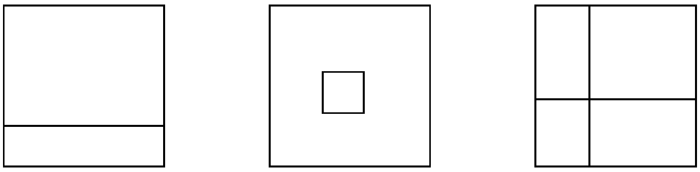
EE 376A/Stat 376A
Information Theory

Handout #26
Tuesday, February 24, 2009

Solutions to Homework Set #6

1. **Images.**
Consider an $n \times n$ array x of 0's and 1's . Thus x has n^2 bits.



Find the Kolmogorov complexity $K(x \mid n)$ (to first order) if

(a) x is a horizontal line.

(b) x is a square.

(c) x is the union of two lines, each line being vertical or horizontal.

Solution: Images.

(a) The program to print out an image of one horizontal line is of the form

```
For 1 ≤ i ≤ n { Set pixels on row i to 0; }
Set pixels on row r to 1;
Print out image.
```

Since the computer already knows n, the length of this program is $K(r|n) + c$, which is $\leq \log n + c$. Hence, the Kolmogorov complexity of a line image is

$$K(\text{line}|n) \leq \log n + c. \tag{1}$$

(b) For a square, we have to tell the program the coordinates of the top left corner, and the length of the side of the square. This requires no more than $3 \log n$ bits, and hence

$$K(\text{square}|n) \leq 3 \log n + c. \tag{2}$$

However, we can save some description length by first describing the length of the side of the square and then the coordinates. Knowing the length of the side of the square reduces the range of possible values of the coordinates. Even better, we can count the total number of such squares. There is one $n \times n$ square, four $(n-1) \times (n-1)$ squares, nine $(n-2) \times (n-2)$ squares, etc. The total number of squares is

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6} \approx \frac{n^3}{3}. \quad (3)$$

Since we can give the index of a square in a lexicographic ordering,

$$K(\text{square}|n) \le \log \frac{n^3}{3} + c. \quad (4)$$

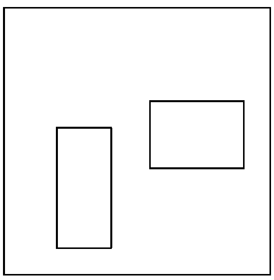
(c) In drawing a line, we must specify two pieces of information: whether it is horizontally or vertically oriented, and also the position of the line. The former requires one bit, and the latter requires no more than $\log n$ bits. Hence,

$$K(\text{pair of lines}|n) \le 2(1 + \log n) + c' = 2\log n + c. \quad (5)$$

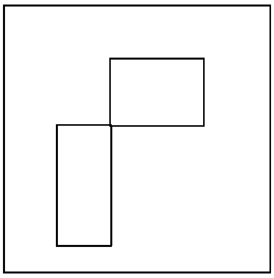
In all the above cases, there are many images which are much simpler to describe. For example, in the case of the horizontal line image, the image of the first line or the middle line is much easier to describe. However most of the images have description lengths close to the bounds derived above.

2. **Kolmogorov complexity.**
Assume n very large and known. Let all rectangles be parallel to the frame.

(a) What is the (maximal) Kolmogorov complexity $K(x|n)$ of the union of two rectangles on an $n \times n$ grid?



(b) What if the rectangles intersect at a corner?



(c) What if they have the same (unknown) shape?

(d) What if they have the same (unknown) area?

(e) What is the minimum Kolmogorov complexity of the union of two rectangles? That is, what is the simplest union?

(f) What is the (maximal) Kolmogorov complexity $K(x|n)$ over all images (not necessarily rectangles) on an $n \times n$ grid?

Solution: Kolmogorov Complexity.

First note that $K(\text{a single point on the screen}|n) \approx 2\log n + c$, since it takes $\log n + c$ bits to describe each of the two coordinates of the point.

(a) To specify one rectangle, we need to describe either the coordinates of two opposite corners (X, Y) and (X', Y'), or one corner (X, Y) (say the upper-left hand corner) and the length and width (L, W) of the rectangle. Hence, specifying a rectangle requires 4 numbers, each of which is $\leq n$, and therefore we need $4\log n + c$ bits for one rectangle, for a total of $8\log n + c$ for two rectangles.

We have not used the fact that the length and width of the rectangle are not independent of the position of the lower left corner – for example, if the lower left corner is near the NE corner of the square, the length and width of the rectangle have to be small. The number of distinct rectangles is thus not n^4, but rather, $\binom{n}{2}^2 = \left(\frac{n(n+1)}{2}\right)^2$. To see this, notice that in specifying the two corners of a non-degenerate rectangle, you must specify four values, two of which are distinct horizontal coordinates, and two of which are distinct vertical coordinates. The number of ways to choose two horizontal coordinates is $\binom{n}{2}$, and similarly for the number of ways to choose two vertical coordinates. Hence, the number of distinct rectangles is $\binom{n}{2}^2$. We could thus save some description length by enumerating all possible rectangles and then simply indexing a particular rectangle in a lexicographic ordering of this list. However, the key term remains unchanged when comparing $(n(n+1)/2)^2$ with n^4.

(b) Assuming two rectangles meet at a corner, we need to only describe 3 corners instead of 4. Let these corners be denoted by C_1, C_2, and C_3. Without loss of generality, we could let C_1 correspond to the corner at which two rectangles intersect. We then draw the rectangle with opposite corners C_1 and C_2, and also draw the rectangle with opposite corners C_1 and C_3. Hence,

$$K(x|n) \approx K(3 \text{ points}|n) \approx 6 \log n + c.$$

(c) We need to describe the upper-left and lower-right corners of one rectangle, and the corner of the other. Hence,

$$K(x|n) \approx K(3 \text{ points}|n) \approx 6 \log n + c.$$

(d) If the rectangles have the same area, then we can describe one rectangle fully and the other rectangle by only one corner and one side (the other side can be calculated). Thus,

$$K(x|n) \approx K(3 \text{ points}|n) + K(1 \text{ side}|n) = 7 \log n + c$$

(e) The simplest case is that the union of two rectangles is the whole $n \times n$ grid. And the Kolmogorov complexity is just a constant.

(f) An image is a specification for whether each pixel is black or white. Since there are n^2 pixels in the image, 1 bit per pixel, the maximal Kolmogorov complexity is $n^2 + c$ bits.