A Review of Cognitive Dynamic Systems and Its Overarching Functions

Waleed Hilal
Department of Mechanical Engineering
McMaster University
Hamilton, ON, Canada
hilalw@mcmaster.ca

Alessandro Giuliano
Department of Mechanical Engineering
McMaster University
Hamilton, ON, Canada
giuliana@mcmaster.ca

Stephen A. Gadsden
Department of Mechanical Engineering
McMaster University
Hamilton, ON, Canada
gadsden@mcmaster.ca

John Yawney
Adastra Corporation
Toronto, ON, Canada
john.yawney@adastragrp.com

***Abstract*—Cognitive dynamic systems are a new field of physical systems inspired by several areas of study such as neuroscience, cognitive science, computer science, mathematics, physics and engineering. Building on Fuster's paradigm, a system is considered cognitive when it is capable of five fundamental processes to human cognition: the perception-action cycle, memory, attention, intelligence and language. With these capabilities, a cognitive dynamic system can sense its environment, interact with it, and learn from it through continued interactions. The goal of this paper is to provide a thorough review of the cognitive dynamic system framework, along with its theory, applications, and its two special functions: cognitive control and cognitive risk control.**

Keywords—cognitive dynamic systems, cognitive control, cognitive risk control, cognitive radio, cognitive radar, cognitive internet of things, smart systems

I. Introduction

Cognitive dynamics systems (CDS) are a new class of systems which combine knowledge from several disciplines such as neuroscience, cognitive science, computer science, mathematics, physics and engineering [1]. The goal of CDS is to provide a framework for dynamic systems to augment them with cognitive capabilities, allowing them to sense, interact and learn from their environments. A dynamic system is considered to be cognitive when it can carry out five fundamental processes which are critical to human cognition. As defined by Fuster's paradigm, these processes are the perception-action cycle (PAC), memory, attention, intelligence, and language [2].

Interest in the field of CDS has been growing at an increasing rate in recent years thanks to the seminal work carried out by Haykin on cognitive radio and cognitive radar [3] [4]. These two applications are among the earliest examples of CDS, stoking research into the theory and design of the CDS framework [1]. With cognitive radio, the primary goal is to solve the spectrum scarcity issue by providing the means for radio systems to access underutilized bandwidths. Cognitive radar, on the other hand, has been proposed as a means of providing improved accuracy and reliability in remote-sensing applications [2]- [4].

In essence, the CDS framework can be broken down into two special functions: cognitive control (CC) [5] and cognitive risk control (CRC) [6]. In the former, the limitation of current adaptive controllers and neurocontrollers when faced with unmodeled dynamics or unstructured environments are addressed. Specifically, CC is additive in nature, meaning that it is augmented to existing system designs by introducing a new state known as the entropic state. The entropic state is based on the notion of an information gap that must be controlled alongside a system model. The second special function, CRC, expands on the CC architecture to account for the risks associated with the uncertainties that are faced by a system, and bring them under control. Such risks may include security threats frequently encountered by physical systems, like cyberattacks on smart grids on jammers acting on radar systems.

The first goal of this paper is to provide a detailed background on the theory and the architecture behind the CDS framework and its two special functions CC and CRC. The second goal is to present a short, structured overview of the current state of the field by reviewing recent literature on applications within the CDS framework. We aim to present a thorough discussion on the methodologies, key findings, experimental results, and limitations of the surveyed literature. Finally, we offer insight into the most promising areas for future research efforts in this field. This paper is organized as follows: Section II of this paper presents a background the CDS framework, theory and architecture. In Section III, cognitive control is introduced as the first special function of the CDS framework, with a review of relevant applications of this architecture. Similarly, Section IV introduces CRC as the second special function of the CDS framework, with a short review of applications and recent advancements. Finally, concluding remarks and suggestions for future researchers are discussed in Section V.

II. Cognitive Dynamic Systems

A. Perception-action Cycle

There are two parts to any CDS: the perceptual and the executive. On the right-hand side of Fig. 1, the perceptual component or perceptor is located, whereas the executive or cognitive controller is located on the left-hand side [7]. Depending on the CDS application, the perceptor is in charge of directly observing the system and the environment using appropriate sensors. During perception, for example, a Bayesian estimator might be utilised, which computes the posterior of a system's state in each PAC and extracts relevant

2022 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS) | 978-1-6654-8684-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/IEMTRONICS55184.2022.9795764

information from what is experienced. A feedback connection transmits the perceptor's extracted relevant environmental information to the executive, which is then tasked with conducting cognitive or physical actions on the environment or system based on this knowledge [7].

The executive's cognitive efforts are designed to continuously improve the information extracted by the perceptor in following cycles. As a result, the executive indirectly observes the environment through the perceptor and acts on the information obtained, completing the PAC with a global feedback loop. In order to indirectly impact the system's perception, cognitive actions are frequently applied to the surroundings, such as increasing the lights in a dark room [5]. The physical state in this scenario contains the positions of objects that are not changed by light. Other forms of cognitive operations, such as altering the system's own sensors or actuators, can be performed alone on the system. The adaptation of a transmitted waveform in a cognitive radar system is an example of this. Furthermore, by adding a new component to the state controller's cost function called the entropic state, cognitive actions can be used to impact state-control actions [5].

The cognitive controller, as shown in Fig. 1, is in charge of making decisions about the aforementioned cognitive operations in the executive, based on the entropic state established by the perceptor [8]. However, all of the distinct types of cognitive actions are not necessarily present in a given problem. A cognitive radar system, for example, can assess the target's states without being able to physically manipulate them because it only executes cognitive actions on its own actuators and the environment [5]. The application of reinforcement learning (RL) for the cognitive control agent is the mechanism underpinning executive decision-making.

B. Memory

The cognitive process of memory occupies its own physical space in a CDS, as shown in Fig. 1, in three forms: perceptual, executive, and working memory. Both sorts of memories have slightly different functions and responsibilities, but the primary purpose of equipping a CDS with memory is to allow for the acquisition and storage of long and short-term information [2]. A CDS can learn from its past experiences in terms of action and perception with access to this data, resulting in enhanced performance and resilience.

According to the CDS concept, perceptual memory should have a hierarchical structure with multiple layers [7]. The goal of this setup is to perform perceptual abstraction of incoming inputs or measurements in order to represent the essence of an object, event, or experience, similar to how the human memory system works. Relevant information is kept whereas irrelevant information is eliminated, allowing for long-term memory in the perceptual component of the CDS [9]. In response to feedback information from the perceptor, executive memory serves a dual function with perceptual memory. The executive memory stores long-term cognitive actions made by the cognitive controller based on input from the perceptor and can be utilised as a reference for future cognitive activities. A new policy that considers both long-term and short-term experiences is produced by adding the output of the executive memory to the cognitive controller and incorporating it into future policies [8]. The executive memory effectively keeps track of the cognitive controller's action space in a probabilistic fashion.

The working memory's function is to reciprocally couple the executive and perceptual memories, acting as a short-term memory interface between the two inside the CDS [4]. The cognitive controller may carry out its actions in each PAC in a synchronised manner as a result of this integration, and in summation, memory's general job is to continuously learn from and model the behaviour of the environment and the CDS's action space [8].

C. Attention

Unlike the PAC and memory, which have their own physical locations in the CDS, the cognitive process of attention reveals itself through computational mechanisms inside the framework. There are two types of attention: perceptual and executive attention, which are both based on localised cycles and feedback linkages in their respective sections of the CDS [5]. Their responsibilities include prioritisation of activities and effective resource allocation, and they work closely with and are driven primarily by the presence of memory. This is accomplished in the perceptor, for example, by a variety of strategies that can be utilised to filter out unnecessary input using previously stored characterizations of the environment. On the executive side, attention uses the well-known explore-

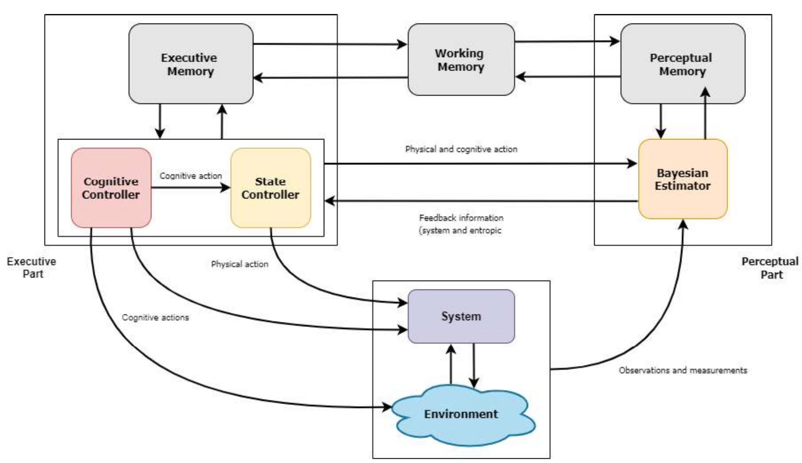

Fig. 1 Block diagram of the basic structure of a cognitive dynamic system and its architecture

exploit tradeoff to help the RL learn from and plan cognitive actions for future cycles, primarily by reducing the action space of the memory for the RL to consider based on the relevance to the perceived information in each cycle [4] [8].

D. *Intelligence*

Intelligence, like attention, does not have its own physical location in a CDS. It, on the other hand, draws on all prior cognitive processes like memory and attention and integrates them with the PAC to promote computational intelligence through efficient decision-making [2] [10]. Intelligence has an impact over the entire CDS, and its power and effectiveness in information processing are generated from leveraging all of the system's feedback loops, both global and local. As a result, intelligence plays a critical part in the CDS framework in terms of optimal decision-making in terms of the controller's actions on the system or environment of interest [7].

III. Cognitive Control

A. *Overview of Cognitive Control*

Cognitive control is a paradigm that was first introduced in 2012 by Haykin et al. [5] and is additive in nature rather than a replacement system design paradigm. CC can increase the usage of computational resources and lower the correctly defined risk functional for the work at hand by complementing state-control paradigms such as adaptive control and neurocontrol. However, before defining CC, the concept of an information gap must be established, which is related to the risk associated with a policy or action. It is thus possible to define CC with the goal of minimizing the information gap.

The following is a useful summary of the concept of the information gap [5]: From noise-affected data, available information is retrieved and transformed from the measurement space to the information space. The accessible information is then partitioned into useful and redundant information based on the task at hand. In addition, sufficient information must be defined as the information required to complete the task at hand while reducing risk; relevant information is thus the intersection of available information and sufficient information. Furthermore, the information gap can be defined as the difference between sufficient information and relevant information.

Quantifying and reducing the information gap necessitates the development of a task-specific metric – this concept equates to a new state that must be managed. The state of a dynamic system represents the basic information characterising the system's conditions at a given point in time, and the state trajectory, or change in state over time, describes the system's behaviour. The state, on the other hand, can only be obtained through noisy measurements, which necessitates a perception process in order to determine a posterior distribution of the state using a Bayesian generative model. The information gap is the difference between the maximal relevant information in the posterior distribution and the required statistics for a specific job. this quantity is also known as the entropic state, which gets its name from Shannon's entropy [5] [11]. As a result, thinking about a two-state model of a CC system, composed of a state-space model, which describes the evolution of the system state over time, and an entropic-state model, which quantifies the information gap given the posterior computed by perception, seems intuitive. According to statistical differences in the environment, both models may change from one cycle to the next. It's also worth emphasising that the entropic state is simply the feedback information sent on to the cognitive controller, and that CC is merely a paradigm for minimising the entropic state [5].

The mathematical paradigm of RL is concerned with learning the best possible actions purely through positive and negative reinforcement or rewards. In CC, RL is in charge of ensuring that the entropic state is reduced after each cognitive cycle and establishing a policy in a particular environment that is driven solely by rewards [8]. The entropic reward is defined as the entropic-state decrement after each consecutive cycle. It can be anticipated using a Bayesian filter if the environment is modelled. Thus, learning and planning are two independent ideas in RL for CC, the former using real values of the entropic reward for a particular action and the latter using the predicted entropic reward from the Bayesian filter [8] [12]. It is worth noting that RL can only learn once for each PAC's chosen action; but, RL can prepare for arbitrary number of simulated future cycles and actions [5]. The number of actions that can be performed during planning is limited by considerations such as computing effort and cost, as well as time limits that need planning to be done before a single PAC finishes.

B. *Related Works in Cognitive Control*

1) *Tracking Radar*

The authors of [9] create a cognitive radar system and install a cognitive controller within it. The goal of the research is to show how powerful the cognitive controller's information processing capabilities are, as well as the system's potential tracking performance increases as a result. The literature provides details on the parameters of relevance, such as the measurement and system noise covariances, as well as relevant state-space and entropic-state models for the study's applicability. The cubature Kalman filter (CKF) [13] is used to estimate the perceptor's state covariance matrix, which is then utilised to compute the entropic-state.

The system is stated as having 382 distinct cognitive actions (the number of different transmit-waveform library combinations), each of which affects the measurement noise covariance matrix during a cycle of the PAC. Three different scenarios are investigated in the first trials; the first is the absence of CC on the system (fixed radar waveform). In the second situation, the cognitive controller learns but does not plan; the algorithm merely remembers the entropic reward values from the previous step. The third and final scenario introduces planning by using an explore-only method in the planning phase, which means the learning process is repeated for three different cases: the exploration of only one, two, or three random cognitive actions in each cycle of the PAC.

To reduce the impacts of randomness in the author's experimental simulations, 50 cycles were run across 1000 Monte Carlo runs [9]. Because the total number of cycles is substantially fewer than the number of possible cognitive activities (50 vs. 382) [9], performance in entropic-state reduction in the scenario without planning is not much better than in the absence of CC. However, in the scenario with varying numbers of cognitive actions per cycle, it was discovered that even just one random cognitive action in the planning phase, which is a fraction of the total number of possible cognitive actions, is sufficient to demonstrate a four-order-of-magnitude improvement in entropic-state reduction [9]. In addition, the cases of two or three random cognitive actions showed the same drop in entropy but with faster convergence. In comparison to standard fixed waveform techniques, the planning process in CC was shown to dramatically improve the entropic-state of the model.

Further simulations in the study attempted to see how three distinct CC algorithms, such as dynamic optimization, Q-learning, and the authors' newly suggested algorithm, which combines Q-learning with learning and planning processes, affected the results. The suggested approach, which was configured to plan for three cognitive acts, outperformed both Q-learning and dynamic optimization in terms of minimising the entropic-state while also having a lower computational load [9]. The suggested technique achieved an entropic state value of $10^{0.4}$ in a 250-cycle trial, compared to about $10^{0.7}$ for both Q-learning and dynamic optimization [9]. Finally, the authors point out that while the Q-learning technique is computationally tractable, it can be inefficient in terms of performance. As a result, it may be useful to focus future research on developing algorithms designed specifically for this purpose.

2) Communication-Based Train Control

Communication-based train control (CBTC) systems are automated train control systems that use bidirectional train-ground wireless communications to assure the safe and efficient running of rail vehicles. These solutions aid in the better usage of railway network infrastructure while also improving customer service. However, there are concerns with train-ground communications and train control, which are usually treated as different topics in the literature.

Recent studies have explored combining many concerns into a single problem with the goal of overcoming each challenge using a CC-inspired technique, as in [14]. The authors employ the entropic state to objectively explain the packet delay and drop of information exchanged between the train-ground connection and the train control centre [14]. Wireless local area networks (WLAN) are widely employed in urban rail transit systems around the world as a medium for train-ground communication [15]. The linear-quadratic cost is utilised as a performance metric for train control, and Q-learning is then used to find the best policy based on this metric and the entropic state. In order to characterise high-speed railway and Rayleigh fading, the wireless channels are modelled as finite-state Markov chains with various state transition probability matrices. The CC model is in charge of ensuring that wireless communications and handoffs are reliable and uninterrupted, guaranteeing that the current train receives accurate information about the front train. As a result, the authors believe that adopting CC to increase communication between the train and the control centre will result in a more robust control of CBTC systems in terms of acceleration, deceleration, speed, distance, and emergency brake profiles [14].

With the proposed approach, more reliable velocity management between the system's front and back trains was proved through experimental trials with measurements retrieved from antennas on a train placed within a tunnel and subsequent MATLAB simulations [14]. When compared to alternative control policies like the semi-Markov decision process (SMDP) and greedy policies, which showed tiny disturbances in the difference in front and back train velocities, CC showed completely smooth and significantly safer behaviour. Furthermore, when compared to handoff delays of one second using the SMDP and greedy policies [14], the handoff delay with CC was significantly reduced to 0.2 seconds, half of the train response time parameter. Finally, while looking at the failure rate of the CBTC system under various policies, it is obvious that the CC approach proposed is the most effective due to its 99.78 percent availability. In comparison to the SMDP policy, which has a 10^{-2} unavailability rate, and the greedy policy, which has a 10^{-1} unavailability rate, this figure leads to unavailability rates of the order of 10^{-3} using CC [14]. Overall, the results demonstrate the usefulness of the proposed approach; however, the authors advise that more research is needed to investigate more advanced train-ground communication technologies, such as relaying, in order to increase the performance of CBTC systems.

3) Smart Grid Control

Oozeer and Haykin [16] suggest a CDS as a supervisor for smart grid networks utilising a CC method, as shown in Fig. 2. The authors present a new method for calculating the entropic state that is customised to the smart grid application, and they use it to create a control-sensing mechanism that can recognise and detect incorrect data from sensor measurements in the grid network. Bad measurements caused by erroneous readings, broken hardware components, or power system disruptions can result in a cascade of domino effects that obstruct the state estimation process and can degrade the performance of ordinary control systems [16].

The direct current (DC) state estimator is regarded as the environment in which the CDS acts in the author's suggested framework because it is the recipient of measurements in the network. To classify the observables from the environment, a generative model based on the cumulative sum (CUSUM) is used, followed by a Kalman filter (KF) filter to produce updated estimates for future cycles. The cognitive controller is thus in charge of learning and planning (as illustrated in Fig. 2), as well as providing the network with the ability to prioritise and disregard specific measurements for optimal state estimation by customising the weights assigned to each sensor or metre. By functioning in the opposite direction and independently of the PAC, the shunt cycles facilitate planning. With the use of the memory system in the perceptual and executive sections, this cycle engages both the perceptor and the executive to account for all planned prospective actions in each PAC [16]. The Bayesian Upper Confidence Bounds (Bayes-UCB) algorithm is used to optimise the system's newly tailored entropic state and provide a means for the cognitive controller to learn the best policy of actions, in this case, measurements weights, which are stored in working memory and applied to the system [16].

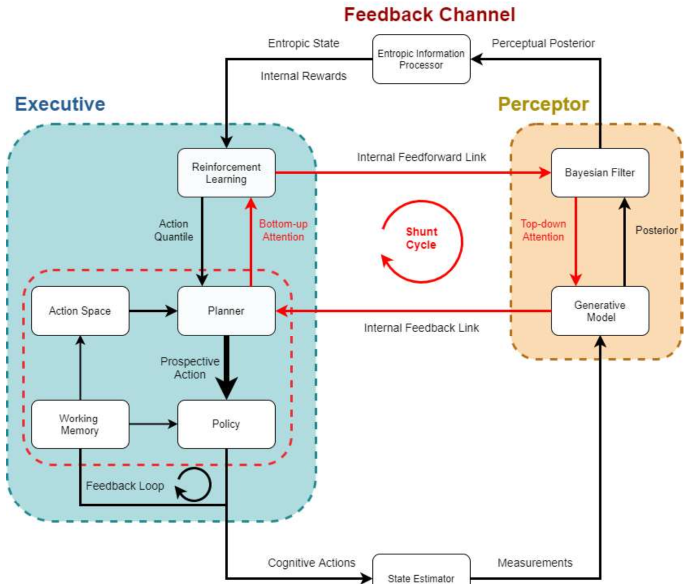

Fig. 2 Structure of a CDS with cognitive control as the supervisor of a smart grid network.

Experimental simulations on a 4-bus network are used to evaluate the suggested CC approach's performance in detecting and correcting erroneous data by changing measurement weights of various metres in the network. With CC, the system was shown to respond dynamically, selecting the optimum collection of metres to obtain readings from at the same time and efficiently assigning the best weight to each measurement for optimal state estimate [16]. Only a few PACs are required when a metre malfunctions in order to learn from the situation and adapt by lowering the weight of the faulty metre.

It is also demonstrated that the cognitive controller's weight assignment to the various measurements is done in such a way that it adapts to the probabilistic characteristics of noisy signals, and that the mean squared error (MSE) of the estimates obtained with the cognitive controller is much lower than without CC [16]. Furthermore, the authors demonstrate that when dealing with cyberattacks such as fake data injection assaults (FDI), the entropic state can be used as a metric to detect such attacks. However, it is noted that the model's structure must be expanded to include CRC in order to effectively deal with and reduce the risk associated with these types of attacks, which the authors address in later works [17] that will be discussed in section IV.

A disadvantage of the suggested framework in dealing with inaccurate measurement data is that it is not scalable to real smart grid networks, which typically have thousands of metres. The rationale provided is that performing an inverse calculation during state estimation is computationally expensive. We believe that machine learning techniques, notably a neural network, might be used to speed up processing and reduce the complexity of the approach. Regardless, the proposed model was more accurate, less prone to false positives, and cost less to compute than existing detection algorithms proposed in [18]. Finally, when scaling up to larger networks, the Bayes-UCB algorithm in the proposed CC model is expected to encounter challenges in terms of response time in determining optimal configurations in the face of metre malfunctions [16]. In this situation, it may be possible to reduce the algorithm's response time by fine-tuning it and increasing the algorithm's sensitivity.

IV. Cognitive Risk Control

A. *Overview of Cognitive Risk Control*

The CRC model adds a subsystem to the executive side of the CC model to allow for more complex reasoning, which necessitates the coining of a new term, the classifier, as shown in Fig. 3. In this illustration, the subsystem is configured in version II, which entails a disturbed cognitive action to the classifier rather than directly to the physical system. The executive memory, in turn, selects a collection of past acts or experiences for the classifier. There are also two pairs of switches, as indicated in Fig. 3, switches 1 and 2, and switches 3 and 4. Switches 1 and 2 are open in version II, preventing the controller from acting directly on the physical system and providing feedback to the executive memory. Instead, as previously stated, a perturbed cognitive action is delivered to the classifier, which is then in charge of making decisions by selecting a past experience that most closely fits the supplied perturbed cognitive action and then updating executive memory [19]. Switches 1 and 2 are closed, but switches 3 and 4 are opened, when the physical system is working without ambiguity, or under version I. In this case, the controller has direct access to the physical system and can change the executive memory.

The classifier's disrupted cognitive action is of probabilistic origin, and the executive memory's projected past events are similarly probabilistic because they are picked at random from its own action space. The Bayesian paradigm is used as a mechanism of decision-making in recognition of these truths, laying the stage for CRC [19]. For each of the past experiences in the specified set, the probability of the perturbed action's

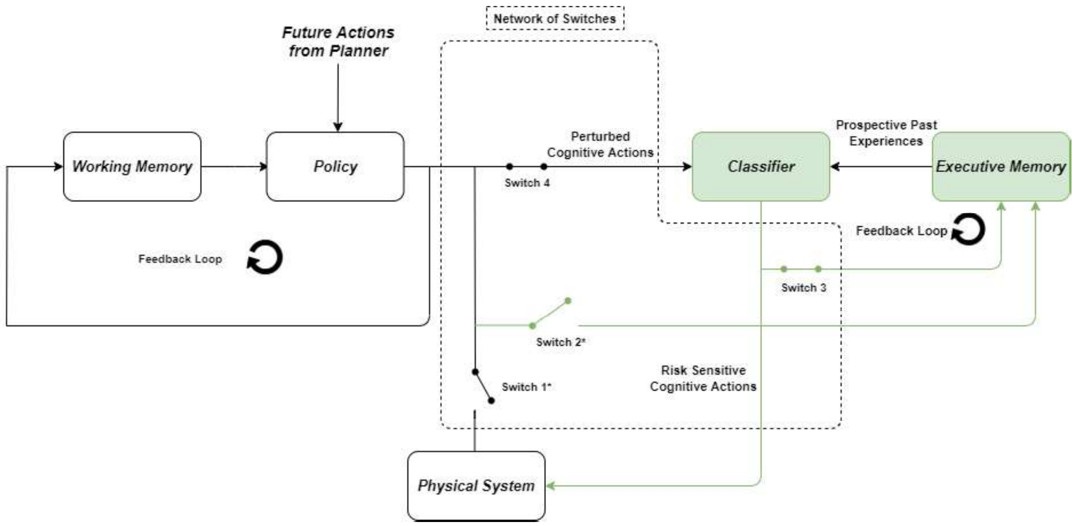

Fig. 3 Structure of the subsystem that is tasked with dealing with risk in the CRC architecture.

posterior given a past experience is determined using Bayes' rule. The risk-sensitive cognitive activity directed to the physical system is then defined as the experience with the highest likelihood [19].

Task-switch control (TSK) is a CRC framework function that exploits the presence of pairs of switches and controls their configuration depending on the presence or absence of uncertainty. The entropic rewards from the feedback channel serve as the foundation for defining TSC methods. The entropic reward in CRC can either be positive or negative, and it can never be zero. These two qualities are crucial in characterising TSC: positive rewards indicate the absence of uncertainty, while negative rewards imply uncertainty [19]. The choice of function to compute the rewards, as well as the tuning of design parameters in the chosen function, are essential considerations when applying CRC to physical systems using this approach. To summarise, when there are no uncertainties, the entropic reward must be positive, so switches 1 and 2 are closed, but switches 3 and 4 are open. When there are uncertainties, the entropic reward must be negative, therefore switches 1 and 2 must be opened and switches 3 and 4 must be closed.

B. Related Works in Cognitive Risk Control

1) Radar and Communications

Feng and Haykin published the first experimental research employing the CRC framework in [20]. The paradigm is studied and used in a cognitive vehicular radar system for self-driving cars by the authors. Recognizing the hazards posed to autonomous vehicles in the presence of uncertainty, the authors attempt to improve the performance of vehicular radar systems in such dangerous situations. The literature discusses the architectural structure of the CRC adapted to the problem of transmit-waveform selection in vehicular radar systems, as well as a simple vehicle-following scenario. A host vehicle is going forward in the stated scenario, and ahead of it is a target vehicle moving in the same direction, both of which are defined by their own velocities and accelerations. Details on state-space dynamics and modelling of the scenario are provided, and we refer the reader directly to the literature in [20] for these specifics. The purpose of the proposed model is to deal with risky events caused by other physical entities robustly when applied as the supervisor for transmit-waveform selection in the radar system.

The authors remove the Bayesian generative model from the perceptual element of the CDS in their work because, in the case of automotive radars, observables are typically taken in a fashion that can be directly processed by the Bayesian filter [20]. As a result, the Bayesian filter has been relocated to the bottom of the perceptor, and the entropic-information processor has been added to take its place and preserve the feedforward link. Aside from that, the suggested work follows the same structure as Fig. 4. The KF is chosen as the Bayesian filter to represent the vehicle-following scenario, and it is formulated according to the transmit-waveform option, which mixes the linear frequency modulated (LFM) waveform with Gaussian amplitude modulation. Invoking Shannon's information theory [20], the entropic state is determined using the filtered posterior from the KF as input. The entropic state uses a defined function to determine entropic or internal incentives, which it subsequently passes on to the executive. The internal incentives are sent through a defined function in the CRC framework's TSC mechanism, which is then subjected to particular conditions and thresholds formulated in the literature to determine the existence or absence of uncertainty. The rest of the methodology follows the standard CC framework described in earlier sections, which is also depicted in Fig. 4 by the red dashed boundary.

The suggested CRC model with Q-learning for RL was compared to alternative systems, such as a radar with fixed transmit-waveform (FTW), the CC framework, and merely Q-learning on its own for waveform creation, using experimental simulations [20]. The root mean squared error (RMSE) was calculated against each model's five states: the velocity and acceleration of the host vehicle, the longitudinal distance between the host and target vehicle, and finally the velocity and acceleration of the target vehicle. Based on simulation findings, the suggested model has the lowest RMSE for each state in the depicted qualitative graphs, with CC and Q-learning performing similarly [20]. These findings show that regardless of the algorithm used, the executive's learning algorithms will

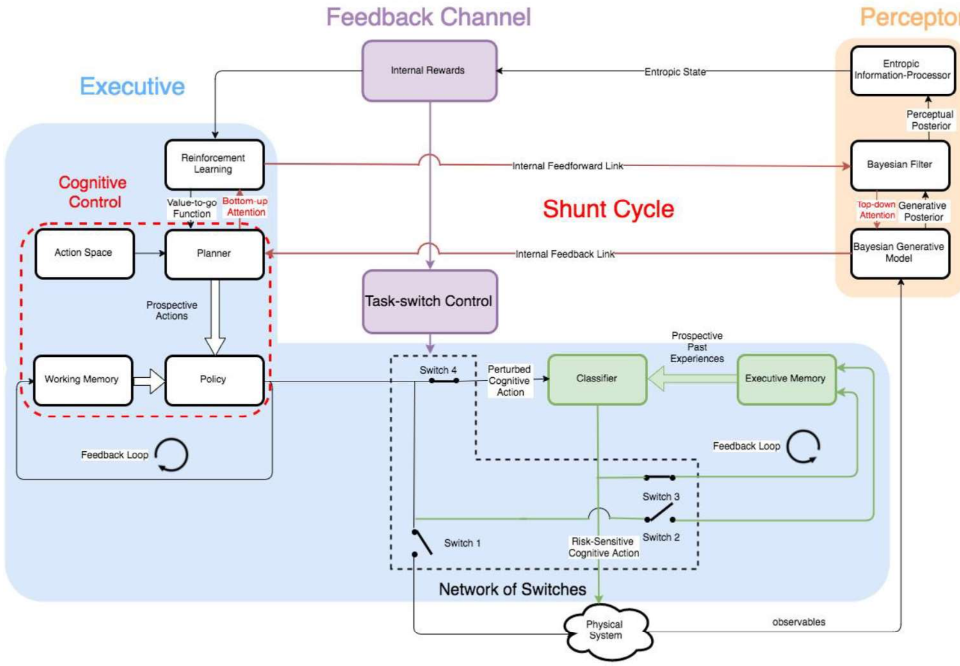

Fig. 4 Structure of the CRC framework. Green coloured elements represent the newly introduced risk-sensitive subsystem.

result in better decision-making and action choices. During the trial, the authors add a structural uncertainty term to the system model to create a dangerous scenario in the experiment, which lasts less than a second. In this circumstance, the Q-learning and CC algorithms were unable to adjust to the uncertainty, resulting in substantial spikes and erratic behaviour in the RMSE, which took upwards of eight seconds to recover from. However, the proposed CRC model relative to the other approaches was only slightly affected in terms of RMSE and recovered within a matter of two or three seconds at most. Overall, the model achieved impressive results and was also deemed a promising alternative to traditional approaches in handling uncertainties and risky events in vehicular radar [20].

These research have culminated in the most current work of Feng and Haykin in [21], in which the authors propose merging cognitive vehicular radar (CVR) and vehicle to vehicle (V2V) communications with a coordinated CRC (C-CRC) model that bridges both systems together. C-CRC investigates the benefits of mutual aid by utilising information emanating from one of the systems that may be useful to the other. In addition, unlike earlier investigations, a nonlinear target-tracking model is used, and the analysis is done with a CKF [21]. The formula for the interference measure in V2V communications in this approach includes tracking findings as well as other practical aspects inferred from those results, such as vehicle motion and channel availability. The CVR also relies on communication system data to determine whether it is operating on a one-vehicle or two-vehicle model, with the latter indicating the presence of a second vehicle engaged in target tracking [21]. Each system in the C-CRC model has its own TSC mechanism and is implemented with a risk-sensitive subsystem. Furthermore, the TSC in each system plays a part in determining what information should be transferred from one system to the next for their dual system.

The authors show that the suggested C-CRC model outperforms previous radar techniques such as FTW and Q-learning in minimising the peak RMSE achieved in tracking longitudinal distance when faced with uncertainty by up to 70% and 67 percent, respectively. Although the typical CRC design in [22] had equivalent performance, it was still 41 percent worse to the C-CRC in terms of RMSE peak reduction. C-performance CRC's is improving across the board, including tracking performance in terms of the utility of power selection in vehicle and jammer communications, total regret from channel selection, and finally, user utility.

However, the ability of V2V communications to keep up with busy networks in specific locations or conditions has been noted in the literature. With less spectrum opportunity, user utility decreases while jammer utility grows, according to further examination of studies relating to the effects of channel availability on power and channel selection [21]. This scenario also leads to a greater regret measure for the host vehicle, a reduced multi-armed bandit (MAB) related reward, and increased channel switching costs. As a result, vehicle networks with several entities sharing available wireless resources in a local area present fascinating and practical V2V performance issues that demand additional research. The authors also highlight that security vulnerabilities in large-scale adversarial CAV networks would be investigated in the future as part of their research efforts.

2) *Cybersecurity in Smart Grids*

In [17], Oozeer et al. improved on their CC technique for smart grid attack detection from [16] by proposing an upgraded CRC-enabled model capable of also defending against such assaults. The entropic state was utilised as a metric in the initial experiments to detect the existence of FDI attacks, which was signified by the entropic state dropping below a predetermined threshold, setting the stage for TSC in the extended CRC version of the model. When an FDI assault is detected and TSC is triggered in the extended framework, the cognitive controller is deemed inactive, while CRC is activated to protect against these attacks [17].

The authors note that the action space involved in this scenario involving CRC differs from CC, recognising that FDI attacks try to produce deviation in specific states to trigger a cascade of incorrect control decisions. Unlike the cognitive controller, which has an action space of possible measurement weights, CRC includes picking tuning parameters to be applied to the DC system's configuration matrix [17]. However, if an assault has been discovered, the predictor or classifier must first recognise the states that are at risk. As explained in the literature [17], the affected states are recognised by whether they exceed the maximum deviation allowed by a formulation based on each estimate's mean recorded in the perceptual memory. Following that, once the attacker states have been identified, the planner must carefully pick tuning parameters in the columns of the DC system's configuration matrix corresponding to the impacted states without interrupting the estimation of other states [17]. Each shunt cycle is dedicated to resolving the hazards associated with one of the states at a time during this planning phase of operation, and a new reward connected with a specific action in the cycle is determined. The Bayes-UCB algorithm, as suggested in [16], is then in charge of optimising the policy in such a way that it prioritises actions that will return current attack states to a condition that is closest to past perceptual memories. Similar to CC, the actions that get the highest quantile from the Bayes-UCB algorithm are stored in the working memory and applied once the shunt cycles have expired. Once the impacted states have been restored to acceptable levels, the risk is considered controlled, and no further actions will be taken in those columns of the system configuration matrix. Finally, once the attacks have been determined as having finished, a mechanism is implemented by supplying the TSC with memory and a watchdog timer that restores the system configuration to its original state and marks the end of the current PAC [17].

The experimental simulations used in the author's research are comparable in configuration to those used in their prior investigations [16], which used IEEE 4-bus and 14-bus networks. The literature shows how the cognitive controller and CRC can operate together in the 4-bus network to bring FDI attacks under control once they've been introduced to the system. The network configuration matrix of the system is detailed in the study along with other pertinent parameters, and it is mentioned that the simulations run for a total of 2000 seconds while allowing for 15 shunt cycles in each PAC for learning and planning. The action space for CRC consists of 63 different tuner values, each of which can be tuned with a specific range of values for relevant columns of the network configuration matrix. In the 4-bus network, three states are measured, and an assault is launched on the first two states 1000 seconds into the simulation, lasting 300 seconds. The phase angles of the desired states are shifted by predefined values to imitate FDI attacks [17]. Before the attacks, the CRC only needs 20 cycles to get the estimations or measurements for the impacted states under control and restored to a tolerable threshold [17]. When the attacks are over, the suggested model continues to run under CRC for another 39 cycles, according to the authors, because the model ensures that conditions for matching current and previous events are met. The altered network configuration matrix is then restored to its original condition before TSC and CRC are triggered.

However, one of the proposed studies' disadvantages is that when other forms of FDI attacks, such as the slowly evolving ramp attack, are used, the detection time can be altered and increased [17]. Furthermore, scaling up the architecture to larger and more realistic networks will necessitate more shunt cycles in each PAC, creating major processing resource and efficiency difficulties. This problem is exacerbated by the fact that other types of FDI attacks, such as developing ramp attacks, necessitate a longer sample time for the DC estimator to overcome them. Another use of this CDS that the author's investigations have not looked into is not just identifying attacked states, but also identifying which sensors or metres have been attacked. Finally, there is a suggestion in the literature that applying a predetermined threshold on the absolute estimated error of the estimated readings might be used to identify the attacked metres in a network [17].

V. Concluding Remarks

Cognitive dynamic systems, as well as their two particular functions, cognitive control and cognitive risk control, were comprehensively examined in this study. This work is the first attempt to compile data from the voluminous literature published in this young and evolving topic. The goal of this study's methodology was to encourage and facilitate additional research into cognitive dynamic systems. We did so by alerting the reader about advancements in each specialized field, outlining the benefits and limits of the surveyed literature, and providing suggestions and directions for future research. Finally, the contents and outcomes of this survey will serve as a foundation for future research, and will hopefully be useful to other academics working in this fascinating new subject.

VI. References

[1] S. Haykin, Cognitive Dynamic Systems: Perception-action Cycle, Radar and Radio, Cambridge, UK: Cambridge University Press, 2012.

[2] S. Haykin, "Cognitive Dynamic Systems," *International Journal of Cognitive Informatics and Neural Intelligence,* vol. 5, no. 4, pp. 33-43, 2011.

[3] S. Haykin, "Cognitive Radio: Brain-Empowered Wireless Communications," *IEEE Journal on Selected Areas in Communications,* vol. 23, no. 2, pp. 201-220, 2005.

[4] S. Haykin, "Cognitive radar: a way of the future," *IEEE Signal Processing Magazine,* vol. 23, no. 1, pp. 30-40, 2006.

[5] S. Haykin, M. Fatemi, P. Setoodeh and Y. Xue, "Cognitive Control," *Proceedings of the IEEE,* vol. 100, no. 12, pp. 3156-3169, 2012.

[6] S. Haykin, "The Cognitive Dynamic System for Risk Control [Point of View]," *Proceedings of the IEEE,* vol. 105, no. 8, pp. 1470-1473, 2017.

[7] S. Haykin and J. M. Fuster, "On Cognitive Dynamic Systems: Cognitive Neuroscience and Engineering Learning From Each Other," *Proceedings of the IEEE,* vol. 102, no. 4, pp. 608-628, 2014.

[8] S. Haykin, A. Amiri and M. Fatemi, "Cognitive control in cognitive dynamic systems: A new way of thinking inspired by the brain," in *IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning*, Orlando, FL, 2014.

[9] M. Fatemi and S. Haykin, "Cognitive Control: Theory and Application," *IEEE Access,* vol. 2, pp. 698-710, 2014.

[10] S. Haykin, "Cognitive Dynamic Systems," *Proceedings of the IEEE,* vol. 94, no. 11, pp. 1910-1911, 2006.

[11] C. E. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal,* vol. 27, pp. 379-423, 623-656, 1948.

[12] R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, Cambridge, MA: MIT Press, 2018.

[13] I. Arasaratnam and S. Haykin, "Cubature Kalman Filters," *IEEE Transaction on Automatic Control,* vol. 54, no. 6, pp. 1254-1269, 2009.

[14] H. Wang, F. R. Yu, L. Zhu, T. Tang and B. Ning, "A Cognitive Control Approach to Communication-Based Train Control Systems," *IEEE Transactions on Intelligent Transportation Systems,* vol. 16, no. 4, pp. 1676-1689, 2015.

[15] L. Zhu, F. R. Yu, B. Ning and T. Tang, "Cross-Layer Handoff Design in MIMO-Enabled WLANs for Communication-Based Train Control (CBTC) Systems," *IEEE Journal on Selected Areas in Communications,* vol. 30, no. 4, pp. 719-728, 2012.

[16] M. I. Oozeer and S. Haykin, "Cognitive Dynamic System for Control and Cyber-Attack Detection in Smart Grid," *IEEE Access,* vol. 7, pp. 78320-78335, 2019.

[17] M. I. Oozeer and S. Haykin, "Cognitive Risk Control for Mitigating Cyber-Attack in Smart Grid," *IEEE Access,* vol. 7, pp. 125806-125826, 2019.

[18] R. Xu, Z. Guan, L. Wu, J. Wu and X. Du, "Achieving Efficient Detection Against False Data Injection Attacks in Smart Grid," *IEEE Access,* vol. 5, pp. 13787-13798, 2017.

[19] S. Haykin, J. M. Fuster, D. Findlay and S. Feng, "Cognitive Risk Control for Physical Systems," *IEEE Access,* vol. 5, pp. 14664-14679, 2017.

[20] S. Feng and S. Haykin, "Cognitive Risk Control for Transmit-Waveform Selection in Vehicular Radar Systems," *IEEE Transactions on Vehicular Technology,* vol. 67, no. 10, pp. 9542-9556, 2018.

[21] S. Feng and S. Haykin, "Coordinated Cognitive Risk Control for Bridging Vehicular Radar and Communication Systems," *IEEE Transactions on Intelligent Transportation Systems,* pp. 1-16, 2020.

[22] S. Feng and S. Haykin, "Cognitive Risk Control for Anti-Jamming V2V Communications in Autonomous Vehicle Networks," *IEEE Transactions on Vehicular Technology,* vol. 68, no. 10, pp. 9920-9934, 2019.

[23] S. Soatto, "Actionable Information in Vision," in *IEEE 12th International Conference on Computer Vision*, Kyoto, Japan, 2009.

[24] R. A. Fisher, "On the mathematical foundation of theoretical statistics," *Philosophical Transactions of the Royal Society of London, Series A,* vol. 222, pp. 309-368, 1922.

[25] M. Li and P. Vitányi, An Introduction to Kolmogorov Complexity and Its Applications, 3rd ed., New York, NY: Springer-Verlag, 2008.

[26] J. J. Gibson, "The Myth of Passive Perception: A Reply to Richards," *Philosophy and Phenomenological Research,* vol. 37, no. 2, pp. 234-238, 1976.

[27] J. J. Gibson, The Ecological Approach to Visual Perception, Classic ed., New York, NY: Psychology Press, 2014.

[28] Y. Niv, "Reinforcement learning in the brain," *Journal of Mathematical Psychology,* vol. 53, no. 3, pp. 139-154, 2009.

[29] P. Dayan and Y. Niv, "Reinforcement learning: The Good, The Bad and The Ugly," *Current Opinion in Neurobiology,* vol. 18, no. 2, pp. 185-196, 2008.

[30] D. J. Surmeier, J. Plotkin and W. Shen, "Dopamine and synaptic plasticity in dorsal striatal circuits controlling action selection," *Current Opinion in Neurobiology,* vol. 19, no. 6, pp. 621-628, 2009.

[31] R. Poovendran, K. Sampigethaya, S. K. S. Gupta, I. Lee, K. V. Prasad, D. Corman and J. L. Paunicka, "Special issue on cyber-physical systems," *Proceedings of the IEEE,* vol. 100, no. 1, pp. 6-12, 2012.

[32] M. Chen, M. Mozaffari, W. Saad, C. Yin, M. Debbah and C. S. Hong, "Caching in the Sky: Proactive Deployment of Cache-Enabled Unmanned Aerial Vehicles for Optimized Quality-of-Experience," *IEEE Journal on Selected Areas in Communications,* vol. 35, no. 5, pp. 1046-1061, 2017.

[33] B. Paden, M. Čáp, S. Z. Yong, D. Yershov and E. Frazzoli, "A Survey of Motion Planning and Control Techniques for Self-Driving Urban Vehicles," *IEEE Transactions on Intelligent Vehicles,* vol. 1, no. 1, pp. 33-55, 2016.

[34] S. Feng and S. Haykin, "Cognitive Dynamic System for Future RACE Vehicles in Smart Cities: A Risk Control Perspective," *IEEE Internet of Things Magazine,* vol. 2, no. 1, pp. 14-20, 2019.

[35] S. Feng and S. Haykin, "Anti-Jamming V2V Communication in an Integrated UAV-CAV Network with Hybrid Attackers," in *IEEE International Conference on Communications*, Shanghai, China, 2019.

[36] S. Feng and S. Haykin, "V2V Communication-Assisted Transmit-Waveform Selection for Cognitive Vehicular Radars," in *IEEE Canadian Conference of Electrical and Computer Engineering*, Edmonton, AB, 2019.

[37] P. Auer, N. Cesa-Bianchi and P. Fischer, "Finite-time Analysis of the Multiarmed Bandit Problem," *Machine Learning,* vol. 47, pp. 235-256, 2002.

[38] J.-Y. Audibert, R. Munos and S. Csaba, "Exploration-exploitation tradeoff using variance estimates in multi-armed bandits," *Theoretical Computer Science,* vol. 410, no. 19, pp. 1876-1902, 2009.

[39] A. Humayed, J. Lin, F. Li and B. Luo, "Cyber-Physical Systems Security—A Survey," *IEEE Internet of Things Journal,* vol. 4, no. 6, pp. 1802-1831, 2017.

[40] X. Yu and Y. Xue, "Smart Grids: A Cyber–Physical Systems Perspective," *Proceedings of the IEEE,* vol. 104, no. 5, pp. 1058-1070, 2016.

[41] G. Liang, S. R. Weller, J. Zhao, F. Luo and Z. Y. Dong, "The 2015 Ukraine Blackout: Implications for False Data Injection Attacks," *IEEE Transactions on Power Systems,* vol. 32, no. 4, pp. 3317-3318, 2017.

[42] Y. Liu, P. Ning and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Transactions on Information and System Security,* vol. 14, no. 1, pp. 1-33, 2011.

[43] J. Hao, R. J. Piechocki, D. Kaleshi, W. H. Chin and Z. Fan, "Sparse Malicious False Data Injection Attacks and Defense Mechanisms in Smart Grids," *IEEE Transactions on Industrial Informatics,* vol. 11, no. 5, pp. 1-12, 2015.

[44] S. Cui, Z. Han, S. Kar, T. T. Kim, H. V. Poor and A. Tajer, "Coordinated data-injection attack and detection in the smart grid: A detailed look at enriching detection solutions," *IEEE Signal Processing Magazine,* vol. 29, no. 5, pp. 106-115, 2012.

[45] W. Sun, F. R. Yu, T. Tang and S. You, "A Cognitive Control Method for Cost-Efficient CBTC Systems With Smart Grids," *IEEE Transactions on Intelligent Transportation Systems,* vol. 18, no. 3, pp. 568-582, 2017.

[46] X. Wang, L. Liu, T. Tang and W. Sun, "Enhancing Communication-Based Train Control Systems Through Train-to-Train Communications," *IEEE Transactions on Intelligent Transportation Systems,* vol. 20, no. 4, pp. 1544-1561, 2019.

[47] L. Lei, J. Lu, Y. Jiang, X. S. Shen, Y. Li, Z. Zhong and C. Lin, "Stochastic Delay Analysis for Train Control Services in Next-Generation High-Speed Railway Communications System," *IEEE Transactions on Intelligent Transportation Systems,* vol. 17, no. 1, pp. 48-64, 2016.

[48] M. L. Tuballa and M. L. Abundo, "A review of the development of Smart Grid technologies," *Renewable and Sustainable Energy Reviews,* vol. 59, pp. 710-725, 2016.

[49] R. Deng, J. Chen, X. Cao, Y. Zhang, S. Maharjan and S. Gjessing, "Sensing-Performance Tradeoff in Cognitive Radio Enabled Smart Grid," *IEEE Transactions on Smart Grid,* vol. 4, no. 1, pp. 302-310, 2013.

[50] X. Fang, Y. Han, J. Wang and Q. Zhao, "A cognitive control approach for microgrid performance optimization in unstable wireless communication," *Neurocomputing,* vol. 355, pp. 168-182, 2019.