# Cyber Physical Human Systems

Anuradha Annaswamy[a], Yildiray Yildiz[b]

[a]*Department of Mechanical Engineering, Massachusetts Institute of Technology, USA*
[b]*Department of Mechanical Engineering, Bilkent University, Turkey*

**Abstract**

Cyber Physical Human Systems (CPHS) is an emerging field where the human, the physical system and enabling cyber technologies are interconnected through complex interactions to accomplish a certain goal. This sharply contrasts with a conventional perspective where the human is treated as an isolated element who operates or uses the system. In this article, we provide a general introduction to CPHS and present representative examples from the recent literature. CPHS is a growing field and currently there are no agreed upon definitions about what constitutes a CPHS or not. Therefore, the article mainly reflects the authors' current understanding of the topic.

*Keywords:*
Cyber-physical human systems, Decision making, Human machine interaction

## 1. Introduction

Systems and control has always been a broad and abstract discipline. It has progressed, thrived, and flourished due to its ability to reinvent, reconfigure, and reorient its vistas of theoretical inquiries and technological advances. As several surveys, reports, and position papers (Murray, 2003; Samad and Annaswamy, 2011, 2014; Lamnabhi-Lagarrigue et al., 2017) have clearly articulated, the need for control systems has been reinforced through societal grand challenges in health and medicine, energy and climate, sustainability and development, and productivity and economic growth, to name a few. New frontiers are continuing to emerge, in application-centric directions such as Internet of Things and Industry 4.0, theoretical directions such as machine learning and data science, and conceptual directions such as Cyber Physical & Human Systems (CPHS). The focus of this article is on the last item, CPHS.

Significant advances in automation, communications, and computing have formed the foundation of the field of Cyber-Physical Systems (CPS). As societal challenges increase, there is an increasing need for human elements to form closer ties with CPS. The underlying interactions and integrations of CPS and human systems are varied, complex, highly challenging, and merit a systematic investigation, forming the underpinning of CPHS.

The common thread between automation and human systems has a long and rich history, and can perhaps be traced to the notion of Cybernetics, if not even earlier. As Wiener postulates in his magnum opus (Wiener, 1948), it is the scientific study of how humans, animals, and machines control and communicate with each other. The fundamental tenets of automation - steer, navigate, and govern - pervade intelligence, resilience, and several other high levels of functioning in an uncertain environment. These same tenets are attributes of a human system as well. One can perhaps view CPHS as a new incarnation of Cybernetics, with more fleshed out roles for humans to play in specific application areas such as ground transportation, aerial platforms, robotics, to name a few.

Let's start with an apt example of a CPHS, an automobile with varying levels of autonomy. Per the SAE J3016 standards, levels 0 through 5 are defined (ASME, 2018), based on increasing levels of automation, with the roles of the human and automation, driver support features, and automated driving features defined. While much of the focus in this taxonomy is on the machine side, the human modeling component or the requisite interface between human and automation are only now beginning to be investigated. As we move from Level 2 to 3 and higher, the CPHS aspects become increasingly important. How the combined human-automation system performance, robustness, and safety can be addressed so that the resulting system exceeds what would be achievable by either a human or an automation acting alone is an important research direction that should be addressed by our research community. A second example is also a transportation one, in air, where the decision makers involve a skilled human (i.e. a pilot), and automation (i.e. an autopilot), and the focus is on their combined interaction in the context of emergencies. Here too, the goal is to design the underlying CPHS to realize a performance that would exceed what may be achievable by the human or automation alone. Several other examples can be drawn, from energy (ex. Participation of consumers allowing power consumption to be flexible and on-demand), robotics (ex. assistance to elders and to differentially abled humans), and manufacturing (congruence between skilled operators and automation).

The examples above illustrate the changing and nuanced nature of interactions between humans and CPS. Humans are no longer passive consumers or actors in these examples; they are empowered decision-makers, and drive the evolution of the technology. While notions such as humans-in-the-loop have been around for a few decades, they are transforming with technological advances. With a parallel development occurring in data analytics and AI research in human decision-making, human-technology co-evolution is likely to lead to a very high impact research. Whether in-the-loop, on-the-loop, for-the-loop, or any combinations thereof, these combined interactions between humans and CPS need to be collectively studied, under the rubric of CPHS. The introduction of humans as a key component in the underlying system necessitates new tools, which are grounded in social sciences. Notions of human resilience have been studied in the area of cognitive systems, and concepts such as Capacity for Maneuver (CfM) and Graceful Command Degradation (GCD) (Woods, 2015) appear to have important roles to play in the design of CPHS. Human decision making under uncertainty has been addressed very elegantly using the notion of Prospect Theory in the noble-prize winning work of Kahneman and Tversky, which too is beginning to be increasingly employed in the energy and transportation systems modeling and control. Intrinsic to several scenarios in CPHS are interactions between multiple decision-makers that need to be strategic, necessitating the deployment of tools from game theory. In addition to strategic decisions, long-term or time-extended decisions may be called for, and may have to be made in the presence of uncertainties. This implies that some form of learning may be useful, which leads to yet another intensely popular tool used of late - reinforcement learning. In this article, we first elaborate on how prospect theory, game theory, reinforcement learning and capacity for maneuver concept are used to model humans. Then, we provide specific examples where these tools are actively employed to model the underlying CPHS.

## 2. The human element in CPHS

The first challenge in CPHS is modeling. Apart from the intricacies of cyber physical system (CPS) modeling, the human element is even more difficult to model in terms of forecasting their behavior. In this section, we introduce three different tools that are used to predict human actions within CPHS setting.

### *2.1. Prospect theory*

Many of the roles that humans play in CPS may be due to discrete actions rather than continuous ones. Typical examples concern a decision-making that consists of choosing one option out of many. In many infrastructures such as in transportation and energy, such decision-making is quite common, and essential to ensure resource allocation and balance of supply and demand. The underlying theory that is often used to model such decision-making is the Expected Utility Theory (EUT) (Ben-Akiva et al. (1985)). Briefly, EUT can be explained as follows: Consider a problem where consumers have several travel modes. Utility theory postulates that they choose a travel mode based on optimization of the corresponding utility $U$ of that travel mode. Suppose in a specific travel mode, $U$ takes on discrete values $u_i \in \mathbb{R}, \forall i \in \{1, \ldots, n\}$ where $n$ is the number of possible outcomes, with $u_i$ and $p_i$ denoting the utility and probability of outcome $i$, respectively, with $\sum_{i=1}^{n} p_i = 1$. Then, one can determine the overall utility function as $U^o = E[U]$, where $E[\cdot]$ denotes the expectation operator and compute $U^o$ as (Von Neumann and Morgenstern (2007))

$$U^o = \sum_{i=1}^{n} p_i u_i \tag{1}$$

An interesting method that has been garnering increased attention of late for decision making in the presence of multiple choices is Cumulative Prospect Theory (CPT) (Tversky and Kahneman (1992); Kahneman and Tversky (2013)). This method is appropriate when the underlying problem involves subjective human decision making in the presence of uncertainty and risk. The behavioral model is described by a value function $V(\cdot)$ and a probability distortion function $\pi(\cdot)$, defined as follows (Tversky and Kahneman (1992); Prelec (1998)):

$$V(u) = \begin{cases} (u-R)^{\beta+} & \text{if } u \geq R \\ -\lambda(R-u)^{\beta-} & \text{if } u < R \end{cases} \tag{2}$$

$$\pi(p) = e^{-(-ln(p))^{\alpha}} \tag{3}$$

In the above, $u(x)$ denotes the utility function for an outcome $x$, and the nonlinearity $V(\cdot)$ introduces an asymmetry depending on whether or not the utility exceeding a certain reference $R$, which can be equated with a gain (if $u \geq R$) and a loss (if $u < R$). The parameter $\lambda$ captures loss aversion and is typically greater than unity, $0 < \beta^+ < \beta^- < 1$ denotes diminishing sensitivity in gains compared to losses, and the parameter $\alpha$ indicates a distorted perception of probability, with low probabilities amplified and high probabilities attenuated. With these two nonlinearities, the subjective utility of the outcome $x$, as perceived by the consumers, is given by

$$U_{PT} = \pi(p)V(u) \tag{4}$$

according to Prospect Theory, rather than the objective utility $pu$.

When multiple outcomes are present, a cumulative approach has to be taken in order to evaluate the distorted probability. Suppose that $k$ out of the $n$ outcomes are losses, and the rest are gains. That is, $u_i < R$ if $1 \leq i \leq k$ and $u_i \geq R$ if $k < i \leq n$, and $F(U)$ is the Cumulative Distribution Function of $U$. Then the perceived probability of outcome $u_i$ with the probability $p_i$ is given by (Von Neumann and Morgenstern (2007))

$$w_i = \begin{cases} \pi\big[F(u_i)\big] - \pi\big[F(u_{i-1})\big], & \text{if} \quad i \in [1,k] \\ \pi\big[1 - F(u_{i-1})\big] - \pi\big[1 - F(u_i)\big], & \text{otherwise} \end{cases} \tag{5}$$

The subjective utility $U^s$ perceived by the passenger within the CPT framework is given by

$$U^s = \sum_{i=1}^{n} w_i V(u_i). \tag{6}$$

### *2.2. Capacity for maneuver concept*

A recent method in modelling of humans under anomalous events utilizes the Capacity for Maneuver (CfM) concept (Farjadian et al., 2016, 2017, 2019). CfM refers to the remaining range of the actuators before saturation, which quantifies the available maneuvering capacity of the vehicle. One example of the CfM is of the form $CfM = u_{\max} - \mathrm{rms}(u(t))$, where $u_{\max}$ and $u(t)$ refer to the maximum available actuator deflection and actuator deflection at time $t$, respectively (Eraslan et al., 2020). It is hypothesized that surveillance and regulation of a system's available capacity to respond to all events help maintain the resiliency of a system, which is a necessary merit to recover from unexpected and abrupt failures or disturbances (Woods, 2018). Below, two different types of human pilot models, both of which use CfM, are explained.

#### *2.2.1. Perception trigger*

Here, the pilot model is assumed to assess the CfM and implicitly compute a perception gain, $K_t$, based on the CfM. The perception algorithm for the pilot is

$$K_t = \begin{cases} 0, & |F_0| < 1 \\ 1, & |F_0| \geq 1 \end{cases} \tag{7}$$

where

$$F_0 = G_1(s)[F(t)], \tag{8}$$

and

$$F(t) = \frac{\frac{d}{dt}(\mathrm{CfM}) - \mu_p}{3\sigma_p}. \tag{9}$$

$G_1(s)$ is a filter introduced as a smoothing and lagging operator into human perception algorithm and $F(t)$ is the perception variable. $\mu_p$ and $\sigma_p$ are the average and the standard deviation of $\frac{d}{dt}(\mathrm{CfM})$. The hypothesis is that the human pilot has such a perception trigger, $K_t$, and when $K_t = 1$ the pilot takes over control from the autopilot (Farjadian et al., 2016).

#### *2.2.2. CfM-GCD Tradeoff*

Here, the pilot is assumed to implicitly assess the available CfM when an anomaly occurs, and decide on the amount of graceful command degradation (GCD) that is allowable so as to let the CfM become comparable to a certain desired value. Degradation of the command, or the control input, can be implemented by setting a certain autopilot parameter by the pilot. In other words, it is assumed that the pilot is capable of assessing the suitable value of a certain autopilot control parameter and input this value to the autopilot following the occurrence of an anomaly (Farjadian et al., 2019).

### *2.3. Game theory and reinforcement learning*

The difficulty of predicting human actions intensifies when the system contains more than one human, which requires factoring in multiple human-human and human-CPS interactions. For example, human interactions are generally not deterministic in nature, which can be projected into CPHS models by utilizing a probabilistic modeling framework. Furthermore, before taking an action, a human generally contemplates other intelligent agents' possible actions and then tries to choose a move that will increase the chances of obtaining the best outcome, which is called "strategic behaviour" (Camerer, 2011). Finally, humans do not always act in an optimal fashion. This final point is important for distinguishing CPHS models from autonomy models, where, based on available information, an algorithm can possibly be designed to react in the most appropriate way.

#### *2.3.1. Game theory*

Game theory studies the interactions between strategic agents. "Players" in a game theoretical setting refer to the entities who can effect the game by their "moves". "Strategy" of a player defines the procedure based on which a player chooses his or her actions. A "Solution concept" is a well established set of rules that are used to predict how a game will unfold. A "Nash Equilibrium" is a solution concept, defined similarly to the equilibrium in system dynamics: When players have no incentive to deviate from their selected actions, the game is said to be in Nash Equilibrium. This means that in Nash Equilibrium, players choose their best actions against each others' actions. There are other equilibrium concepts such as "quantal response equilibrium" where instead of giving the best response to other players' actions, the players choose a probability distribution over their action space where actions with higher expected payoffs have higher probability of being played.

Not all solution concepts predict an equilibrium. For example, "level-k thinking" is a non-equilibrium model, which assigns different levels of "reasoning" for players (Stahl and Wilson, 1995; Costa-Gomes et al., 2009). In this model, the lowest level of reasoning is level-0, which represents non-strategic thinking, simply meaning that the players who reason at this level have a strategy that does not take into account other players' possible actions. A level-1 player, on the other hand, takes the best action assuming that his or her opponents are level-0 players. Similarly, a level-k player responds best to his *belief* that the other players are reasoning at level-(k-1). Therefore, this model assumes an *iterated* best response (Crawford, 2008).

*2.3.2. Reinforcement learning*

In RL, there exists an "agent" capable of exerting "actions" that can change the "state" of the environment where the agent operates. The RL problem can be defined as finding the optimal set of action sequence for an agent to achieve a given goal defined as a function of the environment states, through interaction with the said environment (Sutton and Barto, 2018).

RL uses the idea of a "reward function" to describe the preferences of the agent while its learning to achieve a predetermined goal. A "policy" is defined as a map from states to actions. The task of the RL algorithm is to find a policy that will make the agent maximize a cumulative discounted reward expressed as

$$C = \sum_{t=0}^{\infty} \gamma^t r_t, \tag{10}$$

where $\gamma$ is the discount factor and $r$ is the reward obtained in every step $t$. There are various RL techniques proposed to discover action sequences that will attain the goal of maximizing (10). Almost all of these different methods are based on estimating the "value function", which is the value of being in a certain state, $s$, based on the policy, $\pi$. A value function is given as

$$V^{\pi}(s) = E_{\pi}\left\{ \sum_{k=0}^{\infty} \gamma^k r_{t+k} | s_t = s \right\}, \tag{11}$$

where $E$ refers to expected value. The "action value function", is defined as the value of taking a certain action $a$, in a given state $s$, and is defined as

$$Q^{\pi}(s,a) = E_{\pi}\left\{ \sum_{k=0}^{\infty} \gamma^k r_{t+k} | s_t = s, a_t = a \right\}. \tag{12}$$

The aim of RL is to find the optimum policy $\pi^*$ that will maximize the value function. The optimal value function is written as $V^{\pi^*}$ and all its state values are larger than or equal to that of all other value functions that are created by policies different than the optimal policy. This can be represented as $V^{\pi^*}(s) \geq V^{\pi}(s), \forall\pi, \forall s$. Similarly, we can write $Q^{\pi^*}(s,a) \geq Q^{\pi}(s,a), \forall\pi, \forall(s,a)$, for the optimal action value function. Once the optimal action value function is found, the following method can be used to select the best action in each state.

$$\pi^*(s) = \arg\max_a Q^{\pi^*}(s,a). \tag{13}$$

The process of finding the optimal policy is called "training". During training, the agent observes the states and produces an action based on the observed states. This action influences the environment and results in a new set of states. These new states are evaluated by the reward function and a reward signal is formed. The agent uses this signal to update the action value function, and the next cycle starts with the new action.

One of the RL methods that played a significant role for the success of RL is "Q-learning" (Watkins, 1989). In Q-learning, an incremental estimate of the optimal action value function is realized using the following update rule.

$$Q_{k+1}(s_t,a_t) = Q_k(s_t,a_t) + \alpha\Big(r_t + \gamma \max_a Q_k(s_{t+1},a) - Q_k(s_t,a_t)\Big). \tag{14}$$

*2.3.3. Merging game theory and reinforcement learning to model human interactions*

For human interactions that last long periods of time using equilibrium based game theoretical solutions may be computationally infeasible. One way around this problem is using the non-equilibrium game theoretical solution concept, level-k thinking, explained in Section 2.3.1. This means that once the "anchoring" level, level-0, is selected, a level-1 human agent's behaviour can be identified as the best response to all the other human actions that are determined by the level-0 policy. Similarly, once level-1 behavior is identified, all the agents in the system are assigned the level-1 policy except the one whose level-2 behaviour is to be found. Therefore, to predict the policy of a level-k agent, all the rest of the agents' policies are set to level-(k-1), which effectively make them a part of the environment whose dynamics are known, and level-k policy is determined as the best response to the rest of the level-(k-1) policy actions. This isolates the level-k policy as the single policy that needs to be computed. Once the difficulty of high computational cost due to multiple decision makers is solved using level-k thinking, the problem reduces to estimating the optimal action sequence of an intelligent agent in a given environment. This problem can be stated as a reinforcement learning (RL) problem by properly defining the states, actions, the environment and the reward function. The algorithm used in obtaining the agent policies, demonstrating the interplay between the game theory and RL, is provided in Algorithm 1 (Albaba and Yildiz, 2019), where $k$ is the maximum desired level.

## 3. Interconnecting humans and cyber physical systems: examples of CPHS

*3.1. Prospect theory examples*

We describe two examples below where Prospect Theory has been explored, the first of which is in power grids, and the second is in ride-sharing services.

**Algorithm 1** Merging RL and Game Theory

Set i = 0
**while** $i < k$ **do**
  Load the level-i policy
  Set cognition levels of all players in the environment other than the learning agent to level-i, i.e. set policies of players to level-i policy
  Place the learning agent in the initialized environment, in which all players are level-i
  Start the training of the learning agent using a reinforcement learning method, through which agent learns how to best respond to level-i players
  Once the training is completed, learning agent becomes a level-(i+1) player
  Save the policy of the learning agent as level-(i+1) policy
  i += 1
**end while**

(1) Suppose there are N consumption units capable of adjusting their loads located at a node in a power grid. This capability may be either due to the fact that the load itself is adjustable (ex. Temperature set-point in a HVAC system), or they include a storage unit that has a flexible charging/discharging schedule. The corresponding utility function of a unit $k$ consists of $U(D_k, S_k)$, where $D_k$ is the quantity of energy that is charged and $S_k$ is the amount of energy discharged. In general $U_k$ depends on the amount of energy imbalance, the price of electricity, the actions of other assets at that node, and their flexibility, and the penalty that the power company may impose for unmet commitment, among others. The argument for using a PT based model rather than EUT is because of the uncertainty introduced by other players in the underlying trading game. The effect of the overall risk-aversion may be modeled through the use of the $(V, \pi)$ tuple described in (2) and (3) using (4). (Wang et al. (2014)).

(2) The second example has been explored in the context of ride-sharing services. Here, the utility function depends on both travel times and the price of the service, and is of the form

$$U = X + b\gamma \tag{15}$$

where $\gamma$ is the tariff from ride offer and is deterministic, and $b$ represents the weighting of the economic cost of the ride in comparison to the travel cost. The time based utility is of the form

$$X = a_1 T_{\text{walk}} + a_2 T_{\text{wait}} + a_3 T_{\text{ride}} + c \tag{16}$$

and is a stochastic quantity, as each of the travel times can vary randomly. As the amount of stochasticity is significant, a prospect theory based approach may be more appropriate. It should also be noted that the utility function is a continuous function, with the corresponding function $U(t)$ varying over an interval $[t_{\min}, t_{\max}]$ which is determined by the bounds of the overall travel. Each $U(t)$ has an associated Cumulative Distribution function $F(U)$. Since there is a significant subjective component in this problem, the passengers who are willing to accept the ride service may perceive the utility $U$ as $V(U)$, and perceive the corresponding probability of $U(t)$ in a distorted manner as well. As the underlying utility function is continuous, one needs to adopt a CPT approach here (Tversky and Kahneman (1992)).

We first present the CPT model for a discrete case, with utility $U_i$ for outcome $i, i = 1, \ldots, n$. Each $U_i$ represents the utility of the ride-sharing service for a time cost $t_i$. For example, the wait time could be 10 minutes, the travel time 30 minutes, and the walk time 5 minutes, leading to a total time of 45 minutes, which has a certain utility to the passenger. The objective utility therefore can be calculated using (1). Using the CPT discussions above, the subjective utility can be calculated using (6). The extension from (6) to the continuous case of $U_R^s$ then becomes

$$U^s = \int_{-\infty}^{R} V(u) \frac{d}{du} \Big\{ \pi \big[ F_U(u) \big] \Big\} du + \int_{R}^{\infty} V(u) \frac{d}{du} \Big\{ -\pi \big[ 1 - F_U(u) \big] \Big\} du \tag{17}$$

One can similarly calculate the distorted probabilities. Suppose we consider a simple case each passenger is assumed to choose between two options, the ride-sharing service and another alternative such as public transportation, which may have a utility $A^o$. The objective probability of acceptance is given by

$$p^o = \frac{e^{U^o}}{e^{U^o} + e^{A^o}} \tag{18}$$

The subjective probability of acceptance is given by

$$p^s = \frac{e^{U^s}}{e^{U^s} + e^{A^s}} \tag{19}$$

where $A^s$ denotes the subjective utility of the alternative perceived by the passenger. The overall prospect theory based passenger model is then quantified using the pair $(U^s, p^s)$. These in turn could be used to determine the dynamic tariff of the ride-sharing service. Such an approach directly accommodates the behavioral model of humans. The reader is referred to Guan et al. (2019) for further details.

### *3.2. Shared Control Architectures*

Shared Control Architectures (SCA) exemplified in this section describe the responsibility sharing between pilots and autopilots in terms of controlling the aircraft. In Section 2.2 two different models of human decision making were discussed, both based on the monitoring of the CfM. In this section, we introduce two different SCAs based on these models (Eraslan et al., 2020).

### *3.3. SCA 1: A pilot with a CfM based perception and a fixed-gain autopilot*

The first shared control architecture can be summarized as a sequence {autopilot runs, anomaly occurs, pilot takes over}. That is, it is assumed that an autopilot based on PD control is in place, ensuring a satisfactory command tracking under nominal conditions. The human pilot is assumed to consist of a perception component and an adaptation component. The perception component consists of monitoring CfM, through which a perception trigger

$F_0$ is calculated using (8). The adaptation component consists of monitoring the control gain in (7), and taking over control of the aircraft when $K_t = 1$. The details of this shared controller and its evaluation using simulations and experiments are given by Farjadian et al. (2016); Eraslan et al. (2020). Figure 1 illustrates the schematic of SCA1.

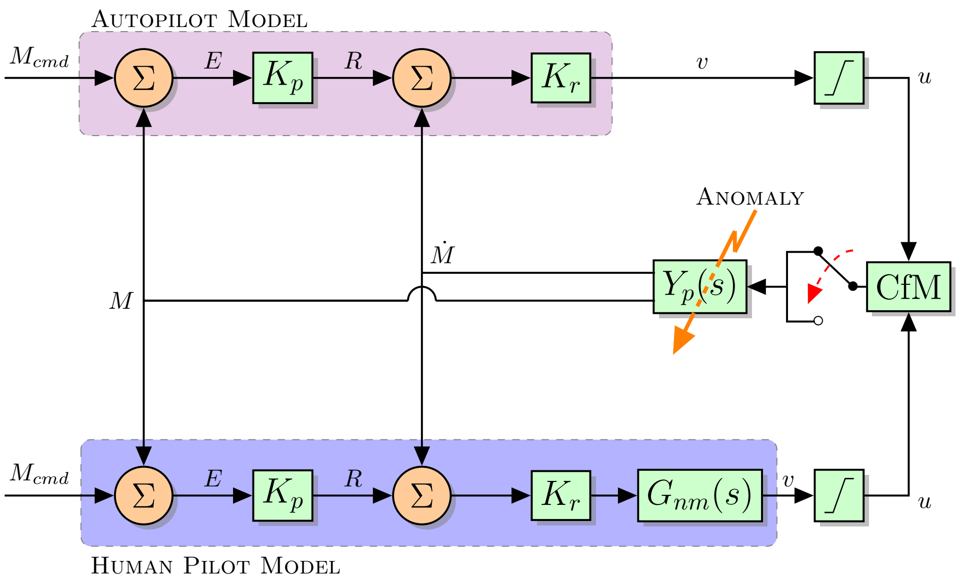


Figure 1: Block diagram of SCA1 (adapted from Farjadian et al. (2016)). The autopilot model consists of a fixed gain controller, whereas the human pilot model comprehends a perception part based on the CfM concept and an adaptation part governed by empirical adaptive laws (Hess, 2016). The neuromuscular transfer function (Hess, 2006), $G_{\mathrm{nm}}(s)$ corresponds to control input formed by an arm or leg. When an anomaly occurs, the plant dynamics, $Y_{\mathrm{p}}(s)$, undergoes an abrupt change by rendering the autopilot insufficient for the rest of the control. At this stage, the occurrence of an anomaly is captured by the CfM such that the control is handed over to the human pilot model for a resilient flight control.

### *3.4. SCA 2: A pilot with a CfM-based decision making and an advanced adaptive autopilot*

In this shared control architecture, the role of the pilot is a supervisory one while the autopilot takes on an increased and complex role. The pilot is assumed to monitor the CfM of the actuators in the aircraft following an anomaly. In an effort to allow the CfM to stay close to a desired value, the command is allowed to be degraded; the pilot then determines a parameter $\mu$ which directly scales the control effort. Once $\mu$ is specified by the pilot, then the adaptive autopilot continues to supply the control input. If the pilot has high situational awareness, he/she provides an estimate of the severity of the anomaly. The details of this shared controller and its evaluation using simulations and experiments are given by Farjadian et al. (2017, 2019); Eraslan et al. (2020). Figure 2 shows the schematic of SCA2.

### *3.5. Airspace in the presence of both manned and unmanned aircraft*

Obtaining airspace models where both manned and unmanned aircraft are present is a necessity for successful integration of unmanned aircraft systems (UAS) into the National

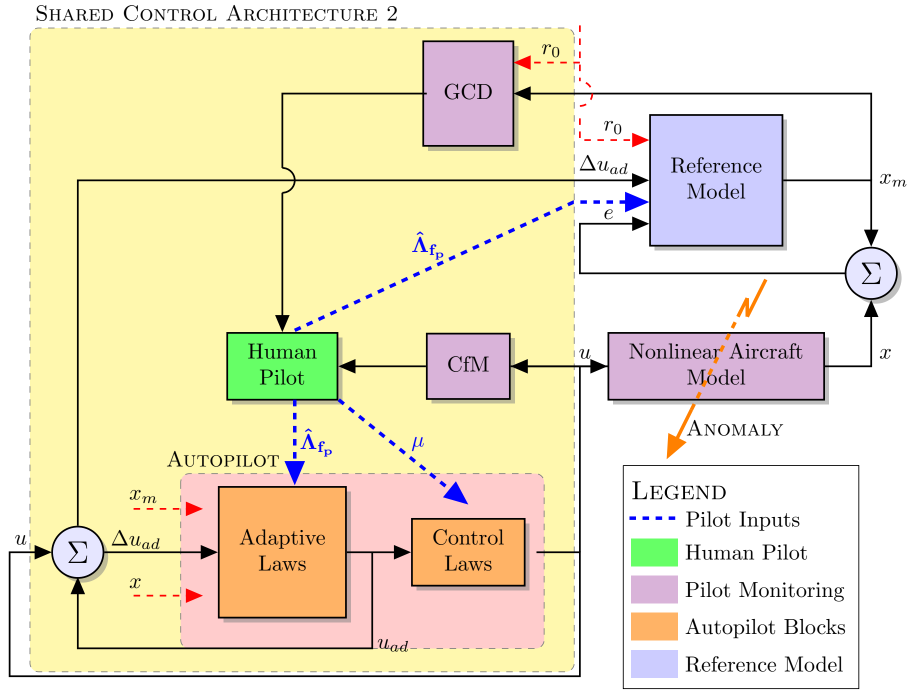


Figure 2: Block diagram of SCA 2. The human pilot undertakes a supervisory role by providing the key parameters $\mu$ and $\hat{\Lambda}_{f_p}$ to the adaptive autopilot. The blocks are expressed in different colors based on their functions in the proposed SCA2.

Airspace System (NAS). Sense and avoid (SAA) algorithms, human-human and human-automation interactions through aircraft dynamics, aircraft, and traffic collision avoidance algorithms (TCAS), together with the communication links make this system a typical example of a CPHS.

Figure 3 shows such an airspace scenario where a UAS (magenta) is assigned to follow certain waypoints (yellow) in a crowded airspace filled with manned aircraft (red). We are interested in how the overall system will evolve in time.

To model this scenario, we need to include aircraft dynamics, TCAS and SAA algorithms, and communication dynamics which may include signal delays. These components of the model are already well understood and documented in the literature. The human interaction dynamics, on the other hand, is the main challenge.

As discussed in Section 2.3, the GT and RL based modeling approach produces policies, which are probabilistic maps from observations to actions, to represent human interactions. Therefore, to obtain these policies, we first need to clearly define the *observation and action spaces* and represent them in a way that is meaningful for the RL algorithm. Once these spaces are explicitly defined, pilot goals and preferences need to be expressed in form of a *reward function*. Obtaining these policies and integrating them with the automation and the physical manned and unmanned aircraft dynamics enables a simulation capability to investigate and study various UAS integration into the NAS scenarios. A detailed description

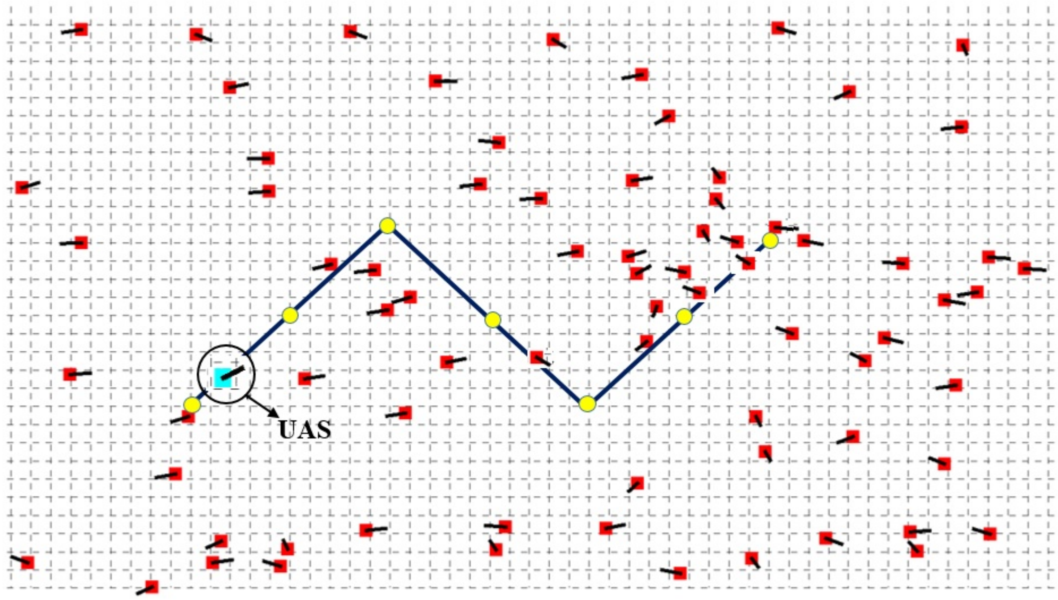


Figure 3: A hybrid airspace scenario where red squares represent manned aircraft and the magenta square is used to indicate an unmanned aircraft (UA). The yellow circles in the 600 km x 300 km airspace are the waypoints assigned to the UA. (Musavi et al. (2016), reprinted by permission of the American Institute of Aeronautics and Astronautics, Inc.)

of how to achieve this together with several simulation studies and validations with a data-based model is given by Musavi et al. (2016).

## 4. CPHS beyond the engineering domain

The role of CPHS is significantly larger. Much of the focus in the topics outlined above was on engineering systems and how we come up with better tools for their analysis and synthesis. The focus needs to be diverted towards social systems as well and how the emergence of CPHS and the varied roles of humans therein introduce new challenges therein, in domains related to ethics, economics, justice, and the like. For example, the topic of ethics been studied for millennia by societies and civilizations, the creation of new socio-technical systems introduces new paradigms for societal interactions and thereby new challenges.

This article would not be complete without bringing in yet another important connection between systems and control and social sciences, which can be grouped under the rubric of dynamics over techno-social networks, DTSN (see Section 5.20 on Social Networks in Lamnabhi-Lagarrigue et al. (2017)). DTSN has its origins in Social Network Analysis which is grounded in social philosophy and psychology and focuses on the examination of social relations, influences, group actions and interactions. By leveraging the intersection between SNA and graph theory, several computationally efficient algorithms have been derived for the analysis of large-scale systems. With rapid changes in technology, these large-scale systems are now beginning to have a strong temporal signature, which implies the need for a dynamical system point of view, leading to DTSN. The focus of this emerging area has a huge range of applicability, with its emphasis on the study of both topological and dynamical properties of the network, properties of resilience and robustness in the presence of natural disasters and cyberattacks that may cause failures in nodes and links in the network. A notable and topical application area is the current COVID-19 spread over the world population.

## 5. Summary

As societal challenges increase, there is an increasing need for human systems to interact with cyber-physical systems in various ways to provide innovative solutions that overcome these challenges. The rubric that addresses the foundations of such emerging interactions and integrations between CPS and human systems forms CPHS. This article provides an introduction to CPHS and representative examples from energy and transportation sectors where such CPHS have begun to emerge.

One of the main challenges in CPHS is modeling of the human element. Given the enormously complex human system that functions over a huge range of timescales to perform a wide range of tasks, there are perhaps as many models of humans needed for as many problems. The models presented in this article therefore are in no way meant to be comprehensive but rather illustrations. We have focused on a few instances where the underlying models attempt to capture the perceptive, strategic, and decision-making abilities of humans in various contexts. In addition, we have tried to illustrate varied concepts and methodologies, which include Prospect Theory, Game Theory and Reinforcement Learning, and Capacity for Maneuver, that may be appropriate for deriving human models in these contexts.

While the applications that we have outlined in this article correspond to the energy and transportation sectors, by no means are they meant to be the only applications of interest in CPHS. Several examples abound in the area of robotics, healthcare, and manufacturing, which have not been covered here (see the ongoing IFAC workshop series on CPHS for ongoing and emerging activities in this area!).

## 6. Acknowledgements

This effort was sponsored by Boeing Strategic University Initiative and the Turkish Academy of Sciences.

## References

Albaba, B.M., Yildiz, Y., 2019. Modeling cyber-physical human systems via an interplay between reinforcement learning and game theory. Annual Reviews in Control 48, 1–21.

ASME, 2018. Sae international releases updated visual chart for its "levels of driving automation" standard for self-driving vehicles. [Available at https://www.sae.org/news/press-room/2018/12/sae-international-releases-updated-visual-chart-for-its-"levels-of-driving-automation"-standard-for-self-driving-vehicles (Accessed at 05/04/2020)] .

Ben-Akiva, M.E., Lerman, S.R., Lerman, S.R., 1985. Discrete choice analysis: theory and application to travel demand. volume 9. MIT press.

Camerer, C.F., 2011. Behavioral game theory: Experiments in strategic interaction. Princeton University Press.

Costa-Gomes, M.A., Crawford, V.P., Iriberri, N., 2009. Comparing models of strategic thinking in Van Huyck, Battalio, and Beil's coordination games. Journal of the European Economic Association 7, 365–376.

Crawford, V.P., 2008. Modeling behavior in novel strategic situations via level-k thinking, in: North American Economic Science Association Meetings.

Eraslan, E., Yildiz, Y., Annaswamy, A.A., 2020. Shared control between pilots and autopilots: Illustration of a cyber-physical human system. IEEE Control Systems magazine, in press. [Preprint available at https://arxiv.org/abs/1909.07834 (Accessed on 05/06/2020)].

Farjadian, A., Annaswamy, A., Woods, D., 2016. A resilient shared control architecture for flight control, in: Proceedings of the International Symposium on Sustainable Systems and Technologies. URL: `http://aaclab.mit.edu/publications.php`.

Farjadian, A.B., Annaswamy, A.M., Woods, D., 2017. Bumpless reengagement using shared control between human pilot and adaptive autopilot. IFAC-PapersOnLine 50, 5343–5348.

Farjadian, A.B., Thomsen, B., Annaswamy, A.M., Woods, D.D., 2019. Resilient flight control: An architecture for human supervision of automation. IEEE Transactions on Control Systems Technology (in press).

Guan, Y., Annaswamy, A.M., Tseng, H.E., 2019. Cumulative prospect theory based dynamic pricing for shared mobility on demand services. IEEE 58th Annual Conference on Decision and Control (CDC) .

Hess, R.A., 2006. Simplified approach for modelling pilot pursuit control behaviour in multi-loop flight control tasks. Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering 220, 85–102.

Hess, R.A., 2016. Modeling human pilot adaptation to flight control anomalies and changing task demands. Journal of Guidance, Control, and Dynamics 39, 655–666.

Kahneman, D., Tversky, A., 2013. Prospect theory: An analysis of decision under risk, in: Handbook of the fundamentals of financial decision making: Part I. World Scientific, pp. 99–127.

Lamnabhi-Lagarrigue, F., Annaswamy, A., Engell, S., Isaksson, A., Khargonekar, P., Murray, R.M., Nijmeijer, H., Samad, T., Tilbury, D., den Hof, P.V., 2017. Systems & control for the future of humanity, research agenda: Current and future roles, impact and grand challenges. Annual Reviews in Control 43, 1 – 64.

URL: http://www.sciencedirect.com/science/article/pii/S1367578817300573, doi:https://doi.org/10.1016/j.arcontrol.2017.04.001.

Murray, R.M., 2003. Control in an information rich world: Report of the panel on future directions in control, dynamics, and systems. SIAM.

Musavi, N., Onural, D., Gunes, K., Yildiz, Y., 2016. Unmanned aircraft systems airspace integration: A game theoretical framework for concept evaluations. Journal of Guidance, Control, and Dynamics 40, 96–109.

Prelec, D., 1998. The probability weighting function. Econometrica , 497–527.

Samad, T., Annaswamy, A., 2011, 2014. The impact of control technology: Overview, success stories, and research challenges. IEEE Control Systems Society. [Available at http://ieeecss.org/sites/ieeecss/files/2019-07/IoCT-FullReport_v2.pdf ($1^{st}$ edition) http://ieeecss.org/general/IoCT2-report ($2^{nd}$ edition)] .

Stahl, D., Wilson, P., 1995. On players' models of other players: Theory and experimental evidence. Games and Economic Behavior 10, 218–254.

Sutton, R.S., Barto, A.G., 2018. Reinforcement learning: An introduction. MIT press.

Tversky, A., Kahneman, D., 1992. Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and uncertainty 5, 297–323.

Von Neumann, J., Morgenstern, O., 2007. Theory of games and economic behavior (commemorative edition). Princeton university press.

Wang, Y., Saad, W., Mandayam, N.B., Poor, H.V., 2014. Integrating energy storage into the smart grid: A prospect theoretic approach, in: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE. pp. 7779–7783.

Watkins, C.J.C.H., 1989. Learning from delayed rewards. Ph.D. thesis. King's College, Cambridge.

Wiener, N., 1948. Cybernetics or control and communication in the animal and the machine. Technology Press.

Woods, D.D., 2015. Four concepts for resilience and the implications for the future of resilience engineering. Reliability Engineering & System Safety 141, 5–9.

Woods, D.D., 2018. The theory of graceful extensibility: basic rules that govern adaptive systems. Environment Systems and Decisions 38, 433–457.