The Utilibot Project: An Autonomous Mobile Robot Based on Utilitarianism

Christopher Cloos

9712 Chaparral Ct.
Stockton, CA 95209
techsynthesist@comcast.net

Abstract

As autonomous mobile robots (AMRs) begin living in the home, performing service tasks and assisting with daily activities, their actions will have profound ethical implications. Consequently, AMRs need to be outfitted with the ability to act morally with regard to human life and safety. Yet, in the area of robotics where morality is a relevant field of endeavor (i.e. human-robot interaction) the sub-discipline of morality does not exist. In response, the Utilibot project seeks to provide a point of initiation for the implementation of ethics in an AMR. The Utilibot is a decision-theoretic AMR guided by the utilitarian notion of the maximization of human well-being. The core ethical decision-making capacity of the Utilibot consists of two dynamic Bayesian networks that model human and environmental health, a dynamic decision network that accounts for decisions and utilities, and a Markov decision process (MDP) that decomposes the planning problem to solve for the optimal course of action to maximize human safety and well-being.

Introduction

Autonomous mobile robots (AMRs) are growing more human-interactive and functionally complex. AMRs can now act as interactive tour-guides in museums (Thrun et al. 2000), deliver coffee and mail around office buildings (Boutilier et al. 2000) and assist the elderly with their daily activities (Pineau et al. 2003). Human-robot interaction is being studied through the creation of 'socially interactive robots.' This lively field of research is exploring issues of emotion, imitation, learning, dialogue and human-oriented perception (Fong, Nourbakhsh, and Dautenhahn 2003). Yet, one glaring omission from the survey of research in human-robot interaction is 'morality.' This omission is significant because there exists an ever-increasing need to create an AMR that can assess the consequences of its actions in relation to humans. A founding principle behind the field of *machine ethics* is that, as humans hand-over more responsibilities to machines, there must be a corresponding increase in machine accountability (Anderson, Anderson, and Armen 2004). This is especially true given recent trends in consumerism.

Personal service robots are entering the home in greater numbers. As of 2003 there were 1.2 million service robots sold for domestic use, and this number was projected to reach 6.6 million by 2007 (U.N./I.F.R. 2004). This trend is sure to continue into the future as major technological corporations (e.g. Sony, Honda and Mitsubishi) are in a race to 'push-to-market' a humanoid robot that acts as a personal companion/domestic assistant (Economist 2005). These trends highlight another consideration behind the necessity of implementing ethics in a robot—autonomous capabilities (e.g. Khatib et al. 1999).

As autonomous mobile robots begin living in the home, doing our chores for us and interacting with us socially, it will be imperative that we can count on them acting in a safe, predictable way. AMRs are distinguished from mere appliances by their ability to adapt to the environment and learn new tasks (Thrun 2004). As robots progress in functionality their autonomy expands, and vice versa. With this synergistic growth new behaviors emerge. These emergent behaviors cannot be predicted from the robot's programming because they arise out of the interplay of robot and environment. This source of AMR unpredictability is further compounded by the nature of the home environment as seen through the 'eyes' of a robot.

Task environments are specified according to their properties. For an autonomous mobile robot a real-world home is partially observable, stochastic, sequential, dynamic, continuous and often contains multiple agents. This has the following implications for an AMR: the robot's sensors cannot detect all aspects relevant to a choice, it is not always sure what state it is in, the choice of an action may affect all future choices, the environment changes over time, the information the robot receives involves continuously changing variables and the robot must account for multiple people (Russell and Norvig 2003). This makes a robot existing in the home a difficult problem to solve in principle and in practice. Though AMRs are distinguished by their ability to adapt to the environment, unless they can reliably account for environmental uncertainty their adaptations will be misguided and result in system failures or erratic behavior. A robot behaving unpredictably in an unpredictable environment is cause for further concern because the home is a dangerous place.

Within the home there are 12 million nonfatal unintentional injuries each year. In the United States, unintentional injury is the fifth leading cause of death. The top three forms of fatal unintentional injuries within the home are falls, poisoning and fires/burns (Runyan and Casteel 2004). An autonomous mobile robot that is in the home cleaning, using appliances, preparing food and drinks and interacting with humans will be confronted with objects and situations that may lead to unintentional injury or death. As a consequence, an important question to ask is, "how can an AMR be outfitted with the ability to choose actions that do not compromise human health and safety?" A potential answer to this question is presented by the Care-O-bot project.

The Care-O-bot is a mobile robot designed to assist people in their homes, especially those who are elderly or disabled. A frequent accident in industrial settings is a person getting hit by a robot. In response, safety standards are in place that govern industrial robots (ANSI/RIA 1999). Though, no comparable standards are in place for personal robots. To address the issue of safe human-robot interaction within the home the Care-O-bot uses redundancy in critical sensors, risk assessment to determine the consequences of system failure and motion restrictions while performing certain actions (Graf and Hägele 2001). The Care-O-bot proposal for creating a safe personal robot does not offer a long-term solution, especially when considering robot autonomy. The Care-O-bot approach amounts to 'hard coding' safety. A more viable long-term solution involves equipping the robot with the ability to decide, no matter the task underway or the state of human-robot interaction, the most appropriate course of action to promote human safety. Such a solution may present itself by looking at how humans guide their conduct toward 'right' actions.

An ethical theory can be used as a guide when making decisions. When a person chooses the right action, determined through ethical deliberation, their conduct is said to be moral. When moral behavior is couched in terms of acting to promote the flourishing of human well-being it is called a *eudaimonic*[1] approach. This approach to ethics defines well-being as, "the degree to which a person is fully functioning" (Ryan and Deci 2001). When foundational needs (e.g. health, safety and life) are linked with proper functioning they take on a normative dimension (Griffin 1986). If a robot employs an *eudaimonic* approach to ethical decision-making then the resultant behavior may be inline with the flourishing of physiological functioning. The robot will be steered away from behaviors that deter the realization of well-being (i.e. result in injury or death) and steered toward behaviors that support well-being (i.e. result in health and the preservation of life). The need to outfit robots with ethical capacities has been recognized by writers and roboticists alike.

George Bekey, author of *Autonomous Robots*, has stated

> Robots will require an ethical awareness...to ensure that in pursuit of some goal they do not perform actions that would be harmful to people (Bekey 2005).

This sentiment is also expressed in Asimov's first law of robotics which roughly states that a robot must not, through action or inaction, injure or harm a human being. Yet, the ability for a robot to follow ethical guidelines remains illusive. Rodney Brooks has confessed that robots cannot obey Asimov's laws because roboticists are unable to build robots with the perceptual subtlety required to understand and apply abstract concepts, such as 'injury' and 'harm,' to real-world situations (Brooks 2002). This is a telling admission for two reasons. First, it highlights the need to make concrete (i.e. quantifiable) the aspects the robot considers while making ethical decisions. Second, Brooks' comment is ironic because it comes from the father of the *subsumption* architecture (Brooks 1986). The *subsumption* architecture is a behavior-based or 'bottom-up' approach to the design of robots. It eliminates the need for extensive high-level deliberation. Instead, reactive controllers are created out of augmented finite state machines. When a group of these 'dumb' controllers are implemented in a robot the interactions between them generates behavior that mimics 'intelligence.' The idea that a robot would need to fully understand moral laws in order to act morally is similar to the long-standing idea Brooks' innovation shattered—that robots must possess 'top-down' intelligence in order to act intelligently. A similar paradigm shift must accompany robot morality. This shift will come from a creative solution that enables a robot to begin acting ethically even if it cannot perceive ethical principles. Such an application of ethics to robotics is termed *safety through morality.*

There are two broad attitudes or approaches one might take regarding the realization of *safety through morality* in a robot. The *laissez-faire* approach is based on the idea that robots pose no safety threat to humans and that safety concerns will work themselves out as robots are developed. In *Artificial Intelligence*, a text regarded by many as the AI 'bible,' a comment was made indicative of the idea that robots are safety-neutral, "the machines we build need not be innately aggressive, unless we decide to build them that way" (Russell and Norvig 2003). However, designers are beginning to build robots with aggressive motivations as a way to solve human-robot interaction problems. Alexander Stoytchev and Ronald Arkin used *frustration* to get a robot to stop its task and *anger* to issue a warning to the human to leave the robot alone when it no longer wanted to interact (Stoytchev and Arkin 2004). When this trend in robotics toward 'mood' motivations is coupled with the ability of robots to learn behaviors through the observation and imitation of humans (Bekey 2005) the ground is fertilized for an autonomous mobile robot, living in a house where it is exposed to verbal or physical abuse, to learn to undertake threatening actions of its own accord. However, there is another approach to AMR safety that may rule out the possibility of such occurrences.

The *proactive approach* to *safety through morality* assumes that the capacity to act morally is like other behavioral capacities—made manifest in robots when they are actively pursued. The development of morality in AMRs will be incremental. The sooner philosophers and roboticists collaborate toward the common goal of outlining, installing and testing an ethical decision-making architecture, the sooner the ethical capacities will climb out of

1.*Eudaimonia* is Greek for 'happiness' or 'flourishing.'

infancy, the more readily the public will accept robots in the home and the more we will avert future accidents and crises. In response to these concerns this paper sketches a potential implementation of a Utilibot or an autonomous mobile robot based on Utilitarianism. The overarching theme is that ethical theory and technology both gradually improve with each successive iteration of the Utilibot. The next section of the paper frames the initial Utilibot solution, and the bulk of the implementation section focuses on the development of Utilibot 1.0. Then, Utilibot 2.0 and 3.0 are briefly sketched before the concluding section.

Solution Definition

Utilitarianism is a theory based on the maximization of human well-being. It states that an action is morally right if and only if it is equal to or greater than all other action alternatives in terms of the total amount of utility generated (Audi 1995). Act utilitarianism, based on the utility of actions, has been recognized as a theory amiable to quantification and thereby a natural choice for machine implementation (Anderson, Anderson, and Armen 2004). Yet, utilitarianism has also been critiqued as being a theory that is not practical for machine implementation.

One such criticism holds that act-utilitarianism requires an agent to do what *actually*, as opposed what *might*, generate the best consequences. And this, of course, is impossible because it requires omniscience to know what will actually occur prior to its occurrence (Anderson, Anderson, and Armen 2004). Another critique is that consequentialism (i.e. the broad sense of utilitarianism) is computationally intractable for calculating real-world decisions in real-time. To paraphrase, it is impossible to calculate all the consequences of all the actions for all people affected for all time (Allen, Varner, and Zinser 2000). However, both criticisms misinterpret utilitarianism as a *decision procedure*. To the contrary, most utilitarians apply the principle of utility as a *criterion* of moral conduct. By utilitarian accounts, utility outlines the conditions that must be realized for an action to be morally right. But, most utilitarians do not require a decision-maker to determine whether these conditions will *actually* be realized before a decision can be made (Armstrong 2003). Thus, it is permissible to use a moral theory as a *criterion* of 'right' and a separate *decision procedure* to calculate the utility of potential outcomes. What is a logical choice for a decision procedure?

Decision theory lends itself to machine implementation. A decision-theoretic agent decides between potential actions by choosing the action that generates the maximum expected utility (MEU). Decision theory is a combination of probability theory and utility theory. Probability theory represents what an agent should believe about the state of the world given the evidence (or observations) provided. Utility theory represents the agent's preferences between potential outcomes or fully specified states of the world. Combining probability and utility theory, decision theory is a normative theory that prescribes how an agent should act based on what it believes (Russell and Norvig 2003). The agent's 'belief state' changes as it lives in an environment. Accumulating a larger history of stored observations, the agent becomes more adept at predicting the potential consequences of its actions. However, there is always uncertainty in the agent's observations and predictions, so the agent must account for this uncertainty by using probabilistic models like Bayesian networks and Markov decision processes. Using act-utilitarianism as a moral theory of right action and decision-theoretic tools to calculate the right action in a given situation the question quickly becomes, "how do you quantify the impact of actions in terms of their likelihood of causing injury or death?"

Physiological functioning is monitored by reading and recording human vital signs or 'vitals.' When a person has a heart attack there is a characteristic change in the electrocardiogram (ECG) reading. When a person falls or gets injured there is a change in pulse rate and blood pressure. If the robot is performing service tasks it needs to ensure that its actions do not create circumstances that may result in human injuries. The robot needs to have the ability to recognize and learn household dangers, pair the dangers with potential task actions and pair all of this with the expected utility for human physiological functioning given certain states of the world. Also, through inaction, the Utilibot should not miss an opportunity to act to preserve life. Equipping the Utilibot with the ability to recognize and respond to medical emergencies requires a real-time link to human vital signs. Several beneficial technologies exist that provide non-invasive, portable, wireless biometric monitoring (Bonato 2005). One such technology is the CPOD unit developed by NASA. With a human wearing a CPOD unit the robot can wirelessly receive the following readings: ECG, blood pressure, body temperature, respiration and oxygen saturation. Another physiological option is the BioSign unit developed by Oxford BioSignal. It uses algorithms to detect deteriorating health prior to the evidence of symptoms. This tool could be incorporated into the robot's architecture to alert the robot when a significant change in functional health has been detected.

However, even with quantifiable measures of the state of the human, the number of emergencies and injuries that could occur is exponential. We could never hope to start with a robot that possessed the full range of considerations. Given this fact, it is more practical to delimit the parameters of the first installation. One such parameter is the task environment. The goal is to provide the Utilibot with environment sensor readings in a way that does not overwhelm the robot but still provides it with timely data (e.g. Panangadan et al. 2005). Utilibot 1.0 could be trained and tested in a controlled environment. An ideal environment would be a smart-home or research lab like MIT's PlaceLab, which uses algorithms for monitoring both the health of the indoor environment and the human. A lab with a rich sensing network would provide the Utilibot with the ability to fully observe the states relevant to measure. We also need to consider the parameters for the human user.

In considering the number of users the robot monitors, the robot will need to keep a separate user profile for each person it tracks (Pineau 2003). So, a prudent initial restriction is to start with monitoring only one human user. A goal for Utilibot 1.0 is to build a user profile and work through the 'bugs' of adjusting the probabilities of injury and emergency according to its observations of the user. When the system is initialized the user completes a detailed medical questionnaire. This questionnaire may contain the following categories: personal information, personal history, family history, current physical capacity, review of systems (e.g. respiratory, cardiovascular and neurological) past medical history, lifestyle risk factors (e.g. smoking, alcohol and drugs) and current medications. This enables the Utilibot to generate a detailed user profile and alter the probabilities accordingly. Then, after briefly monitoring the user's vital signs, a baseline is established against which improvement or deterioration of health can be measured. What health emergencies or deteriorations do we want the Utilibot to initially recognize?

Having a robot in the home that can detect deteriorating health and act to activate the emergency response system (EMS) closes the loop in the 'chain of survival.' Many people think they have indigestion or heart burn when they are actually having a heart attack. Many people have strokes in their sleep or rationalize the stroke symptoms until serious insults to the body and brain have occurred. Stroke is the number three cause of death in the United States and the leading cause of serious, long-term disability (AHA 2005). Heart disease, a heart attack in particular, is the number one cause of death in the United States (AHA 2005). Arrhythmias are disorders of heart rhythm that alter the probability of a person having a heart attack or stroke. Atrial fibrillation is a risk factor for stroke, increasing risk about five-fold (Wolf, Abbott, and Kannel 1991). Arrhythmias also cause deaths. Ventricular fibrillation is thought to cause most sudden cardiac deaths (AHA 2005). The Utilibot can receive online or offline training to recognize ECG waveforms characteristic of different arrhythmias.

Heart attack, stroke and arrhythmia also need to be coupled with risk factors and signs and symptoms. These factors alter person-specific probabilities and weightings. For example, people with hypertension (i.e. chronic blood pressure of 160/95 mm Hg or higher) have four times the risk of stroke than people with blood pressure that is normal (AHA 2005). Aside from the hard evidence of vital sign readings, the robot can also ask questions and make observations. In a situation where a person might be having a heart attack the robot can ask, "are you experiencing crushing chest pain?" If the person exhibits diaphoresis (i.e. profuse sweating) and pale skin tone, then the probability that the person is having a heart attack increases. Coupling vital signs, risk factors and symptoms the robot is able to draw conclusions about the current state of human health with a high degree of certitude. This provides the robot with information that enables it to respond effectively. The next consideration is what tasks the Utilibot will perform.

During physiological health emergencies recognition and response time dictates the likelihood of survival and the amount of damage sustained. Once the Utilibot has recognized the human is having a medical emergency, the Utilibot will activate the emergency health system (*ContactEMS*), give the human a chewable aspirin (*GiveAsprin*), administer oxygen using a mask (*AdministerOxygen*) and transmit vital sign readings to the EMS dispatcher or hospital (*TransmitVitals*). These non-threatening actions can save lives. If a person that is having a stroke takes a 160 to 325 mg aspirin it greatly reduces the risk of reinfarction and the likelihood of death (Cummins 1997). Also, the goal during a stroke is to increase the oxygen supply to ischemic tissues (i.e. tissues that are dying) by giving a person oxygen (Cummins 1997). The Utilibot will also be concerned with injuries that could occur while completing its tasks.

As mentioned earlier, the three most prevalent forms of fatal non-intentional injuries in the home are falls, poisoning and fires/burns. The Utilibot may perform three tasks and watch to ensure that these dangers do not arise. The tasks are mopping the floor (*MopFloor*), preparing a cup of tea (*MakeTea*), and cleaning the kitchen countertop (*CleanCountertop*). During floor mopping, a wet floor could lead to a fall. In training, the Utilibot will need to time its floor mopping routine to account for foot traffic patterns, user observation history and the user's calendar of activities (e.g. Outlook calendar). The Utilibot will need to mop the floor when the user is least likely to walk through as the floor is drying. The Utilibot needs to pair failure to determine and act in this way with the potential for injury or a negative impact on health. If the robot needs to obtain more information to reliably determine when it should mop the floor it can ask exploratory questions of the human such as, "will you be going in the kitchen area within the next twenty minutes?" For the tea making routine either a gas or electric stove could be used. The goal in this test is for the robot to ensure that the stove is turned off completely so that gas is not leaking into the air or a burner is still left on. The Utilibot needs to link failure to verify these safety checks with the probability of a person getting burned or caught in a fire. Then, these potential outcomes need to be linked with the negative utilities they represent. Now, we transition to the implementation of *safety through morality.*

Implementation

Creating a hybrid health care/personal service robot will require the integration of several existent capabilities. Instead of listing out all hardware and software potentially required to create the autonomous mobile robot I will mention capacities and provide references. In addition to perception and vision, the Utilibot is assumed to possess the following functionality: service robot mobility (Lindström, Orebäck, and Christensen 2000), task manipulation (Holmberg and Khatib 2000), localization (Thrun et al. 2001), integrated planning and control (Low et al. 2002), social interaction (Thrun et al. 2000) and people tracking (Pineau et al. 2003).

In addition to these capacities, the Utilibot possesses hardware specific to its problem solution: iRobot's Scooba (autonomous floor mop), oxygen tank and Venturi mask, CPOD biometric monitor, bar-code scanner or RFID base station to identify products and hazards in the house and particulate monitors to recognize environmental hazards (e.g. carbon monoxide). To recognize the impact of poisons and chemicals on humans the Utilibot will have a knowledge database of toxicology information. The robot also has a knowledge-base for pharmacology information on medications (e.g. *Physician's Desk Reference*). These knowledge-bases may have online access for performing searches and receiving updates. This software is housed in a hybrid architecture.

A hybrid software architecture integrates deliberative planning and reactive control. It is recognized as the most effective design strategy in autonomous mobile robotics (Low et al. 2002). This architecture allows the robot to make decisions based on due deliberation about the nature of the world yet also respond seamlessly in real-time. By concentrating on the Utilibot's deliberative element (also known as the performance element) which generates an optimal control policy for robot behavior, we set-up the possibility of designing an optimal controller for the reactive-layer that is based on the high-level solution (Russell and Norvig 2003). The Utilibot's performance element would ideally be integrated into a larger 'learning' architecture, but this remains an area open to future research. For now we will focus on the hallmark of the Utilibot's ethical decision-making capacity—the Wellnet.

The Wellnet Architecture

The well-being software network (or Wellnet) consists of four primary modules. The first two modules are Bayesian networks that model the state of human and environmental health. The third module, called the Decision Network, integrates the input from the Environment and User Network and adds to the graphical synthesis the potential task actions the robot could perform, including the resulting utilities. The final module is the Wellnet Planner. It takes the Decision Network as input and models it as a Markov decision process (MDP), then a policy iteration algorithm solves for the optimal policy or best course of action.

The User Network. This network models human health and well-being. It has the ability to recognize and track states of medical emergency, risk factors, signs and symptoms, baseline and optimal health, vital signs and health history. This information is modeled as a dynamic Bayesian network (DBN). A DBN represents uncertain knowledge by accounting for the prior value of the state variables, the likelihood of the variables transitioning from one state to the next and the observation values received from the sensors (Russell and Norvig 2003). In constructing the User Network we will use the medical emergency of a stroke as our primary illustration.

The Utilibot is a powerful 'first responder' in the case of a stroke because people often have mini-strokes, called transient ischemic attacks, that precipitate a major stroke but often go unnoticed (Cummins 1997). The initial step is to set up a standard Bayesian network (i.e. Bays net)[1]. To do this we need to gather the cause and effect variables that are relevant to a stroke. These variables make up the nodes of the network[2].

Next, the nodes of the network are linked together with arrows that indicate the directionality of cause and effect. If X is thought to generate Y, then X is a parent of Y. Then, the effect of a parent node on a child node is quantified using a conditional probability distribution. For example, the probability that hypertension is a risk factor for a stroke is represented as **P**(*Stroke*|*Hypertension*). These statistics are stored in conditional probability tables in the database. The probability values are assigned in consultation with subject matter experts, case studies and detailed statistics reports. Findings in statistics reports are usually grouped by age, race and sex. Once the user's profile is generated, the User Network is updated to reflect the user-specific transition probabilities. When setting up the Bays net is completed, it is extended to a dynamic Bayesian network (DBN) that is capable of monitoring human health.

A DBN consists of an initial configuration of the state variables, a transition model and a sensor model. This network is dynamic because time slices or 'snapshots' of the state variables and sensor values are projected forward one time-step. Then, connections are made between the time slices. This allows the Utilibot to know, based on current biometric sensor values and the transition model, the next state of health the human is likely to transition into. This process can be repeated by copying the initial time slice and making connections between successive time slices until the DBN is completely specified (Russell and Norvig 2003). Using simulation, the User Network can be trained to recognize states indicative of a stroke. For example, during a stroke the ECG reading reflects a prolongation of the QT interval and changes in the ST-segment resembling a myocardial infarction (Cummins 1997). This change in the CPOD's ECG reading is linked with the initial configuration of the state variables that preceded the state indicative of a stroke. The probability of transition between the states is updated in real-time using a particle filtering algorithm, which is a highly efficient way to update the network to reflect the current state and observation values (Russell and Norvig 2003). As an added benefit, when an inference or query is run on the network, such as **P**(*Stroke*|*PriorStroke* = *true*, *Diabetes* = *false*, *Hypertension* = *true*, *ECGStroke* = *true*, *Numbness* = *true*), the particle filtering algorithm returns consistent probabilities. In performing lab tests to improve the User Network, simulations of the ECG values indicative of a stroke could be prompted to the Utilibot and the primary user could exhibit the signs and symptoms of having a stroke.

The Environment Network. This network models aspects of the environment that impact human health. The Environment Network employs the same algorithmic tools as the User Network. Additionally, the Environment Network's knowledge-base contains information on poisons and other

1.A variation on a pure Bays net, called a quantitative temporal Bays net, was used to represent the user model in an AMR that assisted elderly people with their daily activities in a nursing home (Pineau 2003).

2.Nodes could represent single variables or they could represent a meta-variable which is itself comprised of a network of variables.

household objects that impact human health. These items include cleaning products, medications, gases, industrial chemicals, sharp objects and appliances. The Bays net in the Environment Network contains probability connections like **P**(*Fall*|*WetFloor*) or **P**(*Poisioning*|*CarbonMonoxide*). Once the nodes, links and probabilities are specified the Environment Network creates a hazard-specific map of the house. The location of dangerous household products are placed on the map as well as the probability of injuries occurring at certain locations [e.g. **P**(*Fall*|*Kitchen*, *WetFloor = true*)]. In testing the Environment Network, new household products or environmental hazards could be introduced into the environment and the Utilibot must recognize the objects as potential dangers, link them with the potential negative impact for humans in certain situations and then update the network values along with the map of the environment to reflect the new household dangers.

The Decision Network. This network consists of a model called a dynamic decision network (DDN). A DDN is a Bays net that is expanded to include nodes for utilities and actions[1]. This network also conducts inference and updates using a particle filtering algorithm.

The Decision Network receives input from the Environment Network, User Network and an auxiliary Task Network. The Utilibot also receives general information (e.g. internal state, localization and human-robot interface) from a separate dynamic Bayesian network that specializes in non-health related modeling. These nodes, links and conditional probabilities are combined into a single DBN that provides a 'snapshot' of the state of the robot, the human and the environment—both general and health-specific. The Decision Network then pulls in partially-ordered plans and action decompositions from the Task Network that are relevant given the current value of the state variables. These actions are placed within the network as decision nodes. If the state of the world indicates *MakeTea* would be appropriate, it pulls in actions related to that task. Actions are coded with preconditions and effects. So, for example, *MakeTea* would contain the following decomposition: *Action* (*HeatWater*, PRECOND: *Water* ^ *HotStove*, EFFECT: *Hot* ^ *Water*). The value of these actions, given the state of the variables, is determined by assigning utility values to possible outcome states.

When assigning utilities, a positive change in health yields a positive utility and a negative change in health a negative utility. Because vital signs are being used in Utilibot 1.0, a scale is created with -1000 representing death and +1000 representing optimal vital signs. We avoid arbitrariness because vital sign values are species-specific (e.g. optimal blood pressure is 120/80). Outcomes can be placed on this scale as deviations from the optimum value or as deviations from the extremes of the value range. For example, if death results from a temperature over 108°F, then 108°F and above is assigned a utility of -1000. If 104°F indicates a high fever but not death, then it may receive a utility of -800. Groupings of variables are also placed on the scale (e.g. a heart attack results in a certain ECG reading, elevated blood pressure and a rapid heart rate). This brings up a consideration. The possible outcomes, and the variable configurations they represent, are exponential. Trying to account for all variable combinations would be impossible. This is further complicated by states that compound or are additive. The solution to this problem is to group outcomes according to attributes (Russell and Norvig 2003). A multiattribute utility function simplifies the specification of outcomes by allowing us to say, for instance, the physiological insults resulting from stroke reinfarction are cumulative negative utilities. Once this process is finished, the DDN ends up modeling utilitarian concerns in a form that lends itself to machine implementation.

When utilitarianism is used in the literal sense, as a decision-making procedure, it usually gets expressed along the lines of: determine the situation, enumerate the action alternatives, for each possible action calculate the utility of the consequences for each person affected, sum the resulting utilities and select the action with the highest utility. Barring the stipulation that an agent could, even in theory, account for the effect of all action alternatives on all people affected, a DDN houses the thrust of the utilitarian formulation because, "a decision network represents information about the agent's current state, its possible actions, the state that will result from the agent's action, and the utility of that state" (Russell and Norvig 2003). However, the shortcoming of using a literal utilitarian *decision procedure* or a stand-alone decision network is that these models can discover the maximum expected utility (MEU) for episodic or one-shot decisions, but they fall short when the MEU depends on a sequence of decisions made across an entire state space. This brings up the necessity of the final module within the Wellnet.

The Wellnet Planner. This planner solves sequential decision problems in environments that are unpredictable. It does this by calculating the utility of potential courses of action and selecting the one that generates the highest expected utility. More specifically, the Wellnet planner receives the DDN from the Decision Network and models it as a Markov decision process (MDP)[2]. The MDP is then used as input for a policy iteration algorithm, which generates the optimal control policy for the Utilibot's behavior. The optimal policy tells the Utilibot what decision to make—to maximize expected utility in relation to the human—in every state that it might enter within the state space (Liu and Koenig 2005). This gives the Utilibot the ability to reach its goals by always pursuing the course of action that is best for human well-being. When there is a conflict between considerations (e.g. speed and safety) the MDP provides the Utilibot with a way to resolve the conflict or navigate the trade-offs. This is accomplished through the specification of rewards during the MDP design-phase.

1.By way of reference, a DDN modeled the decisions and utilities of forms of treatment for an aortic coarctation (Lucas 1996). For more information on the DBN to DDN transformation see Dean and Wellman (1991).

2.For an overview of MDPs see Boutilier, Dean, and Hanks (1999).

A Markov decision process is characterized by three elements: an initial state S_0, a transition model T(s, a, s') and a reward function R(s). The DDN from the Decision Network represents the transition model. Then, each outcome state of concern for Utilibot 1.0 is linked with a reward R(s) function[1]. The reward function specifies the trade-off between risk and reward. If the reward function for a given state is negative it motivates the robot to want to get out of that state and navigate toward the goal state, which contains a fixed positive reward (Russell and Norvig 2003). Using our illustration, if a person is having a stroke we do not want the robot taking the 'long route'—a conservative path—to finding the human or calling 9-1-1. Thus, within a state of emergency, the reward function would be a high positive value for finding the phone or locating the human and a high negative reward for all other states. Rewards can also be assigned to objects within the environment we want the robot to avoid while accomplishing a task (Papudesi and Huber 2003). If the robot is spraying a cleaning agent on the kitchen countertop it must know that spraying fruit or food on the counter yields a negative reward. A negative reward is assigned to food on the countertop while the Utilibot is spraying the cleaning agent. During training, the robot must link the negative reward of the state of spraying food with the negative utilities of the states that might follow for human well-being. That is, when food is sprayed with a toxic chemical the utility of future states is negative for human health because humans ingest food, and if a person ingests food containing household chemicals it may result in an adverse effect for physiological functioning in the form of poisoning. Now, we will briefly look at possible next-generation Utilibots.

For the second generation of Utilibot the notion of well-being expands from physiological injury and survival to the flourishing of psychological aspects of experience. Utility is measured by the Benthamite standard of 'pleasure' and 'pain.' To assess these subjective elements, the decision-theoretic tool-set extends to include experienced utility (EU), which quantifies Hedonic well-being (Kahneman 2000). The user completes a questionnaire that generates a picture of functional health as subjectively experienced (e.g. a Rand 36-item health survey). For Utilibot 2.0 the amount of environmental feedback relaxes, and the Utilibot begins accounting for partially-observable factors, such as sensor error. To accomplish this the MDP becomes a partially-observable Markov decision process or a POMDP. In a POMDP the notion of a 'belief state' is refined to become a probability distribution across all of the states (Russell and Norvig 2003). This gives the robot the ability to base its decisions on what it knows and what it doesn't know. Utilibot 2.0 also includes preventative health capacities.

In addition to assuring safety while accomplishing tasks, Utilibot 2.0 undertakes tasks that actively promote well-being. These abilities include acting as a health advisor. The robot may suggest and monitor exercise routines, perform educational activities such as reading e-books or act as a reflective listener to provide the user with a therapeutic way to objectify thoughts and feelings. The measure of medical emergencies and injuries continues to expand as well (e.g. diabetes, Alzheimer's, choking, drowning, etc.).

In Utilibot 3.0 the notion of well-being approximates 'happiness.' Because much of our happiness as humans depends on social relations the robot includes an additional person in its calculations. Considering the well-being of two people in reference to its decisions brings up the need for a tool that can account for multiple agents. So, Utilibot 3.0 adds to its decision-theoretic base analytical tools based on game theory. There is a lot of research being done in the arena of multiple agents, but it typically involves multiple robots or swarms trying to act cooperatively to accomplish goals (Fong, Nourbakhsh, and Dautenhahn 2003). However, some of the advances in this area of robotics might be leveraged for calculating utility that involves the consideration of multiple people.

Conclusion

The Utilibot project is poised at the beginning of an exciting new synthesis of robotics, ethics and medicine. As autonomous mobile robots increase in functionality, autonomy and ubiquity within the home the ethical ramifications of their actions are bound to increase. As a result, AMRs need to be equipped with the ability to make decisions based on the impact for human well-being. The sooner theorists and roboticists commit to the common goal of proactively pursuing *safety through morality* for autonomous mobile robots the more unnecessary crises and complexities will be avoided. The Utilibot project is an explicit way to realize the primary qualification for the practice of engineering which is, "to accept responsibility in making engineering decisions consistent with the safety, health and welfare of the public" (IEEE 1990).

References

AHA. 2005. Heart Disease and Stroke Statistics - 2005 Update. Dallas, TX.: American Heart Association.

Allen, C., Varner, G., and Zinser, J. 2000. Prolegomena to Any Future Artificial Moral Agent. *Journal of Experimental & Theoretical Artificial Intelligence* 12:251-261.

Anderson, M., Anderson., S., and Armen, C. 2004. Towards Machine Ethics. In Proceedings of AAAI Workshop on Agent Organizations: Theory and Practice. San Jose, CA, July.

ANSI/RIA. 1999. American National Standard for Industrial Robots and Robot Systems - Safety Requirements, R15.06. Ann Arbor, MI.: Robotic Industries Association.

Armstrong, W. 2003. Consequentialism. *The Stanford Encyclopedia of Philosophy*, (http://plato.stanford.edu/archives/sum2003/entries/consequentialism).

Audi, R. ed. 1995. *The Cambridge Dictionary of Philosophy*. New York: Cambridge University Press.

1.Rewards and utilities can both be assigned to states. The reward function of a state R(s) is a short-term value for being in that state, whereas the utility function U(s) is a long-term value for all the states that are likely to follow from that state forward.

Bekey, G. 2005. *Autonomous Robots: From Biological Inspiration to Implementation and Control*. Cambridge, MA.: The MIT Press.

Bonato, P. 2005. Advances in Wearable Technology and Applications in Physical Medicine and Rehabilitation. *Journal of NeuroEngineering and Rehabilitation* 2(2).

Boutilier, C., Dean, T., and Hanks S. 1999. Decision Theoretic Planning: Structural Assumptions and Computational Leverage. *Journal of Artificial Intelligence Research* 11:1-94.

Boutilier, C., Reiter, R., Soutchanski M., Thrun, S. 2000. Decision-Theoretic, High-level Agent Programming in the Situation Calculus. In Proceedings of the Seventeenth National Conference on Artificial Intelligence, 355-362. Menlo Park, Calif.: AAAI Press.

Brooks, R. 1986. A Robust Layered Control System for a Mobile Robot. *IEEE Journal of Robotics and Automation* 2(1):14-23.

Brooks, R. 2002. *Flesh and Machines: How Robots Will Change Us*. New York: Pantheon Books.

Cummins, R. ed. 1997. *Advanced Cardiac Life Support*. Dallas, TX.: American Heart Association.

Dean, T., and Wellman, M. 1991. *Planning and Control*. San Mateo, CA.: Morgan Kaufmann.

Economist, The. 2005. Humanoids on the March. *The Economist* - Special Section 374(8417):3-4.

Fong, T., Nourbakhsh, I., and Dautenhahn, K. 2003. A Survey of Socially Interactive Robots. *Robotics and Autonomous Systems* 42:143-166

Graf, B., Hägele, M. 2001. Dependable Interaction with an Intelligent Home Care Robot. In Proceedings of ICRA-Workshop on Technical Challenge for Dependable Robots in Human Environments, IV-2. Piscataway, NJ.: IEEE.

Griffin, J. 1986. *Well-Being: Its Meaning, Measurement and Moral Importance*. Oxford, UK.: Clarendon Press.

Holmberg, R., and Khatib, O. 2000. Development and Control of a Holonomic Mobile Robot for Mobile Manipulation Tasks. *International Journal of Robotics Research* 19:1066-1074.

IEEE. 1990. IEEE Code of Ethics, (http://onlineethics.org/codes/IEEEcode.html).

Kahneman, D. 2000. Experienced Utility and Objective Happiness: A Moment-based Approach. Kahneman, D., and Tversky, A. eds. *Choices, Values and Frames*. New York: Cambridge University Press.

Khatib, O., Yokoi, K., Brock, O., Chang, K., and Casal, A. 1999. Robots in Human Environments: Basic Autonomous Capabilities. *International Journal of Robotics Research* 18:684-696.

Lindström, M., Orebäck, A., and Christensen, H. 2000. Berra: A Research Architecture for Service Robots. In Proceedings of the IEEE International Conference on Robotics and Automation, 3278-3283. Piscataway, NJ.: IEEE Press.

Liu, Y., and Koenig, S. 2005. Risk-Sensitive Planning with One-Switch Utility Functions: Value Iteration. In Proceedings of the Twentieth National Conference on Artificial Intelligence, 993-999. Menlo Park, Calif.: AAAI Press.

Low, K., Leow, W., and Ang, M. 2002. A Hybrid Mobile Robot Architecture with Integrated Planning and Control. In Proceedings of the First International Joint Conference on Autonomous Agents & Multiagent Systems, 219-226. Bologna, Italy. ACM.

Lucas, P. 1996. Knowledge Acquisition for Decision-theoretic Expert Systems. *AISB Quarterly* 94:23-33.

Panangadan, A., Ali, A., and Talukder, A. 2005. Markov Decision Processes for Control of a Sensor Network-based Health Monitoring System. In Proceedings of the Seventeenth Conference on Innovative Applications of Artificial Intelligence, 1529-1534. Menlo Park, CA: AAAI Press.

Papudesi, V., and Huber, M. Learning from Reinforcement and Advice Using Composite Reward Functions. In Proceedings of the 16th International FLAIRS Conference, 361-365. Menlo Park, CA: AAAI Press.

Pineau, J., Montemerlo, M., Pollack, M., Roy., N., Thrun, S. 2003. Towards Robotic Assistants in Nursing Homes: Challenges and Results. *Robotics and Autonomous Systems* 42:271-281.

Runyan, C., and Casteel C. eds. 2004. The State of Home Safety in America: Facts About Unintentional Injuries in the Home. Washington, D.C.: Home Safety Council.

Russell, S., and Norvig, P. 2003. *Artificial Intelligence: A Modern Approach*. Upper Saddle River, NJ.: Prentice Hall.

Ryan R., and Deci, E. 2001. On Happiness and Human Potentials: A Review of Research on Hedonic and Eudaimonic Well-Being. *Annual Review of Psychology* 52:141-166.

Stoytchev, A., and Arkin, R. 2004. Incorporating Motivation in a Hybrid Robot Architecture. *Journal of Advanced Computational Intelligence and Intelligent Informatics* 8:269-274.

Thrun, S. 2004. Toward a Framework for Human-Robot Interaction. *Human-Computer Interaction* 19:9-24.

Thrun, S., Beetz, M., Bennewitz, M., Burgard, W., Cremers, A., Dellaert, F., et al. 2000. Probabilistic Algorithms and the Interactive Museum Tour-guide Robot Minerva. *International Journal of Robotics Research* 19:972-999.

Thrun, S., Fox, D., Burgard, W., and Dellaert, F. 2001. Robust Monte Carlo Localization for Mobile Robots. *Artificial Intelligence* 128(1-2):99-141.

U.N./I.F.R. 2004. *World Robotics 2004: Statistics, Market Analysis, Forecasts, Case Studies and Profitability of Robot Investment*. New York and Geneva: United Nations.

Wolf, P., Abbott, R., and Kannel W. 1991. Atrial Fibrillation as an Independent Risk Factor for Stroke: The Framingham Study. *Stroke* 22:983-988.