Prof. Raghuveer Parthasarathy
University of Oregon; Fall 2007
Physics 351

Computer Programming: A Brief Introduction

DUE DATE: *The exercises **P1-P3** at the end of this reading assignment are due by **2:30 pm** on **Wednesday, September 26, 2007**. I recommend handing them in at the start of Wednesday's class, but I've moved the due date to the afternoon in case there are small, last-minute questions you'd like to ask during office hours. (Don't even consider starting the assignment at 1:30 on Wednesday!) The assignment can be dropped into a box outside my office at any point before 2:30 on Wednesday. As noted in the syllabus, all the "programming exercises," which start with "**P**," are graded separately from the other problem sets.*

The uses and abuses of computer programming

A computer program consists of instructions to a computer to perform a series of mathematical operations. These instructions and operations can be designed to model some physical situation, and so computer programs are often used as tools by scientists and engineers.

REASONS *NOT* TO WRITE A COMPUTER PROGRAM. There are many reasons *not* to resort to programming, and no shortage of examples of pointless computational work. Science rests on the pillars of experiment and theory. An experiment tells us how nature behaves – for example watching a pendulum swing back and forth shows that there is periodic motion, and that this periodicity depends in particular ways on parameters such as the length of the pendulum (L) and the gravitational acceleration (g). Theory aims to provide quantitative, predictive models, whose truth is tested by their relation to experiments – for our pendulum, we can deduce that the period is $2\pi \sqrt{(L/g)}$. Suppose instead of formulating an analytic theory, we turned to a computer program that modeled Newton's equations of motions for a particular pendulum configuration and told us the period. We would have an answer, but would we have an *understanding* of pendulum motion? Would we have to run our computer simulation again and again for every different pendulum length we cared to consider? What if we want to build something more complex than a single simple pendulum – would we have learned anything about the components that make up our complex device, or would we have to simulate every tiny piece of it starting from Newton's laws of motion? These issues will become less abstract later on, but for now think of the following analogy: One could use a phrasebook and a translating dictionary to communicate with native speakers of a foreign language. As anyone who has tried it will tell you, this is more painful and less satisfying than actually learning the language. Learning a language implies a real *understanding* of its structure and vocabulary that adds depth to our usage of it. Similarly, learning about science through experiment and theory leads to an *understanding* that is deeper than that gained from computer simulations.

Even worse: Despite their popular "infallible" image, computer simulations are plagued with issues of accuracy. There are (*i*) "human" inaccuracies due to programming mistakes and, more perniciously, errors intrinsic to the ways in which they (*ii*) model numbers and (*iii*) simulate reality. As a simple example of (*ii*), consider the following relations: $x = \sqrt{2};\, y = x^2 - 2$. Of course, we know analytically that y should equal zero. However, if I program this into my computer, it claims $y = 4.4 \times 10^{-16}$. Why? Because √2 is an irrational number and cannot be exactly represented by a finite number of bits or digits. Therefore, calculations involving √2 necessarily contain errors due to the computer's truncation of the representation (1.41412...). Whether these errors are important or not depends on the task at hand. As a conceptual example of (*iii*), consider a pendulum bob that moves through continuous space and time, but is, necessarily, modeled on a finite grid of spatial and temporal points (e.g. by calculating its position every 10 microseconds). The fineness of the grid depends on the available computational resources, and the grid resolution limits the accuracy of the model. You'll show the effects of this explicitly in a few weeks. Issues of accuracy have led to all sorts of incorrect conclusions in scientific research.

One should also keep in mind the maxim "garbage in, garbage out:" any computer program merely implements the instructions programmed into it, i.e. the sequence of calculations prescribed by the model. Whether this model is wrong or right, the computer will calculate regardless – its reaching a conclusion is no measure of the accuracy of the underlying model. An astrologer, convinced that my happiness depends on the configuration of the moons of Jupiter, can write a computer simulation of all their complex, coupled orbits that is "correct" in terms of its astronomical output, but that is nonetheless nonsense.

REASONS TO WRITE A COMPUTER PROGRAM. Despite the above warnings, there are many situations in which computer programming is a valuable tool. Computers excel at performing repetitive, "brute-force" calculations, and these often arise in the modeling of physical systems and in the analysis of experimental data. (We'll see specific examples later.) In experiments, computers can often be programmed to control instruments (done routinely in the Parthasarathy Lab, for example, to simultaneously address millions of mirror-like pixels that sculpt laser light). Computers are also useful for visualizing data.

Exercises

The past few years have seen several "high profile" retractions of scientific papers from prestigious journals due to computational errors. Briefly scan each of the three articles and related documents below, and for each, **write a few sentences** describing the retraction and the nature of the error – for example, which of the "classes" of errors discussed above the mistake *may likely* correspond to. (No one besides, perhaps, the authors knows what the mistakes *really* involve.) Don't worry if you don't understand the scientific content of the papers, especially Set #1. You're not expected to. Try to glean a sense of the issues of accuracy involved. (Knowing what you *don't* need to read is a useful skill. If it takes you more than 20 minutes for any set of papers, you're probably reading too much.) *On-line links to the papers can be found on the course web site* – you may need a UO IP address or the UO proxy server to access them. If you have trouble, ask for help at the Science Library. Incidentally, having to retract one's papers is generally disastrous for one's career, though not as disastrous as not retracting a paper that's wrong.

P1. X-Ray crystallography (3 pts). X-ray crystallography is one of the most important techniques in materials science and structural biology. As we'll discuss next quarter, the diffraction of x-rays from atoms in a crystal reveals the atomic and molecular structure. Paper 1a is a retraction of several papers (including papers 1d and 1e, *which you don't have to read*); Paper 1b is a News feature on the retraction, and Paper 1c is a short letter to the editor. Paper 1f, by UO's own Brian Matthews, renowned in the field of crystallography, casts doubt on the story of Paper 1a – reading it is optional, but you might find it interesting.

1a. Chang, G., Roth, C. B., Reyes, C. L., Pornillos, O., Chen, Y.-J., and Chen, A. P. Retraction. *Science* **314**, 1875 (2006). [*http://www.sciencemag.org/cgi/reprint/314/5807/1875b.pdf*]

1b. Miller, G. Scientific publishing. A scientist's nightmare: software problem leads to five retractions. *Science* **314**, 1856-7 (2006). [*http://www.sciencemag.org/cgi/reprint/314/5807/1856.pdf*]

1c. Miller, C. (Letter) *Science* **315**, 459 (2006). [*http://www.sciencemag.org/cgi/reprint/sci;315/5811/459b.pdf*]

1d. Chang, G. and Roth, C. B. Structure of MsbA from E. coli: a homolog of the multidrug resistance ATP binding cassette (ABC) transporters. *Science* **293**, 1793-800 (2001). [*http://www.sciencemag.org/cgi/content/full/293/5536/1793*]

1e. Reyes, C. L. and Chang, G. Structure of the ABC transporter MsbA in complex with ADP.vanadate and lipopolysaccharide. *Science* **308**, 1028-31 (2005). [*http://www.sciencemag.org/cgi/content/full/308/5724/1028*]

1f. Matthews, Brian W. Five retracted structure reports: Inverted or incorrect? *Protein. Sci.* **16**, 1013-1016 (2007). [*http://www.proteinscience.org/cgi/reprint/16/6/1013*]

P2. Colloidal Physics (4 pts). These papers deal with the interactions of colloidal (few micrometer and smaller) particles. Experimentally, it was recently found that like-charged particles (all positive or all negative) can attract one another under certain conditions. The explanation remains controversial. Paper 2a is a simulation of the "DLVO model" of interactions, which has been very successful in describing colloidal physics. ("Adaptive finite element method" in paragraph 2 refers to the type of spatial grid used to model the system.) Paper 2b describes an analytic theory (not a simulation) of the consequences of DLVO interactions – just read the first two paragraphs. Paper 2c is a short simulation paper – read the first few paragraphs.

2a. Bowen, W. R. and Sharif, A. O. Long-range electrostatic attraction between like-charge spheres in a charged pore. *Nature* **393**, 663-665 (1998). [*http://www.nature.com/nature/journal/v393/n6686/pdf/393663a0.pdf*]

2b. Neu, J. C. Wall-Mediated Forces between Like-Charged Bodies in an Electrolyte. *Phys. Rev. Lett.* **82**, 1072-1074 (1999). [*http://prola.aps.org/pdf/PRL/v82/i5/p1072_1*]

2c. Gray, J. J., Chiang, B., and Bonnecaze, R. T. Colloidal particles: Origin of anomalous multibody interactions. *Nature* **402**, 750 (1999). [*http://www.nature.com/nature/journal/v402/n6763/pdf/402750a0.pdf*]

P3. Planetary mechanics (2 pts). Our solar system's giant planets have large obliquities (obliquity being the tilt of the rotational axis with respect to the orbital plane). Paper 3a reports a simulation of planetary mechanics that purports to explain its cause, and 3b is the author's retraction of 3a. (You need only read the first paragraph – the abstract – of 3a. You may find that the paper is very sparse on details of its methods.)

3a. Brunini, A. Origin of the obliquities of the giant planets in mutual interactions in the early Solar System. *Nature* **440**, 1163-5 (2006). [*http://www.nature.com/nature/journal/v440/n7088/pdf/nature04577.pdf*] Supplementary information: [*http://www.nature.com/nature/journal/v440/n7088/extref/nature04577-s1.pdf*]

3b. Brunini, A. Retraction: Origin of the obliquities of the giant planets in mutual interactions in the early Solar System. *Nature* **443**, 1013 (2006). [*http://www.nature.com/nature/journal/v443/n7114/pdf/nature05298.pdf*]