COLLEGE OF NATURAL SCIENCES
THE UNIVERSITY OF TEXAS AT AUSTIN

Department of Computer Sciences • 1 University Station C0500 • Austin, Texas 78712-0233 • (512) 471-7316

May 22, 2008

Dear Search Committee:

I write this letter in support of Paolo Bientinesi who is being considered for a Juniorprofessor (W1) in Algorithmically-Driven Code Generation for High-Performance Computing Architectures in the faculty of Mathematics, Computer Science and Natural Sciences of RWTH Aachen University. Paolo is my strongest student to date and undoubtedly among the strongest that our department has produced in the almost 20 years that I have been at UT-Austin. He has all the skills and attributes that will make him a great faculty member: he is brilliant, has employed unconventional innovation in his solutions, has demonstrated initiative on many fronts, has the drive and ambition to succeed, has proven leadership qualities, and communicates well. He has had all experiences that prepare one for an academic career: He has independently initiated new research, formulated innovative solutions to difficult problems, coauthored numerous papers, provided valuable input on grant proposals, served on departmental committees, mentored fellow students, reviewed journal and conference papers, and taught courses. I would describe him as an ideal "bridge candidate", namely a researcher who integrates many different subjects within computer science and mathematics while in return making contributions to those fields. He is already a recognized leader in the field of Algorithmically-Driven Code Generation and High-Performance Computing Architectures. And he is a nice guy as a bonus. Any top-10 department will be greatly enriched with Paolo as a faculty member. He has my unconditional support.

Paolo was already an experienced researcher when he arrived at UT. Originally, he was a visitor to Prof. Bajaj's visualization group. He applied to both the Computational and Applied Mathematics (CAM) and the Computer Science Ph.D. programs and was offered admission to both. He chose to enter the CS program because he felt that he already had solid credentials as a mathematician and wanted to expand his experience and understanding on the computer science side of computational science. Upon entering our program he changed his focus of interest from visualization to numerical analysis, high-performance computing, and formal derivation methods. Initially, he pursued a project that overlapped my own interests and those of Prof. Inderjit Dhillon, leading to a paper on parallelization of the so-called "Holy Grail" algorithm that was the topic of Inderjit's dissertation. That work has been published in the SIAM Journal on Scientific Computing (SISC), the top journal in the field, and has also been presented in talks and poster sessions at a number of workshops. This work was prominently mentioned in grant proposals by the (Sca)LAPACK group led by Jim Demmel and Jack Dongarra and an adaptation of the approach has been incorporated in ScaLAPACK. Google Scholar shows that the paper, published in 2005, has already been cited 24 times. This project established his credentials in high-performance library development, numerical analysis, and parallel computing. For many it would have evolved into a dissertation, but for Paolo it was merely a warmup exercise.

Paolo's dissertation research pursued a more fundamental contribution to computer science: How to systematically derive loop-based algorithms from the mathematical specification of the input and output. The systematic, constructive derivation of correct algorithms was, for example, identified in Dijkstra's 1972 Turing Award Lecture as the way programming should inherently be approached as a science. The problem has been that for loop-based algorithms and/or complex problem domains the constructive development of correct code has been elusive. What Paolo's research shows is that within the problem domain of dense linear algebra, this goal is very much within reach. Even five years ago this problem domain was considered to be much too complex for formal derivation to be feasible (after all, dense linear algebra libraries like LAPACK comprise hundreds of thousands and even millions of lines of very nasty Fortran code). Paolo's research has shown this problem to now be tangible. To me, this is a clear indication that he has made a true contribution to science: He has exposed the system behind the knowledge that an expert uses to painstakingly construct libraries. He has taken what was an art and transformed it into a science. He has demonstrated that what was thought to be difficult is actually relatively simple, when approached in the right way with the right notation. This work has led to numerous journal and conference papers, the most significant of which is the paper "The Science of Deriving Dense Linear Algebra Algorithms", published in 2005, which Google Scholar shows to have been cited 48 times already.

Paolo has taken the systematic derivation of algorithms one step further and has made it mechanical. Some Formal Methods researchers in our department have been mildly critical of people like Dijkstra for the following reason: only a verification of correctness that is performed mechanically can be truely trusted and it is thus mechanical verification via Automatic Theorem Provers that can guarantee correctness. Note that Paolo's approach, as well as the approach advocated by Dijkstra, is actually different: both advocate the constructive derivation of correct algorithms rather than an a postiori verification. However, the point about mechanization is clearly valid. It is Paolo's mechanical derivation tool, which is the centerpiece of his dissertation, that demonstrates and validates the systematic nature of the derivation process. The tool is astonishingly powerful: Very respected numerical analysts like Prof. Bo Kagstrom and Dr. Isak Jonsson painstakingly develop individual algorithms for libraries that target dense linear algebra operations in control theory. Paolo's tool can mechanically generate numerous (often dozens of) algorithms and implementations, many of which outperform the manually crafted algorithms and implementation. This part of Paolo's research also is not just a significant theoretical achievement: it has very practical implications. A journal paper on this subject is still in preparation.

Paolo has shown that dense linear algebra operations are particularly amenable to systematic derivation of algorithms (now that we finally understand the problem). However, numerical issues make the problem much harder than ordinary derivation of correct algorithms. Correctness in the presense of round-off error is a tricky business. As part of his dissertation Paolo made progress towards the systematic derivation of numerical stability analyses hand-in-hand with the derivation of the algorithms that are correct in the absense of round-off error. Prior to this stability analyses were performed by a few experts with particular skill and endurance. Paolo's dissertation provides initial evidence that his formal derivation techniques and notation can be extended to stability analyses. In my opinion, the eventual completion of this research will be such an accomplishment that several of us advised him against making it his dissertation topic. Our reasoning was that it would be hard to make a more significant contribution to science, and that a dissertation that is one's most significant contribution will always cast a shadow over subsequent work (which will always be inferior by comparison). Paolo compromised by only including preliminary results towards this goal in his dissertation. Much work needs to be done in this area, which will provide him with many interesting challenges for years to come. (Note that he will likely pursue this more as a hobby than as the primary topic of his future research.) A journal paper on this subject is almost ready for submission to the SIAM Journal on Matrix Analysis and Applications.

There is a second way in which the derivation of dense linear algebra operations presents a special challenge: the user community demands that algorithms and implementations be high performing on increasingly complex computer architectures with increasingly complex multi-level memory hierarchies and increasing numbers of processors and/or cores. The arrival of multi-core and future many-core architectures makes this issue even more urgent. One advantage that Paolo's research already brings to the table is that his mechanical tool generates families of (all?) loop-based algorithms and their implementations, so that the highest performing member of the famuly can be chosen for a given architecture. But how to choose? While some have given up on the notion of being able to understand the behavior of such complex systems (e.g., the ATLAS project and those who advocate cache-oblivious algorithms) within the scope of dense linear algebra libraries we believe it to be possible to create highly accurately models of performance. Very preliminary research of Paolo's, not yet published, related to this shows a lot of promise and hints at yet another research topic on which he will be able to make a major impact: performance modeling. We believe that his mechanical system can be extended to perform highly detailed performance analyses for this important problem domain. (Note that he will likely pursue this more as a hobby than as the primary topic of his future research.)

Numerical analysts and computational scientists are a demanding audience and making a scientific contribution to computer science is not enough: the research must make a dent in their bottom line, which invariably involves sparse linear algebra computations. I would argue that the value of computer science in this area is often (erroneously) judged by its impact on this bottom line and by the market share of the resulting tools. There are three ways in which I respond to this issue: First, Paolo's dissertation research represents a fundamental contribution to computer science and it should be appreciated as such. Second, the dense linear algebra libraries that result, with functionality that encompasses the Basic Linear Algebra Subprograms (BLAS) and LAPACK, are widely used for subproblems that arise within sparse solvers. Finally, for several years Paolo and I have been contemplating sparse direct solvers, which do impact the bottom line more directly. Recently, with the arrival of Dr. Victor Eijkhout at UT-Austin and in collaboration with researchers in the UT-Austin Institute for Computational Engineering and Sciences (ICES), we have put this project on the front burner. In a nutshell, we are exploring how to allow problem-specific information, especially regarding the hierarchy of successively refined domains, to be communicated to the sparse linear algebra library, which has the potential of greatly decreasing the amount of computation that must be performed. I refer to Paolo's Statement of Research for details. The work was recently funded by NSF and we hope that Paolo will continue to play a critical role in this project, regardless of where he finds employment.

In reading the above description of Paolo's work, a natural question becomes how often "Paolo" needs to be replaced by "we" or "the FLAME project". Let me address this directly. The research described above is part of the Formal Linear Algebra Methods Environment (FLAME) project of which I am the PI. While there were papers that we wrote on the derivation approach before Paolo started to contribute, key insights that he contributed allowed this project to be elevated to the next level of understanding and, ultimately, mechanization. I consider him an equal partner with myself in every aspect of management, innovation, and execution of this project. The large group of additional collaborators in the project play a vital, but less fundamental, role. Much of the research described above represents work that is almost solely owned by Paolo. Paolo made key contributions to the systematic approach to deriving algorithms: He is the one who suggested organizing the derivation around Hoare triples, stating assertions that then motivated the (constructively correct) statements of the algorithm. He is the one who observed how the operations can be mathematically stated so that loop-invariants (the state of variables before and after the execution of the loop) can be systematically identified, which solved the (thirty year old) key obstacle to constructively deriving loops. It was by making the derivation mechanical (which is entirely his project) that many other details of the approach were also identified. Although I was trained as a numerical analyst, stability analysis was never a strength or love of mine. The systematic derivation of stability was a project that he identified as critical to the approach and is entirely his. The performance analysis is something we identified as interesting, but of which he then immediately took ownership (collaborating closely with another student, Field Van Zee, and a research associate, Kazushige Goto). The sparse direct work mentioned above has been a close partnership. To summarize, Paolo has been a key contributor to all parts of the FLAME project, providing leadership, innovation, drive, and labor.

Since his graduation, Paolo has been working with Prof. Xiaobai Sun on algorithms related to digital signal processing. This research is more mathematical while also requiring a detailed understanding of the architectures of the future, like IBM's Cell processor and General Purpose GPUs. I defer details to Paolo's Statement of Research and to Prof. Sun's letter of recommendation. What this collaboration with Dr. Sun shows clearly is Paolo's ability to move between the worlds of mathematics, applications, and computer science, as an interdisciplinary researcher.

It is hard to directly compare Paolo to other junior researchers in the field. On the mathematics side, he is as strong as any of the Ph.D. students that our very highly acclaimed Computational and Applied Mathematics (CAM) program has produced. Students who graduate from that program have typically moved on to top departments or key positions at the government labs. I have been present when Paolo discusses interfaces between h-p adaptive finite element applications and libraries, and it is clear that he is intimately familiar with the mathematics and practical aspects of such applications. As a numerical analysis candidate, he is mathematically as strong as any recent graduate in that area. I would rank him much stronger than Richard Vuduc who is several years more senior, graduated from Berkeley, did a postdoc at LANL (?), and is now an Assistant Professor at Georgia Tech. He is also stronger than Lek-Heng Lim, who is now the Morrey Assistant Professor in the Mathematics department at Berkeley. Almost without exception, numerical analysis candidates are really mathematicians with some appreciation of computer science. Paolo is unique in his exceptional understanding of the computer science issues that affect the area. But he has additional proven research skills in computational geometry, formal methods, algebraic computation, and parallel computing. It would be wrong to pigeon-hole him as a numerical analysis or scientific computing candidate. The strength of Paolo's dissertation relative to others in our department who graduated last year is illustrated by the fact that it is Paolo's dissertation that was nominated by the department for the annual ACM Best Dissertation award.

I was specifically asked by Prof. Behr to comment on Paolo's ability to teach and lead a group. Paolo's interests are inherently intertwined with his interests in teaching. Many of the mechanical tools he developed with Mathematica were developed in part to be able to illustrate methodologies with animations. He was often asked to guest lecture for classes while he was a graduate student here. He is a good speaker (although not quite as animated as the prototypical Italian). After his Masters degree, he was drafted into the Italian army, where he was in charge of educating (if I recall correctly) about 300 cadets. This involved leading a team of teachers, themselves recruits, which helped him hone his leadership skills. While he was my graduate student, he acted more like a postdoctoral student who took charge of guiding and supporting his fellow graduate students. He has all the qualifications to become a highly effective faculty member.

Paolo has proven himself to be ready for a faculty position at a top department. He has demonstrated much more depth and breadth than other top students at a similar stage of their careers. He has proven ability to work in an interdisciplinary setting. He would be a particularly good choice since he has established connections in Europe (UT-Delft, Univ. Jaume I (Spain), Cerfacs, Manchester Universtity, Oxford University, to name a few). He will be an asset to any top department.

Best Regards,

Robert A. van de Geijn
Professor of Computer Sciences
Member, UT-Austin Institute for
Computational Engineering and
Sciences