**Graduate Research Plan Statement**

Name: Brandon Barker

Application Link: <https://astrobarker.github.io/essays/nsf_grfp_research.pdf>

Personal Statement Link: <https://astrobarker.github.io/essays/nsf_grfp_personal.pdf>

**Research Proposal – Abstract:**

Core-collapse supernovae (CCSNe) are the spectacular explosions that accompany the deaths of massive stars. CCSNe have been the subject of ongoing research for decades but the explosion mechanism is still not fully understood. Proper treatment of the CCSN problem requires all areas of physics; in particular, general relativistic gravity, complex neutrino transport, turbulent magnetohydrodynamics, and nuclear physics. CCSNe are the primary engines of galactic chemical evolution; many of the elements heavier than H and He are synthesized here, those necessary to life in particular. Furthermore, a complete view of CCSNe is necessary for understanding the compact objects that arise from core-collapse, such as the binary neutron stars and stellar mass black holes that have been detected by Advanced LIGO and Advanced Virgo[1].

**Research Proposal – Intellectual Merit**

As current high-fidelity CCSN simulations lack the ability to properly predict electromagnetic (EM) emission, I propose to investigate the explosion mechanism driving CCSNe by further developing the current CCSN simulation capabilities to provide accurate EM predictions. While yielding far more insights into the processes at work in the CCSN, it will also allow for comparison with observational astronomy. This collaboration of full multimessenger signals is a yet untapped resource that could give new insights into the explosion mechanism. The proposed project is broken down into three stages: (i) upgrading our nuclear physics, (ii) getting 1D LC and EM information, and (iii) going to multiple dimensions. My background in stellar astrophysics and computational methods makes this project a natural next step.

Upgrading our nuclear physics. In the hot interior of massive stars, material is said to be in nuclear statistical equilibrium (NSE), meaning that forward and backward reactions are balanced such that elemental abundances are given by relatively simple statistical relations. Current high-fidelity CCSN simulations assume that NSE is satisfied throughout the entire star. While this is a good approximation in the interior regions of interest most pertinent to the explosion, it breaks down quickly at large radii so that only the central regions can be accurately modelled. I will work to transition the equation of state (EOS) to the non-NSE regime in the FLASH code. This will allow for whole-star simulations and more accurate nucleosynthesis calculations. Inclusion of the outer regions of the star will also allow for modelling of the neutrino-driven wind – a supersonic outflow of stellar material powered by neutrinos emitted from the core of the star. This wind is proposed as a possible site of heavy element nucleosynthesis. We can then begin to meaningfully study the full nucleosynthetic signatures of CCSNe.

Getting 1D LC and EM information. The ultimate goal of the study of the CCSN explosion mechanism is the ability to make predictions and understand observations. Armed with a more realistic EOS and accurate nucleosynthesis, I will study the EM signals emitted during a CCSN. To achieve this, I will use a new model for driving 1D explosions that includes the crucial effects of turbulence and convection and map simulation data into SuperNu[2], a multi-D Monte Carlo radiation transport code, to produce the EM signals. This is imperative as, to date, we have only one observation of a CCSN that includes any signals other than electromagnetic. The FLASH code is already capable of handling the gravitational wave (GW) and neutrino emission, so this extension will provide complete predictions of the multimessenger signals. With this, we can begin to make direct connections between physical conditions of the explosion and what is observed. An understanding of how, for example, uncertain nuclear physics affects the electromagnetic signals is crucial to the success of the CCSN problem. This will allow us to compare our findings to observations and, for the first time, connect observations of CCNSe with details of the progenitor stars.

Going to multiple dimensions. The final goal of this project is the ability to run fully 3D CCSN simulations that for the first time include a proper treatment of the non-NSE EOS and EM information. This project will push the frontiers of current high fidelity 3D simulations and greatly enhance their explanatory and predictive powers. Due to the extreme computational resources required of these simulations, the simulations would begin during years 3-4 using computing allocations available to Dr. Sean Couch, the PI of the Michigan State University (MSU) research group. At this stage of the project, I will have the ability to study the full range of multimessenger signals from CCSNe including EM, GW, and neutrino signals in addition to the nucleosynthetic signatures of the explosion. With all of this in hand, we can make accurate and meaningful predictions of how the various physics that go into the CCSN impact the explosion, and how that in turn affects the observations.

**Research Proposal – Broader Impact**

The work presented here will result in the advancement of our understanding of the CCSN explosion mechanism, galactic evolution, and ultimately, the origin of the elements including those that comprise life. The results of this work will promote constructive collaboration between theoretical stellar astrophysicists and observational astronomers. EM data produced through this project can be compared against observational data as a benchmarking tool and may also be used by observers to inform future studies. This project aligns with the goals of the DOE’s Scientific Discovery through Advanced Computing (SciDAC) initiative, as well as the National Strategic Computing Initiative, in the push to exascale computing. MSU houses the Joint Institute for Nuclear Astrophysics, National Superconducting Cyclotron Laboratory, Facility for Rare Isotope Beams, and a new Department of Computational Mathematics, Science and Engineering and as such it the ideal campus for this interdisciplinary work. In an effort to reach first generation students, I will create a chapter of Ask A Scientist at MSU utilizing my connections with the national organization. To best leverage the available resources, I plan to create collaborations with existing programs at MSU such as the 4-H Michigan Extension, MSU Science Fair, and first generation student mentor program. I will travel to rural Michigan schools to show first generation and low income students that a college education and career in science are options for them. Through this chapter of Ask A Scientist, these communities can continue to benefit after the conclusion of my graduate studies. Astronomy has amazing potential to transform both lives and communities1 and the NSF GRFP would give me the resources necessary to begin my career while using my research as a tool for change.

Name: Kyle T. David

Application Link: <https://drive.google.com/file/d/1ObuHEI6GgO4_Qf7S_fJTDV6HbFXYaPM4/view>

Personal Statement Link: <https://drive.google.com/file/d/1CuzD7jntZrg5OEzYV3qF62dCF7_9Zv_H/view>

**Research Proposal – Introduction**

Arrow worms (chaetognaths) are a phylum of free-living predatory marine plankton. They are the second most abundant zooplankton group and represent a significant proportion of marine biomass1  . Despite their abundance and ecological significance, arrow worms are very poorly understood. Charles Darwin2 commented on the “remarkable... obscurity of their affinities” and in the 174 years since, arrow worms have been placed in many different bilaterian groups, including nematodes, annelids, molluscs, crustaceans, arachnids, and chordates1  . Most modern molecular analyses place arrow worms within protostomes (Fig. 1), but a consensus has not yet been reached3,4,5. Internal relationships within the phylum are similarly ambiguous1,6. I will broadly sequence arrow worm transcriptomes to determine relationships within and outside the group and use these transcriptomic data to elucidate the evolution of development within animals.

Aim 1: Infer a Robust Chaetognath Species Tree. The inclusion of arrow worm transcriptomes to a larger protostome dataset will add significant power to phylogenetic analyses and resolve evolutionary relationships that have confounded biologists for hundreds of years1 .

Aim 2: Explore the Origins of Bilaterian Development. Though considered protostomes by most modern researchers, arrow worms possess deuterostome-like development (enterocoely, secondary mouth formation, radial cleavage). This makes arrow worms uniquely positioned to explore novel questions on the origins of development in bilaterians. If arrow worms are indeed the sister-group to protostomes (Fig. 1A), it is likely that deuterostomous development was present in the bilaterian ancestor. Alternatively, if arrow worms are instead nested somewhere within protostomes (Fig. 1B-D), it is likely these features are an example of convergent evolution.

**Research Proposal - Methods**

The Halanych Lab has an established history of collecting and sequencing transcriptomes of non-model marine invertebrates. Our lab has sequenced and annotated 59 transcriptomes listed on NCBI’s Sequence Read Archive (SRA). I will collect specimens of at least one representative from each of the 11 arrow worm families recognized by the World Register of Marine Species. I will be able to accomplish this through a previously established relationship with Dr. Janet Voight, an Associate Curator at the Field Museum of Natural History, who has access to these groups as well as with samples previously collected from my lab. I also aim to participate in the Graduate Research Internship Program (GRIP) available to GRFP fellows, which would allow me the opportunity to intern at the Smithsonian under Dr. Jon Norenburg in order to study and sample their collections. It may be necessary to collect from the field as well, which will be possible through research cruises like the Icy Inverts Antarctica Cruises with which my lab has a history of participation. RNA samples will be extracted, prepared, and sequenced through previously validated Halanych lab protocols3  . The generalized bioinformatics pipeline is represented in Figure 2. I will use the skills I have learned from my recent participation in the Workshop on Molecular Evolution to infer maximum likelihood and Bayesian gene and species trees while using a variety of model assumptions and parameters in a comparative approach. Several deuterostomes (sea urchin, acorn worm, mouse, human) will serve as an outgroup.  Figure 2. Simplified bioinformatics workflow for species and gene tree inference

**Research Proposal – Intellectual Merit**

There is currently only a single arrow worm sequence on the SRA. This project will increase the amount of genetic data for this poorly understood group by an order of magnitude. A well resolved tree will also provide a phylogenetic framework for understanding the evolution of several key features in animal evolution and provide evidence for the ancestral bilaterian state. Arrow worms are known to have many unique features including lamellar photoreceptors7 and mosaic hox genes8 in addition to a putative whole genome duplication event9 . Increasing the availability of coding sequences in this group will allow myself and others to explore expansions/losses of several significant gene families (e.g., opsin and hox genes) and test for evidence of whole genome duplication within this enigmatic group.

**Research Proposal – Broader Impacts**

Results will be disseminated widely to expert (i.e., publications, symposia, talks) and non-expert (i.e., Skype a Scientist, outreach events, for details see Personal Statement) audiences. Through connections already established with faculty, I will also be able present my work as a guest lecturer through Auburn University’s Alabama Prison Arts + Education Project, which provides pre-college classes to prisoners. In 2016, the New York Times reported that inmates who participate in college programs have a 4% re-offence rate, creating a 500% return on investment in prison education initiatives. Alabama law does not allow prisoners to take remote classes meaning courses must be run on-site and in-person, something that would only be possible for me to participate in with GRFP support. All assembled transcriptomes and raw reads from this project will be made publically available on the SRA. I am committed to open-source software and will continue to upload all scripts required to reproduce analyses to my public repository (github.com/KyleTDavid). I will also mentor students through the NSF Research Experience for Undergraduates (REU) program, of which my lab is a participating member in computational biology. Students will receive a primer in basic programming skills and an introduction to phylogenomic workflows, as well as an opportunity to pursue independent projects.

Name: Nate Weinman

Application Link: <https://drive.google.com/file/d/1gjzHrwICVaKbSoAjNOEPkdXKwtdejx65/view>

Personal Statement Link: <https://drive.google.com/file/d/1MwGxOmYkasQTnJ0znHy4iQRGq61I2j1s/view>

**Research Proposal – Motivation**

Though more students than ever are graduating with Computer Science (CS) degrees [1], a major gap still exists between the skills of these graduates and the skills needed for industry success [2]. Two major components of this gap are programming skills (e.g. code style) and computational thinking (e.g. extending algorithms). My research will enable teachers to close these gaps by building a toolkit to integrate better-targeted problem types than those currently in use.

Programming assignments in CS classrooms, primarily code tracing and code writing, lack the granularity to target students’ Zones of Proximal Development (ZPD, i.e. challenging but solvable problems). Code tracing questions involve reading and understanding code, but they may not trigger students’ mental models of concepts because they are tedious or too easy [3, 4]. Code writing questions are a large leap from code tracing, conflating many programming skills into a single solution (e.g. design, computational thinking, programming, code style) and supporting a wide space of disparate approaches and solutions.

Extrapolating from earlier results [3], I propose the use of Parsons problems (e.g. [5]) to help students remain in their ZPD while learning computational thinking and programming through advanced CS concepts. Parsons problems involve unscrambling chunks of code, often individual lines, from a target program into a correct solution. Such problems provide similar learning gains to code writing problems – often 30% faster – for introductory assignments [3]. My research will enable teachers to integrate Parsons problems into existing curricula to better attain their teaching outcomes by providing a toolkit for these new problems.

**Research Proposal – Preliminary Work**

Within my Ph.D., I have been studying how Parsons problems can be applied to improve CS pedagogy. In Spring 2018, I ran an exploratory between-subjects study which found that when students begin by solving a Parsons problem instead of writing code, they are more able to generate multiple alternative solutions, addressing a skill gap in the “ability to generate alternate solutions” [2]. In Fall 2018, I ran a within-subjects study, followed by a structured interview, teaching algorithms to students who had not taken an algorithms class. They were taught two algorithms by either writing code from a pseudocode specification (i.e. a language-independent solution) or by solving a Parsons problem. Students with a range of skills noted that Parsons problems let them focus on the logic of the algorithm, engaging them in computational thinking. Students found writing code from pseudocode to be either too complex, distracting their focus, or too trivial, not engaging with the algorithm. These results suggest that Parsons problems can support a variety of students in practicing their computational thinking.

**Research Proposal – Approach**

In my thesis research, I will explore how we can leverage new problem types, e.g. Parsons problems, to supplement existing teaching tools for advanced CS concepts. I will run longitudinal studies on these new problem types by partnering with UC Berkeley professors who teach relevant classes with hundreds of students. I will then synthesize the results from these studies to develop a toolkit for teachers to help teachers easily adapt existing material into Parson problems and achieve their curricular desires.

RQ1: How can Parsons problems support the development of computational thinking with algorithms? Based on my Fall 2018 study, I am now exploring how to help students further focus on computational thinking with new Parsons problems that flexibly blend pseudocode with code within or between problems. By using pseudocode, students are constrained to rely less on syntax and more on computational thinking. I will measure students’ learning gains of the taught algorithms as well as their performance on interview-style algorithm questions.

RQ2: How can Parsons problems improve programming ability by teaching programming idioms? One major risk of supplementing existing assessments with Parsons problems is that it could hurt students’ programming skills by reducing their time spent practicing writing code. To overcome this pitfall, I will explore how Parsons problems can improve code writing by teaching programming idioms (i.e. common code and design patterns such as finding the five largest numbers in a list). There is a strong connection between students’ ability to write well-styled code and their ability to select and apply appropriate programming idioms [6]. However, complex idioms are often not explicitly taught. Results from my Spring 2018 study indicate that Parsons problems could support students exploring multiple solutions to a problem using different idioms, helping them compare when they are effective to use. They could also expose students to a range of problems where an idiom is applicable, helping students learn how it can be applied. I will run a formative study to better understand how students learn and select idioms. I will measure ABC scores of solutions, a well-established metric for code complexity, to evaluate students’ ability to efficiently apply idioms [6].

RQ3: How can instructors easily integrate Parsons problems into their teaching? Even if these new problem types are found an effective teaching tool, it must also be straightforward for instructors to generate and integrate them into the classroom. I will interview professors to explore the range of teaching methods used in classes and homework. These interviews will guide the design of a system to give teachers a powerful tool to target specific learning goals with more problem types. For example, a teacher could use think-pair-share in class to encourage students to share their problem-solving strategies by having students discuss which line of pseudocode should be placed next in a Parsons problem. Or, the large corpora of student solutions could automatically generate Parsons problems for teachers to modify and use.

**Research Proposal – Resources**

UC Berkeley provides rare access to collaborate with teaching professors in large, innovative classrooms. Here, my research will change how thousands of students learn CS.

**Research Proposal – Intellectual Merit**

My work will enable further research of teaching tools throughout the CS curriculum. My proposal explores how students learn and use computational thinking and programming idioms in complex problems through problem types and assessments. While there is a plethora of research on teaching introductory CS concepts, there is minimal research on how to teach advanced CS concepts, which are closer to real-world needs. The results of my work will empower teachers to create new resources to help students learn complex CS concepts.

This research will be the first to explore the effectiveness of Parsons problems beyond introductory courses, creating new interactions and contexts for Parsons problems. This will inspire applying Parsons problems in new ways: incorporating them into more domains in CS curricula, using them for post-school learning such as API tutorials or system documentation, or new situations where engaging with multiple solutions is beneficial.

**Research Proposal – Broader Impact**

My research is inspired by a desire to make CS concepts more accessible. Code writing problems are “one of the most significant reasons for giving up” by online learners in introductory classes [3]. These techniques will help improve learners’ self-confidence in these areas, with the aim of reducing impostor syndrome and attrition, as a step towards making CS programs more inclusive.

The results of this research will be disseminated within top-tier publications in HCI and CS Education. My software engineering experience enables me to make the toolkits developed over the course of my research robust and publicly available. Together, these will help researchers and content creators make knowledge more accessible to a diversity of audiences.