Computer Ethics: A Practical Approach towards Solving Ethical Dilemmas

Osama Shata
Qatar University
PO Box 2713
Doha - Qatar

ABSTRACT
Ethical problems are part of our daily lives. Individuals face ethical dilemmas in all disciplines such as business, medicine, engineering, and computer science. Hence teaching computer ethics became a mandatory course in most computer science curriculums. The syllabus of such a course usually introduces students to ethical theories such as Utilitarianism, Kantianism or Divine. Most of these theories and which have been developed to deal with ethical situations, are mainly the result of the work of philosophers. Hence, many computer science students look at the computer ethics course and these theories as intruders to the discipline because they are not applied. Computer science students are used to practical approaches towards problem solving and which depend on algorithms, procedures and well-defined steps. Meanwhile computer ethics theories introduce proportional and subjective solutions and which may differ from an individual to another for the same situation or dilemma. In this paper we introduce a computational approach towards solving ethical dilemmas. Our approach depends on using well-defined algorithms and scientific theories and which may be seen as practical and applied by computer science students, and hence closer to their modes of studies.

General Terms
Computer Ethics, Ethics.

Keywords
Computer ethics, algorithms.

1. INTRODUCTION
Ever since the initial appreciation of philosophy, ethics has earned a central and critical importance amongst philosophers thinking and rightfully so. Ethical thinking is deeply associated with every aspect of human behavior and, thus, affects mostly every area of philosophical endeavor. Having situations which constitute ethical dilemmas are undeniably ancient ones, and are also a common practice which individuals experience on daily basis. However, despite the numerous studies conducted in the field of ethical inquiry, by a multitude of scholars, most suggested treatments of ethical issues remain patently primitive and largely unchanged since antiquity. How does one then explain the inability of even the most veteran of ethical scholars to consider a systematic methodology of tackling ethical dilemmas? And could such a methodology even exist? This paper will explore the possibility of realizing such a methodology while adopting a computational and an applied approach. This will be accomplished by first rationalizing the existence of ethics and articulating a valid utility for it, then coming up with a practical definition (which explicitly defines ethics with regards to various ethical theories), then introducing our suggested model.

2. THE UTILITY OF ETHICS
What purpose do ethics, ethical behavior, or even morality serve? The most simplistic and most general utility of Ethics (a term which we will take to mean the branch of philosophy that studies how members of a certain society or organization should act in different situations with respect to some believes and / or principles) is that it serves to establish the best behavior a person should undertake in order to honor a set of principles.

This definition of the utility of ethics follows directly from the fact that multiple ethical theories exist, and that they all differ on their core belief of what should be the ultimate goal of human behavior. In the theory of Divine Command, people should behave according to the will of God. The theory of Utilitarianism or Consequentialism dictates that the goal of individuals' actions should be the attainment of happiness for the largest numbers and the reduction of pain or misery. Virtue Ethics perceive ethical behavior to be those actions, which a supremely virtuous individual would engage in, in any given situation. Thus, it is clear that the theories are many, but the pattern is clear, and so are the agreements as to what role ethics play in the realization of each theory's respective goal [1].

Thus, strong emphasis should be placed on what Ethics mean in our context. Ethics for us mean the process for ascertaining the best behavior to best honor a set of principles. It does not imply that we are searching for the best set of principles, as such an attribute would undoubtedly be subjectively allocated and would never receive universal acceptance. Such an endeavor is beyond the scope of this paper, and best left for the realm of philosophy.

3. A SUGGESTED MODEL FOR SOLVING ETHICAL DILEMMAS

3.1 Introduction
Naturally, the question that follows is whether the procedure or methodology for ascertaining the best course of action in all situations, in order for an individual to best manifest the core beliefs of his / her ethical model, is structurally similar across the different ethical theories. In order to best answer this question, let us elaborate and analyze the details of this procedure.

In every given situation, a person has a finite number of possible actions that s/he can undertake in response to this situation. Each of these actions can lead to one or more possible consequences, depending on some environmental variables. The degree of ambiguity and uncertainty increases

as we descend through the levels of this logical tree of possible consequence. However, the most immediate consequence is generally the one that is best understood and often acknowledged to be the most important. This fallacy is of significance, but will be dealt with later; the intention here is to model the thinking process of the average person, who possesses some ethical standard.

The person in question here is faced with two entities that need to be evaluated: the action and most immediate consequence. For example, a person in a public place has heard, by accident, a couple of people planning for a murder. The person is now in a position to continue eavesdropping in order to attain information that would assist in stopping the crime, or to just change place or ignore listening. For the sake of argument, let us assume this person considers eavesdropping to be unethical, in respect to his / her particular ethical model, but s/he also considers murder to be unethical. Now, s/he considers the action of eavesdropping to be ethically wrong, but the consequence – the prevention of murder – to be a good thing. The key element that will dictate the outcome of this is the person's perception of the magnitude, either of goodness or harm, of the action and the consequence. The methodology for dealing with particular scenario has been clarified, so it is possible now to move on to the next. The other type of scenario is when a person is placed in a position where one of the alternatives is in fact not considered unethical. We can consider the exact same situation from the perspective of a person whose ethical model does not view eavesdropping as a violation of his/her ethical code. To this individual, the ethical action is both clear and trivial. It should be apparent to the reader, from the binary treatment of these situations in terms of ethical and unethical, that two more scenarios remain. Those two scenarios would be where either the consequence is unethical and the action is unethical, or where the action is ethical, but the consequence is unethical. In the first scenario, there's clearly no motivation to undertake an action that is both unethical and would lead to an unethical consequence. The second scenario is, in fact, the mirror of the very first scenario. This would be the scenario where the person chooses not to eavesdrops and thus a murder happens, and, once again, the perceived magnitude is the key to this dilemma.

Considering that, while until now, no reference to any particular ethical theory was a given, nevertheless, it was possible to outline a satisfactory thought process for tackling an ethical problem, it is evident that a universal methodology can be conceived of for dealing with ethical problems with respect to multiple ethical models.

3.2 Description of the suggested model.

Having attained a satisfactory definition of Ethics and its utility, one can engage in a closer analysis of its mechanics. This paper will start this by delving deep into the analysis of what causes an action to be considered ethical or unethical by individuals. The first, and most conspicuous, aspect of differentiating the ethical from the unethical is the influence of the key principles in the ethical theories. As an example, consider the theory of divine command. If a person believes that s/he is divinely mandated not to steal, this will play a role in his/her ethical evaluation of every situation. Of course it is not always trivial to discern whether an action is in line with certain principles. The areas begin to grow more and greyer as unfamiliar scenarios are introduced. For a person that believes stealing to be unethical, it is not clear whether violation of intellectual property should be considered stealing. However, it is obvious the two are related in some manner. While most people would sense that the two are related, it is not directly obvious why that is so. One answer might reside in the fact that the two lead to similar consequence. Stealing causes honest working people to lose what they have rightfully earned, and, similarly, so does the violation of intellectual property. Thus, one key point to take into consideration when treating ethics systematically is that the evaluation of how actions manifest the principles of an ethical theory should also extend to the immediate consequence and indeed all subsequent ones.

The second aspect of the mechanism of the introduced model is to inquire into the determination of the most ethical alternative in a given situation. Two fundamental problems arise when one attempts to think about this issue systematically: How does one know what all the possible alternatives are, and how realistically can one predicts the chain of consequences triggered by a certain alternative.

The first issue appears to be particularly daunting and arguably no perfect solution exists. The core of this problem rests on the fact that an individual may not be aware of all the environmental variables in every situation, and, thus, some solutions never occur to him/her. This problem does appear to exist in popular Game Theory as well [2,3], and to illustrate this fact, let us consider the problem of the Prisoner's Dilemma [4] and its so-called optimal solution. The Prisoner Dilemma is a classic scenario used in Game Theory to demonstrate rational decision making based on mathematical calculations of the best possible outcome. Suppose the police have apprehended two individuals that are accused of armed robbery and grand theft auto. The police does not possess enough evidence to convict the two of the robbery, but they can be convicted of grand theft auto and jailed for two years. Consequently, offers are made to both of the prisoners, while in separate rooms, now if one of them was to confess and implicate the other partner, then s/he would go free and the other partner would get 10 years. If they were both to confess, then both would receive 5. If neither was to confess, then both would get 2 years for grand theft auto. To ascertain the best possible course of action, according to Game Theory, one would attach magnitudes to each outcome and then construct a matrix to compare each the merits of each action. It follows that both prisoners are better off confessing and implicating their partner. Surely, however, there are other alternatives besides either confessing or refusing. What if the prisoner manages to escape captivity and be rendered free without confessing or refusing, what if it is possible to bribe the chief of police and let both go, etc… The list of alternatives can potentially be endless, and the shortcomings of this model are clear, and, in fact, they are expected, since to be able to construct a perfect solution for this problem, one would need nothing short of universal omniscience.

Having clarified the problematic issue, there is a need to find an approximate solution. It follows that, since the root cause of the inability to perceive all possible alternatives is the lack of knowledge of all relevant factors, whatever alternatives one can conceive of are directly related to what s/he may know. This conclusion appears true enough as it is evident in everyday life. One can easily observe that children or individuals who lack experience or knowledge are not consistently able to see all the possible alternatives in a situation. In fact, those individuals who can see more alternatives than others are labeled as gifted and talented. Another significant consequence of individuals short coming as humans in this area is that their inability to see all possible alternatives effectively means that sometimes they won't be

able to perceive of the best alternative/ solution to an ethical problem. This fact is clearly as true in Ethics as it is in any other area of human thought.

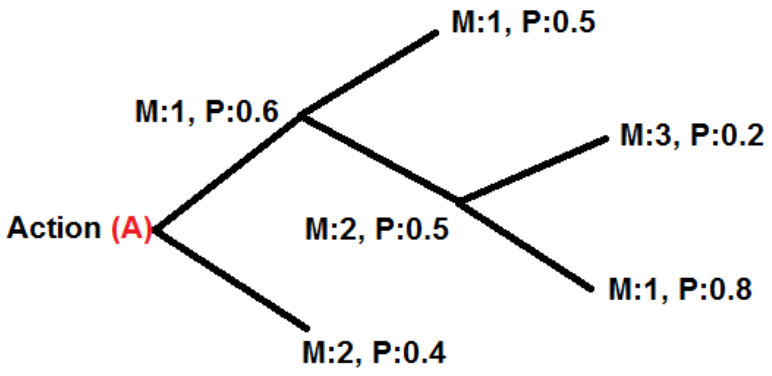
The paper introduces next the second issue, which is how realistically can one predicts the chain of consequences triggered by a certain action. Again, the answer is 'not fully', and it is for similar reasons. Chaos theory indicates that the flap of a butterfly's wing can potentially cause a hurricane after enough time. Applying this to Ethics, it is possible conclude that an action that appeared to have an immediate ethical consequence, and could even have subsequent ethical consequences, could potentially be the cause of an unintended ethical violation of colossal magnitudes. This shortcoming can also be attributed to the lack of knowledge about the future and the outcome of an action in the long-term. It should also be noted that ones' inability to predict the future will undoubtedly cause him/her to erroneously evaluate alternatives, leading him/her to dismiss the truly optimal alternative and mistake sub-optimal alternatives for optimal ones.

The final component left to discuss in the suggested model is the degree of magnitude of each outcome, which needs to be utilized to compare each outcome and alternative. First there is a need for a system for rating the ethicality. This rating system will consist of 4 levels, ranging from extremely unethical (4) to neutral (2) to ethical (0). The most optimal alternative will be the one with the lowest score and, conversely, the least desirable alternative will be the one with the highest score. Additionally, because it is possible that each action and consequence might lead to more than one outcome, then there is a need to introduce a probability factor. The probability factor will be a number between 0 and 1 and which indicates how likely this consequence to follow from the preceding consequence or action. Since it is necessary that at least one event takes place, the sum of the probability factors for the outcomes of one alternative or consequence on each level of tree must be 1.

Now the suggested model is complete, it is possible to observe the complete picture. Begin with establishing a set of ethical principles, from which the criteria for gauging each alternative and its outcomes will be derived. Then enumerate and define the conceivable alternative(s). Each alternative will then be rated from 0 to 4 based on how ethical the action itself is. Next, one should predict the first level of consequences for each alternative and assign each consequence a rating for how ethical it is. Lastly, repeat this process until ceasing to be able to predict new outcomes. Now the structure of suggested model will be a logical tree with a number of branches for each alternative and then a number of branches from each alternative designating the possible consequences and a number of branches from each consequence indicating the consequences that might follow. The next step is then to calculate the optimal course of action by going down every branch and calculating the scores of the consequences that follow. Whenever one consequence / alternative has more than one consequence that follows from it, the score of each consequence that follows will be multiplied by the probability factor, and any consequences that follow from those consequences will also be multiplied by the probability factor. Figure (1) illustrates the formula.

After calculating the magnitude of each action, the most optimal action will be that with the least magnitude.

There is another major logistical point worth discussing, which is the database that will record the consequences of the various actions. Beyond simply calculating the most ethical actions at some situation, it is also of significant utility to store the series of actions and consequences created for that scenario, in order to allow for reusing it. This is clarified by observing that different situations might lead to the same consequence and that the predicting powers of individuals using this algorithm differ. As time goes by, this will allow the database of predictions about actions and their consequences to constantly evolve to become more comprehensive and more accurate.



Total Magnitude for Action (A) = 0.6(1+ 0.5(1) + 0.5(2 + 0.2(3) + 0.8(1))) + 0.4(2) = 2.72

Fig 1: Calculating magnitude of an action

4. CONCLUSION

This paper has presented a new model for solving ethical dilemmas. The model may be suitable and appealing to

students in science-based majors (e.g. Computer Science, Engineering, … etc.) because of its applied and practical nature (e.g. uses probability). Hence, it differs from other models and which mainly depend on philosophical underlying theories. However, the model may be generalized and used for solving ethical dilemmas by any individual. The presented model has a well-defined algorithm and allows even the least knowledgeable individuals to both predict, to a large extent, the outcome of their actions and ascertain which action leads to the outcome that is most compatible with their ethical view. When this idea is applied to other usable ethical theories (e.g. utilitarianism, which is mainly concerned with the well-being of the majority of individuals in a society) then the presented model could potentially lead to unprecedented prosperity for the masses, as decisions are constantly being optimized for their benefit. In addition, it would definitely contribute to other research areas such as the automation of ethical thinking, and the computerized delegation of outcome prediction and enumeration of alternatives [5]. The presented model can be further expanded and combined with artificial intelligence techniques for potential research on approaches that would allow computers to predict outcomes and assign alternatives, which then would yield nothing shorter than the complete automation of ethical thinking.

4. ACKNOWLEDGMENTS

This publication was made possible by NPRP grant #09-1205-2-470 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the author.

5. REFERENCES

[1] Jackson, A.M. 2011. SCIENCE OF ETHICS: A User's Guide for Modern Humans. [cited 4/27/2012]; Available from: <http://www.webring.org/l/rd?ring=rot;id=1;url=http%3A%2F%2Fwww.arthurmjackson.com%2FMentor.html>.

[2] Verbeek, B. and Morris, C. 2010. "Game Theory and Ethics", *The Stanford Encyclopedia of Philosophy (Summer 2010 Edition)*, Edward N. Zalta (ed.), [cited 3/20/2012]; Available from URL = <http://plato.stanford.edu/archives/sum2010/entries/game-ethics/>.

[3] Ross, D. 2011. "Game Theory", *The Stanford Encyclopedia of Philosophy (Fall 2011 Edition)*, Edward N. Zalta (ed.), [cited 4/10/2012]; Available from URL = <http://plato.stanford.edu/archives/fall2011/entries/game-theory/>.

[4] Kuhn, S. 2009. "Prisoner's Dilemma", *The Stanford Encyclopedia of Philosophy (Spring 2009 Edition)*, Edward N. Zalta (ed.), [cited 4/20/2012]; Available from URL = <http://plato.stanford.edu/archives/spr2009/entries/prisoner-dilemma/>.

[5] Mclaren, B.M. 2006. Computational Models of Ethical Reasoning: Challenges, Initial Steps, and Future Directions. IEEE Intelligent Systems. [cited 2/12/2012]; Available from <http://doi.ieeecomputersociety.org/10.1109/MIS.2006.67>.