CHAPTER XIII. FAILURE ANALYSIS

XIII.A. Discussion

The most conclusive evidence that the fundamental problem of the Grand Challenge was not software engineering or artificial intelligence but system integration, is *failure analysis*. To the author's knowledge, no comprehensive failure analysis was performed. DARPA did not perform or publish a comprehensive failure analysis, and stated, when questioned by the author ([237]):

> DARPA did not publish any reports on Grand Challenge 2004 results. Our goal was, and remains, sparking interest and encouraging innovation in autonomous vehicle technology. It is up to the entrants in Grand Challenge to determine why no vehicle finished the course. We are confident that the $2 million prize for Grand Challenge 2005 will be adequate incentive for many teams to do just that. Grand Challenge 2004 teams are perhaps the best source of information regarding vehicle performance.

DARPA later amended their reply ([238]):

> Our October 22 response to your question failed to mention the final report from Grand Challenge 2004 ... which might be of interest to you. It is available in the "Highlights" section at the bottom of DARPA's homepage...

Although DARPA reported results following the 2004 GCE, insufficient technical detail was reported to determine the cause of failure ([30] and [3], pp. 8 - 9), or the information reported by DARPA did not agree with information reported by the teams. For example:

- Team 2004-01

Team 2004-01 passed on their turn on the first day of the 2004 QID ([78]), and terminated within the starting chute area on the last day of the 2004 QID ([79]). Team 2004-01 was not selected to participate in the 2004 GCE ([80]). DARPA stated only that Team 2004-01 "terminated within the starting chute area" ([79]). However, in private communication with the author the Team 2004-01 team leader attributed the cause of the problem to an unknown system integration failure caused by "severe lack of time" ([239]).

- Team 2004-24

Team 2004-24 was selected as a semifinalist to participate in the 2005 NQE. On September 30, 2005, DARPA stated that Team 2004-24 "did not pass their technical inspection" and "were not permitted to conduct a run" ([240]). However, Team 2004-24 did not report failing a technical inspection. At 2108 on September 30, 2005, Team 2004-24 stated: "[The challenge vehicle] had a good day today, breezing through the dynamic inspection after finding a final irritating bug that kept us from starting the dynamic run earlier in the day." ([241]). Team 2004-24 reported the team made four attempts to complete the 2005 NQE, including an "easier course", but was not selected to participate in the 2005 GCE ([242]).

- Team 2004-25

DARPA attributed the failure of Team 2004-25 to exit the starting chute as ([30]):

> `[Team 2004-25] - Vehicle brakes locked up in the start area. Vehicle was removed from the course.`

and ([3], p. 8):

> `[Team 2004-25] – The vehicle brakes locked up in the start area; the vehicle was removed from the course.`

However, in private communication with the author the Team 2004-25 team leader attributed the cause of the problem to "human error" ([243]).

As a result, the author is not confident sufficient technical detail was reported by DARPA to determine the cause of failures encountered by teams participating in the 2004 GCE, and concluded the published record did not report sufficient technical detail to determine the cause of failure for most teams which participated in the 2004 or 2005 GCE. Several teams reported the results of formal or informal failure analysis via the Journal of Field Robotics. The author considers these records generally reliable.

XIII.B. Analysis

The author reviewed the results of formal or informal failure analysis reported via the Journal of Field Robotics to determine the causes of failure before, during, and after the 2005 GCE. To evaluate the assertion that the fundamental problem of the Grand Challenge was not software engineering or artificial intelligence but system integration, the author divided the failures into categories depending on their *type:* "system integration", "controlling intelligence", or "other".

The author then determined if the failures were *preventable* through the use of effective simulation. The most common approach used by teams participating in the 2004 and 2005 GCE may be described as "mixed" or "composite", where a client-server

architecture was used to join disparate elements in a distributed architecture. Each of these elements can be reproduced through the use of high-fidelity simulation. Therefore, the use of *effective* simulation would have allowed a team to simulate failure of these elements by introducing noise or errors, and test and evaluate the ability of the controlling intelligence to adapt to, and respond to, failure.

The author classified all failures of the controlling intelligence as preventable because the controlling intelligence was, in general, entirely dependent on external input which could be simulated with high fidelity. However, not all system integration failures were preventable. Teams participating in the 2005 GCE were held to a high standard: system integration failures were considered preventable if they could have been identified through the use of effective simulation, even if this would have required the team to implement a method for performing adequate test and evaluation in simulation.

The DOD's interest in autonomous ground vehicle technology is to enable the controlling intelligence to replace a human driver. As a result, the author first established the expectation that the driving ability of the challenge vehicle's controlling intelligence should equal or exceed that of a human driver. Based on this equivalency, the author reasoned that system integration failures resulting in the same or substantially similar outcome for both the controlling intelligence and a human driver were not preventable, and would not have been identified through the use of effective simulation. Also, the author considers failures not preventable if the failure could not be identified without adequate test and evaluation of the challenge vehicle in the field.

For example, undiagnosed engine problems are not predictable, because unexpected mechanical malfunctions may occur with either a controlling intelligence or human driver driving a vehicle. Components wear and malfunction unpredictably. As a result, undiagnosed engine trouble was classified as unpreventable.

However, suddenly swerving into a wall because of “GPS drift” was predictable. Although reliance on dashboard GPS may cause a human driver with no advance knowledge of the road or terrain to conclude that a road exists where no road exists, no human driver would suddenly swerve into a wall because “GPS drift” caused it to be incorrectly reported as traversable terrain. As a result, this problem was classified as preventable.

XIII.B.1. Team 2005-02

Team 2005-02 selected a purpose-built platform for their challenge vehicle. See Table XVI. During testing prior to the 2005 GCE, the platform selected by Team 2005-02 failed. Team 2005-02 stated: “One of the rear shocks snapped and the engine and frame dropped onto the rear drive shaft and odometer gear. The sudden stop also caused the front sensor cage struts to snap and the sensor cage collapsed forward. The causes of the failures were determined and the system was redesigned and rebuilt in approximately one week.” ([50], p. 618).

Team 2005-02 failed to complete the 2005 GCE due to "a 20-foot position error [which] caused a corresponding shift of the boundary smart sensor that eliminated the actual sensed road as an option to the planner." ([50], p. 622).

The author considers the platform selection failure an unpreventable other failure, and the GPS drift failure a preventable system integration failure.

XIII.B.2. Team 2005-04

Team 2005-04 stated: "In the first set of robotic operations, [the challenge vehicle] tried five times back-and-forth operations before it overcame an obstacle. In the second set of robotic operations, [the challenge vehicle] was terminated. However, neither DARPA report nor [the challenge vehicle's] race log indicated that [the challenge vehicle] had had any collisions or gone off the road in the GC05 race. Also, [the challenge vehicle] was still totally driveable and stayed in the middle of the road when it was terminated. We guess it was because of the slowness in [the challenge vehicle's] gear shifting. Since it took up to 1 min for [the challenge vehicle] to shift its gear position in some situations, which might be intolerably slow, the second set of robotic operations might present the illusion that [the challenge vehicle] came to a halt and thus caused the termination." ([51], p 741). Team 2005-04 later stated: "We were not provided with official reasons of the termination." ([51], p 741).

The author was unable to determine if a failure occurred, or if Team 2005-04 was able to determine the cause of failure, if a failure occurred. As a result, the author did not include Team 2005-04 in the summary results presented later in this section.

XIII.B.3. Team 2005-05

Team 2005-05 stated: "[The challenge vehicle] crashed on three significant occasions: Twice during NQE trials and once during the GCE." ([170], p. 550). Team 2005-05 described the causes of the two failures that occurred during the 2005 NQE as follows:

- "...one of the vertical ladars had been repositioned and miscalibrated (due to a missing decimal point)" ([170], p. 550).
- A path planner "...failed to properly validate all the possible candidate trajectories and ended up selecting a degenerate trajectory containing two sharp 180º turns" ([170], p. 551). As a result, the challenge vehicle drove into a concrete barrier.

Team 2005-05 failed to complete the 2005 GCE due to "static memory over-allocation" and stated: "[The challenge vehicle] had made experimental autonomous runs of 10 miles or so, but had never made a continuous overland journey on the scale of the GCE. Furthermore, an endurance trial which consisted of driving for long periods around a track would probably not have uncovered this bug." ([170], p. 551).

The author considers the path planner failure a failure of the controlling intelligence, but the other two failures to be preventable system integration failures.

XIII.B.4. Team 2005-06

Team 2005-06 successfully completed the 2005 GCE. Team 2005-06 stated: "Several issues were discovered after analyzing the vehicle during a postrace inspection." ([28], p. 524). Team 2005-06 described the failures that occurred during the 2005 GCE as follows ([28], pp. 524 - 525):

- "The vehicle's steering was severely out of alignment."
- "The ABS was displaying intermittent failures that caused the brakes to behave in an erratic fashion."
- "The logging system crashed after 28 miles."
- "...an error in the path planning algorithms ... caused them to time out when faced with sections of the route with extremely wide lateral boundaries."

The author considers the steering and braking failures other failures, and the logging system failure a preventable system integration failure. When evaluating the causes of the steering and braking failures, Team 2005-06 stated they were assumed to be the result of rough terrain. The author concluded these failures were not preventable.

Team 2005-06 indirectly attributed their failure to place first or second during the 2005 GCE on extreme lateral boundary offset. Team 2005-06 stated: "...the director of DARPA said later that if we hadn't had a bug where we slowed down in the dry lakebeds, we would have either beaten [Team 2005-16] or been very, very close to [the Team 2005-16 challenge vehicle]. The bug meant we went from 30 miles an hour to two miles an hour on all the dry lakebeds. We'd never tested in an area 100 feet wide like that." ([31]).

The 2004 RDDF defines 12 segments with lateral boundary offset exceeding 50 ft. See paragraph II.C.7.d. As a result, the author concluded Team 2005-06 should have expected to encounter areas of extreme lateral boundary offset and considers the error in the path planning algorithms a preventable system integration error.

XIII.B.5. Team 2005-09

Team 2005-09 described the cause of multiple failures during the 2005 NQE in the "hand-off from planning and to reactive modes" as: "GPS loss" ([52], p. 832). The author considers this a preventable system integration failure.

Team 2005-09 reported test and evaluation using "a ball cap covered with tin foil over the GPS unit, effectively killing its signal" to induce "GPS loss" to diagnose this

problem ([52], p. 832). The author considers this supports a conclusion that this was a preventable system integration failure.

Team 2005-09 failed to complete the 2005 GCE, and stated the challenge vehicle detected occasional dust clouds as transient obstacles, which ultimately caused the challenge vehicle to veer off course where it was unable to continue because "the lasers could not differentiate between weeds and large rocks" ([52], p. 835). The author considers this a preventable system integration failure.

XIII.B.6. Team 2005-12

Team 2005-12 failed to complete the 2005 GCE due to "a bug in the obstacle tracking code, as obstacles were never entirely cleared from the list of tracked obstacles when passed. Tracking the position of thousands of irrelevant obstacles overwhelmed the processor, and starved critical code." ([183], p. 752). The author considers this a preventable system integration failure.

In addition, Team 2005-12 described several failures that occurred while attempting to evaluate the challenge vehicle's performance after the failure due to the bug in the obstacle tracking code was corrected ([183], p. 753):

- "a communications failure between the GPS unit and the guidance computer".
- The "vehicle blew out its left front tire at the base of the pass on the descent, following a collision with a small sharp rock" because the challenge vehicle's stereo camera pair "could not detect small but crucial features of this size". As a result, the "front wheels were also jarred out of alignment".
- Team 2005-12 also reported "three hardware failures would have ended a fully autonomous attempt of the course: A communications cable came loose, the steering position encoder became jammed with sand, and the vehicle's spare tire, installed to replace the old left front tire, was eventually destroyed by the terrain".

Team 2005-12 reported insufficient technical detail to evaluate the cause of the communications failure reported. The author considers the failure of the stereo camera pair to reliably detect an obstacle which disabled the challenge vehicle, communications cable failure, and steering position encoder failure preventable system integration failures. The author considers the failure of the challenge vehicle's spare tire an unpreventable other failure.

XIII.B.7. Teams 2005-13 and 2005-14

Although Teams 2005-13 and 2005-14 successfully completed the 2005 GCE, Team 2005-14 reported a failure due to an undiagnosed engine problem ([24], pp. 501 - 502) and a gimbal failure ([24], p. 502). The gimbal housed the challenge vehicle's Riegl LMS-Q140i.

The author considers the failure due to an undiagnosed engine problem an unpreventable other failure, and the gimbal failure a preventable system integration failure.

XIII.B.8. Team 2005-15

Team 2005-15 described the cause of a failure that occurred during the 2005 NQE as follows: “The GPS receiver incorrectly reported its measurements−valid to within 10 cm. Instead, the position measurement was off by 10 m to the north and east, and the velocities were reported as pure zeros causing the localization algorithm to crash. The addition of a few simple lines of code ignored these false messages...” ([133], p. 595). The author considers the GPS receiver failure a preventable system integration failure.

Team 2005-15 failed to complete the 2005 GCE and stated: “Although the vehicle was capable of operating at higher speeds, the obstacle detection system could not process the data reliably beyond 11 m/s. ... After an analysis of the recorded data, the cause of failure is fairly certain. Shortly before [the challenge vehicle] went off road, it lost all LIDAR data. Soon thereafter, it lost all vehicle state data. The LIDAR and the internal sensors were connected via USB hubs to the processing computers. Speculation is that one of the following faults occurred and ended [the challenge vehicle's] day: USB hubs lost power−terminating the connection between the computer and sensors, or the USB hubs overheated and ceased to function.” ([133], p. 595).

The author considers the USB hub failure an unpreventable system integration failure, even though the effect was indistinguishable from complete loss of sensors.

XIII.B.9. Team 2005-16

Team 2005-16 stated: “The primary measure of system capability was 'MDBCF'−mean distance between catastrophic failures. A catastrophic failure was defined as a condition under which a human driver had to intervene. Common failures involved software problems... occasional failures were caused by the hardware, e.g., the vehicle power system. In December 2004, the MDBCF was approximately 1 mile. It increased to 20 miles in July 2005. The last 418 miles before the National Qualification Event were free of failures; this included a single 200-mile run over a cyclic testing course. At that time, the system development was suspended, [the challenge vehicle's] lateral navigation accuracy was approximately 30 cm. The vehicle had logged more than 1,200 autonomous miles.” ([25], pp. 685 - 686).

Team 2005-16 successfully completed the 2005 GCE, placing first with a time of 06:53:58 hours. Team 2005-16 reported insufficient technical detail to evaluate the cause of the failures reported by the team. As a result, the author did not include Team 2005-16 in the summary results presented later in this section. However the author considers the aggressive test and evaluation reported by Team 2005-16 to have been a factor

contributing to their success and to support a conclusion that system integration was the real Grand Challenge.

XIII.B.10. Team 2005-17

Team 2005-17 described the causes of the failures that occurred during the 2005 NQE as follows ([196], pp. 574 - 575). The challenge vehicle:

- "climbed the hay bails [*sic*] when approaching the tunnel due to a rounding error in the path planner". Team 2005-17 had previously identified this as a potential cause of failure.

- "suddenly stopped after going through the tunnel due to a GPS/INS-related failure" caused by "loss of GPS signal in the tunnel".

- "did not compete in Run 3 because of a mechanical failure" caused by a "broken transmission".

- "ran into the last car on the final stretch in Run 6" which was not detected by the challenge vehicle's LIDAR sensors because "the second car was in the blind spot of the top lidar".

The author considers the failure caused by a broken transmission to be an unpreventable other failure. The author considers the failures caused by rounding error, loss of GPS signal, and a blind spot in the challenge vehicle's LIDAR sensors to be preventable system integration failures.

Team 2005-17 failed to complete the 2005 GCE due to a problem related to the failure caused by the broken transmission noted above, and stated: "After transmission failure, [Team 2005-17] did not calibrate the actuators [*sic*] correctly. Its '1' was mapped to a position that the transmission could not physically reach. When the vehicle was put into pause mode, to engage the brakes the levers had to be moved to '1.' However, this position could not be reached and the motor controller continued to attempt to move it. In the process, for 45 min, the motor was fed its peak current, a current it can withstand only for short duration. That caused the motors on the actuators to burn out." ([196], p. 576).

If the motor had burned out due to routine use and wear, the author would consider the failure due to actuator calibration an unpreventable other failure, but because the controlling intelligence could not distinguish between the actual position of the actuator and the target position of the actuator and consequently burned out the motor, the author considers this a preventable system integration failure, however difficult it would be to simulate in practice.

The author acknowledges it was not an expected failure mode of the vehicle, and that it is unreasonable to expect a team to willfully sabotage its own entry to determine

the impact of re-assembling the vehicle incorrectly after component failure. However, the author asserts an installation or repair procedure would have prevented this problem.

XIII.B.11. Team 2005-18

Team 2005-18 failed to complete the 2005 GCE and described the causes of several failures that occurred during the 2005 GCE as follows ([54], pp. 807 - 808):

- "...[the challenge vehicle's] ultimate demise was rooted in its incorrect state estimates (due to poor GPS signals)".
- "midrange LADAR sensor failures" caused by entering "an error mode from which they cannot recover".
- "the lack of a system-level response to such failures".
- "high speeds assigned to long-range sensor data, even in the face of state uncertainty."

The author considers the lack of system-level response to failures encountered by Team 2005-18 during the 2005 GCE and high speeds assigned to long-range sensor data to be failures of the controlling intelligence, but the incorrect state estimation due to "poor GPS signals" and midrange LIDAR sensor failures to be preventable system integration failures.

XIII.B.12. Team 2005-19

Team 2005-19 failed to complete the 2005 GCE and described the causes of two failures that occurred during the 2005 GCE as follows ([198], p 649):

- A "GPS receiver experienced a jump of approximately 2 m, and thereafter it reported an apparent error of more than 1 m" due to "re-acquisition of the OmniSTAR HP signal, which [the challenge vehicle] had lost approximately 175 s earlier".
- The "attitude estimator's pitch estimate to pitch derived from GPS velocity" caused the challenge vehicle to "localize LIDAR measurements incorrectly".

The author considers both failures preventable system integration failures.

XIII.B.13. Team 2005-21

Team 2005-21 completed the 2005 GCE course, but was not successful. Team 2005-21 did not report failures encountered by the team via the Journal of Field Robotics. As a result, the author did not include Team 2005-21 in the summary results presented later in this section.

XIII.B.14. Teams 2005-22 and 2005-23

Teams 2005-22 and 2005-23 described the cause of a failure that occurred during test and evaluation prior to the 2005 GCE: “A common experience for [Teams 2005-22 and 2005-23] GPS and inertial-based positioning systems is the 'GPS pop'... This occurs when, after running on inertial-only positioning, the GPS/ INS regains the GPS signal. The perceived position of the vehicle instantaneously jumps from the INS-computed location to the GPS-based position.” ([59], pp. 723 - 724).

Teams 2005-22 and 2005-23 failed to complete the 2005 GCE and stated: “Both vehicles failed due to mechanical problems, rather than poor navigation decisions. [The Team 2005-22 challenge vehicle] drive engine stalled when it briefly slowed to an idle, and [the Team 2005-23 challenge vehicle's] on-board generator shut down due to a suspected false low-oil reading.” ([59], p. 726).

The author considers the “GPS pop” failure a preventable system integration failure. Teams 2005-22 and 2005-23 reported insufficient technical detail to evaluate the cause of the mechanical failures reported by the teams. Although Teams 2005-22 and 2005-23 reported what occurred, the teams did not report sufficient technical detail to determine if the failures were preventable. As a result, the author did not include the mechanical failures reported by Teams 2005-22 and 2005-23 in the summary results presented later in this section.

XIII.C. Results

Tabulated results are presented in Table LXIX, and summarized below:

- Three of 32 failures reported (9 percent) were failures of the controlling intelligence,
- Twenty-three of 32 failures reported (72 percent) were system integration failures, and
- Six of 32 failures reported (19 percent) were other failures.
- Twenty-three of 32 failures reported (72 percent) were preventable.

XIII.D. Conclusions

The author concluded the majority of failures reported by teams before, during, and after the 2005 GCE were system integration failures which were preventable through the use of effective simulation. As a result, the author considers adequate test and evaluation a key factor.