

Second Midterm

CSC 242

10 May 2002

Write your **NAME** legibly on the bluebook. Work all problems. Best strategy is not to spend more than the indicated time on any question (minutes = points). Open book, open notes.

1. Grammar: 30 mins.

We all remember: *Wenn Fliegen fliegen hinter Fliegen, Fliegen fliegen Fliegen nach.* And of course *Die Männer, die vor dem Schokoladen Laden Laden laden, laden Ladenmädchen zum Tanze ein.* (These Germanic noun capitalizations make parsing these sentences easier, *nicht?*)

A. Make up a context-free phrase structure grammar with a lexicon made up only of transitive verbs, adjectives, and nouns that will parse (or generate) sentences of the form:

cats fight dogs.

troy cats fight dogs.

cats fight albany dogs.

troy cats fight albany dogs.

cats love dogs albany residents hate.

syracuse cats hate albany dogs albany residents love.

rochester cats rochester dogs hate love syracuse cats albany dogs hate.

....

But which do not include sentences of the form:

cats fight.

big troy cats fight dogs.

the rochester cats fight large syracuse dogs.

...

(Note that we don't need no stinkin' capitalization.)

B. Now consider your grammar with the following lexicon. Noun: buffalo (as in bison). Adjective: buffalo (as in from Buffalo, NY). Verb: buffalo (as in baffle, frustrate). How many parses does your grammar give for the following sentence?

buffalo buffalo buffalo buffalo buffalo.

C. What is the longest sentence your grammar can generate (or parse) consisting entirely of repetitions of the word "buffalo"? Considered as a model of actual English usage, is the embedding property of these phrase structure grammars a problem? If not why not, and if so what do you suggest we do about it?

D. Consider the sentences

I dropped the lightbulb on the table and it broke.

I dropped the anvil on the table and it broke.

What mechanisms do NLU systems provide to resolve the reference of “it” in sentences like the above?

Answer:

A. I think something like this should work:

S = ModN V ModN

ModN = SubN | SubN NClause

NClause = SubN ModN V | SubN V

SubN = N | A N

N = cats | dogs

V = fight | love | hate

A = troy | syracuse | albany | buffalo

B. I get three, with the main verb being the 2nd, 3rd, or 4th buffalo.

C. Infinite. Definite problem because the famous linguistic touchstone, the “native informant”, will not agree an infinitely long sentence (or even a deeply embedded one) is grammatical. What to do, I dunno...one would have to check out some linguistics books. Clearly some sort of quick hack would be to put a limit on the number of times you invoke certain embedding rules like the 2nd and 3rd mutually recursive ones above.

D. Lexical categories and information could encode some of this common sense knowledge, but also logical rules about object properties and the associated inferences may have to be used.

2. Vision and Linear Systems: 30 mins

A. What does the convolution theorem for Fourier transforms tell us and why is it useful?

B. What does the sampling theorem for Fourier transforms tell us and why is it useful?

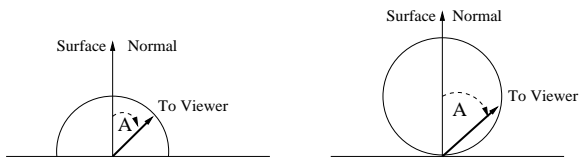
C. (see D.) What are three sorts of variation in an image that we might want to ignore for the purposes of object recognition? What algorithms or methods could implement such “invariance” calculations?

D. (see C.) On the other hand, how could each of the “irrelevant” variations in your answer to C be used to tell you something about the object? Be technical and specific.

E. On page 729 your text states that a Lambertian surface appears equally bright to an observer in any direction. That is, the number of photons per second impinging on a retinal receptor is constant for an observer looking at the surface from any angle. (Just how bright is given by the equation in the book, but that’s irrelevant to this question.).

Now I have another text that says that a Lambertian surface is one for which the photon flux, or amount of light (that is the number of photons per second emitted from the surface in any direction) is proportional to the cosine of the angle between the surface normal and the emittance direction. So one text says that the brightness is constant with angle A and the other says that the flux varies with A . In the figures, the arrow measuring brightness follows the semicircle and stays constant size, the arrow measuring flux follows around the circle and goes from 0 to 1 and back to 0 as the angle goes from 90 degrees to zero and back to 90.

Since these are textbooks, they have to be correct, so how do you reconcile these statements?



Length of Arrow is
Brightness in its
Direction

Length of Arrow is
Flux in its Direction

Answer:

A. convolution in time (spatial) domain is multiplication in (spatial) frequency domain, which yields elegant ways to think about the operation of linear operators. Also, since the fast FT is fast, the conv. th. allows convolution, usually an expensive (n^2 in one dimension) operation, to be done quickly ($n \log n$) by swapping domains.

B. If you sample faster than the Nyquist frequency (twice the highest freq in the signal), you can reconstruct a continuous (but band-limited) function with a finite number of discontinuous samples. Basic to digital image science.

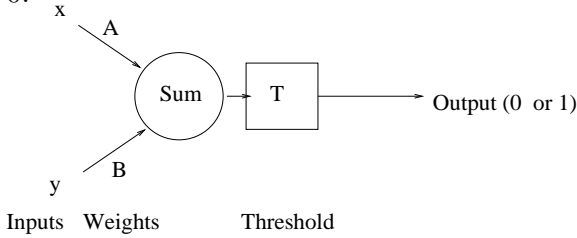
C. Shading, texture variation, perspective distortion (size variation with distance), color appearance with different colored lights, etc... To beat shading you could filter or edge-find to remove slowly-varying components. Or you could reason backward from assumptions of light source location and object surface reflectance uniformity. Texture variation could maybe be beat by something like “affine invariants”, that is extracting geometric properties of the texels and their arrangements that are invariant to size, rotation, and skew. Color constancy is something humans are pretty good at, and ways to get it vary from trying to guess the color of the light source(s) to assuming all light sources and reflectances are made up of a small number (3) of basis functions. Also the differences of colors between the various patches in the scene can be less variant than the colors themselves... this is basis of Land’s color theory. Size-invariant measures can beat down the “shrinks with distance” phenomenon, as would an inverse camera model.

D. Shape from shading, shape from texture: one approach is the equivalent of a reflectance function mapping properties (for shading, just brightness, for textures, maybe some other properties) to a small subset of surface normals. Then some regularization and boundary condition assumptions can allow for iterative solution. Using location or size in an image to infer distance (above or smaller means farther) is quite common.

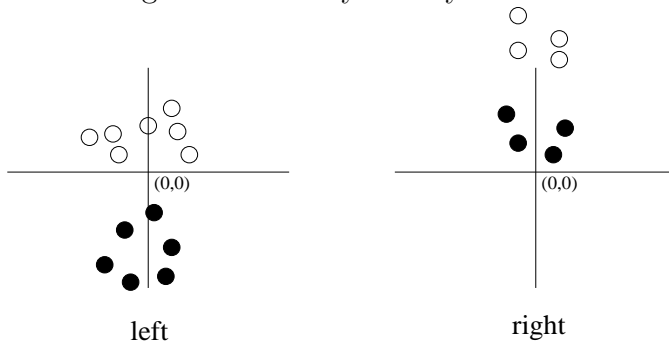
E. The foreshortening of the surface when you see it from more oblique angles means that each receptor in your eye gets photons from a larger chunk of surface. In fact the size of the surface goes as $1/\cosine$, while the flux is going as the cosine of that angle. These effects cancel, leading to the uniform brightness you experience.

3. Perceptrons: 40 mins

A. You have a perceptron with a threshold permanently fixed at zero (0) and two inputs, one for the x -coordinate of a data point and one for the y coordinate. It looks like this with T fixed at 0.



Can it discriminate between the filled and hollow data points in the data shown below for the left and right cases? Why or why not in each case?



B. You have a similar perceptron to the one in part A, only it can have a non-zero threshold. In fact you know that the weight on its x input is A , the weight on its y input is B , and its threshold is T . Specify the decision surface (in this case, the line) it implements. You can use any form of the line equation you want.

C. Now let's see a perceptron learn. Recall that if the perceptron's summed weight is over the threshold its output is 1, else 0. Again you have a similar perceptron, this time with weight 2 on the x input and weight 1 on the y input. The threshold is fixed at 2. The learning rate α is 0.2.

You give it input (2,2) (x input and y input both 2). It happens that this point is "in" the class to be detected, so the correct thresholded perceptron output is 1. Apply the perceptron learning algorithm – What happens to the weights?

D. Starting with the same perceptron as in part C (weights 2 and 1, threshold fixed at 2), you give the input (1,1). This point is NOT in the same class as point (2,2), so the perceptron's output should be 0. Apply the perceptron learning algorithm – What happens to the weights?

E. How many times do you have to present the input (1,1) and train the perceptron before it classifies (1,1) correctly?

Answer: A. The decision surface is fixed to go through the origin (as you can tell by the answer to part B, or figure out intuitively). Thus in the "left" case, the clusters on either side of the $y = 0$ axis can be discriminated but not the two lying on the same side of that axis.

B. $Ax + By + T = 0$ (or $y = (-A/B)x + (-T/B)$).

C. Nothing changes if there is no error in output.

D. The initial decision surface is $2x + 1y - 2 = 0$. The input $I = (1, 1)$ causes error signal of $E = -1$, and the perceptron learning formula being $w \leftarrow w + \alpha IE$ means that 0.2 gets subtracted from each of the two weights. At weights $1.8x + .8y$ the wrong thing still happens, but the sum is dropping by 0.4 each time you give this data, so by the third iteration you have weights of $1.4x + .4y$ which satisfies the equation and learning stops. Of course in real life you'd be able to adjust the

threshold as well with the same formula.

E. Three