Lecture 36: Integrals and AI

Can we build an artificial calculus teacher?

Machines assist us already in many domains: heavy work is done by **machines and robots**, accounting by **computers** and fighting by **drones**. Lawyers and doctors are assisted by artificial intelligence. There is no reason why teaching is different. The **web** has become a "gigantic brain" to which virtually any question can be asked or googled: "Dr Know" in Spielberg's movie "AI" is humbled: enter symptoms for an illness and get a diagnosis, enter a legal question and find previous cases. Enter a calculus problem and get an answer. Building an **artificial calculus teacher** involves calculus itself: such a bot must connect dots on various levels: understand questions, read and grade papers and exams, write good and original exam questions, know about learning and pedagogy. Ideally, it should also have "ideas" like to "make a lecture on artificial intelligence". But first of all, our AI friend needs to know calculus and be able to generate and solve calculus problems. [1]

Generating calculus problems

Having been involved in a linear algebra book project once, helping to generating solutions to problems, I know that some calculus books are written with help of computer algebra systems. They generate problems and solutions. This applies mostly to drill problems. In order to generate problems, we first must build **random functions**. Our AI engine "sofia" knew how to generate random problems with solutions. Random functions are involved when asked "give me an example of a function". This is easy: the system would generate functions of reasonable complexity:

Call the 10 functions $\{\sin, \cos, \log, \exp, \tan, \mathrm{sqt}, \mathrm{pow}, \mathrm{inv}, \mathrm{sca}, \mathrm{tra}\ \}$ **basic functions**.

Here $\mathrm{sqt}(x) = \sqrt{x}$ and $\mathrm{inv}(x) = 1/x^k$ for a random integer k between -1 and -3, $\mathrm{pow}(x) = x^k$ for a random integer k between 2 and 5. $\mathrm{sca}(x) = kx$ is a scalar multiplication for a random nonzero integer k between -3 and 3 and $\mathrm{tra}(x) = x+k$ translates for a random integer k between -4 and 4.

Second, we use addition, subtraction multiplication, division and composition to build more complicated functions:

A **basic operation** is an operation from the list $\{f \circ g, f + g, f * g, f/g, f - g\ \}$.

The operation x^y is not included because it is equivalent to $\exp(x \log(y)) = \exp \circ (x \cdot \log)$. We can now build functions of various complexities:

[1] In the academic year of 2003/2004, thanks to a grant from the Harvard Provost, I could work with undergraduates **Johnny Carlsson**, **Andrew Chi** and **Mark Lezama** on a "calculus chat bot". We spent a couple of hours per week to enter mathematics and general knowledge, build interfaces to various computer algebra systems like Pari, Mathematica, Macsyma and build a web interface. We fed our knowledge to already known chat bots and newly built ones and even had various bots chat with each other. We conceptionally explored the question of automated learning of the bots from the conversations as well as to add context to the conversation, since bots needs to remember previous topics mentioned to understand some questions. We learned how immense the task is. In the mean time it has become business. Companies like **Wolfram research** have teams of mathematicians and computer scientists working on content for the "Wolfram alpha" engine. Having recently seen a group at work here in Cambridge on Mass Av, I guess they generate probably in one day as much content as our Sofia group could do in a week for our "pet project".

A **random function** of complexity n is obtained by taking n random basic functions $f_1, \dots, f_n$, and n random basic operators $\oplus_1, \dots, \oplus_n$ and forming $f_n \oplus_n f_{n-1} \oplus_{n-1} \cdots \oplus_2 f_1 \oplus_1 f_0$ where $f_0(x) = x$ and where we start forming the function from the right.

1 **Visitor:** "Give me an easy function": Sofia looks for a function of complexity one: like $x\tan(x)$, or $x + \log(x)$, or $-3x^2$, or $x/(x-3)$.

2 **Visitor:** "Give me a function": Sofia returns a random function of complexity two: $x\sin(x) - \tan(x)$, or $-e^{\sqrt{x}} + \sqrt{x}$ or $x\sin(x)/\log(x)$ or $\tan(x)/x^4$.

3 **Visitor:** "Give me a difficult function": Sofia builds a random function of complexity four like $x^4 e^{-\cos(x)}\cos(x)+\tan(x)$, or $x-\sqrt{x}-e^x+\log(x)+\cos(x)$, or $(1+x)(x\cot(x)-\log(x))/x^2$, or $(-x + \sin(x+3) - 3)\csc(x)$

Now, we can build a random calculus problem. To give you an idea, here are some templates for integration problems:

A **random integration problem** of complexity n is a sentence from the sentence list { "Integrate $f(x) = F(x)$", "Find the anti derivative of $F(x)$", "What is the integral of $f(x) = F(x)$?", "You know the derivative of a function is $f'(x) = F(x)$. Find $f(x)$." }, where F is a random function of complexity n.

4 **Visitor** "Give me a differentiation problem". **Sofia:** Differentiate $f(x) = x\sin(x) - \frac{1}{x^2}$. The answer is $\frac{2}{x^3} + \sin(x) + x\cos(x)$.

5 **Visitor**: "Give me a difficult integration problem". **Sofia**: Find f if $f'(x) = \frac{1}{x} + \big(3\sin^2(x) + \sin(\sin(x))\big)\cos(x)$. The answer is $\log(x) + \sin^3(x) - \cos(\sin(x))$.

6 **Visitor**: "Give me an easy extremization problem". **Sofia**: Find the extrema of $f(x) = x/\log(x)$. The answer is $x = e$.

7 **Visitor**: Give me an extremization problem". **Sofia**: Find the maxima and minima of $f(x) = x - x^4 + \log(x)$. The extrema are

$$-\frac{\sqrt{\left(9+\sqrt{3153}\right)^{2/3} - 8\sqrt[3]{6}} + \sqrt{8\sqrt[3]{6} - \left(9+\sqrt{3153}\right)^{2/3}\left(1 + 6\sqrt{\frac{2}{9+\sqrt{3153}-8\sqrt[3]{6\left(9+\sqrt{3153}\right)}}}\right)}}{2 2^{5/6}\sqrt[3]{3}\sqrt[6]{9+\sqrt{3153}}}.$$

The last example shows the perils of random generation. Even so the function had decent complexity, the solution was difficult. Solutions can even be transcendental. This is not a big deal: just generate a new problem. By the way, all the above problems and solutions have been generated by Sofia. The dirty secret of calculus books is that there are maybe a thousand different type of questions which are usually asked. This is a reason why textbooks have become boring clones of each other and companies like "Aleks", "demidec" etc exist which constantly mine the web and course sites like this and homework databases like "webwork" which contain thousands of pre-compiled problems in which randomness is already built in.

Automated problem generation is the "fast food" of teaching and usually not healthy. But like "fast food" has evolved, we can expect more and more computer assisting in calculus teaching.

Be assured that for this course, the problems have been written by hand (I sometimes use Mathematica to see whether answers are reasonable). Handmade problems can sometimes a bit "rough" but hopefully some were more interesting. I feel that it is not fair to feed computer generated problems to humans. It is possible to write a program giving an answer to "Write me a final exam", but the exam would be uninspiring.

Corner detection

How do we detect corners in pictures? This is necessary to understand pictures, drawings. It might also be needed to see whether a given function is reasonably shaped. There should not be too many "wiggles" for example. There are various techniques to measure that. One of the best methods in computer vision uses the notion of **curvature**:

Given a function $f(x)$, define the **curvature** as
$$k(x) = \frac{f''(x)}{(1+f'(x)^2)^{3/2}} \, .$$

Is is a measure on how much the curve is bent at the point $(x, f(x))$. Positive curvature means the curve is concave up, otherwise concave down.

8 For a quadratic function $f(x) = x^2$, we have $\kappa(x) = 1/(1+x^2)$. We see that the curvature is maximal at the lowest part of the parabola.

9 For the function $f(x) = \sqrt{1-x^2}$, we have $f'(x) = -2x/\sqrt{1-x^2}$ and $f''(x) = -(1-x^2)^{-3/2}$. We have $(1+f'(x)^2) = 1/(1-x^2)$ and $k(x) = -1$.

10 **Problem:** Find the curvature for the graph of $f(x) = x^5/5 - x$. Where is the curvature maximal?

Here is a cool theorem:

If we integrate up the curvature along a graph of f on $[a, b]$ so that we always travel with constant speed on the graph, we get the angle difference $\beta - \alpha$.

Proof. $f'(x)$ is the slope of the curve and $g(x) = \arctan(f'(x))$ the angle. We have $\alpha = g(a) = \arctan(f'(a))$ and $\beta = g(b) = \arctan(f'(b))$. The fundamental theorem of calculus tells $\int_a^b g'(x)\, dx = g(b) - g(a)$. But $g'(x) = f''(x)/(1+f'(x)^2)$. If we travel so that $\sqrt{1+f'(x)^2} = 1$ then this is curvature.

It follows that if we integrate up curvature along the boundary of a region in the plane, we get 2π. This is a simple version of the Gauss-Bonnet theorem called Hopf Umlaufsatz.

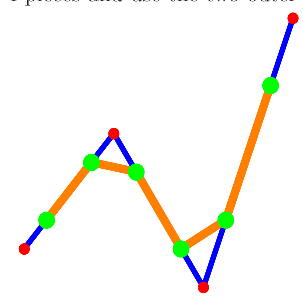
Connecting the dots

We want to connect points $P_1, \dots, P_n$ by a smooth graph. This "connecting the dots" problem is quite frequent. Our brain does this automatically. We need to see a few glances to "see" the motion of an object and predict where it will end. We need to connect dots if we drive a car, if we interpret a picture etc. On a more abstract level, we need to connect dots in the landscape of ideas whenever we solve a problem. We want to go from A to B and need to construct intermediate steps.
Here is a simple method found by G. Chaikin in 1974 [2] which generates a smooth curve through a few points.

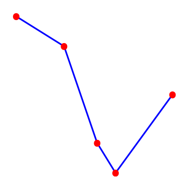
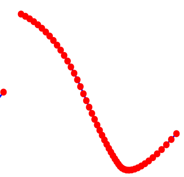
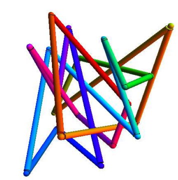
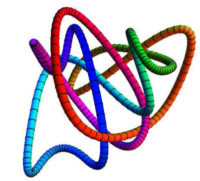
[2] G. Chaikin, An algorithm for high speed curve generation. Computer Graphics and Image Processing 3 (1974), 346-349.

Given a sequence of n points $P_1, ..., P_n$ define a new sequence of $2n-2$ points $R_2, ..., R_{2n-1}$ by
$$R_{2i} = \frac{3}{4}P_i + \frac{1}{4}P_{i+1}, \qquad R_{2i+1} = \frac{1}{4}P_i + \frac{3}{4}P_{i+1}$$
for $i = 1, ..., n-1$.

One such a step defines a **Chaikin step**. The limiting curve is called the **Chaikin curve** defined by the original points. The picture should explain how we get the new points from the old ones: divide each segment into 4 pieces and use the two outer points to get new points.



The Chaikin steps produce a smooth curve approximating a given set of points.

The method can be used for example to study the complexity of **random knots**. To answer the question stated initially:

AI will assist teachers in the future and help them to be more efficient.

Homework

1 A calculus bot wants to build a differentiation problem by combining log and sin and exp. Differentiate all of the 6 combinations $\log(\sin(\exp(x)))$, $\log(\exp(\sin(x)))$,$\exp(\log(\sin(x)))$, $\exp(\sin(\log(x)))$, $\sin(\log(\exp(x)))$ and $\sin(\exp(\log(x)))$.

2 Four of the 6 combinations of log and sin and exp can be integrated as elementary functions. Do these integrals.

3 Find the curvature of the sin curve at $x = 0$, $x = \pi/2$ and $x = 3\pi/2$.

4 Draw the points $(0, \sin(0)), (\pi/2, \sin(\pi/2)), (\pi, \sin(\pi), (3\pi/2, \sin(3\pi/2)), (2\pi, \sin(2\pi))$ and connect them with lines. Now do Chaikin iteration for 2 generations on paper.

5 Make a list of the 10 most important integrals you currently not know, look the integrals up and list the table here.