

Multi-agent safety invariant set.

Huaiyuan Teng

Background

There are many smart agents $\{A_1, A_2, A_3 \dots A_n\}$ living in a shared space. Each agent is capable of controlling its own behavior by:

$$\dot{x}_i = f_i(x_i, u_i) \quad (1.1)$$

For easier representation, also define its integration:

$$F_i(x, u, t) \triangleq \text{state after controlling } x \text{ using } u \text{ for time } t \quad (1.2)$$

For each two of them, certain pair of state is dangerous, known as crush. We use the following Boolean function to denote that:

$$Crush_i^j(x_i, x_j) \quad (1.3)$$

Agents are not capable of communicating with each other. However, they can observe the state of any agent in the system.

Given such a system, this article is proposing a strategy for those agents to peacefully live in the space, by applying control supervisor based on safe invariant set.

For each agent, its control u_i is limited by supervision set R_i :

$$u_i \in R_i(x_1, x_2, \dots x_n) \quad (1.4)$$

In the following pages, the word "regulation" will be used to represent such supervision set.

Principle 1. Non-correlated regulation (NCR)

The author thinks that one agent is never capable of considering all the other agents in a correlated manner at real time. The computational power of a single agent is limited, such that it can only regulate its behavior using following form:

$$R_i(x_1, x_2, \dots x_n) = \cap_{j=0, j \neq i}^n R_i^j(x_i, x_j) \quad (2.1)$$

First, for each other agent A_j , A_i works out a two-body regulation R_i^j considering only the states of A_j and A_i . Then A_i chooses its final regulation to be the intersection of all such two body regulations.

Principle 2. Basic agent right.

There can be many different results of R_i^j considering the abundance of the agents. No matter what they are, the intersection of them in eq (2.1) cannot be empty, or otherwise there would be no choice for A_i to make.

In fact, each other agent may cast certain regulation upon A_i , but they are not clever enough to think in A_i 's perspective. To achieve the non-emptiness of A_i 's final regulation, one solution is to make the contract of "basic agent right":

For each agent A_i , there exist a control known as the "basic agent right", denoted as b_i . A_i shall be always allowed to take $u_i \in B_i(x_i)$ as its next control. On the other hand, A_i 's control needs to be regulated such that it won't harm the basic right of any other agents. (B_i can be estimated by all agents.)

In the later pages, we will show that the peaceful space can be achieved if every NCR agent observes the principle of basic agent right.

Notion of backup safety.

Given the fact that A_i 's control is regulated by certain strategy $B_i(x_i)$, there would be an obtainable forward reachable set:

$$FR_i(B_i, x_i, t) \triangleq \{x_f \mid A_i \text{ keep use } u_i \in B_i(x), \text{ and can reach } x_f \text{ at time } t\} \quad (3.1)$$

Here the strict definition is not given, since there would be a lot of words.

We further define the collision of two forward reachable set:

$$Crush_i^j(x_i, x_j, B_i, B_j) = \exists t \geq 0, s. t.$$

$$\exists x_i^c \in FR_i(B_i, x_i, t), x_j^c \in FR_j(B_j, x_j, t), Crush_i^j(x_i^c, x_j^c) \quad (3.2)$$

The straight forward understanding of eq (3.2) is: if A_i follow strategy of B_i and A_j follow strategy of B_j , whether it is possible for a future crush to happen.

For some two agents A_i, A_j in the n-agent system, by instinct one necessary condition for "eventually no crush between A_i and A_j for all possible trajectories starting from x_i, x_j " is "backup safety":

$$BS_j^i(x_j, x_i) = BS_i^j(x_i, x_j) \triangleq !Crush_i^j(x_i, x_j, B_i, B_j) \quad (3.3)$$

Instinct of necessity: since there are many other agents in the space, it is possible that A_i 's final regulation set get maximumly squeezed by other agents, such that $\forall t > 0$, A_i has nothing to do but to take the backup choice from B_i . The same goes for A_j . If such situation

happens (such situation can always happen no matter how the agents are regulated), eq (3.3) is equivalent to the definition of “eventually no crash between A_i and A_j , starting from x_i, x_j ”. (Elements in B are not allowed to be further regulated according to the description of “basic right”)

So the guidance of the design is that eq (3.3) must hold true for all selection of i, j to avoid the risk of collision.

Notion of moral bottom line

To avoid a complicated story, we further limit the dynamic to discrete time axis.

$$x_i(k+1) = F_i(x_i(k), u_i(k), dt) \quad (4.1)$$

As mention before, A_i, A_j must try to keep eq (5) all the time since from their perspective doing so is necessary for collision prevention.

Suppose $BS_i^j(x_i(k), x_j(k)) = \text{true}$ at moment k . To keep such property at $k+1$, A_j must satisfy:

$$R_j^i(x_j, x_i) \subseteq M_j^i(x_j, x_i) \triangleq \{u_j | \forall b_i \in B_i(x_i), BS_i^j(F_i(x_i, b_i, dt), F_j(x_j, u_j, dt))\} \quad (4.2)$$

From the principle 2 basic agent right, A_j must respect the basic right of A_i . It is possible for A_i to choose b_i at k , thus A_j must constraint its action in $M_j^i(x_j, x_i)$ so as to avoid potential backup safe regression. The constraint of M_j^i , known as A_j 's moral bottom line over A_i , must also be enforced to everyone in the space to rule out the risk of collision.

$$BS_i^j(x_i, x_j) \rightarrow M_j^i(x_j, x_i) \neq \emptyset \quad (4.3)$$

Proof of eq (4.3): by showing $B_j(x_j)$ is subset of $M_j^i(x_j, x_i)$.

First order freedom

Knowing that each agent is at least regulated by moral bottom line, there is one design of regulation set that can be sufficient for the purpose of peaceful living:

$$R_i^j(x_i, x_j) \triangleq \{u_i | \forall u_j \in M_j^i(x_j, x_i), BS_i^j(F_i(x_i, u_i, dt), F_j(x_j, u_j, dt))\} \quad (5.1)$$

It is not hard to prove that:

$$B_i(x_i) \subseteq R_i^j(x_i, x_j) \subseteq M_i^j(x_i, x_j) \quad (5.2)$$

So here such regulation R_i^j observes moral bottom line, while also protects A_i 's own basic right.

When A_i regulates its control $u_i \in R_i^j(x_i, x_j)$ defined by (5.1) and A_j observes (4.2), it can be proved that for any step k :

$$BS_i^j(x_i(k), x_j(k)) \rightarrow BS_i^j(x_i(k+1), x_j(k+1)) \quad (5.3)$$

The backup safe invariance between A_i and A_j is then achieved.

$R_i(x_1, x_2, \dots, x_n)$, as the intersection of R_i^j , is always respecting other's basic right, while protecting A_i 's own basic right, using (5.2), and can be proved to be not empty if $BS_i^j(x_i, x_j)$. Together with (5.3), the design is done.

Multiple possibilities of basic right

In practice, it is hard for one to tell the basic right of an unfamiliar agent. For A_i the choice of B_j varies, depending upon the background of A_j , or culture, ideology... However, the existence and unique-ness of B_j still stands. A_i may anticipate a list of possibilities for $B_j \in G_j^i = \{B_j^{i1}, B_j^{i2}, \dots, B_j^{im}\}$, while not knowing what exactly B_j is.

The moral bottom line becomes the intersection of enumerating through the other agent's basic right list:

$$M_j^i(B_2, B_1, x_j, x_i) \triangleq \{u_j | \forall b_i \in B_1(x_i), !Crush_i^j(F_i(x_i, b_i, dt), F_j(x_j, u_j, dt), B_1, B_2)\} \quad (6.1)$$

$$R_j^i(x_j, x_i) \subseteq M_j^i(x_j, x_i) = \cap_{B_1 \in G_j^i, (skip\ unsafe\ B_1, B_j)} M_j^i(B_j, B_1, x_j, x_i) \quad (6.2)$$

Some anticipation B_1 can lead to empty result of $M_j^i(B_j, B_1, x_j, x_i)$, which shall be skipped in the calculation of moral bottom line. (B_1 can be ruled out by assuming the current situation is backup safe).

With A_j observes (6.2), the action set of A_i is then defined as:

$$R_i^j(B, x_i, x_j) \triangleq \{u_i | \forall u_j \in M_j^i(B, B_i, x_j, x_i), !Crush_i^j(F_i(x_i, u_i, dt), F_j(x_j, u_j, dt), B_i, B)\} \quad (6.3)$$

$$R_i^j(x_i, x_j) \triangleq \cap_{B \in G_j^i, (skip\ unsafe\ B_i, B)} R_i^j(B, x_i, x_j) \quad (6.4)$$

It is not hard to show that basic right is protected:

$$B_i(x_i) \subseteq R_i^j(x_i, x_j)$$

For the observation of moral bottom line:

$$R_i^j(x_i, x_j) \subseteq M_i^j(x_i, x_j)$$

For each single enumeration of B , $R_i^j(B, x_i, x_j) \subseteq M_i^j(B_i, B, x_i, x_j)$ can be proved. The intersection operation will simply keep that property.

If $B_j \in G_j^i$ and $B_i \in G_i^j$, we have:

$$R_i^j(x_i, x_j) \subseteq R_i^j(B_j, x_i, x_j)$$

$$M_j^i(x_j, x_i) \subseteq M_j^i(B_j, B_i, x_j, x_i)$$

The safe invariance property in eq (5.3) can be proved.

Precedence

In the engineering practice of autonomous driving vehicle (ADV), one important feature is to observe the precedence according to the specific traffic condition.

More right to be bestowed on to certain special agents. Such action is used to maximize the overall benefit.

In the section of “multiple possibilities of basic right”, it can be seen that one agent's observation of another can be described as a list of basic rights. Among the list, some are ruled out according to current state situation (For e. g. not backup safe). In general, one vehicle's basic right shall be determined by local traffic rule.

A natural way is to require others to treat additional right as if it is basic right if possible.