Route selection method with ethical considerations for automated vehicles under critical situations

Balázs Németh

Abstract—Route selection for automated vehicle is a challenging task under critical situations, i.e., which can lead to serious or fatal injury. Due to the high risk of a dangerous collision with its consequences on human participants, the selection of the possible router poses ethical problems. This paper proposes a multi-layer route selection strategy for automated vehicles under critical situations, in which ethical considerations have been incorporated. The selection strategy contains a graph-based quantitative evaluation layer and a layer with qualitative evaluation based on the ethical principles. The operation of the route selection method through some examples is illustrated.

I. INTRODUCTION AND MOTIVATION

In the last decade the complexity of automated systems has been significantly increased, which posed various theoretical challenges for artificial intelligence methods and control theory. It affects not only the technological problems, but it also affects broader context, i.e., social, political, ethical, philosophical and theological fields [1]. Due to the trend of automation in transportation systems, the challenges by automated driving features with reduced human participation have been induced [2].

Although the most popular challenge in this field is the decision of automated vehicle under critical situations, increasing disparity between developed/undeveloped regions [3], data privacy [4], military applications [5], economic impact [6], and the problem of alienation [7] are also crucial challenges. Therefore, examination of ethical consequences of automation is a current research field, which supplements the technical oriented viewpoint of engineering research [8]. In the context of automated vehicles, the studies of Massachusetts Institute of Technology in the framework of Moral Machine is outstanding. The goal of this project is to analyze the decision under critical traffic situations [8]. The research has pointed out that human priorities and moral preferences can depend on individual, cultural and demographic characteristics. Nevertheless, the consequences of the results can provide a psychological roadblock to automated vehicles, because the requirements of vehicle consumers and of lives saved maximization can be in conflict [9].

A further control-oriented contribution in the connection of automated vehicles and ethical problems can be found in [10]. In this paper selected ethical considerations in the control design through Model Predictive Control (MPC) formulation are incorporated. The cost-based ethical principles (i.e., consequentialist viewpoints) through the objective function are formed, and the rule-based ethical principles (i.e., deontology) in the constraints of the MPC optimization problem are involved. Another approach [11] distinguishes the automated vehicle control to be selfish or altruistic, which can be interpreted on a global traffic-oriented level. Through the tuning of the vehicle control different characteristics can be achieved. Moreover, a further way of considering ethical decisions is their consideration through learning process on samples. In this way, the consideration is indirect, because the samples contain the behavior of the driver in itself. For example, imitation learning is an efficient tool to incorporate in the driver model the human driving style through neural networks [12]. The learning of human driving style and decisions poses the problem of acceptance of driver decision, even if it is problematic from ethical point of view.

The presented contributions of the topic show that incorpotation of ethical principles in the control design of automated vehicles is challenging. Its first difficulty is arisen by appropriate ethical laws, which are acceptable from the aspect of consumers, groups of society, economically and technically reliable. Nevertheless, if it is possible to find a set of acceptable ethical laws, second, it can be difficult its formulation in the control synthesis. Ethical challenges of automated vehicles mainly from the viewpoint of philosophy are examined [13], [14], [15]. Some theological works also deal with these challenges, especially trolley problem, (see e.g. [16]), but a solution, which in practice can be used by a control engineer, has not been presented yet, to the best of the author.

This paper focuses on the ethical problem of route selection under critical situations of automated vehicles, which can lead to serious or fatal injury. The goal of the paper is to provide a route selection algorithm, in which ethical considerations have been incorporated. This paper is under the assumption that biblical texts can be efficient sources of ethical laws, which is confirmed by the Jewish and Christian theological and philosophical traditions [17]. The contribution of the paper is a multi-layer route selection algorithm, which incorporates in quantitative and qualitative evaluation layers. In the qualitative evaluation a graph-based method for the prediction on probability of critical conflict is involved. In the qualitative evaluation some Christian ethical
principles are incorporated, with which a systematic method for the analysis of the possible routes is created.

The paper is organized as follows. The quantitative evaluation method with the formulation of graph-based route selection algorithm in Section II is presented. Section III proposes ethical considerations, which in the qualitative evaluation layer of the route selection method have been incorporated. Section IV illustrates the results of the route selection algorithm for some critical scenarios and finally, the paper is concluded in Section V.

II. QUANTITATIVE EVALUATION: GRAPH-BASED ROUTE SELECTION ALGORITHM

In this section the quantitative analysis of the possible routes is presented. The aim of the quantitative analysis is to explore the probability of collision on a given route and to estimate the risk of a fatal accident.

The presented method is based on a graph-based route selection algorithm, which can be effectively used for vehicle control problems under multi-vehicle context. In this paper, the summation of the method is presented, the details of the computations can be found in [18]. First, the graph-based algorithm the probability of collision \( P_f \) on the routes is estimated. Second, as an addition to the original method, the risks of accidents with serious injuries or fatal occasions \( P_f \) on the routes based on the analysis of [19] are evaluated.

The purpose of the route selection algorithm [18] is to explore the probability of collision for the automated vehicle, considering the predicted motion of the surrounding vehicles. In the method the possible routes on the forthcoming road section in longitudinal and lateral directions are divided equidistantly. The aim of the selection algorithm is to choose route and velocity for the automated vehicle. For achieving this goal, a directed graph \( G = (V,E) \) on the predicted road section is built. Vertices \( (V) \) of the graph represent the possible route points and velocity profile of the vehicle. Vertices are connected by edges \( (E) \), which are route and acceleration profile for the vehicle. The aim of directing the graph is to constraint the route and the vehicle motion, i.e., physically reliable routes are represented by the graph. An example on the graph with the possible routes in Figure 1 is found. The vertex of the graph in the origin represents the current position of the vehicle, \( x \) and \( y \) axes are related to the longitudinal and lateral motion, while \( z \) is the longitudinal acceleration. The graph is closed in one vertex at the end of the predicted road horizon, which expresses the destination of the vehicle.

The probability of collision on each vertices of the graph is evaluated. The evaluation requests the prediction of the automated vehicle and the prediction of the surrounding vehicles, pedestrians, etc., [18]. An illustration on the result of the evaluation for a fixed velocity of the vehicle in Figure 2 is found. In the illustration red coloured rectangles represent high probability of collision, and green rectangles are related to low probability of collision. The rectangles are fitted to \( G = (V,E) \), and thus, the risks of each possible routes can be evaluated. In the presented example it is requested to perform a lane change by the automated vehicle, due to parking vehicles on the side of the road.

In case of accidents with serious injuries or fatal occasions the probability of collision with the examination on risk of mortality must be extended. In road traffic context, the work of Doecke et al. [19] in this paper is used. This study provides an analysis on the relationship between impact velocity and the risk of serious injury, based on a wide-range database on vehicle crashes. As an example, it has been presented that 90\% risk of serious injury on 99km/h impact velocity in head on situation, on 110km/h impact velocity in side situation and on 128km/h in rear situation is achieved. It means that through the velocity of automated vehicle and the surrounding vehicles, pedestrians the value of risk \( P_f \) can be estimated. Thus, the probability map for route selection under critical situations from \( P_f \) and \( P_f \) is formed, i.e., the total probability of critical conflict \( P \) for each rectangle is:

\[
P = P_f \cdot P_f.\]

At the end of the quantitative evaluation of the routes, the route on the graph based on the probability of critical conflict is selected. Weights are defined for each edges of the graph, where the value of the weight is equal to \( P \).
Thus, it is necessary to find the route which guarantees the minimum probability of a critical conflict on the graph. It has been aided by the Dijkstra algorithm, see [20], whose role is to find the shortest path from the initial vertex to the target vertex. At the beginning of the algorithm all of the vertices are called unvisited, which form the unvisited set. Moreover, the distances of all unvisited vertices from the initial vertex are considered to be infinite, except the initial vertex, which is related to the current position and velocity of the autonomous vehicle, and it has 0 distance. During the algorithm each neighbour \( j \) of the current \( i \) vertex is examined. It means that the weights on each edge between \( i \) and \( j \) vertices is added to the current route between \( i \) and the initial vertex. If the target vertex is marked as visited, or if all of the vertices from the unvisited set are removed, the algorithm is stopped. The algorithm yields the shortest path with the minimum distance between the initial and the target vertices, where the distance \( D \) is defined as

\[
D = \sum_{d=0}^{M-1} P(d, d+1),
\]

where \( M \) is the number of vertices in the route.

III. QUALITATIVE EVALUATION: ETHICAL CONSIDERATIONS IN ROUTE SELECTION

In this section the qualitative evaluation of the route selection is proposed. First, the ethical principles are formulated and second, the incorporation of the principles in the qualitative evaluation method is proposed.

A. Formulation of ethical principles

In this paper four fundamental ethical principles are provided, which in the qualitative evaluation on route selection are incorporated. The following principles on Christian ethics are based, especially connected to the ethical viewpoints of Protestantism.

Designing safe vehicle control is a consequence of divine mandates

The requirement against vehicle control systems to be safe has been generally known as a control objective. Nevertheless, it goes beyond the mathematical forms and thus, it is a social requirement. Moreover, the task of designing vehicle control systems to be safe can also be approached from Christian theology; i.e., it is a consequence of the first divine mandate.

Originally, the terms of mandates the world-famous theologian, Dietrich Bonhoeffer in his book *Ethics* has been formed [21]. He defined four divine mandates based on the Bible, i.e., work, marriage, government, and church. These terms in wide contexts at the last decades have been interpreted, e.g., in political [22], in social, family [23] and in environmental crisis [24]. Similarly, the divine mandate of work also can be interpreted for control design. Since the work, as divine mandate is based on the call of God, high-quality profession has legitimacy in Christian context. Thus, in the work of control design, it leads to the goal of achieving safe and high performance control operation. In the context of route selection, it means that the results of the quantitative evaluation are fundamental for the selection. Under normal, i.e., non-critical situations, the results of quantitative evaluation due to the resulted collision-free motion profile is suitable. Nevertheless, under critical situations, i.e., collision with serious or fatal injury, quantitative evaluation provides an initial solution on the problem. Thus, the initial solution through qualitative evaluation is improved.

Avoiding instability in the context of critical situation

Avoiding unstable motion of systems is a fundamental goal for all control systems. Furthermore, instability using biblical and theological analogies in the context of human life and social behavior can also be interpreted, see [25], [26]. In [25] analogies between instability and sin, death are provided, with which the goal of guaranteeing stable operation of systems reaches an extended interpretation.

In this paper the extended interpretation of instability is used. It has two consequences in the context of route selection problem. First, the goal of the route selection under critical situations is to avoid involving new participants into the situation. For example, if a fatal accident by an irregularly driving vehicle ahead of the automated vehicle is occurred, it is prohibited to select lane changing, which results in fatal collision with another vehicle. Second, a fundamental challenge of automation is that it can increase the social gap [3], e.g., expensive vehicles with high-quality automated technologies can be reached by wealthy consumers, which further improves their quality of life and their probability of survival under a fatal situation. Therefore, the second interpretation of avoiding instability in the context of automated vehicle route selection is as follows. It is prohibited to prefer saving lives of passengers compared to further participants in the critical situation, if serious injury or fatal situation cannot be avoided. Although this ethical consideration goes against the mainstream concept of consumers in Europe or in North-America (see the analysis on individualism in [8]), it can facilitate the social equality in automated driving.

Forming objective function can lead to the problem of good and bad

In complex real world scenarios of automated vehicles, especially in fatal situations, it is impossible, or at least difficult to form objectives for automatic control systems, which fulfill the specifications of biblical ethics. Under simplified world scenarios, the objectives may be easily formed. For example, in a control design for cruise control, the reward in the reinforcement learning process can be formed, which leads to the acceptable motion of the vehicle in signalized intersections, i.e., considering red and green light for stopping and moving. As a counterexample, classification of good and bad instances to achieve an appropriate decision tree for lateral vehicle control systems is not trivial [27]. Nevertheless, complex scenarios with human participation
in a multi-vehicle situation can pose extremely difficult evaluation of the control objectives.

Forming objective function in complex real world scenarios of automated vehicles, especially in the problem of route selection for handling fatal situations, leads to the problem of good and bad. Although it is out of scope for control engineering science, on the level of biblical texts and theological thought the problem may be handled. From biblical viewpoint, knowing of good and bad, as absolute categories, is in the authority of God. Its reason is that absolute good is the part of his own character, e.g. Jesus states "No one is good but God alone." (Mark 10:18b). Nevertheless, human has limited knowledge on good and bad. Its limitation roots in the fall of Adam and Eve, see "the Lord God said, 'See, the man has become like one of us, knowing good and evil'" (Gen 3:22) or "Give your servant therefore an understanding mind to govern your people, able to discern between good and evil, for who can govern this your great people?" (1 King 3:9). Finally, Apostle Paul reflects also on the problem of good and bad, i.e., "I can will what is right, but I cannot do it. For I do not do the good I want, but the evil I do not want is what I do." (Romans 7:18-19). This text reflects to the discrepancy between having knowledge on good and acting in a good way.

Thus, the automotive control engineer in the forming of objective function meets with the problem of good and bad. It is analogous to the main challenge of managers, who also deal with the problem of good and bad, as defined by the founder of modern management, Drucker [28]. Nevertheless, decision on good and bad has divine character, and thus, an absolute good decision is impossible. It leads to the next ethical fundamental assumption.

Random can have legitimacy under fatal situations

In critical situations the automated system must make a decision, and thus, it is requested to find an algorithm, which results in a good decision. Nevertheless, formulation of a control objective, which results in good automated decision is problematic (see previous subsection) i.e. neither an optimal control, nor an imitation of human behavior can result in absolute good decision. Therefore, in critical fatal situations it can be suggested to use a random decision, especially in the case of automated vehicles a random selection of the vehicle route.

Some biblical thoughts confirm the legitimacy of using randomness for decision. For example, the using Urim and Thummim is a special practice to ask the will of God, see e.g., Numbers 4:27 or Ezra 2:63. Nevertheless, there are some open problems in the exegesis of these texts, see e.g., [29], [30]. The chosen of Matthias to be an apostle also reflects on this old tradition, such as "And they cast lots for them, and the lot fell on Matthias", see Acts 1:26. In spite of the random outcome of lot, it can be acknowledged as decisions of God, such as "The lot is cast into the lap, but the decision is the Lord’s alone.", see Proverbs 16:33. It is an interpretation of random events, which has become the part of some Christian traditions, i.e., it has been handled to be the part of his providence [31]. Although randomness in popular discourse seems to be analogous to unsystematic thought, it also can be formed as part of the nature of God [32]. This thought roots in the omniscience of God and in the creation [33].

B. Formulation of qualitative analysis

In the rest of this section the presented ethical principles in the qualitative analysis are incorporated. The qualitative analysis through an algorithm can be formulated, whose steps are as follows.

1) Quantitative evaluation provides an initial solution on the problem, which is the consequence of the divine maneuvers for designing safe vehicle control. Therefore, the total probability of critical conflict $P$ using (1) for the routes ahead of the vehicle must be computed. It provides information on the danger of the possible route choices. The goal of the route selection algorithm (2) is to find the route, where the cumulative value of $P$ on the selected route has minimum. Nevertheless, the algorithm cannot handle scenarios, where $D$ has minimum, but it contains a serious or fatal injury, while fatal occasion with the selection of another route with higher $D$ can be avoided. Moreover, if all selections lead to serious or fatal injury, further ethical principles, e.g., avoiding instability, must be incorporated. Therefore, the provided map on $P$ is the input for the qualitative analysis.

2) If the route selection problem contains routes with serious or fatal injury, these routes as possible selections have been ignored. Then, for the further routes the selection process (2) is carried out.

3) There are critical situations, when all of the possible routes lead to serious or fatal injury, i.e., all of them have high $P$ value sections. In these cases the further ethical principles have high importance, as detailed below.

4) Avoiding instability leads to avoid involving new participants into the situation. Thus, the possible route choices in the further part of the process are ignored, which involve new participants. This decision is based on the first interpretation of avoiding instability.

5) All of the further possible routes result in serious or fatal injury. Due to the second interpretation of avoiding instability, and due to the problem of good and bad, the social characteristics or other properties of the participants cannot be considered. Under critical situations, which necessarily leads to serious or fatal injury, the random choice on the selected route is ethically established. Clinical therapies pose similar challenge, when random choice is part of decision under critical situations, see e.g., [34], [35]. Consequently, random choice is selected for route selection only, but only in these critical situations.
IV. ILLUSTRATION EXAMPLES

The proposed qualitative and quantitative evaluation-based route selection process in this section through some examples is illustrated. During the examples, it is considered that the automated vehicle is in a critical situation, where the qualitative analysis for the route selection is requested.

The example on the critical situation in Figure 3 is illustrated. In the example a pedestrian suddenly crosses the road, which induced a critical situation with the automated vehicle. In the example it is considered that the automated vehicle due to its short distance from the pedestrian is not able to stop without collision. The automated vehicle has three route choices for handling the problem, i.e., it moves straight and the pedestrian is hit, it moves left into the opposite lane, which leads to a crash with vehicle 2, or it leaves the road to the right, which leads to a frontal impact collision with a tree on the roadside.

![Fig. 3. Critical situation of the automated vehicle](image)

In the first critical scenario the automated vehicle has high speed (90 km/h), while vehicle 2 is considered to have low speed (40 km/h). The illustration of the result of the quantitative evaluation for the characteristics of $P$ is found in Figure 4. In this scenario the high speed of the vehicle leads to a serious or fatal injury for the pedestrian, i.e., the mortality of the pedestrian is very probable (80%). Similarly, the collision with the tree on the roadside occurs serious injury for the passengers of the automated vehicle (75%). Nevertheless, a lane change to the left and a side impact collision with the slower vehicle 2 can significantly reduce the risk of a dangerous collision. Therefore, in this situation the straight and right turn routes are ignored and the route with left lane change is selected.

![Fig. 4. Illustration on the first scenario](image)

In the second critical scenario not only the automated vehicle has high speed (90 km/h), but also the speed of vehicle 2 to 120 km/h is increased. The map for the characteristics of $P$ in Figure 5 is illustrated. In spite of the previous scenario, the collision with vehicle 2 leads to the most dangerous collision with $P = 90\%$. It is resulted by the high mortality rate at frontal impact collisions [19]. Consequently, in this scenario all of the routes can lead to serious or fatal collision with high probability. Therefore, in this scenario the qualitative analysis has high importance in the route selection. A route selection based on the quantitative evaluation and a selection process with the minimization of $D$ (see (2)) results in changing lane to the right. Together with the qualitative selection process it leads to an ethical decision (first interpretation) results to ignore lane change to the left. This decision guarantees that new participants, i.e., the passengers of vehicle 2, are not involved in the critical situation. Second, the route of the automated vehicle is resulted by a random choice between straight motion and leaving the lane to the right. Since the probabilities of selecting straight motion and leaving the road are the same, it provides almost the same probability of survival for the pedestrian $100 - 0.5 \cdot 80 = 60\%$ and for the passengers of the automated vehicle $100 - 0.5 \cdot 75 = 62.5\%$, which is the consequence of the the principle avoiding instability (second interpretation). However, in case of an individualistic ethic in the automated vehicle, this disproportion can be compensated.

![Fig. 5. Illustration on the second scenario](image)

Remark: Although the selected ethical problem of automated vehicles in the popular culture of artificial intelligence is well-known and several papers deal with it, there is still some unrealistic characteristics from engineering viewpoint (see also [36]). First, in case of automated vehicles it is supposed that if there is a critical fault in the vehicle (e.g. actuator or sensor), it cannot be started, or it operates with reduced functionality. Thus, due to the advanced fault-detection algorithms and fault-tolerant control system, the risk of a fatal accident is significantly reduced [37], [38], [39]. Second, the goal of the automated vehicle control is to avoid the possibility of critical maneuvering, i.e., the motion of the vehicle far from critical vehicle dynamic situations must be kept [40], [27]. Consequently, the risk of an accident caused by an automated vehicle (e.g. through an actuator fault) is
very low, but not zero.

V. CONCLUSIONS

This paper has proposed a multi-layer strategy for route selection under critical situations of automated vehicles. The novelty of the paper has been the formulation of ethical principles in a qualitative analysis, which into the route selection algorithm has been taken part. Through illustrative examples the benefits of incorporation ethical principles in the engineering problem has been shown.

The further challenge of the method is to extend the algorithm for more complex critical situations, e.g., route selection in highway situations with lots of participants. First, it requires well-established statistics on accidents for extending the quantitative layer. Second, the consequences of ethical principles in traffic situations must also be interpreted.

REFERENCES


