Research update, and a question for you

1 message

From Ben Kotrc <kotrc@fas.harvard.edu>
To Lazarus David <david.lazarus@mfn-berlin.de>
Date Thursday, March 24, 2011 12:18:42 PM
Subject Research update, and a question for you

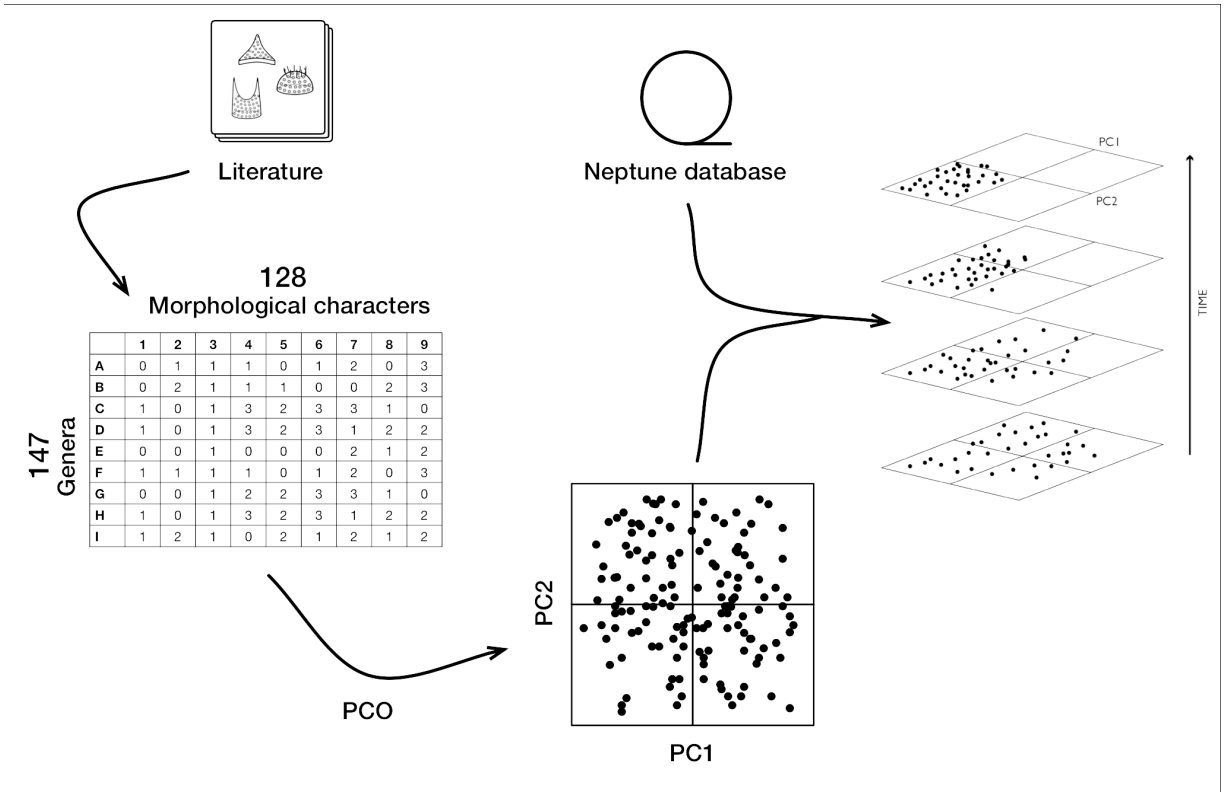
Hi Dave,

Here is a ~~short~~ much too long update (sorry!) on what I've been up to research-wise! [Feel free to skip down to the "Question for Dave" section if/when you get bored.] Over the past couple of months I've been working on two projects. One is a diatom morphospace, the other is prep work for the lineage-based follow-up to our rad silicification study.

Diatom project

With the diatom morphospace project I'm taking a discrete character approach, kind of like Mike Foote's crinoid morphospace from back in the day. I've amassed genus-level morphological descriptions for those 147 diatom genera that are in the Neptune database, and have distilled those into 128 binary and multi-state characters. I'm in the process of coding those into the various character states for each of the genera; once that's done the idea is to reduce those 128 dimensions down into 2D (using PCO), which makes a visual representation of that morphospace. Then the idea is to use Neptune to provide genus ranges, to resolve the occupancy of that morphospace through time.

Here's a picture of the process:



	1	2	3	4	5	6	7	8	9
A	0	1	1	1	0	1	2	0	3
B	0	2	1	1	1	0	0	2	3
C	1	0	1	3	2	3	3	1	0
D	1	0	1	3	2	3	1	2	2
E	0	0	1	0	0	0	2	1	2
F	1	1	1	1	0	1	2	0	3
G	0	0	1	2	2	3	3	1	0
H	1	0	1	3	2	3	1	2	2
I	1	2	1	0	2	1	2	1	2

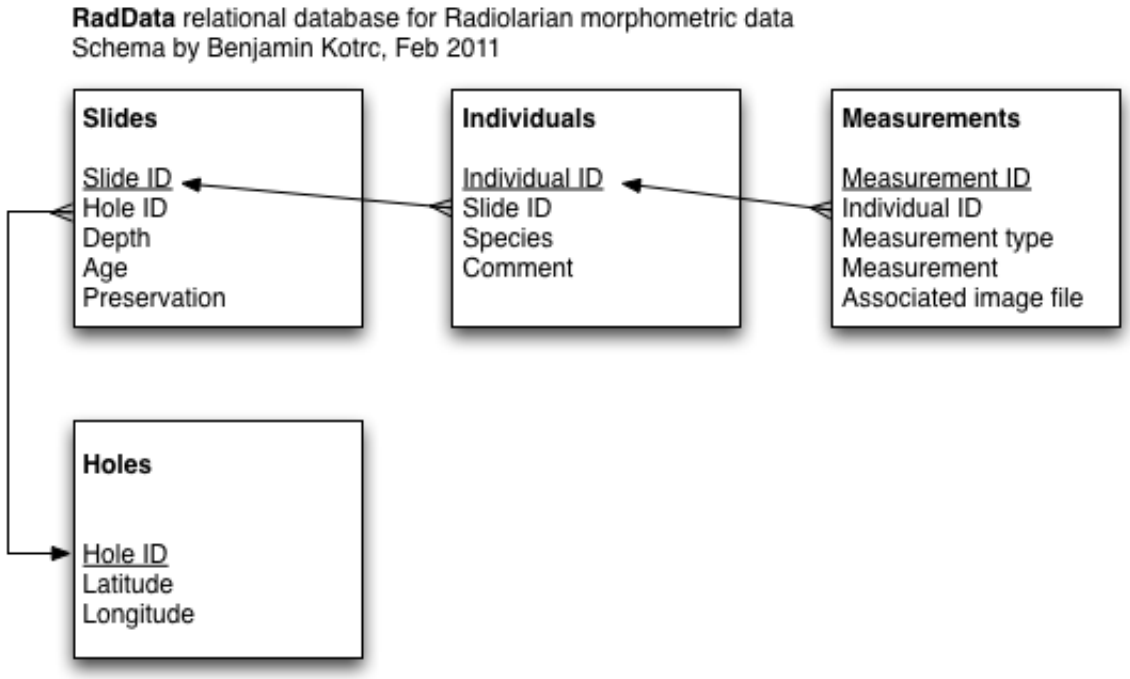
Radiolarian project

For the rad lineage project I flew down to DC a few months back to borrow some slides from the MRC (thanks to Brian Huber, who was really really nice and even showed me around their fantastic new marine biology exhibit he helped design!). I think we emailed back and forth a bit at the time I was choosing the slides; I used the MRC database and Neptune together to find the best candidates that are most likely to contain the lineages I've chosen—Diartus-Didymocyrtis, Artophormis, Stichocorys, Centrobotrys and Phormocyrtis. There are about 100 slides which should give me ~ a 20 slide time resolution per lineage, for a first pass.

For each of the lineages I made some geometric models—same principle as what we did assemblage-wide, but with more detail now, since the morphologies within lineages and species are much more constrained (lots of combinations of partial cones, spheres, cylinders, etc). That allowed me to determine which parameters I'll need to measure for each species—and since they're different geometric models in different species, different

measurements will need to be made.

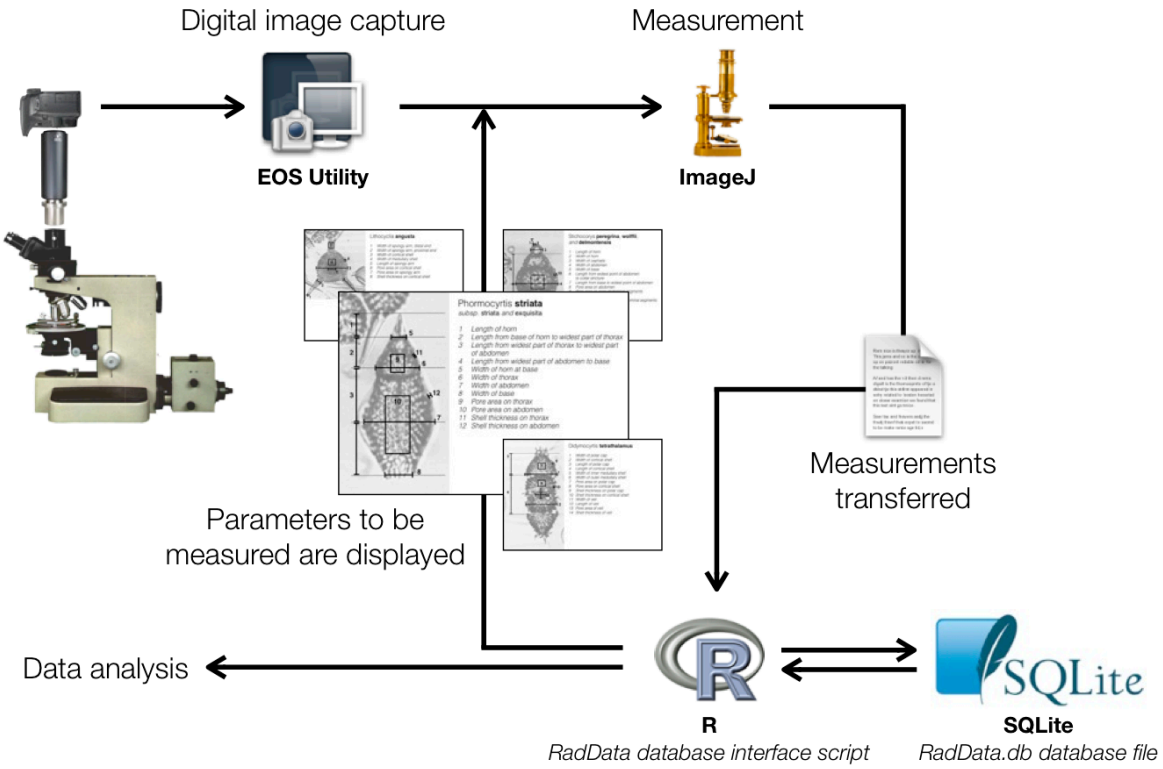
To manage all these measurements more intelligently, I followed your recommendation to flee the icy (=freeze-prone) clutches of Excel and set up a proper measurement protocol and data management system. I chose to store the data in a SQL database (and spent a week or so teaching myself SQL in order to do that!)—here's the schema for what I built (so exciting to show this to someone who'll actually understand it!!):



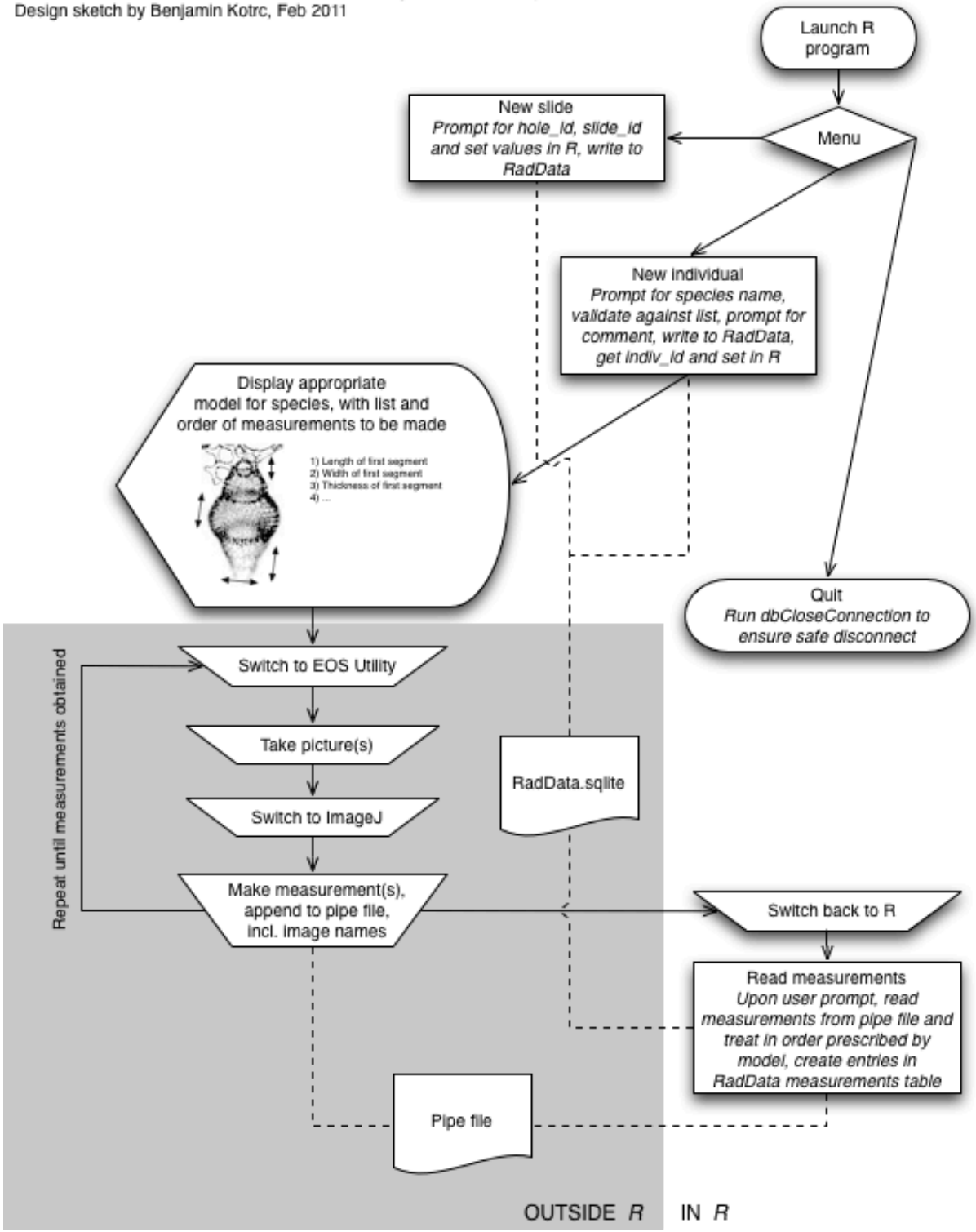
SQL is obviously too clunky to interact with through the command line for adding lots of data, so I built the database into a broader data collection protocol and interface based around R, which has a pretty easy-to-install SQLite engine available as a package.

Basically how it works is that I take images through the new scope camera we bought, a Canon EOS 500D/Rebel T1i, using the software Canon supplies. This is pretty neat because it allows for live (~20ish fps) on-screen preview at high resolution. Once a picture is taken it's opened automatically in ImageJ (another nice feature in the Canon software does this), where measurements are made and written to a text file along with the magnification information and name of the .JPG file associated with each measurement, using some ImageJ macro magic. This text file acts as a conduit to shuttle data between ImageJ and R (maybe not super elegant, but it was the easiest way I could think to do it—and it works!). The R interface provides a little graphic alongside this process, telling me which measurements I need to make, and the order in which to make them, depending on which particular species I'm looking at. Once I've made all the measurements, I prompt the R script to read those measurements in from the text file and it parses them into a series of SQL INSERT statements.

There's a bit more to it with various other ImageJ macros and R functions that make sure the scale is set and stored for each new image that's taken, keep track of what measurement I'm on, make sure the right number of measurements are in the text file, automatically calculate pore area proportions, etc—but that's the basic story. Here's a graphical cartoon representation of the protocol:



If you want to see a slightly more technical, under-the-hood look at how this works, here's my design sketch for the interface. It's not a proper flow-chart but it gives a basic idea of how the R script works and plays with the image acquisition and measurement steps. I've added several other functions to the main menu of the interface, including a "review" feature to see the last set of measurements added, a "backup" feature that writes all the tables to tab separated text files (just in case there's some sort of corruption at some point!). Anyway, feel free to ignore this part if it doesn't make sense.



As you can see both the diatom and rad projects are pretty close to "data collection" phase, which is exciting. My next goal is to crank through the diatom character matrix and get some results on that project.

Question for Dave!!

The radiolarian project is also almost ready for making measurements—I was testing the measurement protocol yesterday and was about to calculate the pixel scale for each objective on the scope when I realized I couldn't get an image out of the 63x objective on our lab scope. I asked Andy for help and he couldn't make it work either—neither on the lab scope, nor on the one in his office, which has a similar 63x objective as well as a 100x objective. We tried a couple of different slides, with no high-mag success. His best guess was that the working distance of those objectives was too short, and that we might need to buy another objective that's better suited for rad work.

There is a decent 40x Olympus objective (DPlan 40PO 0.65, 160/0.17) that seems like it would work well for the length/width/pore area measurements, but I don't think the resolution is really good enough for the shell thickness measurement.

The objectives that haven't been giving us joy are oil immersion objectives, one is a Leitz PL APO 63/1.40 (*160/0.17), the others are another 63x

and a 100x (I think they're also Leitz, but they're locked into Andy's office at the moment so I can't go and check the exact specs). Both the lab scope and Andy's scope are Leitz Orthoplan (one looks like it's a 50s or 60s model, the other I would guess 80s).

So the questions I was hoping to ask are (1) are we doing something wrong? Should these objectives work fine with rad slides? And if not, (2) can you recommend a better objective we could buy?

Alright. In my usual fashion I've gone on for far too long. Let me know if you have any suggestions about the objectives! Also, if you think the measurement protocol might be of use for you in a project some day, let me know. All the code is fully commented and I can prepare a list of all the steps that would need to be taken to adapt it to making a different set of measurements—it would need some tinkering, but not an insane amount.

Best regards from Cambridge (where it snowed again two days ago!!),
- Ben.

--
Ben Kotrc
Sent with Sparrow