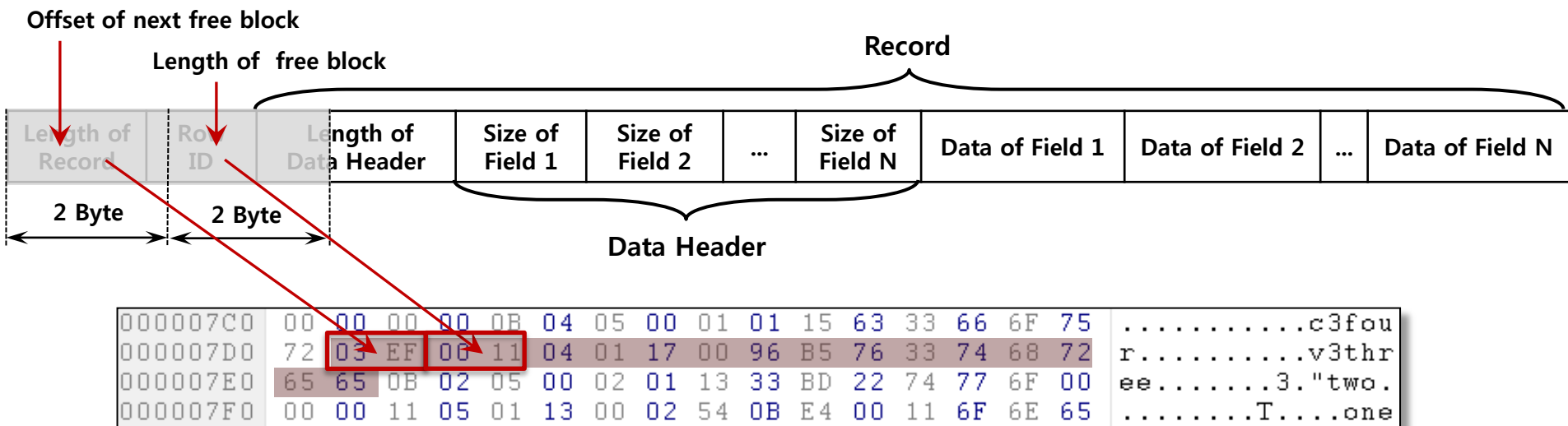
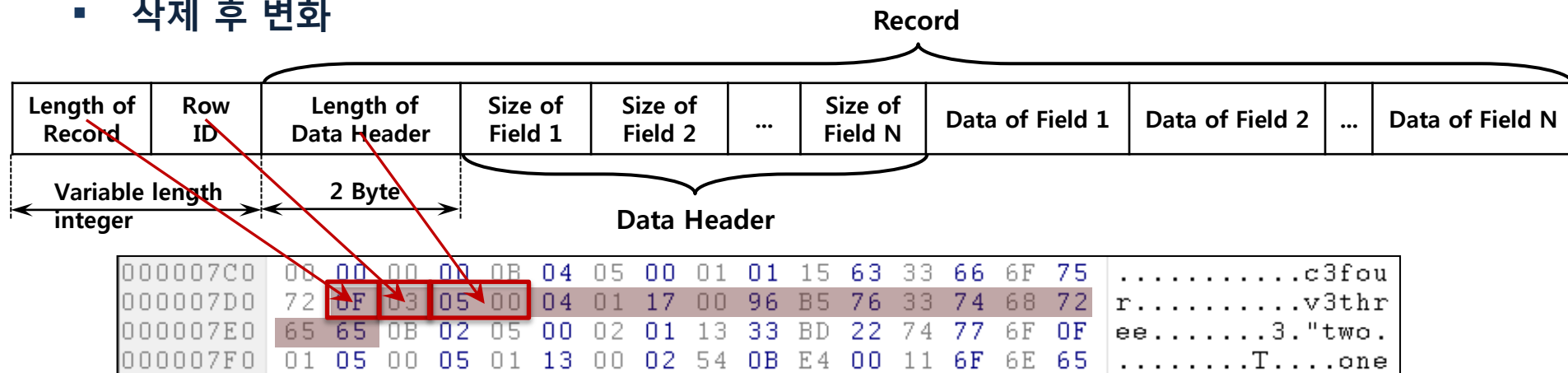
- Target data

000007C0	00	00	00	00	0B	04	05	00	01	01	15	63	33	66	6F	75	.....c3fou
000007D0	72	03	EF	00	11	04	01	17	00	96	B5	76	33	74	68	72	r.....v3thr
000007E0	65	65	0B	02	05	00	02	01	13	33	BD	22	74	77	6F	00	ee.....3."two.
000007F0	00	00	11	05	01	13	00	02	54	0B	E4	00	11	6F	6E	65	.....T....one



## Inspect Free Area

### 삭제 후 변화





## Inspect Free Area

### ▪ Values of Data header

Value	Data Type	Data Size
0	NULL	0
N (N=1-4)	Signed Integer	N
5	-----	6
6		8
7	IEEE float	8
8-11	Reserved	
N>12 (N:even)	BLOB	(N-12)/2
N>13 (N:odd)	TEXT	(N-13)/2

### ▪ Target data

- 스키마의 데이터 헤더와 일치하는 점을 찾은 지점에서 복원 진행
- 스키마가 복잡할 수록 오탐은 적어짐.

000007C0	00 00 00 00 00 00 00 00	04 05 00 01 01 15 63 33 66 6F 75	.....c3fou
000007D0	72 05 EF 00 11 04 01 17 00 96 B5 76 33 74 68 72		r.....v3thr
000007E0	65 65 0B 02 05 00 02 01 13 33 BD 22 74 77 6F 00		ee.....3."two.
000007F0	00 00 11 05 01 13 00 02 54 0B E4 00 11 6F 6E 65		.....T....one



Q&A