Coding Theory – I

I. Binary channels (High complexity)

1) *Basic question:* You have 2^{nR} possible messages you wish to transmit to someone. How many bits does it take for you to represent each message uniquely?
2) *Toy question* 1*:* Suppose you have to transmit one of two possible messages over a *fairly noisy channel* – less than half of the bits could get *flipped*. How would you proceed?
3) *Toy question* 2*:* Suppose your channel is not quite as noisy as the one in the previous problem – say strictly less than one-fifth of bits get flipped. How many distinct messages can you transmit on such a channel, with a guarantee that the decoder is able to accurately reconstruct the transmitted message? *(Be careful – deep waters!)*
4) *Definition – Hamming distance:* Consider the following model which is a generalization of the above examples – suppose you have a set of possible messages $\{m_1, \ldots, m_{2^{nR}}\}$, (for $R < 1$) only one of which you actually wish to transmit at any point in time. Each possible message m_i has an *encoding* x_i as an n-bit *codeword*. A certain fraction p of the transmitted bits get flipped on the noisy channel, so that pn of the n received bits in the received codeword y_i differ from the corresponding locations in the transmitted codeword x_i. Can you generalize your encoding/decoding schemes from the previous two problems to this model? That is, what would your decoding criteria be? Based on this, what should your criteria be for choosing the encoder's *codebook*?[1]
5) *Combinatorial numbers, Stirling's approximation [2], entropy function [3]:* **Try this one at home** : Prove that for large n, the number of length-n bit vectors that have Hamming distance at most pn from a given length-n bit vector is approximately[2] $2^{nH(p)}$. Here $H(p)$ is the *binary entropy function*, defined as $-p\log_2 p - (1-p)\log_2(1-p)$, and is strictly positive for $p \in (0, 1/2)$.
6) *Gilbert-Varshamov (GV) codes [4]:* Are you now ready to improve on your bound in 3)? In fact, in general can you prove that you can transmit approximately $2^{n(1-H(2p))}$ messages, *i.e.,*, at *rate* $1 - H(2p)$, without error over a channel that flips at most pn out of every n bits?
7) *Complexity of GV codes:* What is the encoding complexity of GV codes? What is the decoding complexity?
8) *Hamming bound [6]:* **Try this one at home:** Try proving an upper bound (non-achievability bound) to complement the lower bound (achievability bound) in 6). A "relatively simple" bound called the *Hamming bound* (see wikipedia) is based on "sphere packing", and shows that no rate more than $1-H(p)$ is achievable.[3]

II. "Large alphabet" channels (Low complexity)

Suppose instead of a binary channel, you have a *q-ary* channel – that is, you are allowed to input any number from the set $\{0, 1, \ldots, q-1\}$ into the channel. A fraction $1-p$ of the time, what you receive from the channel is what the transmitter transmitted. A fraction p of the time, however, what the encoder transmits is corrupted to some other arbitrary symbol from the set $\{0, 1, \ldots, q-1\}$.

1) Vandermonde matrices [7]: **Try this at home:** The $n \times m$ *Vandermonde* matrix G, for $n \geq m$, is defined as

$$G = \begin{bmatrix} 1 & r_1 & r_1^2 & \ldots & r_1^{m-1} \\ 1 & r_2 & r_2^2 & \ldots & r_2^{m-1} \\ \vdots & & & & \vdots \\ 1 & r_n & r_n^2 & \ldots & r_n^{m-1} \end{bmatrix},$$

with each r_i distinct from every other.[4] Prove that it has the following property – every $m \times m$ sub-matrix is invertible.

[1] Additional reading for HAROLD (Hypothetical Alert Reader of Limitless Dedication) – *Hamming distance* [1]

[2] In what sense is this "approximately" correct?

[3] For HAROLD, the best known current upper bound, called the MRRW bound, is over 30 years old, and is given in [5]. Whether either the GV bound or the MRRW bound or neither is tight is still an open question.

[4] HAROLD, be careful here – this condition implies a lower bound on the "alphabet-size" required for Reed-Solomon codes to exist, and hence this section is only for "large" alphabets.

2) *Reed-Solomon codes [8]:* Suppose your message has rate $1-2p$, that is, for every n channel uses, you wish to transmit $q^{n(1-2p)}$ messages. To do so, you perform the following encoding. You write your $m = n(1-2p)$ symbols as a q-ary column vector $\mathbf{u}$ of length m. You then left multiply $\mathbf{u}$ with a $n \times m$ Vandermonde matrix G, to obtain the length-n vector $\mathbf{x} = G\mathbf{u}$. Prove that for any distinct $\mathbf{u}$ and $\mathbf{u}'$, the number of locations in which the corresponding $\mathbf{x}$ and $\mathbf{x}'$ differ is at least $2pn$. Based on this, propose a decoding scheme for Reed-Solomon codes, so that any fewer than pn symbol errors can be tolerated.
3) *Complexity of RS codes:* What's the encoding complexity of RS codes? What's the naïve decoding complexity of RS codes? *For HAROLD, here's a fact [9] you can* **read at home** *– there exists a computationally efficient algorithm that decodes Reed-Solomon codes in time $\mathcal{O}(n^2)\log^3(q)$ (in fact, even more efficient algorithms exist).*
4) *Singleton bound [10]:* **Try at home:** Can you prove that in fact the best rates that are achievable on a q-ary channel where a fraction p of the symbols are corrupted, is $1-2p$?
5) *Modeling binary channels via q-ary channels:* Can you think of a means of modeling binary channels as 1-ary channels? What would be the advantage in terms of complexity? What would be the disadvantage in terms of achievable rate?

III. Concatenated codes for binary channels: High rate with low complexity

We saw in the last two sections that GV codes have good rates for binary channels, but have high computational complexity, whereas RS codes have low computational complexity, but bad rates for binary channels. The natural question is whether there's any way to combine these two techniques to preserve the desirable characteristics of both.

1) *"Short" GV codes:* Suppose you use a GV *inner code* with *block-length* $\log(n)$ instead of n. What is the design complexity? Encoding complexity? Decoding complexity? Does this indicate a possible "divide-and-conquer" strategy to transmit cn bits over n channel uses over a channel that flips up to pn bits? What could be a problem with a naïve such strategy?
2) *Code concatenation [11], [12]:* Can you think of a strategy, using an RS *outer code*, that could compensate for the drawbacks highlighted in the previous problem? What rates could you hope to achieve via such a strategy? What would be the design/encoding/decoding complexity of such a code?
3) *Random bit-flips/Forney's code construction for the Binary Symmetric Channel (BSC) [13]:* **Try at home (HARD)** Instead of being guaranteed that at most pn out of every n bits are flipped, suppose one is instead told that the probability of a bit-flip is p. How does this BSC differ from the one we've been analyzing thus far? Could the BSC be better than that previous channel? Could it be worse? Can you prove that you can efficiently communicate at a rate of approximately $1-H(p)$ over a BSC, and that this is the best possible rate?

IV. Further reading

If what you heard today whetted your appetite for Coding Theory and Information Theory, read [14], [15].

References

[1] http://en.wikipedia.org/wiki/Hamming_distance
[2] http://en.wikipedia.org/wiki/Stirling's_approximation
[3] http://en.wikipedia.org/wiki/Entropy_(information_theory)
[4] http://en.wikipedia.org/wiki/Gilbert-Varshamov_bound
[5] http://people.csail.mit.edu/madhu/FT01/scribe/lect9.ps
[6] http://en.wikipedia.org/wiki/Hamming_bound
[7] http://en.wikipedia.org/wiki/Vandermonde_matrix
[8] http://en.wikipedia.org/wiki/Reed-Solomon_error_correction
[9] http://courses.csail.mit.edu/6.440/spring08/scribe/lec10.pdf
[10] http://en.wikipedia.org/wiki/Singleton_bound
[11] http://en.wikipedia.org/wiki/Concatenated_error_correction_code
[12] http://people.csail.mit.edu/madhu/FT04/scribe/lect10.pdf
[13] http://en.wikipedia.org/wiki/Binary_symmetric_channel
[14] R. Roth, *Introduction to Coding Theory*. Cambridge University Press.
[15] T. Cover and J. Thomas, *Elements of Information Theory*. John Wiley and Sons, 2nd Ed., 2006.