# Social Psychology Inspired Reinforcement Learning Framework for Conflict Management in Connected Vehicles

Heena Rathore, Senior Member, IEEE, Yash Kumar Singhal, Student Member, IEEE, George Kodimattam Joseph, Member, IEEE

Abstract—In any connected network, resource scarcity, perceived road blocks, and incongruent objectives can potentially ensue conflicts among stakeholders. In the existing literature, trust has been cited as a crucial component in effective conflict management. Besides trust, empathy, and social intelligence play decisive roles in enhancing cooperation, encouraging information sharing, and promoting problem solving. In this paper, we discuss the three major components of conflict management and propose a computational model, which is inspired from social psychology for conflict management in connected vehicles (CV). Our mathematical algorithm focuses on three factors, namely trust, empathy, and social intelligence that are learned via social interactions among vehicles to ensure safety of vehicles and passengers. The triad of trust, empathy, and social intelligence is used to aid reinforcement learning (RL) for obtaining the optimal q-values and rewards in the shortest duration of time in the CV network. We have examined how the three factors influence the learning process and analysed their conflict management potentials. Results show that the proposed model is 118:18% more efficient than the trust-only-based RL algorithm.

Index Terms—Trust, social intelligence, empathy, conflict management, connected vehicles, reinforcement learning.

## I. Introduction

Conflicts, at multiple levels, might emerge in any network over a period of time as different stakeholders interact with each other [1]. Since ancient times, conflict resolution has played a key role in how human civilizations have evolved [2]. Even to this day, harmonious coexistence of nations faces great challenges [3]. In a social context, examples of networks include families, circles of friendship, or web of professional acquaintances. We refer to these stakeholders that participate and negotiate their way through their networks as parties. Conflict management is a practical and methodical approach to help parties find an amicable solution for a problem appearing on domestic, professional, social, or political fronts [4]. Some models have been proposed over time to explain how systems sustain cooperation when presented with conflicts and challenges [5]. A prerequisite to conflict management is effective negotiation, which builds upon teamwork and collaboration [6].

Connected vehicles (CV) are designed to communicate with each other and with infrastructure to enhance safety on the road. However, conflicts can arise when multiple vehicles try to access the same resources or space particularly at intersections [7]. These challenges arise due to the lack of understanding of the measuring environment and the limited accuracy of data sources, leading to data conflicts across diverse sources [8]. These challenges can ultimately affect traffic navigation systems and the safety of passengers. To address these challenges, building trust and reputation of CVs has emerged as a topic of interest [9], [10]. By improving the reliability and accuracy of information exchange, trust and reputation models can reduce the potential for accidents and improve the overall safety of the CV system [11].

Trust is the foundation for establishing and maintaining any relationship [12]. It is the conviction, existing among the parties of relationship, of mutual benevolence, and additionally, a social capital which ensures a solid social structure [13]. Absence of trust and understanding results in conflict, and furthermore, hampers any attempts of negotiation. Furthermore, paying attention to others' point of view in decision-making promotes cooperation and harmony [14]. In non-social circumstances, reinforcement models provide persuasive descriptions of feedback-based learning, but social interactions frequently involve inferences of others' attribute traits [15], which are not captured with taking trust as the only parameter. Therefore, the concepts of trust, empathy and social intelligence that are adopted from social psychology are incorporated in reinforcement learning (RL) framework to evaluate the interactions among entities. We believe that our work is significant since all the above three attributes together are not modelled in the state-of-the-art algorithms. Furthermore, we use these factors to reach to an optimal conflict state for a constructive and mutually beneficial outcome [16].

The paper introduces a model, which focuses on the triad of trust, social intelligence and empathy, to achieve conflict free actions in CV networks. The organization of the paper is as follows: Section II examines the status of contemporary research on conflict management. Section III provides insights on concepts such as trust and understanding and elucidates major implications of the concepts. Additionally, the section provides an argument that conflicts may be avoided by focusing on root causes and proposes a framework for minimizing conflicts. Section IV presents results of the proposed algorithm in CV networks and a comparative analysis of the results. Finally, in Section V, we present the conclusions drawn and the scope for future work.

H. Rathore is with the Department of Computer Science, Texas State University Texas, 78249 USA e-mail: heena.rathore@ieee.org

Yash Kumar Singhal and George Kodimattam Joseph are with Indian Institute of Technology, Roorkee and Jodhpur, India respectively.

Manuscript received March, 2022

This article has been accepted for publication in IEEE Transactions on Cognitive and Developmental Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCDS.2023.3279817



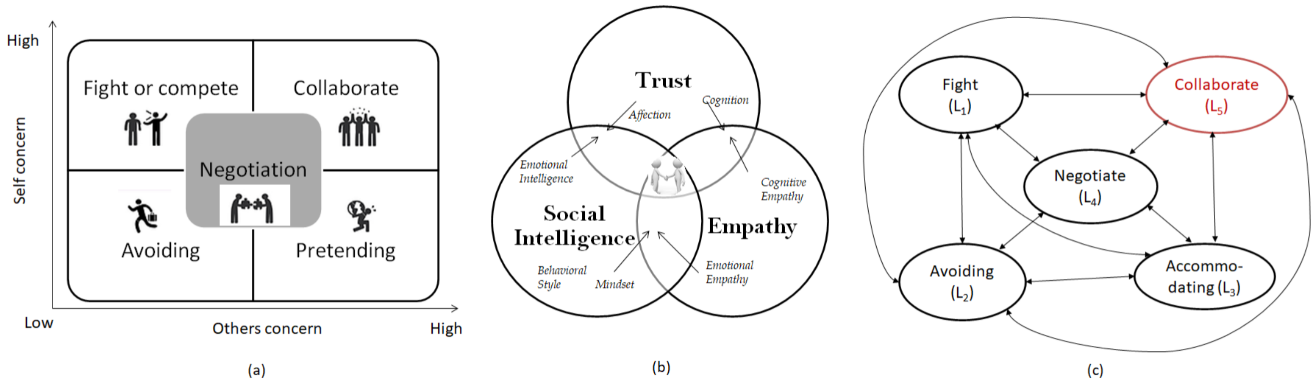



Fig. 1: (a) Conflict management strategies based on five states. (b) Computational Model for Conflict Management. (c) State diagram that illustrates transition between different strategies in the conflict management process.

## II. Background and Related Works

### A. Related Works in Social Psychology

Occurrence of conflicts is unavoidable in any relationship among individuals. This is because each individual manifests differences in physical constitution, behavioural patterns, priorities, and preferences. To ensure harmony in relations, it is imperative to explore possible ways of managing conflicts and it is unwise to overlook conflicts or tolerate it for long. Achieving mutual trust and understanding is central to managing conflicts, and furthermore, crucial in ensuring collective welfare any organization. This may be easier said than done since individuals may have dissimilar objectives that they prefer to achieve, which can be difficult to achieve simultaneously. Major types of conflicts that appear in multiple levels of social network are individual or social clashes, instrumental clashes, circumstantial clashes and resource conflict [17]-[18].

Irrespective of the source and the nature of the conflict, we believe, a timely intervention is the best way to manage, and if possible, to resolve a conflict. Delayed attempts of resolution might claim more resources and leave marks of grudge and resentment. These remnants can lead to latent hostility among the parties that may manifest itself in other completely unrelated contexts [19]-[20]. The five strategies, considering inter and intra party responses, proposed by Floyd [21], for resolving conflicts are described below and graphically presented in Figure 1(a) and (c). The first three strategies have deep psychological bearings and are closely related to the Fight, Flight, Freeze responses, often shown both by humans and animals.

Fight is a strategy when the parties fight and aggressively compete for resources without any consideration for emotions and feelings of others.

Avoidance is a strategy when the parties try to move away from the conflicting situation, possibly as if "kicking the can further down the road," in the hope that time will resolve the problem.

Accommodation is the strategy when, in order to maintain harmony, the parties acknowledge conflicts to be normal or deemed existential.

Collaboration is a strategy which, if handled properly, has a higher probability of mutual welfare as the parties try to work together to achieve a common goal [22].

Negotiation is a strategy that lies midway between accommodation and collaboration, wherein the parties compromise on the issue and resources.

Any technique that manifests the virtue that all affected parties are happy about the solution is probably the best one for resolving conflicts. It is prudent for the involved parties to negotiate on common grounds and clearly express their willingness to collaborate.

### B. Related Works in Conflict Management in CV

In this section we present related works of managing conflicts in CV networks. In most of the available state of the art studies, conflicts in CV are studied at unsignalized intersection. For instance, the authors in [23] proposed a differential game theory approach based on collision risk assessment algorithm along with Gaussian potential field approach to prevent conflicts at multilane intersections. They implemented the strategy on a hardware testing platform taking three and five vehicles at the intersections and modelled the behavior with aggressiveness parameter. A subset of recent works, which validates data shared in CV using information obtained from vehicle's onboard sensors, focuses on trust estimation workflows [28]-[29], [47]. However, these trust and reputation models do not capture other social principles namely empathy and social intelligence which can be measured through peer-reported estimates. Moreover they have poor interpretability when under unforeseen circumstances such as errors, it is hard to track down the root cause analysis [23].

This paper addresses the above listed shortcomings related to conflict management in CV and proposes a RL framework. RL [26] can be used to implement conflict mitigation because it can enable vehicles to learn from past experiences and make decisions based on the current situation [48]. RL can also handle complex and dynamic environments, and uncertain and incomplete information. By using RL, CV can learn to make better decisions in conflict situations, which can improve safety and efficiency on the roads. The authors in [24] pro-posed the non-cooperative game hypothesis with Q-learning to improve the decision making and address the driving conflicts of CV. The authors use semi-cooperative Nash Q-learning and Stackelberg Q learning which improved the traffic condition by 3.50% and 13.32% respectively as compared to constant strategy. Similarly, authors in [25] have proposed a distributed conflict free cooperation for multiple connected vehicles at intersections. They transform the vehicle cluster intersection to a platoon based on the conflict relationship of different traffic movements. They used linear feedback controller along with matrix decomposition to organize the movements of the vehicles. The state of the art techniques in this domain do not discuss the different types of resolution techniques that can be applied and can help in building a recommendation system [27].

This paper, proposes a better conflict management algorithm which incorporates trust, empathy and social intelligence factors in Q-learning based RL algorithm. In our earlier work [30]-[31] we used a distributed ledger technique based on a directed acyclic graph to mitigate the effects of data tampering attacks in CV. Social pyshcology factors such as ability, integrity, and benevolence of the vehicles were used to calculate the trust factors, which in turn determine the reputation of the vehicle [35]. We postulate that a similar approach can be used for conflict resolution in a CV network with additional cooperative factors such as empathy and social intelligence.

## III. Proposed Work: A Framework for Conflict Management

In this section, we propose a framework for conflict management that focuses on the principles of trust ($T$), social intelligence ($SI$), and empathy ($E$).

### A. Building Blocks for Conflict Management: Trust, Social Intelligence, and Empathy

1) Trust: is the conviction prevailing among the parties that each one is benevolent to others. This conviction serves as the foundation for both establishing and retaining all relationships. Stated otherwise, no relationship can be established or retained in the absence of trust. Trust is a subject-dependent and dynamic phenomenon [13], and cognition and affection are the two building blocks that construct trust [32] and understanding [33] (See Figure 1(b)). Cognition is the ability of human being to observe the environment, learn from it, and make decisions. Cognition depends on the following factors:

- Memory which captures the success or failure of past interactions and relationships
- The perceived value and status of the relationship
- The qualities possessed by the parties

$T$ is a dynamic and persistent variable which increments through the gathering of cognition [33]-[34]. Affection is induced and nurtured by both the cadence and process of performing interactions. Higher levels of trust can be achieved if cognition and affection are present throughout interactions. Nevertheless, stable and intense presence of affection may invalidate the need for cognition.

2) Social Intelligence: is another major trait that helps individuals cooperate with each other to improve relationships [36]- [37]. The presence of social intelligence facilitates better adaptation through a favorable mindset that promotes emotional sharing and collaborative behavior. $SI$ primarily depends on three grounds, such as emotional intelligence, behavioral style, and mindset:

- Emotional intelligence is the capability to improve the relationship skills by strengthening interactions between the parties
- Behavioral style is the characteristic to improve on personal skills to strengthen one self optimally in accordance to others
- Mindset signifies the ability for self-adapting and self-evolving, learning from previous interactions

In practical situations, emotional intelligence can be built by affection, mindset can be molded through cognition, and behavioral style can be improved by self learning and adaptation.

3) Empathy: is another major building block that shapes relationships [38]. It is the disposition to identify and understand others' emotions and feelings and to consider their needs and concerns. It comprises of two aspects, such as cognitive empathy and emotional empathy, where cognitive empathy signifies mental disposition and emotional empathy denotes emotional sharing. The two aspects complement each other and are indispensable for perfect amalgamation of collective existence. The cognitive aspect makes individuals capable to understand each other and emotional empathy promotes social bonding among parties. In practical situations, cognitive empathy and emotional empathy can be correlated to trust cognition and trust affection respectively.

The optimum relationship appears when there exists a perfect congregation of $T; SI;$ and $E$ as shown in Figure 1(b). The following equation represents the relationship between the three factors and conflict management (CM).

$$CM = f(T; SI; E) \quad (1)$$

For the present work, we have given equal importance to all the three factors. This is further explained in Section III (b). As elucidated in the previous section there are five strategies, such as to fight, avoid, accommodate, collaborate, and negotiate, for conflict resolution which is intended to maintain collaboration, harmony, and happiness among all parties. We can make the below hypothesis for all these states:

- Fight: In this state, since the parties compete with each other, there is negligible affection, behavioral style, and cognition.
- Avoiding: Avoiding is a state when the parties change their behavioral style. Here, improving $SI$ for a team is easier as compared to $T$ and $E$. $SI$ of a team depends on affection, mindset, and behavior existing within the network. Among the three elements, behavior of the team plays significant role in improving $SI$. Behavior does not affect the $T$ and the $E$ of the team, and therefore all the resources from behavior are useful for $SI$.
- Accommodating: This is a state which manifests efforts from the parties to accommodate each other. Here, it is

This article has been accepted for publication in IEEE Transactions on Cognitive and Developmental Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCDS.2023.3279817



comparatively easier for the network to gain $T$. Cognition helps in understanding others' emotional values. In this scenario, any conflict that might occur among the units of the network that were working together successfully can be mitigated by regaining and leveraging $T$.

Negotiation: Negotiation is a scenario in which the parties attempt to compromise so that the optimal level of cooperation is achieved.

Collaboration: Collaboration is the state which the parties the aim for working together to achieve a goal.

## B. Reinforcement Learning

People learn the environment by making decisions and receiving feedback, a process characterized by $RL$ models in which agents learn by taking actions which leads to positive consequences. In nonsocial circumstances, reinforcement models provide persuasive descriptions of feedback-based learning, but social interactions frequently involve inferences of others' attribute traits, which may be independent of their reward value. $RL$ is a field of machine learning which studies how rational agents take actions in any given environment with an intention to maximize cumulative reward. Q-learning algorithms are a class of methods [39] where we determine a value of an action in a particular state. In other words, the goal is to learn optimal Q-values for each state and action pairs. Value function over states defines how an agent goes through the entire action space (which contains all the possible list of actions) using the below equation:

$$V(s) = \max_a \left( R(s; a) + \sum_{s'} P(s; a; s') V(s') \right) \quad (2)$$

where, $s$ = specific state from the entire state space (conflict)

$a$ = specific action from the action space which enables moving between different states

$s'$ = state to which the parties transition to, starting from state $s$

= discount factor

$R(s; a)$ = a reward as a function of state $s$ and action $a$ which coule be linear piecewise, continuous, or discontinuous [42].

$P(s; a; s')$ - the probability of moving from state $s$ to state $s'$ with action $a$

$\sum_{s'} P(s; a; s') U(s')$- expectation of the situation that the parties incurs randomness.

$V(s)$- Value of being in a particular state.

The Q-function takes the optimal action $a$ in state $s$ and is defined as the expected return starting from $s$, taking the action $a$ defined below.

$$Q_t(s; a) = Q_{t-1}(s; a) + [R(s; a) + Q_{max}(s'; a') \quad Q_{t-1}(s; a)] \quad (3)$$

Q-values are also known as the action state values which proves as an estimation of how good it is to take a particular action at some specific state, hence it defines the state action pair. Q-value is thus used iteratively to improve the learning of an agent. Kindly note while the Q-value does indeed talk about how attractive a particular state action pair is, it does not guarantee the agent will take decision on the basis of the q-value as it may try to explore instead. In our experimentation, the states are chiefly the values of Empathy, Trust and Social Intelligence which is grouped with the possible actions that one can take (i.e. "Fight", "Avoiding", "Negotiation", "Accommodate", "Collaboration").

In this algorithm, predicted state $\hat{S}(t)$ is defined as

$$\hat{S}(t) = [T(t); SI(t); E(t)] \quad (4)$$

$T$, $SI$ and $E$ parameters are defined as trust, social intelligence and empathy which is described in next subsection. With given $T$, $SI$ and $E$ parameters we seek to locate the action which gives us the maximum reward. The actions which gives us the maximum reward with the particular value of these parameters is thus chosen and is cached by the algorithm for the later use in calculating the Q values.

## C. Unification of All the Above Findings

The mapping of trust ($T$), social intelligence ($SI$), and empathy ($E$) to $CV$ use case is as follows.

$T$ is analogous directly to the confidence that vehicles have in each other in knowing that they will be supported by the other vehicles in the network. This can be facilitated by a long history (memory) of honest (correct) information sharing between vehicles [33].

$SI$ is analogous to accepting published driving rules and guidelines to decision making (or actions). It is similar to the human cognitive system which completes the perception-action cycle of decision making through the diverse processes of social learning [43]-[44].

$E$ is analogous to fair and ethical decision making in connected networks. Authors in [38] have described empathy as the ability of someone to change their opinion of a node based on what other nodes think of that node.

In the proposed algorithm, ($T$) is primarily based on the opinion consensus established between the pair of agents as described in [33]-[34], by taking into account the factors of benevolence ($B$) and integrity ($I$). The algorithm keeps track of the previous consensus between the pair of agents and therefore at every iteration a benevolence parameter is calculated which is dependent on the mean of the previous consensus ($X'$) and current distance between the pair of agents ($x$).

$$B_t = e^{(x \quad X')} \quad (5)$$

If the benevolence parameter is below a certain threshold positive value is incremented between the agents that affects the integrity factor. We define positive outcome using the inter-vehicle distance model, also called headway, as described in [46].

$$p_t = p_t + 1 \text{ if } \quad < \qquad n_t = n_t + 1 \text{ otherwise} \quad (6)$$

where $= \max_{i;j \in N}(d_{ij})$; $i = j$ and $d_{ij}$ is the distance value between the $i^{th}$ and $j^{th}$ objects. is the threshold value chosen for to define safe driving, where $> 0$. $= 0$ implies a fatal crash, state fight. Later integrity is defined as:

$$I_t = \frac{p_t}{p_t + n_t} \quad (7)$$

The final trust value is calculated by taking the weighted mean of (B) and (I) factors where $w_1$ and $w_1$ is equal to 0.5 .

$$T(t) = w_1\ B{=}5 + w_2\ I \tag{8}$$

B is divided by 5, so that the value does not become not a number. Meanwhile the benevolence parameter calculated is added to the consensus list maintained by the agent and thus changing the benevolence parameter which is used to update in the future iterations.

Factor (E(t)) is formulated as described in [38] by assuming the pair of agent where one agent has some cost defined namely \b". This particular agent if it agrees to cooperate with the other agent it will pay an altruistic cost namely \c" from \b", while the other agent will get the benefit of \b". If however the first agent refuses to cooperate, the first agent will pay no cost while the other agent will get no benefit. However, being a cooperative environment it is possible the cooperative act is erroneously executed with an error value denoted by \E1". Similarly the observer in mishap might assign a bad reputation instead of good one with error \E2". According to [38], if the initial level of empathy observed exceeds a given E SH value the population will tend to move towards the complete empathy which is given by parameter E1. Similarly, if the initial empathy observed is below the E SH value, the population on the contrary will evolve towards a lack of empathetic situation with the empathy level of 0. The authors hypothesis the value of cooperation which here is denoted by value \b" and \c", directly influence the population empathy as higher the cooperation lesser would be the E SH value and hence higher the chances of population empathy.

SI(t), the third variable used in the simulation, is calculated by taking the average of the memory, perception and language parameters as described in [40] and is passed through the sigmoid function to constraint the output within a range of 0 and 1.

$$SI(t) = \frac{1}{1 + e^{(P_t + M_t + L_t)}} \tag{9}$$

The equation is motivated by the fact that SI is directly proportional to each of the individualistic parameters namely perception (P), memory (M), and language(L) [45]. The principal activities accomplished for CV are as follows:

- Informed object identification via perception using cameras. This also includes dynamic (time-variant) representation of the environment via perception (using radars, lidars)
- Language is the sharing of vehicle information using basic safety messages among other vehicles.
- Retrieving prior experience using information stored in memory. Making intelligent decisions employing historical information and learning algorithms to adapt to new scenarios.

Based on the above learning through T; E and SI, conflict state is defined. Figure 2 shows the mapping of states and actions with respect to T; E and SI. We say when vehicle has high T; E and SI values, it is in collaborate state and the reward is given as 1. Algorithm 1 describes the model parameters and the state/action relationship:

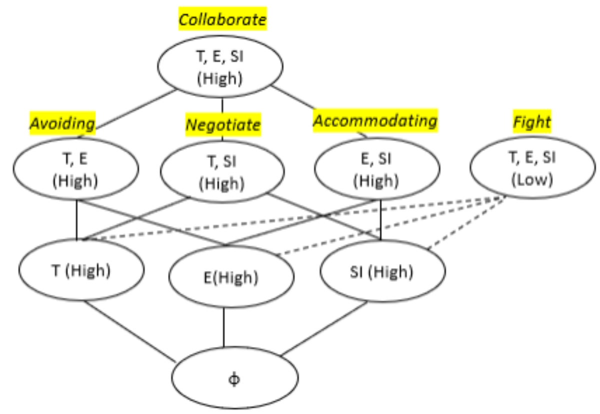



Fig. 2: Mapping of states and actions

**Algorithm 1:** Model Parameters and State/Action Relationship

```
Input: States of the two parties, Th= threshold set as 0.5
Output: State and Action
Data: Particular state of Trust (T), Social Intelligence (SI) and Empathy (E)
while true do
    if T, SI, & E is < T and at = Fight then
        then Rt = -1
    else if T, E is > Th, and SI is < Th and at = Avoiding then
        then Rt = 0.5
    else if E and SI is > Th, and T is < Th and at = Accommodating then
        then Rt = 0.5
    else if T, SI is > Th, and E is < Th at = Negotiate then
        then Rt = 0.5
    else if T, SI, E are > Th and at = Collaborate then
        then Rt = +1
    else if T, SI, E are > Th and at = Avoiding or Negotiate or Accommodate then
        then Rt = 0.75
    else Rt = 0
```

## IV. Social Psychology Inspired Reinforcement Learning for Conflict Management in Connected Vehicles

In this section, we apply the proposed algorithm to CV network, wherein vehicles may use information received from their own sensors and messages received from other vehicles. Additionally, vehicles may also rely on their prior experience,

This article has been accepted for publication in IEEE Transactions on Cognitive and Developmental Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCDS.2023.3279817



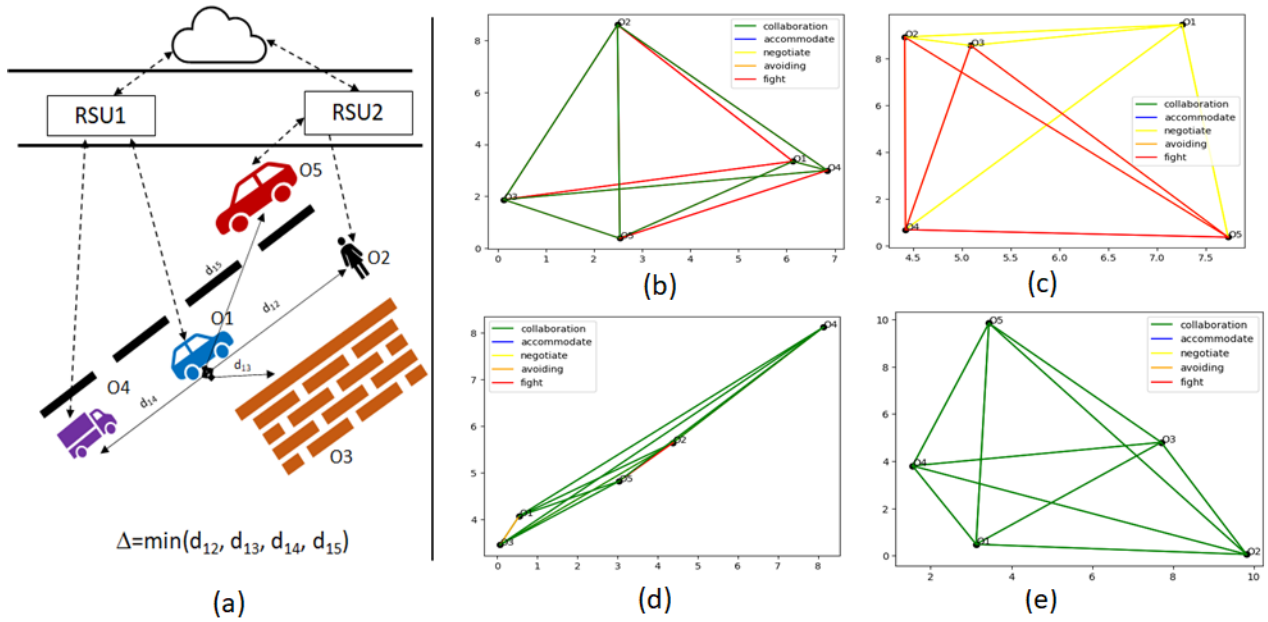



Fig. 3: (a) Environment containing a car (Object1) with multiple passengers, a person (Object2) who suddenly decides to cross the road, a brick wall (Object3) to the right of Object1, a high-speed truck (Object4) behind Object1 and a car (Object5) in the opposite lane, (b) Only E, (c) Only SI, (d) Only T, (d) All three factors

using historical data stored in the form of transactions in memory.

## A. Analytical Description

In this section, the conflict management model introduced in the previous section is incorporated into a computer algorithm that identifies optimal levels of the factors for facilitating conflicts management. In Figure 3(a), the environment is described in terms of objects: pedestrians, vehicles, and structures. The algorithm uses sensors in the car, such as cameras, which perform the act of seeing as done in human driven vehicles; thereby providing static (time-invariant) information of the object such as color and size. It then retrieves apriori information stored in the memory from previous interactions with similar objects to make decisions about objects being seen currently. The decision making process in the algorithm is based on object and environment detection using current and historical information and is built on an inference-level hierarchy [41]. The observation matrix $O(t)$, representing color, size, position, and speed, is defined as:

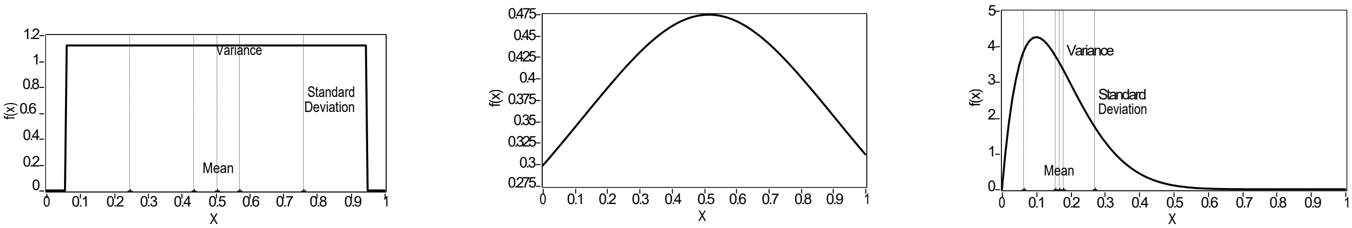
$$O(t) = \begin{bmatrix} \text{"Blue"} & \text{"Medium"} & X_1(t) & V_1(t) \\ \text{"NA"} & \text{"Medium"} & X_2(t) & V_2(t) \\ \text{"Brick"} & \text{"Large"} & X_3(t) & 0 \\ \text{"Purple"} & \text{"Large"} & X_4(t) & V_4(t) \\ \text{"Red"} & \text{"Large"} & X_5(t) & V_5(t) \end{bmatrix}$$

wherein the first row refers to the agent (O1), the second row refers to the pedestrian (O2), the third row refers to wall (O3), the fourth row refers to the truck in same lane (O4) and the fifth row refers to the car in the opposite lane (O5). While color and size are considered to be constant (time-invariant), position and speed are considered as varying temporally. Note that, while position is generally defined using values in the $X$ and $Y$ coordinates, for the sake of notational simplicity, in this paper we define position only with one coordinate $X(t)$.

Let us first consider the case, denoted as Case1, when the agent makes a decision with no factors taken into account. With a static snapshot of the environment, the vehicle does not have sufficient information to prefer one action over the other and will be going straight. Here, the probability distribution function for the action space is uniform as seen in Figure 4(a).

Next, we consider the case, denoted as Case2, when the agent makes a decision exclusively based on either $T$ (See Figure 3(b)) or $SI$ (See Figure 3(c)) or $E$ (See Figure 3(d)). Agent recalls from memory that red cars generally have a history of driving fast. Agent makes a judgement that the red car is not going to slow down and that the pedestrian is not going to back off. The agent makes this decision with an additional assumption that the fast moving truck behind the agent can see the pedestrian crossing the road. When $T$ is present between parties in conflict, it can lead to collaboration because parties may be more willing to work together and share information or can lead to fight, especially if there are underlying power dynamics or if trust is violated in some way. In the case of $SI$, it can lead to negotiation or fight as the possible state depending on how it is applied. Similarly, $E$ can lead to collaboration, avoiding, or fight as possible states because it helps parties understand each other's perspectives. In either case, the probability distribution function is shown in Figure 4(b).

In the proposed work algorithm, denoted as Case3, the agent makes decision based on $T$, $SI$ and $E$. Social intelligence




(a) Uniform Probability Distribution Function (PDF): When none of the factors are taken into account, all actions have equal probability.

(b) Cauchy PDF: When either $T$, $SI$ or $E$ are integrated into predicting state at next time epoch, PDF indicates a higher likelihood of reducing losses if the agent keeps moving straight.

(c) Beta PDF: when $T; SI;$ and $E$ are integrated into predicting state at next time epoch, PDF indicates a higher likelihood of maximizing reward if agents swerves to the left

Fig. 4: $X = 0.5$ denotes the position of the vehicle driving straight. $X > 0.5$ indicates vehicle swerving to the right. $X < 0.5$ indicates vehicle swerving to the left.

allows the agent to communicate with other agents using a commonly understood standard syntax. The agent learns that the red car is actually slowing down, whereas the truck behind the agent is not slowing down or changing direction. In this situation, the agent decides that swerving to the left is the best option that is most likely to yield positive rewards leading to collaborative state (See Figure 3(e)). The probability distribution function of the reward signal is a beta distribution function, as shown in Figure 4(c).

### B. Performance Analysis

The experiment was further conducted in a 3X3 grid world environment with two vehicles. The two vehicles face the opposite direction along the same column of the grid world such that two of them face each other. As the experiment proceeds both the vehicles will travel opposite to each other until there is no square separation between the cars. The vehicle will then take the action to minimize the conflict between them too. The experiments ran for specific loops, each of the loops consist of 100 steps taken. There is also an obstacle situated at the periphery of the grid world which is at (1,0). Following test cases were formulated:

1) Scenario 1: When only one factor is taken into account:

$T$ only: Figure 5(a) illustrates the observation when we controlled the experiment by keeping the $T$ parameter high. The other two parameters were kept low. Here, it is observed that the agent has a stronger preference for a accommodate than any of the other actions. Thus, accommodate action dominates in this case while for the secondary preferences all other action was likely be preferred by the agent.

$E$ only: Figure 5(b) illustrates the observation when we controlled the experiment by keeping the $E$ parameters high. Here, also the agent exhibits a greater inclination towards the "accommodate" action compared to any of the other available actions. The agent is likely to have a preference for the remaining actions as secondary choices.

$SI$ only: Figure 5(c) illustrates the observation when we controlled the experiment by keeping the $SI$ parameter high. Here we observe that the agent prefers negotiate followed by accommodate for the entire simulation.

2) Scenario 2: When two factors are taken into account:

High E, High SI, Low T: Figure 6(a) illustrates the observation when we controlled the experiment by keeping the $T$ parameter low. Here we see the agents tend to prefer accommodate action more than any of the other actions. While for the secondary preferences any of the other four actions are equally like to be preferred by agents. However, it can be seen that the Q-values are higher than the Scenario 1 simulations.

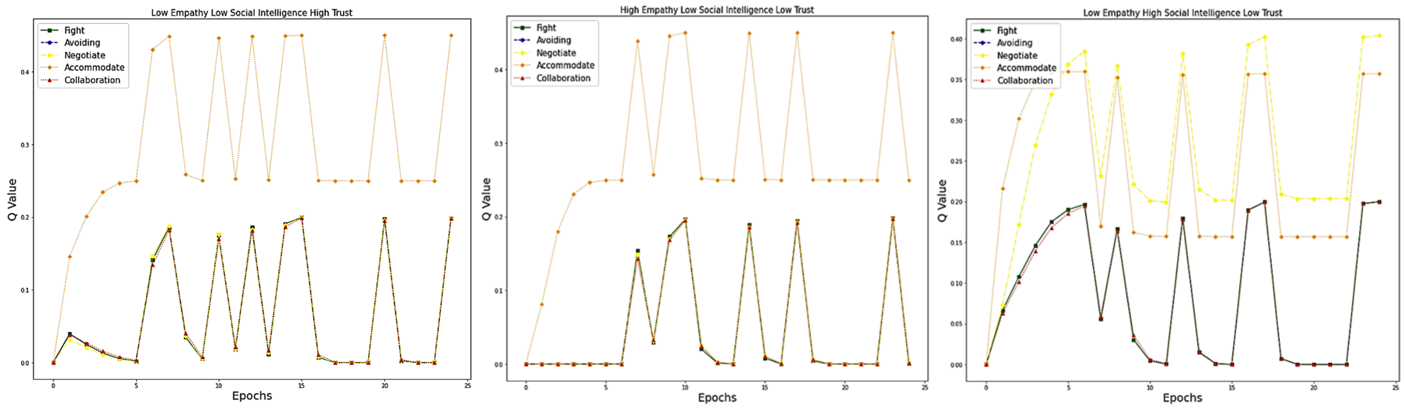
High E, Low SI, High T: Figure 6(b) illustrates the observation when we controlled the experiment by keeping the $SI$ parameter low. While it was difficult initially to discern the actions that the agents prefer, it is observed that the agent starts to prefer negotiate followed by accommodate more than any of the actions by the end of the simulation interval.

Low E, High SI, High T: Figure 6(c) illustrates the observation when we controlled the experiment by keeping the $E$ parameter low. We observed that the accommodate is preferred in every iteration.

3) Scenario 3: When all three are taken into account:

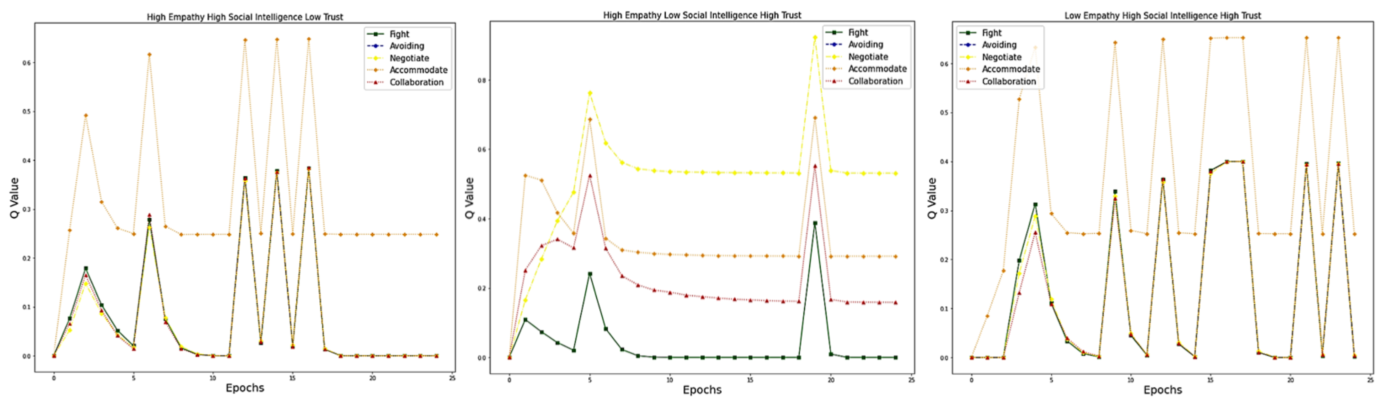
Low E, Low SI, Low T: Figure 7(a) illustrates the observation when we set low values for trust, social intelligence, and empathy and let the algorithm run without any constraints. We observe that as the simulation runs, all the Q-values are below zero and fight is preferred over all other action values.

High E, High SI, High T: Figure 7(a) illustrates the observation when we set high values for $T; SI$ and $E$ and let the algorithm run without any constraints. We observe that as the simulation runs for more epochs, the agents tend to prefer collaboration over the other actions. This is consistent with the above scenarios where the individual parameter of $T$, $SI$, and $E$ were high. To the contrary, if we set any of the variables to low, then we observe undesirable actions such as the tendency to accommodate and no strict preference for collaboration. Thus we conclude that all the variables should necessarily



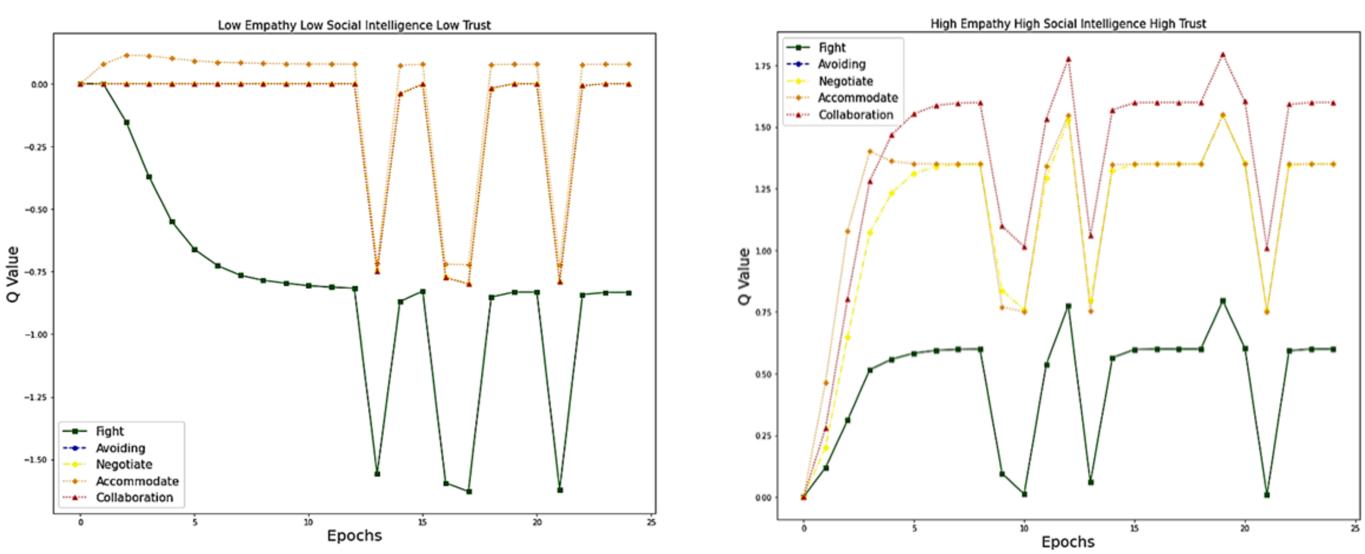

(a) High T only (b) High E only (c) High SI only

Fig. 5: Plots for Q values vs epochs per cycle by taking one factor into account




(a) High E, High SI, Low T (b) High E, Low SI, High T (c) Low E, High SI, High T

Fig. 6: Plots for Q values vs epochs per cycle by taking two factors into account




(a) Low T, Low SI and Low E (b) High T, High SI and High E

Fig. 7: Plots for Q values vs epochs per cycle by taking either none or three factors together

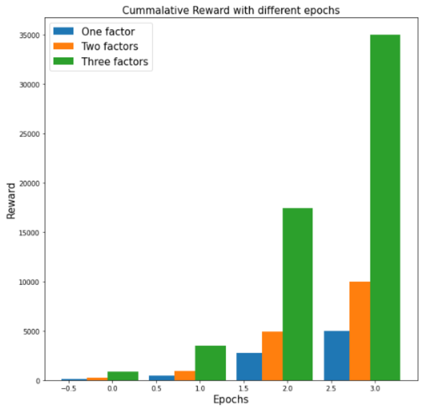



Fig. 8: Cummulative reward with different epoch values when only one, two or all the three factors are taken into account

be included and enhanced for for cooperative decision making and conflict management.

### C. Comparative Analysis

Table I presents the comparative analysis of the three factors. As seen from the table, when all the three factors are considered the Q-values is maximum and the agent tends to select collaborate (the optimum choice under conflict). Furthermore, when none of the factors are taken into account, the q-values are below 0 with the fight state being favoured. Another thing we can observe from the results is that when only one or two factors are considered, the agent tends to select accommodate action with the highest Q-values being 0.4 and 0.8 respectively. Most of the state of the art algorithms are based on one factor that can help in resolving immediate conflict, but are not favoured in the long run. Collaborating is seen as a cooperative approach to conflict resolution because it leads to more sustainable and equitable solutions [16].

TABLE I: Comparative analysis

| No. | Empathy | Social Intelli-gence | Trust | Highest Q values | Conflict State |
|---|---|---|---|---|---|
| 1 | Low | Low | Low | < 0 | Fight |
| 2 | Low | Low | High | 0.45 | Accommodate |
| 3 | Low | High | Low | 0.4 | Negotiate |
| 4 | High | Low | Low | 0.45 | Accommodate |
| 5 | High | Low | High | 0.84 | Negotiate |
| 6 | High | High | Low | 0.65 | Accommodate |
| 7 | Low | High | High | 0.7 | Accommodate |
| 8 | High | High | High | 1.75 | Collaborate |

We also ran the simulations to understand the significance of rewards for different number of epoch values. Figure 8 illustrates how rewards change for various numbers of epochs depending on whether one, two, or all three parameters are taken into account. It is evident from the results that the three factors are 118.18% and 70.27% more effective than only one factor and two factors respectively. Based on these results, we summarize the key advantages of the proposed work as follows:

- A baseline framework for future conflict management to take trust, social intelligence, and empathy principles in resolving conflicts is proposed.
- This work also identifies five styles of conflict resolution: fight, collaborating, negotiating, avoiding, and accommodating based on the state of the agent.
- It shows how the social psychology components influence individuals' choice of conflict style and can help them find a mutually beneficial solution.

## V. Conclusion and Future Work

We have presented a RL based technique for resolving conflicts among multiple entities. The proposed algorithm is based on agents participating in actions by virtue of observations from the environment and attempting to reach the goal of getting a high reward in return. Importantly, a triad of trust, empathy and social intelligence is integrated in RL framework, wherein the environment is specified as the parties under conflict having different states of fight, avoidance, accommodation, negotiation, and collaboration. Results discussed in the paper demonstrate how the number of iterations needed to reach ideal state varies when each of factors are changed. Simulation results show that the proposed method is 118.18% more efficient than the trust only based models in reaching the optimal reward state. Our simulation technique can be further extended to more scenarios and experiments under different contextual demands. The current experimentation involved a couple of agents in the grid world simulated environment. The future scope of the project will include multiple agents in the grid world environments. We will also be adding weights to the each of the factors for future experiments to signify the importance of each. Additionally, static obstacles will also be incorporated. The motivation here is to explore the characteristics of the agents where the environment scope is quite large with more static and dynamic objects.

## Acknowledgment

This publication has been supported by NSF CISE Research Initiation Initiative (CRII) grant #2153510 and #2313351

## References

[1] S.A. Nan, "Conflict resolution in a network society". International Negotiation, vol. 13, no. 1, pp.111-131, 2008.
[2] J. Galtung, "Peace by peaceful means: Peace and conflict, development and civilization", vol. 14, Sage, 1996.
[3] X. Xu, "The Belt and Road, global governance, and China". In The China-led Belt and Road Initiative and its reflections, pp. 265-282. Routledge, 2022.
[4] F. Ettensperger, "Comparing supervised learning algorithms and artificial neural networks for conflict prediction: performance and applicability of deep learning in the field.", Qual Quant, vol. 54, pp. 567–601, 2020.
[5] K. Sigmund, "The calculus of selfishness". Princeton University Press, 2010.
[6] S.J. McGuire et al., "Dyad integrative behavior and negotiation outcomes". Conflict Resolution Quarterly, vol. 39, no. 3, pp.295-314, 2022.
[7] C. Liu, C. -W. Lin, S. Shiraishi and M. Tomizuka, "Distributed Conflict Resolution for Connected Autonomous Vehicles," in IEEE Transactions on Intelligent Vehicles, vol. 3, no. 1, pp. 18-29, 2018.
[8] J. Radak, B. Ducourthial, V. Cherfaoui and S. Bonnet, "Detecting Road Events Using Distributed Data Fusion: Experimental Evaluation for the Icy Roads Case," in IEEE Transactions on Intelligent Transportation Systems, vol. 17, no. 1, pp. 184-194, 2016.

[9] P.V. Prenzel, and F. Vanclay, "How social impact assessment can contribute to conflict management". Environmental Impact Assessment Review, 45, pp.30-37, 2014.
[10] M. Dibaei et al., "Attacks and defences on intelligent connected vehicles: A survey". Digital Communications and Networks, 6(4), pp.399-421, 2020.
[11] Z. Tian, X. Gao, S. Su and J. Qiu, "Vcash: A Novel Reputation Framework for Identifying Denial of Traffic Service in Internet of Connected Vehicles," in IEEE Internet of Things Journal, vol. 7, no. 5, pp. 3901-3909, 2020.
[12] Merriam-webster, http://www.merriam-webster.com/dictionary/trust; accessed on August 2021.
[13] F. Flores, and R.C. Solomon, "Creating trust", Business Ethics Quarterly, vol. 8, no. 2, pp. 205-232, 1998.
[14] N. Masuda, and F. C. Santos, "Social psychology: a mathematical look at empathy". Elife, 8, p.e47036, 2019.
[15] L.M. Hackel, P. Mende-Siedlecki, and D.M. Amodio, "Reinforcement learning in social interaction: The distinguishing role of trait inference". Journal of Experimental Social Psychology, 88, p.103948, 2020.
[16] K.W. Thomas, "Thomas-kilmann conflict mode". TKI Profile and Interpretive Report, vol. 1, no. 11, 2008.
[17] http://www.skillsyouneed.com/ips/conflict-resolution.html, accessed on August 2021.
[18] M. Humphreys, "Natural resources, conflict, and conflict resolution uncovering the mechanisms". Journal of conflict resolution, vol. 49, no. 4, pp.508-537, 2005.
[19] M.J.M.C. Luis, "A assembleia de turma como dispositivo de mediacao e conictos". Master's thesis. Universidade de Lisboa-Faculdade de Psicologia e Ciencias da Comunicacao, 2009.
[20] Y.G. Cheong et al., "A computational approach towards conflict resolution for serious games". In Proc, International Conference on Foundations of Digital Games, no. 6, pp. 15-22, 2011.
[21] K. Floyd, "Interpersonal communication: the whole story", 2009.
[22] M. Hogan, "Learning Together and the Challenge of Collaboration", http://www.psychologytoday.com/blog/in-one-lifespan/201309/learning-together-and-the-challenge-collaboration; 2 December, 2013.
[23] P. Hang, C. Huang, Z. Hu, and C. Lv. "Driving Conflict Resolution of Autonomous Vehicles at Unsignalized Intersections: A Differential Game Approach." IEEE/ASME Transactions on Mechatronics, 2022.
[24] J. Guo, I. Harmati, "Evaluating semi-cooperative Nash/Stackelberg Q-learning for traffic routes plan in a single intersection", Control Engineering Practice, vol. 102, 2020.
[25] B. Xu et al., "Distributed conflict-free cooperation for multiple connected vehicles at unsignalized intersections". Transportation Research Part C: Emerging Technologies, no. 93, pp.322-334, 2018.
[26] R. S. Sutton and A. G. Barto, "Reinforcement learning: An introduction". MIT press, 2018.
[27] D. Chaki, A. Bouguettaya, and S. Mistry,"A conflict detection framework for iot services in multi-resident smart homes". In 2020 IEEE International Conference on Web Services (ICWS), pp. 224-231, 2020.
[28] A. Mahmood et al., "Trust on wheels: Towards secure and resource efficient IoV networks". Computing 104, 1337–1358, 2022.
[29] T. Rosenstatter and C. Englund, "Modelling the Level of Trust in a Cooperative Automated Vehicle Control System," in IEEE Transactions on Intelligent Transportation Systems, vol. 19, no. 4, pp. 1237-1247, 2018.
[30] H. Rathore et al.,"TangleCV: Decentralized Technique for Secure Message Sharing in Connected Vehicles", AutoSec Workshop in conjuction with ACM Conference on Data and Application Security and Privacy (CODASPY), 2019.
[31] H. Rathore, A. Samant, and M. Jadliwala, "TangleCV: a distributed ledger technique for secure message sharing in connected vehicles". ACM Transactions on Cyber-Physical Systems, vol. 5, no. 1, pp.1-25, 2020.
[32] B. McEvily, V. Perrone, and A. Zaheer, "Trust as an organizing principle". Organization science, vol. 14, no. 1, pp. 91-103, 2003.
[33] H. Rathore, V. Badarla, and K.J. George, "Sociopsychological trust model for Wireless Sensor Networks". Journal of Network and Computer Applications, 2015.
[34] H. Rathore, V. Badarla, and S. Shit, "Consensus Aware Sociopsychological Trust Model for Wireless Sensor Network". ACM Transactions on Sensor Networks, 2016.
[35] H. Rathore, S. Sai, A. Gundewar, "Social Psychology Inspired Distributed Ledger Technique for Anomaly Detection in Connected Vehicles", IEEE Transactions on Intelligent Transportation Systems, 2023.
[36] S. Moradpour, B. Taheri, and A.H. Pasha,"Investigating the Relationship between Emotional and Social Intelligence with Mathematical Performance". International Journal of Mental Health and Addiction, vol. 15, no.5, pp.1044-1046, 2017.
[37] D. Goleman, and R. Boyatzis, "Social intelligence and the biology of leadership". Harvard business review, vol. 86, no. 9, pp.74-81, 2008.
[38] A.L. Radzvilavicius, A.J. Stewart, and J.B. Plotkin, "Evolution of empathetic moral evaluation". Elife, no. 8, p.e44269, 2019.
[39] R. S. Sutton, D. McAllester, S. Singh, and Y. Mansour, "Policy gradient methods for reinforcement learning with function approximation," Advances in neural information processing systems, vol. 12, pp. 1057–1063, 1999.
[40] H. Rathore and V. Bhaduria,"Intelligent decision making in autonomous vehicles using cognition aided reinforcement learning", IEEE Wireless Communications and Networking Conference (WCNC), pp.524-529, 2022.
[41] D.L. Hall, and J. Llinas, "An introduction to multisensor data fusion. Proceedings of the IEEE", vol. 85, no. 1, pp.6-23, 1997.
[42] A. Mukherjee, "A Comparison of Reward Functions in Q-Learning Applied to a Cart Position Problem".arXiv preprint arXiv:2105.11617, 2021.
[43] J.M. Fuster, "The cognit: a network model of cortical representation". International Journal of Psychophysiology, vol. 60, no. 2, pp.125-132, 2006.
[44] D. Graves, K. Rezaee, and S. Scheideman, S., "Perception as prediction using general value functions in autonomous driving applications". In 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 1202-1209, 2019.
[45] W.S. Lee, and M. Selart, "When emotional intelligence affects peoples' perception of trustworthiness". The Open Psychology Journal, 8, pp.160-170, 2015.
[46] S. Moon, I. Moon, and K. Yi, "Design, tuning, and evaluation of a full-range adaptive cruise control system with collision avoidance". Control Engineering Practice, vol. 17, no. 4, pp.442-455, 2009.
[47] Z. Liu et al., "LPPTE: A lightweight privacy-preserving trust evaluation scheme for facilitating distributed data fusion in cooperative vehicular safety applications". Information Fusion, 73, pp.144-156, 2021.
[48] J. Guo et al. "TROVE: A context-awareness trust model for VANETs using reinforcement learning". IEEE Internet of Things Journal, vol. 7, no. 7, pp.6647-6662, 2020.

Heena Rathore is an Assistant Professor at Texas State University. Academically, she has worked as an Assistant Professor of Instruction at The University of Texas at San Antonio and as a Visiting Assistant Professor at Texas A&M University, Texarkana. She also worked as a Data Scientist and Program Manager at Hiller Measurements in Austin, Texas. Prior to that, she worked as a postdoctoral researcher for the US-Qatar Joint Collaborative Project between Temple University, the University of Idaho, and Qatar University. Her research interests include cognitive AI, biologically inspired systems, machine learning, cybersecurity.

Yash Kumar Singhal Yash Kumar Singhal is a doctoral candidate in Artificial Intelligence at Indian Institute of Technology Roorkee. He received his bachelor's degree in Computer Science and Engineering at Jaypee Institute of Information Technology Noida. His research interest includes cognitive science, data mining and machine learning.

George Kodimattam Joseph George Kodimattam Joseph is an Associate Professor at School of Liberal Arts at Indian Institute of Technology, Jodhpur. He holds a PhD. in Ethics from IIT Kanpur in 2011. His research focuses on Applied Ethics, Philosophy of Life Sciences, Sociopolitical Ethics, Sustainability Studies, Biomedical Ethics and Ethics and Technology.