00100001 1001

8) 1) 12 elias

$1 + \log_2 12 = 4$

elias gama coding for 4 = 00100

~~⇒ 0001100 elias gama coding~~:

12 = 1100 [00100100] elias delta code for 12

decode

⇒ decode 00100101 using elias delta

⇓

step 1 ⇒ L = 2

2 consider the one that was reached to be the first bit of an integer with value of 2^L, read the remaining L digits of the integer this is the integer M

∴ 5 bits is consider

we also obtain m = 4 (100) in binary

Then 101

⇒ which finally gives

1101 ⇒ 13

2) ~~m = 15 with n = 10~~ m = 10, n = 15

Golomb code = $q = \frac{15}{10} \Rightarrow 1$

q = 2 ⇒ 01

m ⇒ is not the power of 2

⇒ $\lfloor \log_2 10 \rfloor = 3$ bits for first $2^4 - 10 = 6$ values

4 bits for binary rep of rest of r value

where r = 15 − 1×10

= 5

[011 01]

~~4) user query represent the user information need, Types of following form~~

decoding ~~00100101~~ 00001]10   m = 3

q = 4

i = 1   d = 2^2 − 3 = 4 − 3 = 1

r = 1 = 1

1 ≥ 1.

r = 10 ⇒ 2 ⇒ r = 2 − 1 = 1

n = 4 × 3 + 1

= 13

4). Keyword queries: The user expresses his information need with a list of key aiming to find documents that contain some or all the query terms.

The term in the list are assume to be connected with a soft version of the logical AND

example, if one is interested in finding information about web mining, one may issue the query 'web mining' to an IR or search engine system

"web mining is retreated as web and mining.

2. Boolean query: The user can use Boolean operators AND, OR and NOT to construct complex query

For example data or web is a Boolean query: which the request document that contain the word 'data or web'.

A page is return for a boolean query if the query is logically true.

3) phrase queries: query consists of a sequence of word that make up a phrase.

→ phrase normally enclosed with double quote for example "web mining technique and application"

4) proximity query: The proximity query is a relaxed version of the phrase query and

can be a combination of term and phrase

→ Proximity query seek the query terms with in close proximity to each other.

→ The closeness is used as a factor in ranking the returned document or page.

5) Full document query : when the query is full document, the user wants to find other documents that are similar to the query document.

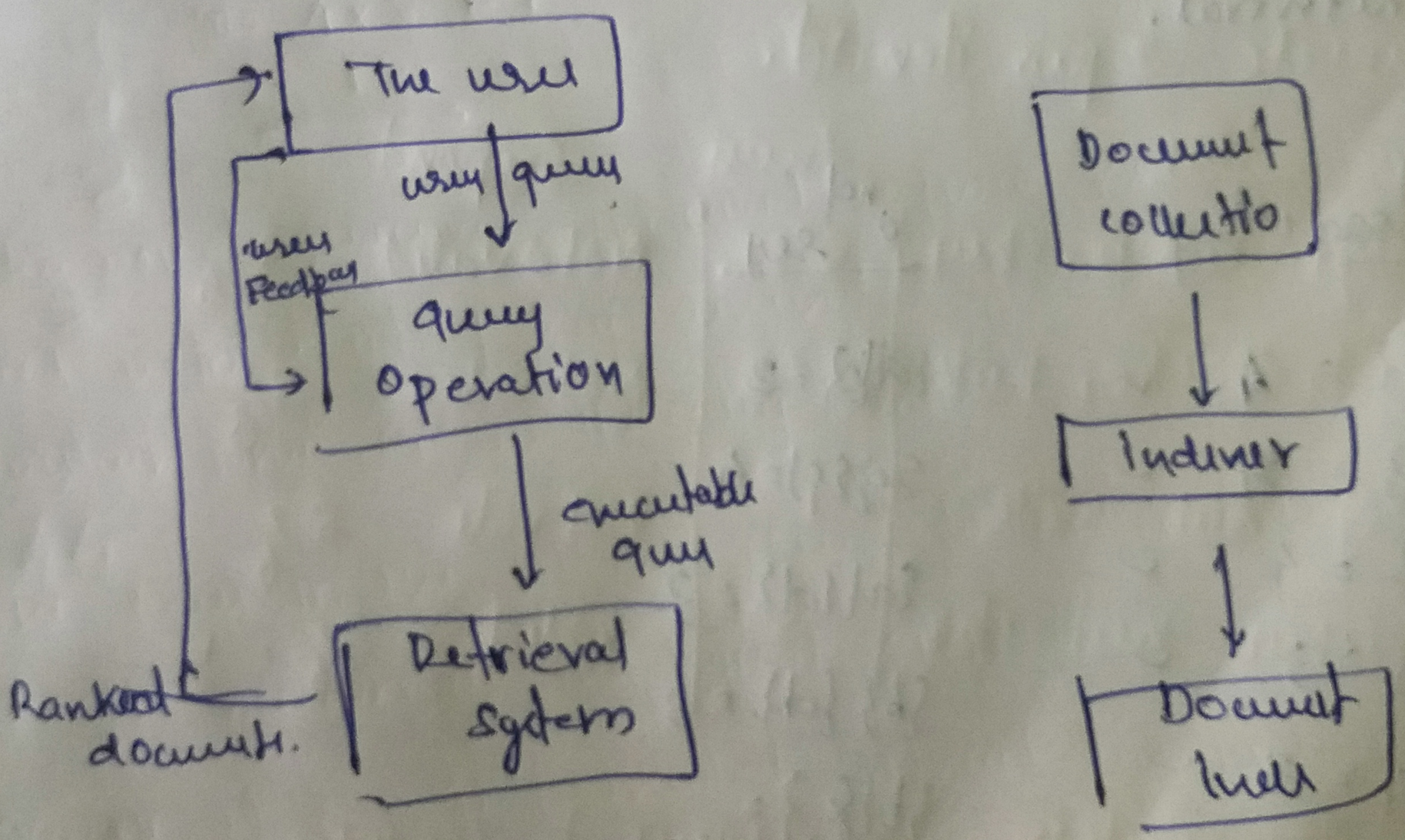
eg of a query page is an example.

4) Natural language question : The user express his/her information need as a natural language question

eg: definition question

→ strong linguistic patter indication of definition sentence.

2) IR System Architecture



1) IR is the study of helping user to find information that matches their information need.

- The user → with information need issue a query (user query) to retrieval system through the query operation modules

- The retrieval module uses the document index to retrieve those documents that contain some query terms compute relevance score for them then rank the retrieved document according to the score.

- The ranked documents are then presented to the user

- The document collection is also called the text database which is indexed by the indexer for efficient retrieval.

a) 1st seq (1st ~~seq~~)

1st seq		2nd seq
{a} = ~~7~~ 4	✓	{a}{b} = 3
{b} = ~~3~~ 4	✓	{a}{c} = 4
{c} = 4	✓	{a}{d} = 3
{d} = 3	✓	{b}{d} = 3
{e} = 3	✓	{b}{c} =
{f} = 3	✓	
{g} = 1	✗	

29

e) Beef

Beef	4	$4/6 = \frac{2}{3} = 60$ ✓
Chicken	= 5	5/6 valid ✓
milk	= 4	4/6 valid
Chum	= 3	3/6 = valid
Clothes	= 3	3/6 valid

* ⇒ Give all possible pair of items

Beef chicken & vice versa	2	3/6 val
Beef milk & vice versa	2	2/6 = 1/3 not vali
Beef chum & vice versa	8	3/6 valid

chicken milk	4	4/6 valid
chicken, cheese	2	1/3 not valid
milk cheese	1	not valid
Beef clothes	1	not valid
Chicken clothes	3	valish 3/6
Clothes milk	3	valid · 3/6

3) genrate 3 pair itemset.

Item	frequency	Support
Beef, Chicken, cheese	2	Notovalid 2/6
Beef, chicke, milk	2	Not valid 2/6
Beef cheese milk	1	Not vali 2/6
Chicken milk clothes	3	valid 3/6 = ✓

4)

Association	frequency	confidence
chicken, milk → clothes	3	3/4 not vali
Chick clothes → milk	3	3/3 valid
clothes, milk → chick	3	3/3 valid

clothes → milk, chicky	3	3/3	valid
milk → clothes chick	3	3/4	Not valid

(6)

∴ final association

Chicken : clothes → milk

- clothes milk → chicken

clothes → Milk, chicken

5) 3

2	238
2	119
2	59
2	29
2	14
2	7
2	3
	1

1110 1110

00000011 11011100