CHAPTER II. THE CASE FOR SIMULATION

II.A. Rationale for the use of simulation

II.A.1. Number of potential participants

Perhaps the best rationale for the use of simulation may be provided by reviewing the interest with which the announcement of the Grand Challenge was greeted by potential participants ([8] and [3]):

DARPA received 106 applications for the 2004 GCE. Eighty-six teams submitted technical proposals by the deadline established by DARPA. Of the 86 technical proposals received, 45 teams proposed autonomous vehicles of interest to the DOD[3].

However, it did not appear as if all 45 teams would have vehicles ready in time to participate in the 2004 GCE. DARPA evaluated the technical proposals for 19 teams as "completely acceptable", and selected these teams for advancement to the next phase of the Grand Challenge. DARPA evaluated the technical proposals for an additional 26 teams as "possibly acceptable" and established a site visit process to determine the final teams[4].

On December 19, 2004, DARPA announced 25 teams from around the United States were selected to participate in the next phase of the Grand Challenge: Qualification, Inspection, and Demonstration (QID). The QID was used to determine the final 20 participants for the Grand Challenge. The 25 teams that passed the technical proposal review process were invited to the QID to take place March 8 through 12, 2004. Twenty-one teams participated. The QID comprised several distinct activities: a safety and technical inspection of the team challenge vehicle[5]; a separate practice area; and a demonstration course which was approximately 1.4-mile long that the vehicle was required to traverse.

The demonstration course allowed DARPA to evaluate the ability of each challenge vehicle to sense a series of static and moveable obstacles representative[6] of those that might be found on the actual 2004 GCE course, and navigate a course described by a series of adjacent waypoints. Each vehicle was ranked according to its overall time to complete the course, and point deductions[7] were taken for impacting obstacles, exceeding established speed limits, or deviating from the established course.

Over a five day period, eight teams completed the 2004 QID course, nine teams partially completed the course, two teams terminated within the starting chute area, and two teams officially withdrew. On March 12, 2004 DARPA announced 15 of 21 teams which participated in the 2004 QID qualified for the 2004 GCE.

In summary, only 15 of 106 applicants were allowed to participate in the 2004 GCE. No challenge vehicle which qualified was able to complete the 2004 GCE. To achieve this result, DARPA effectively eliminated 91 of 106 potential applicants, removing incentive those teams may have had to participate in the Grand Challenge and provide access to "new talent, new ideas, and innovative technologies" or develop autonomous ground vehicle technologies in the areas of "sensors, navigation, control algorithms, vehicle systems, and systems integration".

If the purpose of the Grand Challenge was that stated by DARPA (see Chapter I.), the Grand Challenge was either a marginal success or an abject failure, depending on perspective. From one perspective, the Grand Challenge was a marginal success because

DARPA was able to achieve some of its goals, although to a limited extent. The author considers it likely:

- the number of individuals, groups, or organizations working on autonomous ground vehicle technologies increased during the years before and immediately after the Grand Challenge,
- the Grand Challenge motivated individuals that would not normally work on a "DOD problem", and
- the Grand Challenge resulted in some development of autonomous ground vehicle technologies in the areas of sensors, navigation, control algorithms, vehicle systems, and systems integration.

The author is aware of no evidence which directly supports or refutes these assertions. For example, the author is aware of no survey of the robotics community before and after the Grand Challenge which supports an assertion that the number of individuals, groups, or organizations working on autonomous ground vehicle technologies increased during the years before and immediately after the Grand Challenge, and has since decreased. However, the author considers it unreasonable to assert, given the published record, that some progress has not been made, in particular in the development of autonomous ground vehicle technologies. The difference, to the author, lies in the intention and meaning of words such as "accelerate", or the duration of time during which DARPA expected the Grand Challenge to provide the DOD with access to "new talent, new ideas, and innovative technologies".

From another perspective the Grand Challenge was an abject failure. Although four teams successfully completed the 2005 GCE, the three teams with the best times, including the winner, Team 2005-16, were representatives of a single academic institution in all but name: Carnegie Mellon University. Team 2005-16 did not participate in the 2004 GCE. The Team 2005-16 team leader was a faculty member at Carnegie Mellon University when the Grand Challenge was officially announced on February 22, 2003 and transferred to Stanford University in July, 2003 approximately eight months before the 2004 GCE took place on March 13, 2004.

The Grand Challenge *might* have been a very close competition and a significant number of the teams participating in the Grand Challenge *might* have successfully completed the 2004 or 2005 GCE course, demonstrating proficiency in the skills required to develop an autonomous vehicle. As a result, the DOD *might* have gained increased or lasting access to "new talent, new ideas, and innovative technologies" or DARPA *might* have accelerated the development of autonomous ground vehicle technologies in the areas of "sensors, navigation, control algorithms, vehicle systems, and systems integration" to a greater extent.

By restricting the number of participants to the few teams with experience or sponsorship which were able to field a research platform, DARPA virtually guaranteed the eventual outcome of the 2004 and 2005 GCE: the only team which successfully completed the 2005 GCE and which was not closely tied to Carnegie Mellon University was Team 2005-06, which placed fourth during the 2005 GCE, and emerged as the only disruptive team which participated in either the 2004 or 2005 GCE.

II.A.2. Risk of rollover

Objective evidence supports a conclusion the 2005 GCE course was engineered and "groomed" to be less difficult than the 2004 GCE course to reduce the risk of rollover. Although some teams were aware of the risk of rollover, the author has not encountered an alternate detailed route analysis which indicates a team was aware of the extent to which the 2005 GCE course was engineered by DARPA. Although the author was unable to determine the total cost of team challenge vehicles, published records report costs from $35,000 to in excess of $3 million. As a result, the potential impact due to rollover was significant.

II.A.3. Stopping distance and field-of-view limitations

Review of team technical proposals supports a conclusion the teams had difficulty visualizing the interaction of their challenge vehicles with the environment, with potentially significant consequences, such as challenge vehicles traveling at speeds exceeding their stopping distance or an inability to adequately detect obstacles during a turn or on sloped terrain.

The use of simulation is proposed specifically to address these deficiencies. The use of simulation would:

- support autonomous vehicle development without requiring the sponsor of such research to engineer a course able to be completed by research platforms consistent with the state of the art,
- encourage participation in autonomous vehicle development by individuals and

institutions not having the resources required to develop a research platform,

- minimize the risk of rollover to research platforms until the necessary technologies were developed to enable the controlling intelligence to adequately evaluate the risk,
- allow teams to visualize the interaction of the research platform with the environment,
- provide teams with a way to identify some key factors which does not require procurement of a research platform or sensors to perform test and evaluation,
- increase focus on the development of basic algorithms and strategies,
- provide a way to increase competitiveness by “leveling the playing field”, and
- provide a tool which would help ensure long-term realization of DARPA's stated goals.

Overall, the use of simulation would allow teams to focus on the basic algorithms for using environment and geolocation sensors, and place the focus of autonomous vehicle development on artificial intelligence, not system integration, and “level the playing field” between teams with more experience and those with less experience.

II.B. Selection of the simulation environment

The most common approach to integrating hardware and software in use by teams which participated in the 2004 and 2005 GCE may be described as a “mixed” or “composite” architecture, where disparate, distributed elements were integrated using client-server relationships. These elements can be reproduced through the use of simulation. Therefore, one of the most important considerations for developing an architecture for simulation of an autonomous vehicle was the simulation environment itself. The author developed a list of requirements and desired features of the simulation environment, the first and most important of which was that it be free for academic use, with a preference for Free and Open Source Software (FOSS). Commercial software was not evaluated. Other requirements and desired features of the simulation environment included (in no particular order):

- Cross-platform availability.
- A graphical user interface using OpenGL.
- High-fidelity, rigid-body three-dimensional (3D) physics simulation, including collision detection and 6 degrees of freedom.
- Support for popular image formats and cameras.
- Terrain rendering.
- An active user community and developer base.

A review of available FOSS alternatives revealed the Player Project satisfied the author's requirements, with some caveats. In addition, the Player Project provided other desirable features, such as the ability to use XML files to configure the simulation, and could be extended by the author. As a result, the author selected the Player Project, specifically the applications Player and Gazebo, to complete this research.

II.C. DARPA evaluation of the use of simulation

The author is unaware of any published record that reports DARPA, following the

2004 or 2005 GCE, concluded that high-fidelity simulation was necessary, desirable, or even useful. From one perspective, this is certainly true. By engineering the 2005 GCE course, DARPA was able to create conditions which made it possible for several teams to successfully complete the 2005 GCE. In addition, a number of key factors contributed to team success (herein referred to as “key factors contributing to success” or “key factors”).

However, the fundamental problem of the Grand Challenge was system integration, not autonomous navigation or artificial intelligence, and the cost of fielding a research platform was prohibitive - out of reach for most individuals and even most academic institutions without corporate sponsorship. Although the author was unable to determine the total cost of team challenge vehicles, available evidence supports a conclusion that team challenge vehicles represented a considerable investment in terms of time and material resources.

Via the “Team Resources” section of the archived Grand Challenge 2004 website ([11]), DARPA hosted an “Outside Resources/Links” link to technical resources such as the Carnegie Mellon Navigation Toolkit (CarMeN) and many other libraries, applications, and utilities written to solve portions of the autonomous vehicle development problem. DARPA did not, however, include the the Player Project on the list of technical resources. None of the technical resources to which DARPA referred provided a simulation environment similar to the Player Project.

II.D. Team evaluation of the use of simulation

Via 2005 GCE Standard Question (SQ) 2.5.1[8] DARPA requested teams: “Describe the testing strategy to ensure vehicle readiness for DGC, including a discussion of component reliability, and any efforts made to simulate the DGC environment." Sixteen of 48 teams which participated in the 2004 QID or GCE or 2005 GCE referred to the use of simulation. Six of 48 specifically referred to the Player Project or to a simulation environment similar to the Player Project.

- Team 2005-02

Team 2005-02 stated: "To support bench testing, a simple vehicle simulator component was devised that sends out position- and velocity-related JAUS messages as if the vehicle were moving through an RDDF corridor." ([12], pp. 616 - 617).

- Team 2005-04

Team 2005-04 stated: "Portions of the software were tested on different simulation and emulation environments. Two specific simulation environments were developed for testing obstacle avoidance. One was a simple, flexible 2-D package for initial testing. The second was based on the Player/Gazebo environment and with the 3-D developments made, could actually include terrain configurations from real data." ([13], p. 6).

Team 2005-04 later referred to the use of simulation ([14]), but not specifically to the Player Project.

- Team 2005-05

Team 2005-05 stated: "[The challenge vehicle controlling intelligence] could be driven by real-time sensor data, by a simple simulator, or from previously recorded log data. The simulator was invaluable for debugging the high-level behaviors of the planner, but its models were not accurate enough to tune the low-level controllers. The replay mode allowed us to debug the ladar obstacle filters and the state estimators in a repeatable way, without having to drive the vehicle over and over." ([15], p. 531).

- Team 2005-09

Team 2005-09 referred to the use of simulation as part of their autonomous vehicle development process throughout their technical proposal ([16]), but did not refer to a specific simulation environment. Team 2005-09 did, however, refer to the use of simulation to "fit" the vehicle's performance in simulation to real-world performance: "The behavior of [the challenge vehicle] during the test would then (1) drive refinements to the simulator to more accurately reflect the demonstrations and (2) lead to new improvements in the software." ([16], p. 6).

Team 2005-09 later stated: "When a problem was found or a new phenomenon identified, it was first modeled in the simulation environment. With a simulation of the problem or new phenomenon in hand, the body of operational code was adjusted to deal with it. Once proven in simulation, the robot was field tested to evaluate the changes, and improvements were fed back to the model. A result of the model-build-test approach was that the model grew in fidelity and became a lasting repository of project experience." ([17], p. 835).

- Team 2005-11

Team 2005-11 stated: "[Challenge vehicle] testing included both physical and software-only simulation runs." and "Multiple simulation runs, particularly obstacle avoidance scenarios, were executed prior to field testing." ([18], p. 9). The author does not consider this reference to "simulation" to be a reference to a simulation environment similar to the Player Project.

- Teams 2005-13 and 2005-14

Teams 2005-13 and 2005-14 stated: "In addition to these system tests, [the challenge vehicle] has tested for software endurance via simulation..." and "Planned tests include end-to-end race day simulations..." ([19], p. 15, and [20], p. 15). The author does not consider this reference to "simulation" to be a reference to a simulation environment similar to the Player Project.

Teams 2005-13 and 2005-14 later referred to testing in simulation of control routines developed using Simulink ([21], p. 471).

- Team 2005-15

Team 2005-15 stated: "...we have simulation modules that allow for testing of all other modules, with the exception of the data acquisition modules." ([22], p. 6) and "In the lab environment, we use the GAZEBO toolkit to perform system and vehicle simulations." ([22], p. 11).

Team 2005-15 later stated: "With the use of the Gazebo simulator ... and tools for playing back recorded vehicle data, much of the debugging and development could be carried out on individual laptops; so development work could continue when the vehicle was not available." ([23], p. 582).

- Team 2005-17

Team 2005-17 stated: "A vehicle simulator is included in [the challenge vehicle] software suite. The simulator provides a test environment that emulates the physical environment in which the vehicle operates. Daily builds of the software are tested against a collection of test cases gathered from the real world. Developers perform unit level testing of changes to the software using the combination of the vehicle simulator and visualization tools included in the software suite." ([24], p. 10).

Team 2005-17 later stated ([25], p. 563):

> [The challenge vehicle's simulator] is a physics-based simulator developed using the Open Dynamics Engine physics engine. Along with simulating the vehicle dynamics and terrain, [the simulator] also simulates all the onboard sensors. It populates the same [queues] with data in the same format as the sensor drivers. It also reads vehicle control commands from [queues] and interprets them to have the desired effect on the simulated vehicle.
>
> While [the simulator] is a physics-based simulator, such as Stage ... and Gazebo ... it has two interesting differences. First, [the simulator] does not provide any visual/graphical interface. The visualization of the world and the vehicle state is provided by the Visualizer module, discussed later. Second, [the simulator] also generates a clock, albeit a simulated one, using the [queues].

Team 2005-17 later stated: "By maintaining a system-wide simulated time, [the

Team 2005-17 simulator] is able to create a higher fidelity simulation than that provided by Stage and Gazebo. The computation in the entire system can be stopped by stopping the clock; and its speed can be altered by slowing down or speeding up the clock. This also makes it feasible to run the application in a single-step mode, executing one cycle of all programs at a time, thereby significantly improving testing and debugging." ([25], p. 563).

Team 2005-17 also stated: "Yet, testing in the current generation of simulation environments, such as [the Team 2005-17 simulator], Stage[,] ... and Gazebo ... is quite limited. While these environments are good for doing integration testing, their simulation abilities are quite limited in providing information about how the vehicle may perform in the real world, such as, in different terrains and weather conditions." ([25], p. 577).

- Team 2005-18

Team 2005-18 stated: "Two simulation environments are also used: a dynamic model of the vehicle motion (including traction) that is used for testing without sensory input and a Gazebo simulation environment." ([26], p. 9). Via a footnote on the same page, Team 2005-18 stated: "The Gazebo simulation environment was used relatively lightly due to the team's decision to focus on desert testing."

- Team 2005-19

Team 2005-19 referred to the use of simulation as part of their autonomous vehicle development process throughout the team technical proposal ([27]), but did not refer to a specific simulation environment. Team 2005-19 later referred to "simulated or logged data", "numerical simulation", and "a simulated course" ([28]), but did not refer to a specific simulation environment.

- Team 2005-20

Team 2005-20 stated: “[Team 2005-20] attempted to implement an open source robotic simulation environment to assist in the evaluation of the code prior to running on the robot. This proved to be ineffective since the overhead of the open source package swamped the limited computational resources available for real-time operation. Therefore, the real-time code had to be redone outside the open source environment. The final solution was to develop a simulator utilizing the Team ENSCO developed real-time code. The simulator estimates where the vehicle position would be based on the commands sent instead of reading its position from a GPS device, but is otherwise identical to the software on the robot.” ([29], p. 15).

- Team 2005-21

Team 2005-21 stated: “Modeling and simulation of the [challenge vehicle] was done using ADAMS to determine vehicle performance over various size obstacles and to evaluate steering response at various vehicle speeds.” and “Rockwell also developed a simulation environment that included all of the vehicle dynamics. This simulation was used to test the vehicle control interface, real-time path planner and behavior control. Similar to on the vehicle, a series of waypoint could be executed while avoiding planned obstacles. The 2004 race path was executed several times in this simulation environment to determine if the vehicle could navigate the entire path.” ([30], p. 13).

Team 2005-21 later stated: “A full vehicle model of the truck was created in Advanced Dynamic Analysis of Mechanical Systems (ADAMS) by assembling subsystem models of suspensions, steering, chassis, and tires. A typical NATO Reference Mobility Model (NRMM) obstacle course with over 70 different obstacles of different sizes and shapes was used to evaluate the underbody clearance... The results of this simulation gave an idea about the truck's capability to maneuver through different obstacles at low speeds." ([31], p. 695).

Team 2005-21 participated in the 2004 GCE as Team 2004-23. Team 2004-23 was the only team which participated in the 2004 GCE to refer specifically to the use of a "simulation environment"[9]. Team 2004-23 stated: "A simulation model of the Challenge Vehicle has been developed and the software modules are being tested on the simulation environment." ([34], p. 11).

- Team 2005-22

Team 2005-22 stated: " A vehicle simulator program was also designed to test conditions and situations that would be difficult, if not impossible, for [the challenge vehicle] to encounter in Blacksburg. This program creates a virtual map and sensor data that is relayed to the actual pieces of software that control the vehicle. This simulator, along with information about [the challenge vehicle's] vehicle dynamics, tested the algorithms in a virtual space before ever placing them on the vehicle. It also allowed for testing during conditions where it would normally not be possible, such as at night or times when [*sic*]" ([35], pp. 12 - 13).

Teams 2005-22 and 2005-23 did not later refer to the use of simulation ([36]).

- Team 2005-23

Team 2005-23 stated: "A vehicle simulator program was also designed to test conditions and situations that would be difficult, if not impossible, for [the challenge vehicle] to encounter in Blacksburg. This program creates a virtual map and sensor data that is relayed to the actual pieces of software that control the vehicle. This simulator, along with information about [the challenge vehicle's] vehicle dynamics, tested the algorithms in a virtual space before ever placing them on the vehicle. It also allowed for testing during conditions where it would normally not be possible, such as at night or during heavy rain." ([37], p. 6).

Teams 2005-22 and 2005-23 did not later refer to the use of simulation ([36]).

II.E. Limits on the use of simulation

Although the approach discussed herein was implemented using Player and Gazebo, it is important to recognize limits imposed by the use of simulation. Several teams referred to specific limits on the use simulation:

- Models only approximate real world behaviors

Team 2005-05 stated: "The simulator was invaluable for debugging the high-level behaviors of the planner, but its models were not accurate enough to tune the low-level controllers." ([15], p. 531).

Team 2005-17 stated: "Yet, testing in the current generation of simulation environments, such as [the Team 2005-17 simulator], Stage[,] ... and Gazebo ... is quite limited. While these environments are good for doing integration testing, their simulation abilities are quite limited in providing information about how the vehicle may perform in the real world, such as, in different terrains and weather conditions." ([25], p. 577).

Team 2005-18 stated: "The Gazebo simulation environment was used relatively lightly due to the team's decision to focus on desert testing." ([26], p. 9). Although Team 2005-18 did not state their decision to focus on desert testing was driven by a limitation of Player and Gazebo, the team implied Gazebo did not represent desert terrain with sufficient fidelity for testing.

Based on the author's experience with Player and Gazebo, the extent to which models, including simulated worlds, terrain, and obstacles, approximate the real world or real world behaviors is more dependent on the accuracy of the model and availability of computing resources than on the simulation environment. The author notes teams participating in the 2004 and 2005 GCE may have had neither the time nor incentive to develop accurate models, but considers poor fidelity evidence of a resource allocation decision or a consequence of limited computing resources. The author does not consider sufficient evidence is available to conclude poor fidelity is due to an inherent limit on the use of simulation.

- The use of simulation is computationally intensive

Team 2005-20 stated: "[Team 2005-20] attempted to implement an open source robotic simulation environment to assist in the evaluation of the code prior to running on the robot. This proved to be ineffective since the overhead of the open source package swamped the limited computational resources available for real-time operation." ([29], p. 15).

The author concluded an increase in processing power available to the challenge vehicle controlling intelligence between the 2004 and 2005 GCE was a key factor. The author asserts an increase in processing power may have addressed the limitation identified by Team 2005-20.

- Real time versus “simulated time” simulation

Team 2005-17 stated: “By maintaining a system-wide simulated time, [the Team 2005-17 simulator] is able to create a higher fidelity simulation than that provided by Stage and Gazebo. The computation in the entire system can be stopped by stopping the clock; and its speed can be altered by slowing down or speeding up the clock. This also makes it feasible to run the application in a single-step mode, executing one cycle of all programs at a time, thereby significantly improving testing and debugging.” ([25], p. 563).

Although the author considers this a feature of the Team 2005-17 simulator, it does identify a limitation inherent in Player and Gazebo: the simulation can be paused or slowed by throttling the simulation time step, but not stopped without exiting the simulation environment, and neither Stage nor Gazebo can be run in single-step mode.

II.F. Advantages to the use of simulation

Several teams referred to specific advantages to the use simulation:

- Reproducibility

Team 2005-05 stated: “The replay mode allowed us to debug the ladar obstacle filters and the state estimators in a repeatable way, without having to drive the vehicle over and over.” ([15], p. 531).

- Software development is independent of hardware development

Team 2005-15 stated: “With the use of the Gazebo simulator ... and tools for playing back recorded vehicle data, much of the debugging and development could be carried out on individual laptops; so development work could continue when the vehicle was not available.” ([23], p. 582).

- The use of simulation increases the number of available test environments or conditions

Team 2005-22 stated: “ A vehicle simulator program was also designed to test conditions and situations that would be difficult, if not impossible, for [the challenge vehicle] to encounter in Blacksburg. ... It also allowed for testing during conditions where it would normally not be possible, such as at night or times when [*sic*]” ([35], pp. 12 - 13).

Team 2005-23 stated: “A vehicle simulator program was also designed to test conditions and situations that would be difficult, if not impossible, for [the challenge vehicle] to encounter in Blacksburg. ... It also allowed for testing during conditions where it would normally not be possible, such as at night or during heavy rain.” ([37], p. 6).