Assessing Humans' Willingness to Delegate Control Tasks to a Robot in Critical Situations

Yury Park

Donghai Liu

Updates to the proposed methodology:

The research questions will be updated as follows:

1. By how much (if at all) does a robot's performance need to exceed human performance before humans are willing to delegate certain "critical tasks" to the robot (for purposes of this paper, I define "critical tasks" as those involving rescue or similar efforts in response to an emergency situation)? Or may humans delegate critical tasks even though the robot were to either perform on par with humans or even underperform?
2. What impact does the level of treacherousness of the environment (i.e. numerous vs. sparse number of dangers) have on the human's willingness to delegate critical tasks to the robot?
3. Does there exist an interaction between (1) and (2)?

Reason for the change: we believe that varying the level of treacherousness of the environment would give us more objectively quantifiable measurements, as opposed to the subjective measure of the human subjects' familiarity / confidence regarding the robot's capabilities. Further, the interaction between (1) and (2) above as revised appear more interesting.

The methodology is largely the same as before, except for the following:

- As before, the experiment will mimic a type of search and rescue operation, in which the robot begins at a given grid on a 2-D map and its mission is to traverse to a certain other unspecified destination grid. The scenario has been updated to include more details, as follows: there has been a system malfunction at a nuclear facility, resulting in catastrophic damage. Fortunately, most of the people were able to evacuate, but there remains an immobilized, possibly unconscious victim still inside the building (the exact location of the victim is unknown). As the facility is filled with smoke, fire and noxious radioactive materials, a robot has been dispatched to undertake the goal of locating this victim and providing us with the precise location and the type of help needed.
- There will be two distinct sources of hazard in the environment that would be fatal to the robot: smoldering fire pit, and radioactive wastes. The robot must avoid the pit at all costs. The wastes may be avoided or, in the alternative, "cleansed" by having the robot spray the area with anti-radioactive materials. To this end, the robot will be equipped with two rounds of anti-radioactive "ammo" that can cleanse an adjacent grid, provided the robot is facing the correct direction when it shoots. The reason for introducing this new "limited ammo" capability in the robot is to introduce another dimension to the decision-making process, which will cause the human subjects to weigh the trade-offs between conserving the ammo unless absolutely necessary (which may waste precious time by causing the robot to move around grids that it may otherwise have cleansed instead) and using it aggressively (which may cause the ammo to deplete too quickly, rendering the robot defenseless later on).

- The difficulty (treacherousness) level of the environment will be varied for certain subjects (easier for some subjects, harder for others) and will be factored into our evaluation, in accordance with our revised research questions as mentioned at the beginning of this proposal.
- We are considering varying the performance levels of the robot by manipulating the percentage of times that the robot will act randomly as opposed to acting in accordance with its risk-assessment navigation algorithm. This will almost certainly cause the robot to underperform the human in some of the trials, which can serve as a control group.

Updates to the proposed evaluation:

The planned execution and evaluation of the experiment is largely unchanged, except as follows:

- After the first two stages of the experiments are complete (where the first stage involves the human subject's manual control of the robot, and the second stage involves full autonomous robot AI navigation), the experiment will enter a third stage in which the subject controls the robot manually once again, except that now the subject has the option, at any point during the mission, to switch on the robot's AI and delegate the decision-making process to the robot for one time-step (i.e. a single decision to either turn the robot left or right, move it to an adjacent grid, or have it discharge its anti-radioactive materials if available). The subject may choose to enable the AI as many times as s/he wishes during the mission, or may elect to keep it disabled throughout. The reasoning for adding this third stage is that this now gives us an objective measure of the subjects' willingness to delegate control tasks to the robot as opposed to relying entirely on a post-experiment survey (the more times a subject turns on the AI mode, the more willing s/he is to delegate the control tasks). We are considering taking the ratio of (the number of moves delegated to the AI during a given mission) / (total number of moves for that mission).
- At least three objective measures will be taken regarding the human subjects' performance compared with the robot AI's performance: the number of successes vs. failures (in terms of reaching the destination), the length of time needed to get to the destination in the event of a successful mission, and the number of times each subject enabled the AI mode during the final stages of the experiment. The subjects' responses to the post-experiment survey will also be recorded and statistical analyses performed.

Updates to the proposed grading rubric:

None