**Q1. Explain Gamma Encoding**

Can someone please explain gamma encoding ? It would be nice if you explain using example.

[week2](https://piazza.com/class/j6h1gozq5xh14z?cid=123)

[**edit**](https://piazza.com/class/j6h1gozq5xh14z?cid=123)·[good question](https://piazza.com/class/j6h1gozq5xh14z?cid=123)0

Updated 2 days ago by Hemendra Singh Choudhary

**the students' answer,**

*where students collectively construct a single answer*

Here is the link which shall help you understand better on this topic:

<https://www.csee.umbc.edu/~ian/irF02/lectures/05Compression-for-IR.pdf>

In case if you are still stuck, i can do a dry run of an example

**~ An instructor (Qihao Shao) endorsed this answer  ~**

[**edit**](https://piazza.com/class/j6h1gozq5xh14z?cid=123)·[thanks!](https://piazza.com/class/j6h1gozq5xh14z?cid=123)2

Updated 2 days ago by Mohit Khanna

**followup discussions**

*for lingering questions and comments*

Resolved Unresolved



[**Haibin Huang**](https://piazza.com/class/j6h1gozq5xh14z?cid=123) [2 days ago](https://piazza.com/class/j6h1gozq5xh14z?cid=123)

Correct me if I am wrong.

Decimal 15 as an example:

(1) Floor of log2(15) is 3;

(2) The unary part of 15 will 1+3 = 4, unary code for that is 1110 (4-1 = 3 of "1" followed by "0");

(3) The binary part of 15 will be 15-2\*\*(floor of log2(15)) = 15-8 = 7 and in 3 bits (determined by floor of log2(15)), so it will be 111;

(4) gamma code for decimal 15 will be 1110111.



[**Hemendra Singh Choudhary**](https://piazza.com/class/j6h1gozq5xh14z?cid=123) [2 days ago](https://piazza.com/class/j6h1gozq5xh14z?cid=123)

Right. I got better understanding with Slides shared - <https://www.csee.umbc.edu/~ian/irF02/lectures/05Compression-for-IR.pdf> by Mohit.

Reply to this followup discussion

Resolved Unresolved



[**Kahtan Al Jewary**](https://piazza.com/class/j6h1gozq5xh14z?cid=123) [2 days ago](https://piazza.com/class/j6h1gozq5xh14z?cid=123)

You can also check this link to get a better understanding of how to decode multiple digit gamma codes and get the d-gap sequence:

<https://cs.stackexchange.com/questions/28783/how-to-decode-multiple-digit-gamma-codes-and-get-the-gap-sequence>

In summary, to decode you can follow this process:

1. Read the stream of ones from left to right until you hit zero. Call this count n
2. Consider the zero you reach to be the start of your integer, with a value of 2n
3. Read the remaining n bits of the integer

Example (decoding gamma code for decimal 15 -> 1110111):

1. You have 3 ones before you hit zero, so n=3
2. The zero you reached has a value of 23=8
3. Read the next 3 bits after the zero to have 8+22+21+20=8+4+2+1=15

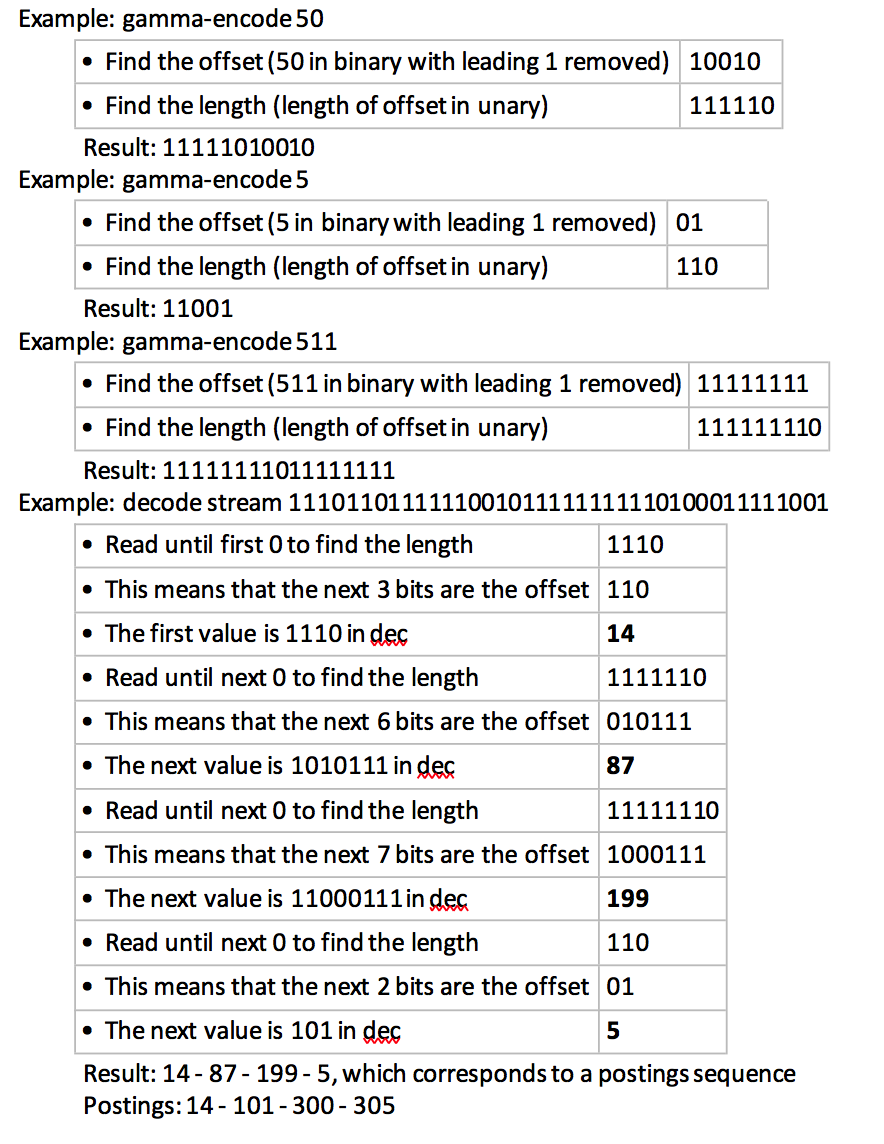
Reply to this followup discussion

Resolved Unresolved

https://d1b10bmlvqabco.cloudfront.net/photos/ixwgidfke2s7u/1490073328_35.png

[**Paco Cruz**](https://piazza.com/class/j6h1gozq5xh14z?cid=123) [12 hours ago](https://piazza.com/class/j6h1gozq5xh14z?cid=123)

Here are some step-by-step examples that I did to understand it correctly. Hope it helps.



**Q2. An easier variant of Gamma encoding**

I found the definition of gamma-encoding of integers in lesson 2.5 unintuitive and hard to understand.  However, I found the representation in this wikipedia article much more lucid in explanation. It is slightly different in how it uses the unary encoding but overall, I thought it is "beautiful" :-).  One can see right away why it works.

https://en.wikipedia.org/wiki/Elias\_gamma\_coding

[week2](https://piazza.com/class/j6h1gozq5xh14z?cid=106)

[**edit**](https://piazza.com/class/j6h1gozq5xh14z?cid=106)·[good note](https://piazza.com/class/j6h1gozq5xh14z?cid=106)1

Updated 4 days ago by Soumya Nanda

**followup discussions**

*for lingering questions and comments*

Resolved Unresolved



[**Xiaoming Ji**](https://piazza.com/class/j6h1gozq5xh14z?cid=106) [4 days ago](https://piazza.com/class/j6h1gozq5xh14z?cid=106)

This is not what has been presented in professor's lecture. We should use the version (https://www.csee.umbc.edu/~ian/irF02/lectures/05Compression-for-IR.pdf) as mentioned in [@98](https://piazza.com/class/j6h1gozq5xh14z?cid=98)

Also notice that log(x) is based on 2 not e(2.71).



[**Keh-Harng Feng**](https://piazza.com/class/j6h1gozq5xh14z?cid=106) [4 days ago](https://piazza.com/class/j6h1gozq5xh14z?cid=106) To be fair the two variants are really the same. The main difference being how the unary part is coded. I suspect coding with just 0 may not be a good idea because one cannot easily determine the number of bits in the unary portion when a variable length binary part also follows. Hence the variant of coding with 1 followed by a 0. There is also a third variant that is basically unary coding with 0 followed by a 1, probably created to address exactly the same thing.



[**Kahtan Al Jewary**](https://piazza.com/class/j6h1gozq5xh14z?cid=106) [2 days ago](https://piazza.com/class/j6h1gozq5xh14z?cid=106)

The course textbook uses unary coding with 0 followed by a 1

Reply to this followup discussion

Resolved Unresolved

https://d1b10bmlvqabco.cloudfront.net/photos/ixwgidfke2s7u/1490073328_35.png

[**Paco Cruz**](https://piazza.com/class/j6h1gozq5xh14z?cid=106) [1 day ago](https://piazza.com/class/j6h1gozq5xh14z?cid=106)

Agreed that for gamma compression we should use the Stanford book reference. But then the Unary encoding in the lectures is different than the one in the Stanford book, so it seems that people pick and choose at random (or at least, without explaining their choices).

Sometimes I wonder how practitioners of a discipline as precise as CS (where ambiguities aren't tolerated when processing an algorithm) can still be so ambiguous when interacting with other humans :)

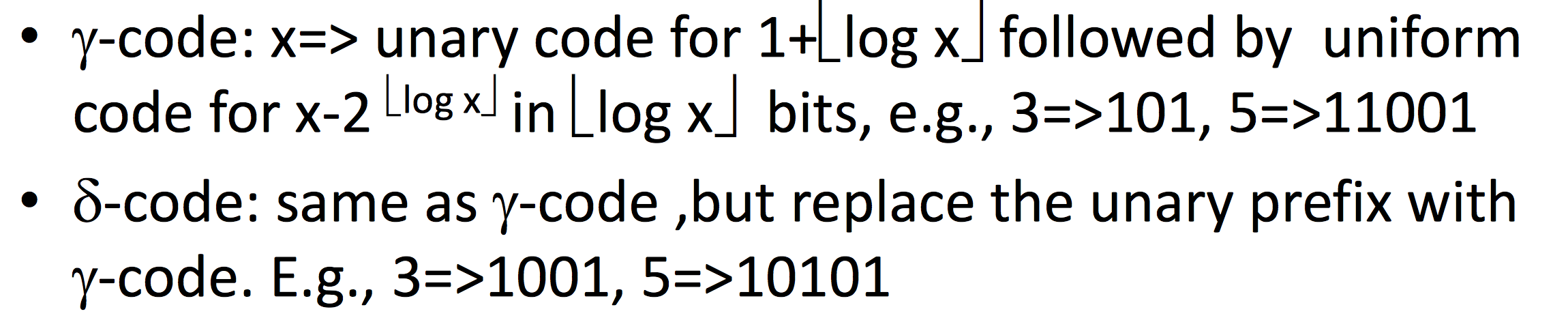


[**Keh-Harng Feng**](https://piazza.com/class/j6h1gozq5xh14z?cid=106) [1 day ago](https://piazza.com/class/j6h1gozq5xh14z?cid=106)

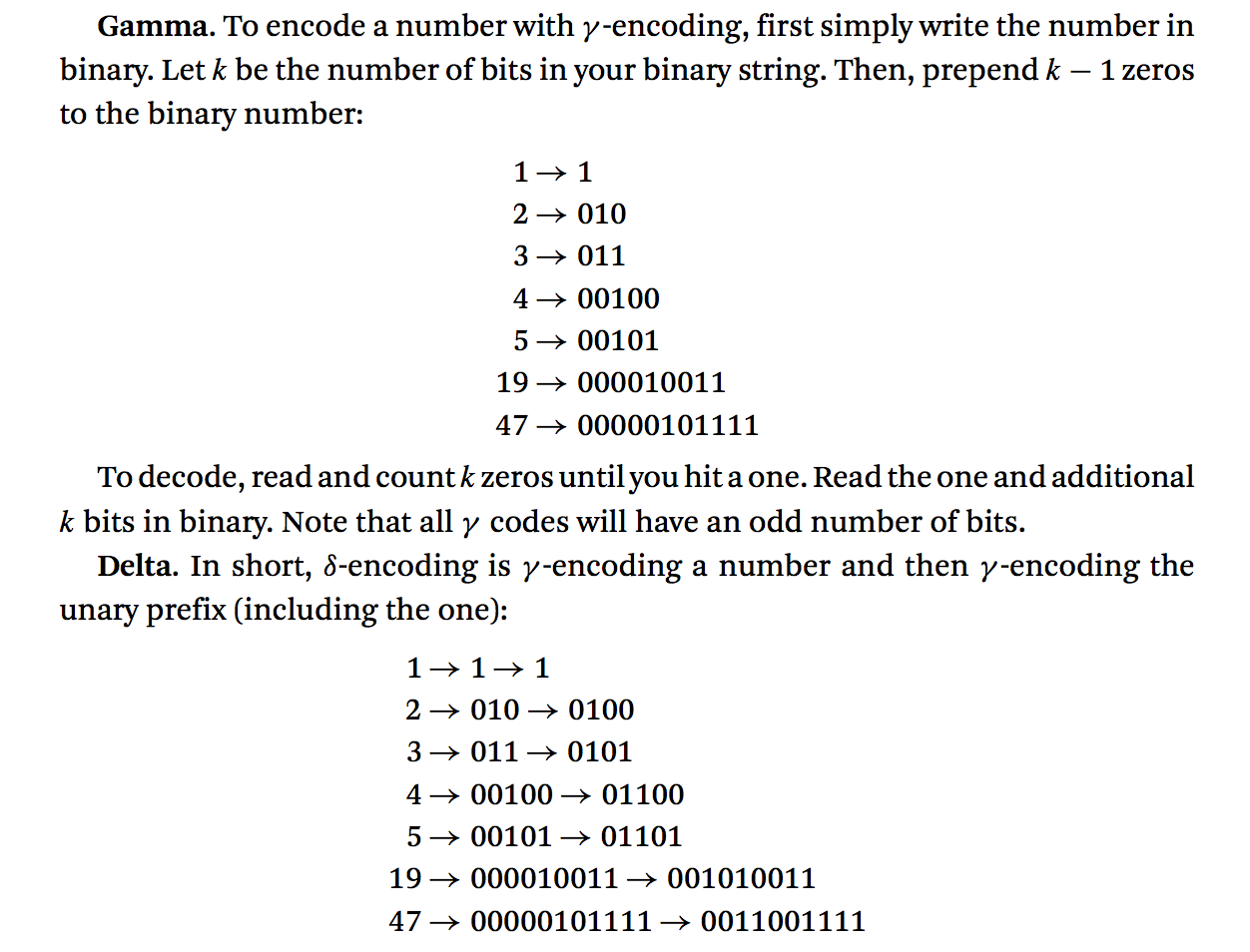
Functionality wise they are all the same though. As long as people using your code know which one you are using and stay consistent with it there shouldn't be any problems. As for answering exam/quizzes for this class, I'd say use the definition in lecture and forget everything else.

**Q3 Gamma code and delta code**

In the lectures, the two codes are defined as



However, in Chapter 8 on page 160 in the book, they are defined differently as below:



Which definition should we consider for these two compression methods?

[week2](https://piazza.com/class/j6h1gozq5xh14z?cid=98)

[**edit**](https://piazza.com/class/j6h1gozq5xh14z?cid=98)·[good question](https://piazza.com/class/j6h1gozq5xh14z?cid=98)1

Updated 5 days ago by Hooman Shirani-Mehr

**the students' answer,**

*where students collectively construct a single answer*

As far as i have read about it in other articles the one explained in lectures is correct

[**edit**](https://piazza.com/class/j6h1gozq5xh14z?cid=98)·[thanks!](https://piazza.com/class/j6h1gozq5xh14z?cid=98)0

Updated 5 days ago by Mohit Khanna

**the instructors' answer,**

*where instructors collectively construct a single answer*

For quizzes, exams etc please follow the lectures!

[thanks!](https://piazza.com/class/j6h1gozq5xh14z?cid=98)0

Updated 2 days ago by Ismini Lourentzou

**followup discussions**

*for lingering questions and comments*

Resolved Unresolved



[**Hadas Niv**](https://piazza.com/class/j6h1gozq5xh14z?cid=98) [5 days ago](https://piazza.com/class/j6h1gozq5xh14z?cid=98)

gamma coding has difference definitions. There is probably not a 'correct' one, but we need to know which definition to use.

for example:

<https://nlp.stanford.edu/IR-book/html/htmledition/gamma-codes-1.html>

<https://en.wikipedia.org/wiki/Elias_gamma_coding>

https://d1b10bmlvqabco.cloudfront.net/photos/j6w8p1mvopk1p8/1504065492_35.png

[**Hooman Shirani-Mehr**](https://piazza.com/class/j6h1gozq5xh14z?cid=98) [5 days ago](https://piazza.com/class/j6h1gozq5xh14z?cid=98)

Agree that gamma coding can have different definitions but the question is that which one we should consider for this course e.g. in the quiz, exam, etc.



[**Keh-Harng Feng**](https://piazza.com/class/j6h1gozq5xh14z?cid=98) [4 days ago](https://piazza.com/class/j6h1gozq5xh14z?cid=98)

Since lecture is mandatory and textbook is optional, I think it's pretty clear the definition listed in the Stanford book should be the one we use. This is also evident in the questions we have encountered so far in the quiz.



[**Vishal Dalmiya**](https://piazza.com/class/j6h1gozq5xh14z?cid=98) [4 days ago](https://piazza.com/class/j6h1gozq5xh14z?cid=98) I am frankly just following the lectures. Hopefullly that is the golden reference to use for this course.

**Q4. Can someone give some more pointers to gamma code frequency?**

gamma code for the term frequency of a certain document

I see some explanation in Lecture 2.5, but need more info...

[week2](https://piazza.com/class/j6h1gozq5xh14z?cid=97)

[**edit**](https://piazza.com/class/j6h1gozq5xh14z?cid=97)·[good question](https://piazza.com/class/j6h1gozq5xh14z?cid=97)0

Updated 5 days ago by Sudheer Kumar Kusuma

**the students' answer,**

*where students collectively construct a single answer*

[Actions](https://piazza.com/class/j6h1gozq5xh14z?cid=97)

Here is the link which shall help you understand better on this topic:

<https://www.csee.umbc.edu/~ian/irF02/lectures/05Compression-for-IR.pdf>

In case if you are still stuck, i can do a dry run of an example

[**edit**](https://piazza.com/class/j6h1gozq5xh14z?cid=97)·[thanks!](https://piazza.com/class/j6h1gozq5xh14z?cid=97)1

Updated 5 days ago by Mohit Khanna

**followup discussions**

*for lingering questions and comments*

Resolved Unresolved



[**Hadas Niv**](https://piazza.com/class/j6h1gozq5xh14z?cid=97) [5 days ago](https://piazza.com/class/j6h1gozq5xh14z?cid=97)

also: <https://nlp.stanford.edu/IR-book/html/htmledition/gamma-codes-1.html>



[**Sudheer Kumar Kusuma**](https://piazza.com/class/j6h1gozq5xh14z?cid=97) [5 days ago](https://piazza.com/class/j6h1gozq5xh14z?cid=97)

good one, able to get encode and decode

Thanks!