Mini Thesis Proposal

Andria C. Schwortz

Exploring the Differences Between Radio-Quiet and Radio-Loud Quasars
and
Expert vs. Novice Differences in Processing Large Databases

Advisors: Dr. Andrea Burrows and Dr. Adam Myers
Department: Physics & Astronomy, University of Wyoming
Committee Members: Dr. Rich Barrans, Dr. Mike Brotherton, Dr. Danny Dale

Science Project

Introduction / Background

The name "quasar" comes from a shortening of the term "quasi-stellar radio source": these objects typically appear as point sources in images of the sky, much like stars, but their actual spectra do not resemble those of stars. In addition, it was later discovered that not all quasars have radio emission, therefore "quasi-stellar object", or QSO, is the more recently coined term for them.

Quasars are very bright sources located in the cores of galaxies far from the Milky Way, or equivalently, existing long ago. The sizes of quasars are smaller than the Solar System, and it is believed that they are powered by hot gas spiraling into supermassive black holes (SMBHs) present in the cores of all galaxies. However not all galaxies possess quasars, and the most popular model is that the merger of two large galaxies is what causes the material to migrate in towards the center of the galaxy where the supermassive black hole is located.

Figure 1. Artist's rendition of a quasar. ESO/M. Kornmesser. http://www.eso.org/public/images/eso1122a/ 29 June 2011.

Because of the distance to quasars, the nuclear region of these galaxies (and in many cases the entire galaxy as a whole) is unresolved, so all the information we have comes from spectra and the brightness at different wavelengths. When we look at quasars in different types of light, we find that in the optical we see something that is extremely bright and hot (presumably the gas as it approaches the black hole). When we look in the radio, we find that some of the sources are bright in radio emission, and we call these radio-loud quasars (RLs). Based on the shape of the radio spectrum, we know that this is due to synchrotron radiation: relativistic charged particles spiraling around a magnetic field. We believe that the geometry is gas jets comprised of outflowing material entangled with magnetic fields. While the physics of the emission from the jets is well-known, the details of how the jets themselves are formed is not well understood, and thus understanding RLs will help us to clarify this picture. In addition, the vast majority of quasars (~90%) have radio emission; the other category is radio-quiet quasars (RQs).

There are two main competing hypotheses regarding the categories of RL and RQ. One says that there are two physically distinct types of quasars, formed in different environments or through different processes, and that each type is constant in time – RLs are always RLs, and RQs are always RQs. The other hypothesis is that all quasars go through phases of being RL and RQ, and that this happens in a periodic fashion and can be described by a duty cycle.

One way to study quasars is to look at their environments. It has been known for a while that RLs live in more dense environments than RQs: if we study individual quasars, we find that there are more nearby galaxies if the quasar is RL than if the quasar is RQ. Since the distribution of luminous matter (that is, galaxies) traces the underling distribution of dark matter, this implies that RLs are found in environments with larger dark matter halos than are RQs. More recently large catalogs of quasars such as the Sloan Digital Sky Survey have allowed us to look at the distribution of multiple quasars – that is, how close are quasars to each other, or how are they clustered together. This is quantified with the correlation function, essentially measuring how many quasars are at a particular distance from each other – see Figure 2 on the right, each blue circle indicates a certain distance from a given quasar. Clustering results from previous authors (e.g., Shen, et al., 2009) also imply that RLs form in more dense environments and thus are intrinsically distinct from RQs due to their formation process.

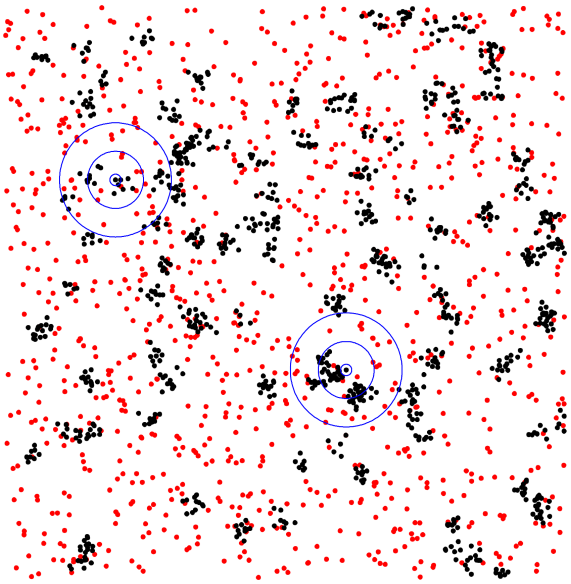

Figure 2. Black dots represent clustered quasars while red dots are randomly placed. Blue circles serve as guides to visualizing clustering. Myers c/o DiPompeo.

Shen, et al. (2009) has also proposed one additional distinguishing characteristic between RLs and RQs: RLs tend to have C_{IV} emission that is less blueshifted than that in RQs. Some of the C_{IV} emission may arise from winds originating from the disk around the SMBH, which would result in a blueshift (that is, the observed motion of these winds is towards the observer because the receding winds are blocked from view by the disk). RLs appear to have weaker blueshifted C_{IV} emission suggesting that although they possess stronger jets than RQs, they may have weaker disk winds than do RQs, or possibly non-existent winds.

If this were the extent of the data however, then we would not have anything to research. In fact though, Richards et al. (2007; in particular, Figure 7) have noted that there exist some quasars without radio emission and thus are traditionally categorized as RQs, but which also do not possess strongly blueshifted C_{IV} emission, a characteristic we expect from RLs. As a result, we are calling these RQRL quasars. We do not at this time know whether these specific objects have clustering tendencies more similar to that of RLs (tightly clustered) or RQs (less tightly clustered).

This project seeks to further explore how the clustering properties of RL and RQ quasars correlate with C_{IV} emission. If we determine that RQRLs are as tightly clustered as are RLs, then this will imply that they could be RLs which are in an "off" portion of their

duty cycle. If we determine that RQRLs are as loosely clustered as RQs, then we will need to seek another reason why they possess blueshifted C_{IV} emission.

Method

Data Release 5 of the Sloan Digital Sky Survey (SDSS DR5) contains approximately 77,000 sources identified as quasars. We will use a subset of uniformly selected DR5 quasars to match our random catalog (i.e. as illustrated by the red points in Figure 2), which will restrict our sample to approximately 38,000 quasars.[1] The DR5 quasar catalog includes information from the VLA FIRST radio survey. Separating our samples based upon whether or not they are detected in FIRST, the DR5 quasar catalog contains 33,279 RQ sources and 2,943 RL sources, which is consistent with previous findings that approximately 10% of quasars are RLs.

Next we will modify the autocorrelation routine from Eftekharzadeh, et al. (in prep), to become a cross correlation routine. Autocorrelation compares each source to another source in the same sample (e.g., comparing RLs to RLs), while cross correlation will compare a given sample to all quasars (RLs to all 36,222 quasars) – this allows us to highlight differences between the RQ and RL samples while leveraging the larger complete sample size. This routine will measure the 3D correlation function, which (described as a possible estimator by Landy and Szalay, 1993) is given by the expression

$$\xi(s) = \frac{DD(s)}{DR(s)} - 1,$$

where *DD(s)* indicates the number of data-data pairs at a co-moving separation of s, and *DR(s)* indicates the number of data-random pairs. In Figure 2, above, the black points indicate data (that is, quasars), while the red points indicate random points. We will measure our error bars using jackknife error sampling, a standard Monte Carlo approach. We will split our region into 16 equal area sections on the sky and recalculate the cross-correlation after successively removing each sample.

We will fit a power-law to the correlation function, and compare to models (also provided by Eftekharzadeh and DiPompeo, both in prep). Depending upon the difference in the amplitudes of the correlation functions, we'll determine whether we have two categories of quasars, RL and RQ, and into which of these two categories the RQRLs fall.

Education Project

Research Question

---

[1] In this study, we will also have the opportunity to use the larger, more recent SDSS Data Release 7 quasar catalog, which will roughly double the size of our uniform sample. This would represent the first published measurement of quasar clustering using DR7

My research questions for this project are "How do college students fit into the framework of novices and experts?" and "How do college students process big datasets for analysis?"

Theoretical Framework

This project will be conducted from the theoretical framework of social constructivism. Constructivism is the viewpoint that students come into the classroom with preconceptions, and they need to actively build new ideas (rather than passively absorbing them) and incorporate what they are learning into their existing ideas. The social aspect means that in this study we take the stance that students do not optimally learn in isolation, they instead learn best with their peers and together construct meaning, and thus we will examine a group learning experience.

Literature Review

The *Handbook of Research on Science Education* (Abell and Lederman, eds., 2013), first published in 2007 and endorsed by the National Association for Research in Science Teaching (NARST), has served as a comprehensive literature review of the many fields of science education research, and as guidance for future work, since its first edition. Throughout the many articles within the *Handbook*, one theme that recurs again and again is that of the transition from novice to expert.

The differences between novices and experts in many subfields of science have already been studied extensively. Bransford, Brown, and Cocking (2000) described six key characteristics distinguishing experts and novices, such as experts seeing a gestalt of the situation which is more than the sum of its parts while novices tend to focus on details, or how experts are able to put a problem into the larger context of their field due to their greater store of content knowledge. In the realm of physics, many researchers (e.g., Larkin, et al., 1980) have studied the differences between how experts and novices approach individual physics problems. For example, it is known that experts are able to distinguish when the shape of an object (i.e., rectangular vs. round) is relevant to the problem (such as when moving down a ramp with friction) as opposed to when it is irrelevant (such as when hanging from a rope) and will choose an approach making use of the relevant factors, while on the other hand a novice is more likely to group all problems with round objects together, regardless of whether the best approach is to use conservation of energy or torque. Because at its root astronomy is a branch of applied physics, one can assume that the differences between how experts and novices approach individual astronomy problems would be similar, but I have been unable to find any studies directly investigating novice/expert distinctions in solving individual astronomy problems.

In addition, astronomy as it exists today is increasingly making use of large sets of data (see for example, the 100,000-plus data set used in the quasar research discussed above). These datasets must be analyzed as a whole, rather than one data point at a time, using extensive computer programs written specifically for the question the researcher wishes to investigate.

With the nation-wide push for STEM integration (for example, in the Next Generation Science Standards, or in NASA's educational goals), big datasets become even more important. The skills that students learn from analyzing astronomical datasets can be applied to everything from other STEM fields, to physical inventory, financial predictions, and human resources management. Because this need for handling big datasets is ubiquitous, it is important that we study the process by which students transition from novices to experts in this field. And yet, as any astronomy researcher can tell you, there exist few or no courses in most astronomy degrees which directly address the use of large sets of data, even at the graduate level.

Methodology

This will be a mixed methods study. Quantitative measurements will include student scores on a pre-/post-test and on the specific task as well. Qualitative methods will explore the process by which students construct meaning from large datasets.

Methods

We will be developing an activity or task wherein students will analyze data approximating that given for SDSS quasars, but of a scope more appropriate for a single classroom or lab period. Students will be sorting the quasars into categories based upon their properties such as radio emission and C_{IV} velocities.

For the quantitative aspect of the study we are aiming for a sample size of N=50 or greater. we will develop a pre-/post-test with questions fitting the required parameters for ANOVA. Qualitative data will include open ended questions from the pre-/post-test, field notes taken during administration of the big dataset task, videos taken of group work during the task, and pre-/post-interviews. The target sample size for the pre-/post-interviews is N=10.

We are currently planning to use this activity in three courses this summer: Astro 101 students at a large research university in the Midwest, online Astro 101 students at a mid-sized junior college in the Northeast, and in-service teachers attending a summer professional development opportunity in the Midwest. As a part of the normal curriculum development process, we will be administering a pre-/post-test to these students and will solicit their feedback on the materials.

The sample for the full study in the Fall will consist of Astro 101 students from a single large research university in the Midwest USA. We are currently in the process of recruiting faculty members and TAs who are willing to have their lab sections participate.

Analysis

We will examine the gains of the students on the quantitative questions, and use ANOVA to determine whether any subsets of the students are distinct from the majority. The qualitative data will be coded for themes, and we will search for evidence of novice/expert characteristics.

Limitations

For the full study in the Fall, we plan to administer the big dataset task a single time in each participating course. Should anything go wrong, we would next be able to administer the task in the Spring of 2015.

To maximize the amount of data obtained, we will be training a few assistants to work with Schwortz on taking field notes. While this will increase the amount of data collected, there may be cross-evaluator differences that need to be taken into account.

At least one of the pilot studies this summer will have Schwortz as the instructor administering the task, possibly introducing a bias since the dataset to be used is one with which she is already familiar.

For the Fall study we plan to recruit other instructors to administer the task, should some students end up taking the task with an instructor with whom they are unfamiliar and others with an instructor with whom they are familiar, this may affect the outcome. In addition, if different groups of students have different instructors administering the task, student results could reflect the instructor's familiarity with the dataset, their teaching proficiency, and their pedagogical content knowledge.

References

Abell, Sandra K., and Norman G. Lederman, ed. 2013. Handbook of Research on Science Education: ProQuest Tech Books. Focal Press.

Bransford, John D., Ann L. Brown, and Rodney R. Cocking. 2000. How People Learn: Mind, Brain, Experience, School. National Academies Press.

Eftekharzadeh, S. *In prep.*

DiPompeo, M. *In prep.*

Landy, Stephen D., and Alexander S. Szalay. 1993. "Bias and Variance of Angular Correlation Functions." The Astrophysical Journal 412 (July 1): 64. doi:10.1086/172900.

Larkin, Jill, John McDermott, Dorothea P. Simon, and Herbert A. Simon. 1980. "Expert and Novice Performance in Solving Physics Problem." Science.

Richards, Gordon T., Nicholas E. Kruczek, S. C. Gallagher, Patrick B. Hall, Paul C. Hewett, Karen M. Leighly, Rajesh P. Deo, Rachael M. Kratzer, and Yue Shen. 2011. "UNIFICATION OF LUMINOUS TYPE 1 QUASARS THROUGH C IV EMISSION." The Astronomical Journal 141 (5) (May 1): 167. doi:10.1088/0004-6256/141/5/167.

Shen, Yue, Michael A. Strauss, Nicholas P. Ross, Patrick B. Hall, Yen-Ting Lin, Gordon T. Richards, Donald P. Schneider, et al. 2009. "QUASAR CLUSTERING FROM SDSS DR5: DEPENDENCES ON PHYSICAL PROPERTIES." The Astrophysical Journal 697 (2) (June 1): 1656–1673. doi:10.1088/0004-637X/697/2/1656.