This is an Accepted Manuscript of an article published by Taylor & Francis in **CYBERNETICS AND SYSTEMS** on 20.01.2020, available online: https://www.tandfonline.com/doi/full/10.1080/01969722.2019.1705553

Postprint of: de Oliveira C., Giustozzi F., Zanni-Merk C., Sanin C., Szczerbicki E.: Stream Reasoning to Improve Decision-Making in Cognitive Systems CYBERNETICS AND SYSTEMS Vol. 51, iss. 2 (2020), pp.214-231

Stream Reasoning to Improve Decision Making in Cognitive Systems

Caterine Silva de Oliveira[1], Franco Giustozzi[2], Cecilia Zanni-Merk[2], Cesar Sanin[1], and Edward Szczerbicki[3]

[1] *Department of Mechanical Engineering, University of Newcastle, Callaghan, NSW, Australia*

(Caterine.SilvaDeOliveira@uon.edu.au, cesar.sanin@newcastle.edu.au)

Address: ES320, Faculty of Engineering and Built Environment, University of Newcastle, Callaghan, NSW 2308, Australia;

[2] *Normandie Université, INSA Rouen, LITIS, Rouen 76000, France*

(franco.giustozzi@insa-rouen.fr, cecilia.zanni-merk@insa-rouen)

[2] *Faculty of Management and Economics, Gdansk University of Technology, Gdansk, Poland*

(edward.szczerbicki@zie.pg.gda.pl)

Stream Reasoning to Improve Decision Making in Cognitive Systems

Cognitive Vision Systems have gained a lot of interest from industry and academia recently, due to their potential to revolutionize human life as they are designed to work under complex scenes, adapting to a range of unforeseen situations, changing accordingly to new scenarios and exhibiting prospective behaviour. The combination of these properties aims to mimic the human capabilities and create more intelligent and efficient environments. Contextual information plays an important role when the objective is to reason such as humans do, as it can make the difference between achieving a weak, generalized set of outputs and a clear, target and confident understanding of a given situation. Nevertheless, dealing with contextual information still remains a challenge in cognitive systems applications due to the complexity of reasoning about it in real time in a flexible but yet efficient way. In this paper, we enrich a cognitive system with contextual information coming from different sensors and propose the use of stream reasoning to integrate/process all these data in real time, and provide a better understanding of the situation in analysis, therefore improving decision making. The proposed approach has been applied to a Cognitive Vision System for Hazard Control (CVP-HC) which is based on Set of Experience Knowledge Structure (SOEKS) and Decisional DNA (DDNA) and has been designed to ensure that workers remain safe and compliant with Health and Safety policy for use of Personal Protective Equipment (PPE).

Keywords: Cognitive Vision Systems, Knowledge Representation, SOEKS, DDNA, PPE compliance, Hazard Control, Stream Reasoning, Industry 4.0

Introduction and Background

Cognitive Vision Systems have gained a lot of interest from industry and academia recently due to their potential to revolutionize human life, as they are designed to work under complex scenes, adapting to a range of unforeseen situations, changing accordingly to new scenarios and exhibiting prospective behaviour (Sanin, Haoxi, Shafiq, Waris, de Oliveira & Szczerbicki, 2018). The combination of these properties aims to mimic the human capabilities and create more intelligent and efficient environments (Vernon, 2006). Contextual information plays an important role when the objective is to perceive the environment and reason such as humans do, as it can make the difference between achieving a weak, generalized set of outputs and a clear, target and confident understanding of a given situation (Chmaj, 2019).

Nonetheless, dealing with contextual information still remains a challenge in cognitive systems applications due to the complexity of reasoning about it in real time in a flexible but yet efficient way. It involves gathering visual and other sensorial information available and translating it into knowledge to be useful. Moreover, past experiences is also an important element of this process and must also be considered if the objective is to improve perception (Gregory, 1973).

In this context, approaches have been proposed aiming to gather and integrate contextual knowledge to prior models, improving learning and reasoning of systems and providing promising guidelines to improve decision making process in various domains (Fritsch, 2003; Bauckhage, Wachsmuth, Hanheide, Wrede, Sagerer, Heidemann & Ritter, 2008; Crowley, Coutaz, Rey & Reignier, 2002). However, in real time industrial applications, when an incident occurs, systems have only a few minutes to forge a representation of the given issue, gather information on the situation, analyse the incident and undertake the correcting actions (Brézillon, 2003). Unfortunately, most of available methodologies cannot guarantee real time performance in such complex situation.

Stream reasoning appeared as an initiative to overcome this issue. Stream reasoning is, in short, the task of continuously deriving conclusions based on the continuous processing of data (Beck, Dao-Tran, Eiter, & Folie, 2018). For application in Cognitive Vision Systems, this approach may enables the integration of rich visual content with sensor data from a variety of sensors in different frequencies for the creation of a context in which a visual event is occurring, deriving useful information, reasoning on it and inferring new knowledge.

The approach proposed in this paper makes use of the stream reasoning presented in Giustozzi, Saunier and Zanni-Merk (2019) work in combination with rich structured visual and non-visual knowledge for applicability in cognitive systems. The objective is to integrate and process all input data in real time, and provide a better understanding of the situation in analysis, therefore improving decision making. The proposed approach has been applied to a Cognitive Vision System for Hazard Control (CVP-HC) which is based on Set of Experience Knowledge Structure (SOEKS) and Decisional DNA (DDNA) and has been designed to ensure that workers remain safe and compliant with Health and Safety policy for use of Personal Protective Equipment (PPE).

This paper is organized as follow: In Section 2, some fundamental concepts are presented, including cognitive technologies, the challenge of representation and management of knowledge in these systems, as well as the stream reasoning to process contextual information for inference of new knowledge. In Section 3 a case study for the case of PPE compliance is presented, including its applicability and design. In Section 4 experimental results achieved so far are discussed. Finally, in Section 5 conclusions and future work are presented.

Fundamental Concepts

In order to offer a more complete view, we briefly introduce concepts that have driven the proposed research as well as the technologies involved.

Cognitive Systems

The use of computer vision techniques can support automatic detection and tracking of objects and people with reasonable accuracy (Han & Lee, 2013; Ciresan, Meier, Masci, Maria Gambardella & Schmidhuber, 2011; Little, Jargalsaikhan, Clawson, Nieto, Li, Direkoglu & Liu 2013; Krizhevsky, Sutskever & Hinton, 2012; Mosberger, Andreasson & Lilienthal, 2013). Visual sensing facilities, such as video cameras can gather a large amount of data, such as video sequences or digitized visual information that, with support of machine learning technologies and powerful machines, can operate in real time (Chen, Hoey, Nugent, Cook, & Yu, 2012). For those reasons, computer vision systems have been a research focus for a long time in surveillance systems, human detection, and tracking.

However, computer vision systems have their own inherent limits, especially those whose task is to work in unidentified environments and deal with unknown scenarios and specifications. Besides the significant improvements in computer vision technologies, they are still challenged by issues such as occlusion or position accuracy; and background changes result in the necessity of adapting the algorithms for different conditions, clients and situations. To date, the creation of a general-purpose vision system with the robustness and resilience comparable to human vision still remains a challenge (Mosberger, Andreasson & Lilienthal, 2013).

In this context, methods incorporating prior knowledge and context information have gained interest. The understanding about scene composition in an image (which set of objects are present) can improve recognition performance about the scene where they are inserted (Zambrano, Toro, Nieto, Sotaquirá, Sanín & Szczerbicki, 2015). For instance, the presence of multiple cutlery items in an image can aid the recognition of a kitchen image. This relationship is held both ways, as contextual knowledge can also offer insights about the function of an object in a scene, reducing the impacts of sensor noise or occlusions (Aditya, Yang, Baral, Aloimonos & Fermuller, 2017). These technologies are known as knowledge-based systems. For instance, an automatic semantic and flexible annotation service able to work in a variety of video analysis with little modification to the code using Set of Experience Knowledge Structure (SOEKS) was proposed in work by Zambrano et al. (2015). This system is a pathway towards cognitive vision and it is composed, basically, by the combinations of detection algorithms and an experience based approximation.

The design of a general-purpose vision system with the robustness and resilience of the human vision is still a challenge. One of the latest trends in computer vision research to mimic the human-like capabilities is the joining of cognition and computer vision into cognitive computer vision. Cognitive Systems have been defined as “a system that can modify its behaviour on the basis of experience” (Hollnagel & Woods, 2005). Although, most experts tend to agree that such systems only exists in theory, that is, systems that can independently process, reason and create in the same capacity as the human brain has not yet been implemented successfully (Cole, 1990).

In this scenario, the concept of Augmented Intelligence, also known as Cognitive Augmentation or Intelligence Amplification (IA) comes into play (Ashby, 1961). For any specific application humans being and machines have both their own strengths and weaknesses. Machines are very efficient in numerical computation, information retrieval, statistical reasoning, with almost unlimited storage. Machines can capture many categories of information from the environment through various sensors, such as range sensors, visual sensors, vibration sensors, acoustic sensors, and location sensors (Yu, Pan, Gong, Xu, Zheng, Hua & Wu, 2016). On the other hand, humans have their own cognitive capabilities which includes consciousness, problem-solving, learning, planning, reasoning, creativity, and perception. These cognitive functions allows humans to learn from last experiences and use this experiential knowledge to adapt to new situations and to handle abstract ideas to change their environment. Therefore, the combination of both human experiential knowledge and information collected by a system can be used to enhance smartness of systems and for improved decision making (Pathak, 2017).

Knowledge Representation for Cognitive Systems

The implementation of cognitive vision systems require the design of functionalities for knowledge engineering (acquisition and formalism), recognition and categorization, reasoning about events for decision making, and goal specification, all of which are concerned with the semantics of the relationship between the visual agents and their environments i.e. context (Vernon, 2006). These functionalities direct cognitive vision systems towards purposeful behaviour, adaptability, anticipation, such as human beings. In this context, knowledge and leaning are central to cognitive vision. To be readily articulated, codified, accessed and shared, knowledge must be represented in an explicit and structured way (Brézillon & Pomerol, 1999). In addition, the choice of a suitable representation greatly facilitates obtaining methods that efficiently learn the relevant information available. Therefore, an appropriate knowledge representation is crucial for the success in designing of cognitive systems.

Nevertheless, most approaches that have been proposed on past years, even though they present some principles for intelligent cognitive vision, they fail in providing a unique standard that could integrate image/video modularization, its virtualization, and capture its knowledge (Sanin & Szczerbicki, 2009). To address these issues an experience-based technology that allows a standardization of image/video and the entities within together with any other information as a multi-source knowledge representation (required for the further development of cognitive vision) without limiting their operations to a specific domain and/or following a vendor's specification has been proposed (de Oliveira, Sanin & Szczerbicki, 2019). This representation supports mechanisms for storing and reusing experience gained during cognitive vision decision-making processes through a unique, dynamic, and single structure called Decisional DNA (DDNA) (Sanin, Toro, Haoxi, Sanchez, Szczerbicki, Carrasco & Mancilla-Amaya, 2012). DDNA makes use of Set of Experience (SOE) in an extended version for the use of storing formal decision events related to image and video. DDNA and SOE provide a knowledge structure that has been proven to be multi-domain independent (Sanin & Szczerbicki, 2008).

Set of Experience Knowledge Structure (SOEKS) and Decisional DNA (DDNA)

The Set of Experience Knowledge Structure (SOEKS) is a knowledge representation structure created to acquire and store formal decision events in a structured and explicit way. It is composed by four key elements: variables, functions, constraints, and rules. Variables are commonly used to represent knowledge in an attribute-value form, following the traditional approach for knowledge representation. Functions, Constraints, and Rules of SOEKS are ways of relating variables. Functions define relationships between a set of input variables and a dependent variable; thus, SOEKS uses functions as a way to create links among variables and to build multi-objective goals. Constraints are functions that act as a way to limit possibilities, limit the set of possible solutions and control the performance of the system in relation to its goals. Lastly, rules are relationships that operate in the universe of variables and express the condition-consequence connection as “if-then-else” and are used to represent inferences and associate actions with the conditions under which they should be implemented (Sanin & Szczerbicki, 2009). Rules are also ways of inputting expert knowledge into the system. The Decisional DNA consists in a structure capable of capturing decisional fingerprints of an individual or organization and has the SOEKS as its basis. Multiple Sets of Experience can be collected, classified, organized and then grouped into decisional chromosomes, which accumulate decisional strategies for a specific area of an organization. The set of chromosomes comprise, finally, what is called the Decisional DNA (DDNA) of the organization (Sanin, Toro, Haoxi, Sanchez, Szczerbicki, Carrasco & Mancilla-Amaya, 2012).

Stream Reasoning Engines

Data streams have become more and more important as a basis for higher level decision processes that require complex reasoning over data streams and rich background knowledge (Stuckenschmidt, Ceri, Della Vall & Harmelen, Milano, 2019). New data is continually being produced by sensors and humans. A stream is a sequence of incrementally available data. Streaming data is dynamic, temporal, spatial and heterogeneous in nature. In order to integrate these data from multiple data sources, the Semantic Sensor Web (SSW) proposed by Sheth, Henson and Sahoo (2008) introduces semantic annotations for describing: (i) the data produced by the sensors, introducing spatial, temporal, or situation/context semantics; and (ii) the sensors and the sensor networks that provide such data. Furthermore, there are also works on defining suitable ontologies for data and sensors to enable both the integration of data from multiple sensor networks and external sources, and reasoning on such data. As an example, the

W3C Semantic Sensor Network Incubator Group developed an ontology to describe sensors and sensor networks, the Semantic Sensor Network Ontology (SSN) (Haller, Janowicz, Cox, Lefranc¸ois, Taylor, Le Phuoc, Lieberman, Garc´ıa-Castro, Atkinson & Stadler, 2018). However, current solutions to perform reasoning on ontologies are limited to work on rather static scenarios.

Stream reasoning appeared as an initiative to perform reasoning over these streams to draw conclusions and make decisions in real-time. Since streams are conceptually infinite, this reasoning has to be done incrementally as new information becomes available. It has been applied in many fields, such as smart cities, to process and understand the information relevant for the life of a city and use it to make the city services run better and faster (Tallevi-Diotallevi, Kotoulas, Foschini, Lecue & Corradi, 2013; Lecue, Kotoulas & Aonghusa, 2012), remote health monitoring, to generate automated and personalized systems for remote patient monitoring (Calbimonte, Ranvier, Dubosson & Aberer, 2017; Shojanoori & Juric, 2013), maritime safety and security, to represent and to perform reasoning over ship trajectories (Santipantakis, Vlachou, Doulkeridis, Artikis, Kontopoulos & Vouros, 2018), semantic analysis of social media, to extend traditional analysis based on graphs enriching the connections between people and concepts with semantic annotations (Ereteo, Buffa, Gandon & Corby, 2009; Mika, 2005), among others.

Sensor data applied to cognitive systems represents an ideal scenario for stream reasoning mainly for two reasons. Firstly, the amount of data collected from sensors is considerable, and it is produced at high (and low) frequencies. Secondly, integrating data coming from different sensors (and from different sensor networks) that measure different properties of the environment is necessary in many settings for deriving useful information such as the detection of abnormal situations or risk situations.

Personal Protective Equipment Safety Compliance

Hazards are present in all workplaces and can result in serious injuries, short and long-term illnesses, or death (Safe Work Australia, 2012). Reports HSE UK report has shown that over 80% of reported workplace injuries are sustained due to a person not wearing correct protective clothing (Health and Safety Executive, 2018). In this context, the verification of PPE compliance becomes essential in the management of safety to ensure the occupational health of workers. Technologies to support its practical and automated implementation have emerged as a need, but the current technologies available still face considerable limitations (Deloy, 2005).

The combination of vision and sensor data together with the resulting necessity for explicit and formal representations builds a central element of an autonomous system for detection and tracking of labourers in workplaces environments. To be able to perform in a variety of plants and scenes, making sure employees remain safe and compliant with Health & Safety policy without the necessity of recoding the application for each specific case scenario, the system must be adaptable and perceive the environment as automatically as possible and change its behaviour accordingly. However, computer vision systems have their own inherent limits, especially those whose task is to work in unidentified environments and deal with unknown scenarios and specifications (de Oliveira, Sanin & Szczerbicki, 2018).

The gaps of current systems may be filled by connecting the probabilistic area of detection of events with the logical area of formal reasoning in a Cognitive Vision Platform for Hazard Control (CVP-HC) (de Oliveira, Sanin & Szczerbicki, 2018). This platform verifies the PPE compliance in variety of video analysis scenarios whilst meeting specific safety requirements of industries (de Oliveira, Sanin & Szczerbicki, 2019).

The proposed system is based on the Set of Experience Knowledge Structure (SOEKS or SOE in short) and Decisional DNA (DDNA) and uses the stream reasoning based approach proposed by Giustozzi, Saunier and Zanni-Merk (2019) to integrate and process all input data in real time, and provide a better understanding of the situation in analysis, therefore improving decision making. The overall scheme of the proposed approach is presented in Figure 1.

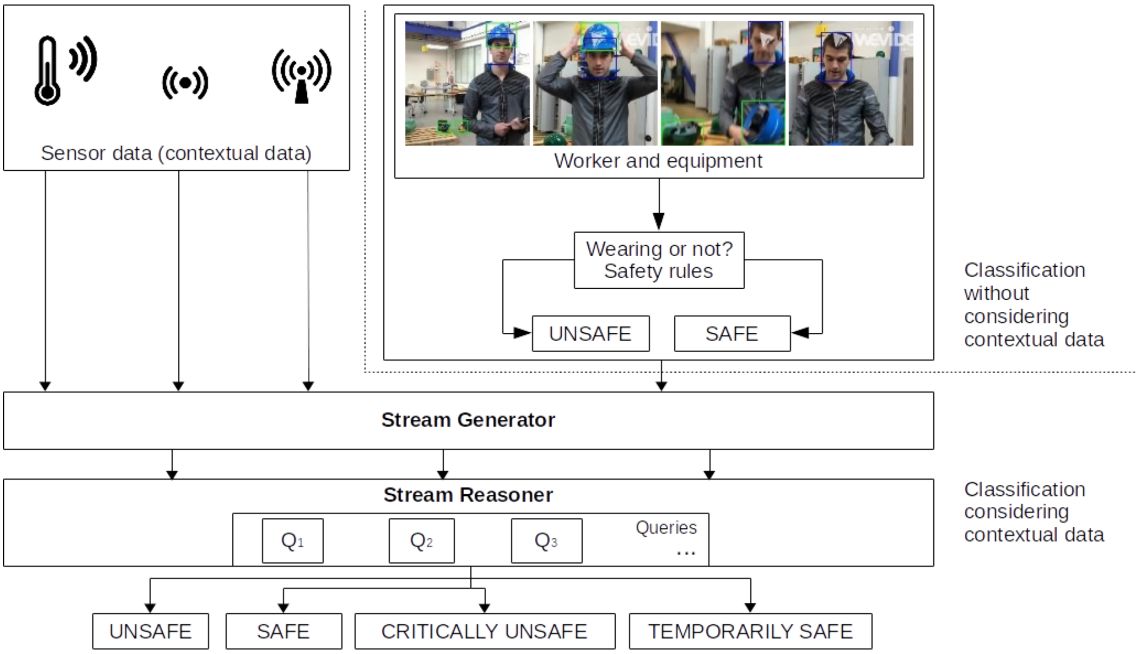


Figure 1. Overall Scheme of the proposed approach.

Automated verification of PPE compliance can be useful in a variety of industries (e.g. Oil & Gas, Manufacturing & Production, Construction, Engineering, Pharmaceuticals, etc.) and applied in a range use case scenarios to ensure employees remain safe (Au, Davalos, Venkatesha, Khurana, Bedros, Mohideen & Cabuz, 2017). Below we exemplify two main applications that the proposed solution can address.

Application 1: Access Control

With cameras positioned above an entrance/exit of a site or facility, the system is able to visually verify that labourers are wearing the protective equipment according to the safety requirements of that industry/area before allowing entry. In case of any equipment being missed at the point of entry, then safety status of the situation can be tagged as *unsafe* and the system will not permit a gate to open and will advise which items must be worn in order to enable access. Once all the mandatory equipment are detected, the status can change to *safe* and the access granted. Figure 2 presents a general overview of applications for access control using visual content.

Sensorial data applied to access control

The visual information from the cameras can be combined with other sensor data collected moments before decision of granting or not the access is given by the system. This extra contextual information about, for instance, crucial required equipment (e.g. oxygen mask when oxygen level read from sensors is critically low) can change the output of system or increase the confidence of the safety status of the scene.

Application 2: Continuous Monitoring

Another solution can address the continuous monitoring of works by the use of cameras covering the site or facility to ensure that employees remain wearing the required PPE in a given context. If labourers remove a required equipment then the system will recognize this in real-time and carry out an action based on a set of given preferences or recommendations. For instance, an alert can be sent directly to the employee or manager for correction on site; the event can be logged for future reports and analysis, etc.

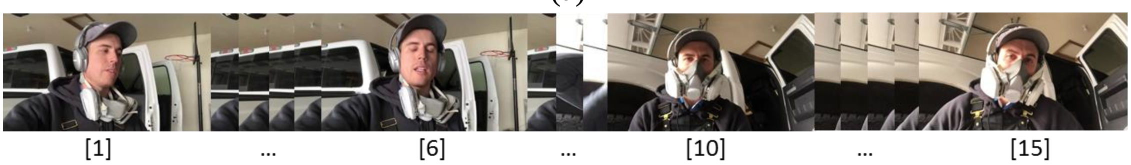
Sensor data applied to continuous monitoring

Such as for access control, sensor data can be combined with visual information from the cameras for a creation of a context and a better understanding of the situation being monitored. For instance, if sensors detect any abnormality, safety status of the scene can be updated accordingly, and workers advised of that for a quick action.

Case Study Scenarios

In this paper we analyse four different situations in which sensorial data might modify the status of safety visual scene in analysis, giving a more suitable output for a given setting. The first scenario checks if employees are wearing respirators to grant access to a restricted area where dust is being monitored. Sensorial data simulating the levels of dust inside that area is given to either increase confidence or to adjust safety status in real time. For the second situation, one more PPE is considered (earmuffs) as well as the measurement of levels of noise in the room. The sequence of frames in analysis for the first two case scenarios are presented in Figure 2 (a) and (b), respectively.

(a)



(b)

Figure 2. Sequences of frames for case scenario 1 (a) and 2 (b).

The third case scenario simulates the continuous monitoring setting. In this case, light intensity is monitored to adjust the necessity (how crucial it is) the wearing or not of high visibility clothing to ensure the safety of the scene. The brightness of images have been manually modified to suit the use case scenario. The last scenario also considers oxygen mask as a required PPE and the safety status of the scene is adjusted according to the oxygen concentration in the area. The sequence of frames in analysis for the third and fourth case scenarios are presented in Figure 3 (a) and (b), respectively.

(a)

(b)

Figure 3. Sequences of frames for case scenario 3 (a) and 4 (b).

Visual Information: SOEKS representation

For the case study in analysis, a set of variables and rules are represented as a Set of Experience Knowledge Structure (SOEKS). SOEKS allows the representation, use, storing and retrieval of visual and non-visual knowledge content together in one single standardized structure (Sanin & Szczerbicki, 2009).

The variables in our system are composed by each image/frame being analysed, bounding boxes containing body parts of workers and each Personal Protective Equipment (PPE) in scene. In addition, we include, as part of the set of variables, the overlap value $\mathbf{O_{I,ppe}}$ between the body part area and corresponding PPE area, and the safety status of the scene, to be assigned according to the set of rules. For this analysis workers are considered wearing the equipment if the overlap value is equal to or superior 0.4. Table 1 shows values of the overlap between head and helmet ($O_{I,\text{helmet}}$) for a sequence of frames and the status of *wearing/not wearing* associated with them.

Table 1. Examples of $O_{I,\text{helmet}}$ and respective wearing/not wearing status.

Frame					
$\boldsymbol{O_{I,\text{helmet}}}$	0.49	0.44	0.40	0.00	0.00
Wearing helmet?	YES	YES	YES	NO	NO

To ensure flexibility and as well as to attend each specific requirements of different industries and scenarios, a set of rules is created. These rules are also a way of allowing expert knowledge to be included in the system reasoning as they can be easily changed and adjusted to attend specific requisites and situations. For this analysis in specific, the following set of rules are considered, each representing a different use case scenario:

Access Control

Rule 1:
IF $O_{I,respirator}$ > threshold
THEN safety_status = SAFE
ELSE safety_status = UNSAFE

Rule 2:
IF $O_{I,respirator}$ > threshold & $O_{I,earmuff}$ > threshold
THEN safety_status = SAFE
ELSE safety_status = UNSAFE

Continuous Monitoring

Rule 3:
IF $O_{I,hi-vis}$ > threshold
THEN safety_status = SAFE
ELSE safety_status = UNSAFE

Rule 4:
IF $O_{I,hivis}$ > threshold & $O_{I,oxygen\ mask}$ > threshold
THEN safety_status = SAFE
ELSE safety_status = UNSAFE

Sensorial Information: Stream reasoning

As mentioned previously, detection of risk situations that may lead to severe incidents in the industrial scenario requires the integration of data from different data sources, with different underlying meanings, different temporal resolutions as well as the need to process these data in real time. Thus, we propose to use stream reasoning to face these issues and add contextual information. In other words, our proposal uses a combination of these approaches to meet the requirements for the detection of risk situations.

The Stream Generator module is mainly responsible for acquiring data from sensors and converting it to RDF streams. It performs semantic annotation of the acquired data, using their corresponding metadata, such as the sensors which made the observation, the observed property, the time, etc. This allows the module to stream out semantic annotated data streams that are then consumed by the Stream Reasoner. The output streams are RDF streams. An RDF stream is defined as an ordered sequence of pairs, where each pair is constituted by an RDF triple and its timestamp t: $(\langle Subject, Predicate, Object\rangle, t)$. An RDF triple is defined as $\langle \mathrm{Subject, Predicate, Object} \rangle \in (I \cup B) \times I \times (I \cup B \cup L)$, where I is a set of IRIs (Internationalized Resource Identifiers), B is a set of blank nodes and L is a set of literals.

Once the data from the distributed and heterogeneous data sources is available in a homogeneous, contextualized and ordered representation, the streams can be explored to generate new information. A set of queries, which combine background knowledge, expert knowledge (rules) and some parts of the streams that are relevant, are registered and executed by the Stream Reasoner over the data streams. These queries represent particular situations to be identified and they include mainly temporal dependencies between observations and/or anomalies. The Stream Reasoner represents queries as query graphs, with query evaluation performed through graph pattern matching over graphs formed by the incoming data streams.

For this component, C-SPARQL (Barbieri, Braga, Ceri, Della Valle & Grossniklaus, 2010) is used to execute queries against the streaming data. This module generates a new classification of the image considering the context information and the rules defined by experts. This classification can be the same as the one generated without considering the context information or not. In this way, the Stream Reasoner itself can be seen as an advanced sensor able to produce high level data, which considers contextual data according to the defined rules. It is important to mention that, the sensorial data has been simulated based on real life values. Below, the sensorial information considered is briefly explained and the thresholds and new range of outputs summarized in Figure 4.

Oxygen Concentration

The concentration of oxygen in our atmosphere is 20.9476 %. When oxygen levels fall below the safe threshold, which is 19.5 percent, health hazards may occur. With only a few breaths of oxygen deficient air, you could fall unconscious and suffocate. Levels below 6% are considered critically low and human body would not survive when exposed to those concentrations. On the other hand, levels above 23.5 % can cause oxidizing free radicals to form, which can attack the tissues and cells of the body and cause muscle twitching.

Dust Level

Dust monitoring is one aspect of air quality that industrial hygienist use to determine the amount of dust particles present in the workplace over a given period of time. It an abstract term that usually comprises measuring several indicators. It is measured in $\mu g/m^3$. In our analysis we consider the PM10 sized particles and as advised by the Australian National Standards for Criteria Air Pollutants (Department of the

Environment and Heritage, 2005), concentrations below $20\,\mu g/m^3$ as safe (yearly average) and above $50\,\mu g/m^3$ as critically unsafe as exceeds the daily maximum limit.

Noise Level

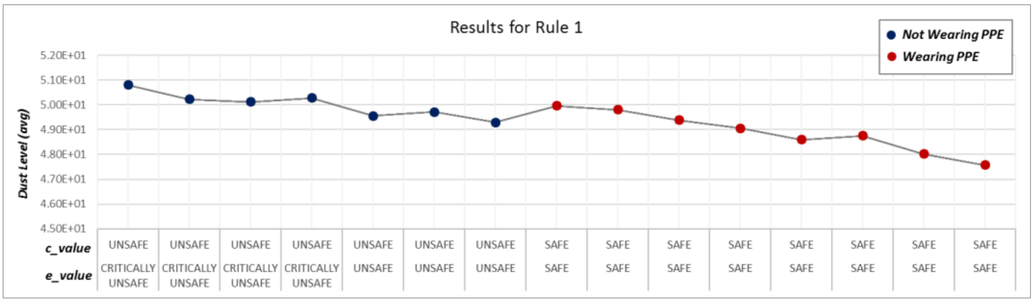
Measuring noise levels and workers' noise exposures is the most important part of a workplace hearing conservation and noise control program. For this analysis we consider the noise exposure limits when Criterion Level is 85 dB(A) (exposure standard for noise in Australia) in a 8 hour shift and 3dB exchange rate to be safe (Safe Work Australia, 2019). We also consider the maximum peak level (140 dB) as a critical level, which workers can't be exposed to as can cause instant damage to hearing.

Light Intensity

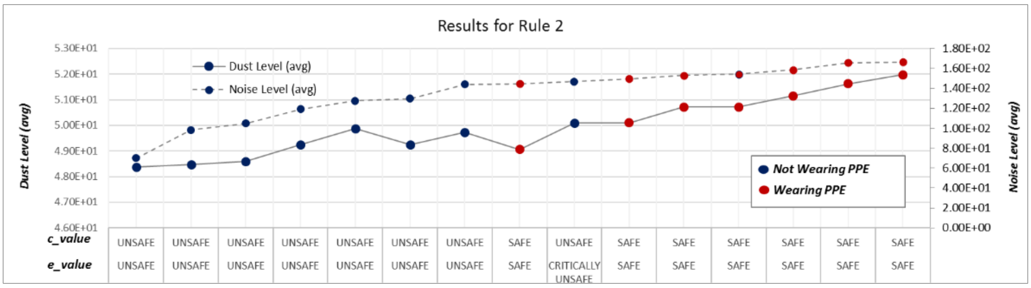
Measuring illumination has become a common practice in workplaces as a way to make sure employee are operating in safe working conditions. The level of light recommended is different for each workspace/activity and can vary from 50 to 20k lux. For normal Drawing Work, Detailed Mechanical Workshops, Operation Theatres, etc, the recommended light level is around *1k lux*, which we consider safe in our analysis.
In addition, we consider *50 lux* as a critical level as it is comparable to a dark day and visibility is considered difficult.

Experimental Results

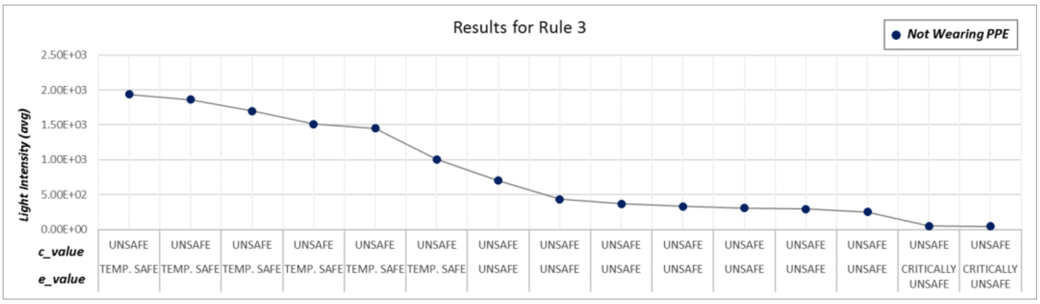
The system has been tested over four different settings of fifteen sequences of frames in which detections of body and/or parts and PPEs has occurred. For each setting sensor data has been added to create a context in which the visual scene is being analysed. The output of system (*c_value* of *safety_status* variable) has, then, been optimised (*e_value*) according to the sensorial information. Figure 4 shows the results for each rule in analysis.



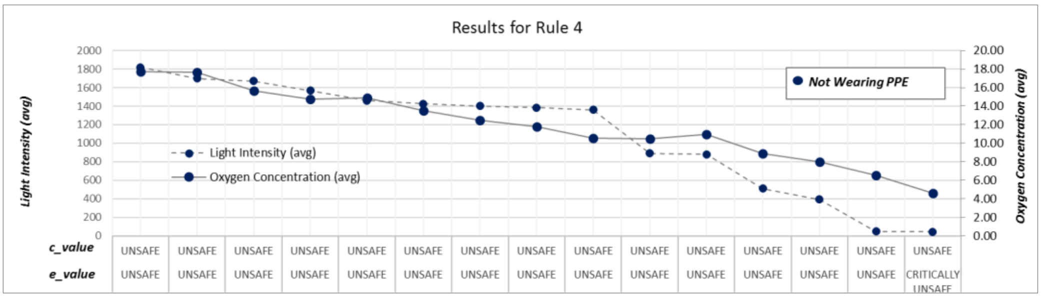
(a)



(b)



(c)



(d)

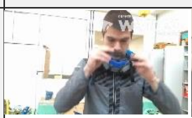
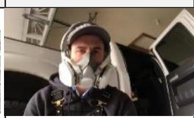
Figure 4. Results for Rule 1(a), Rule 2 (b), Rule 3 (c) and Rule 4 (d).

From Figure 4 we can observe the change in the final outputs (*e_value)* for cases in which the sensorial information provided are in critical levels (in this case we have *UNSAFE* status becoming *CRITICALLY UNSAFE).* This can be used to give the employees an idea of the severity of not wearing the required equipment. In addition, when properties values in analysis are considered normal, even having the set of rules already set up, a momentarily *SAFE* status can be granted. Finally, for the unchanged status, the sensorial information can be used to improve the confidence of safety of the scene in analysis.

One of drawbacks of such approach is when the detector fails to detected the PPE. In this case, when sensorial information is taken in consideration, misleading outputs, such as observation number 9 from Figure 4 (b), can be given, worsening the output status of the scene. This issue can, however, be minimised by retraining of classifiers to improve accuracy, or by considering a higher number of frames per second to determine the presence of not of PPEs and body parts.

Table 2 shows examples of the outputs representing the safety status (*c_value)* of frames for the given rule and the updated value *(e_value)* after adding the sensorial information.

Table 2. Example of Output of system for each given set of rules.

	Rule 1	**Rule 2**	**Rule 3**	**Rule 4**
Frame (or sequence)				
Frame number	004	011	002	015
Required Equipment	Respirator	Respirator and Earmuffs	High Visibility Clothing (Hi-Viz)	High-Visibility Clothing (Hi-Viz) and Oxygen Mask
Wearing PPE?	No	Yes	No	No
Safety Status (*c_value*)	UNSAFE	SAFE	UNSAFE	UNSAFE
Sensorial Information (property)	Dust Level	Dust and Noise Levels	Light Intensity	Light Intensity and Oxygen Concentration
Value(s)	5.03E+01 $\mu g/m^3$	5.07E+01 $\mu g/m^3$ and 1.53E+02 dB	1.86E+03 Lux	44.77 Lux and 4.57 %
Optimised Safety Status (*e_value*)	**CRITICALLY UNSAFE**	**SAFE**	**TEMPORA-RILLY SAFE**	**CRITICALLY UNSAFE**

Conclusion and Future Work

This study has presented a cognitive system enriched with contextual information coming from different sensors and proposed the use of stream reasoning to integrate and process all these data in real time. The proposed approach has been tested on a Cognitive Vision System for Hazard Control (CVP-HC) which is based on Set of Experience Knowledge Structure (SOEKS) and Decisional DNA (DDNA). Four different settings have been analysed and the use of stream reasoning has demonstrated being useful to provide a better understanding of the situation in analysis, therefore improving decision making and ensuring that workers remain safe and compliant with Health and Safety policy for use of Personal Protective Equipment (PPE).

For next steps, more complex scenarios will be explored for the creation of more complex set of rules and a deeper analyse of the results presented for online operation of the system in which the input images and context variables are gathered from video cameras and sensors in real time.

References

Sanin, C., Haoxi, Z., Shafiq, I., Waris, M. M., de Oliveira, C. S., & Szczerbicki, E.: Experience based knowledge representation for Internet of Things and Cyber Physical Systems with case studies. Future Generation Computer Systems, 2018.

Vernon, D.: The space of cognitive vision. In *Cognitive Vision Systems*, 7-24, Springer, Berlin, Heidelberg, 2006.

Chmaj, J.: Contextual Knowledge: Your Key to Building Effective Knowledge Tools, *KM World*, 2019.

Gregory, R. L. (1973). Eye and brain: The psychology of seeing. McGraw-Hill.

Brézillon, P. (2003). Representation of procedures and practices in contextual graphs. *The Knowledge Engineering Review*, *18*(2), 147-174.

Bauckhage, C., Wachsmuth, S., Hanheide, M., Wrede, S., Sagerer, G., Heidemann, G., & Ritter, H. (2008). The visual active memory perspective on integrated recognition systems. *Image and Vision Computing*, *26*(1), 5-14.

Crowley, J. L., Coutaz, J., Rey, G., & Reignier, P. (2002, September). Perceptual components for context aware computing. In *International conference on ubiquitous computing* (pp. 117-134). Springer, Berlin, Heidelberg.

Beck, H., Dao-Tran, M., Eiter, T., & Folie, C. (2018). Stream Reasoning with LARS. *KI-Künstliche Intelligenz*, *32*(2-3), 193-195.

Sanin, C., Toro, C., Haoxi, Z., Sanchez, E., Szczerbicki, E., Carrasco, E.,. & Man-cilla-Amaya, L. (2012). Decisional DNA: A multi-technology shareable knowledge structure for decisional experience. Neurocomputing, 88, 42-53.

C. Sanin, E. Szczerbicki, Experience-based Knowledge Representation SOEKS. Cybernet Sys. 40(2) (2009) 99-122.

Sanin, C., Szczerbicki, E.: Decisional DNA and the smart knowledge management system: A process of transforming information into knowledge. In Techniques and tool for the design and implementation of enterprise information systems, ed. A. Gunasekaran, 149–175. New York: IGI Global (2008).

Han, S., & Lee, S. (2013). A vision-based motion capture and recognition framework for behavior-based safety management. Automation in Construction, 35, 131-141.

Ciresan, D. C., Meier, U., Masci, J., Maria Gambardella, L., & Schmidhuber, J. (2011, July). Flexible, high performance convolutional neural networks for image classification. In IJCAI Proceedings-International Joint Conference on Artificial Intelligence (Vol. 22, No. 1, p. 1237).

Little, S., Jargalsaikhan, I., Clawson, K., Nieto, M., Li, H., Direkoglu, C., ... & Liu, J. (2013, April). An information retrieval approach to identifying infrequent events in surveillance video. In Proceedings of the 3rd ACM conference on International conference on multimedia retrieval (pp. 223-230). ACM.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

Mosberger, R., Andreasson, H., & Lilienthal, A. J. (2013, November). Multi-human tracking using high-visibility clothing for industrial safety. In Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on (pp. 638-644). IEEE.

Chen, L., Hoey, J., Nugent, C. D., Cook, D. J., & Yu, Z. (2012). Sensor-based activity recognition. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Re-views), 42(6), 790-808.

Zambrano, A., Toro, C., Nieto, M., Sotaquirá, R., Sanín, C., Szczerbicki, E.: Video semantic analysis framework based on run-time production rules – towards cognitive vision. J. Univ. Comput. Sci. 21(6), 856–870 (2015).

S. Aditya, Y. Yang, C. Baral, Y. Aloimonos and C. Fermuller, "Image Understanding using vision and reasoning through Scene Description Graph," Computer Vision and Image Understanding, pp. 1-13, 2017.

Hollnagel, E., & Woods, D. D. (2005). Joint cognitive systems: Foundations of cognitive systems engineering. CRC Press.

Cole, G. S. (1990). Tort liability for artificial intelligence and expert systems. Computer/LJ, 10, 127.

Ashby, W. R. (1961). An introduction to cybernetics. Chapman & Hall Ltd.

Yu, Y., Pan, G., Gong, Y., Xu, K., Zheng, N., Hua, W., ... & Wu, Z. (2016). Intelligence-augmented rat cyborgs in maze solving. PloS one, 11(2), e0147754.

Pathak, N. (2017). The Future of AI. In Artificial Intelligence for. NET: Speech, Language, and Search (pp. 247-259). Apress, Berkeley, CA.

Vernon, D.: The space of cognitive vision. In Cognitive Vision Systems, 7-24, Springer, Berlin, Hei-delberg (2006).

Brézillon, P., & Pomerol, J. C.: Contextual knowledge and proceduralized context. In Proceedings of the AAAI-99 Workshop on Modeling Context in AI Applications, Orlando, Florida, USA, July. AAAI Technical Report (1999).

Sheth, A., Henson, C., Sahoo, S.S., 2008. Semantic Sensor Web. IEEE Internet Computing 12, 78–83

Stuckenschmidt, H., Ceri, S., Della Valle, E., Harmelen, F., Milano, P., 2019. Towards expressive stream reasoning. Proceedings of the Dagstuhl Seminar on Semantic Aspects of Sensor Networks.

Haller, A., Janowicz, K., Cox, S.J., Lefranc¸ois, M., Taylor, K., Le Phuoc, D., Lieberman, J., Garc´ıa-Castro, R., Atkinson, R., Stadler, C., 2018. The modular SSN ontology: A joint W3C and OGC standard specifying the semantics of sensors, observations, sampling, and actuation. Semantic Web , 1–24.

Tallevi-Diotallevi, S., Kotoulas, S., Foschini, L., Lecue, F., Corradi, A., 2013. Real-Time Urban Monitoring in Dublin Using Semantic and ´ Stream Technologies, in: The Semantic Web – ISWC 2013, Springer Berlin Heidelberg. pp. 178–194.

Lecue, F., Kotoulas, S., Aonghusa, P.M., 2012. Capturing the Pulse of Cities: Opportunity and Research Challenges for Robust Stream Data ´ Reasoning, in: Semantic Cities @ AAAI.

Calbimonte, J.P., Ranvier, J.E., Dubosson, F., Aberer, K., 2017. Semantic representation and processing of hypoglycemic events derived from wearable sensor data. Journal of Ambient Intelligence and Smart Environments 9, 97–109.

Shojanoori, R., Juric, R., 2013. Semantic Remote Patient Monitoring System. Telemedicine journal and e-health : the official journal of the American Telemedicine Association 19.

Santipantakis, G.M., Vlachou, A., Doulkeridis, C., Artikis, A., Kontopoulos, I., Vouros, G.A., 2018. A Stream Reasoning System for Maritime Monitoring, in: 25th International Symposium on Temporal Representation and Reasoning (TIME 2018), Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany. pp. 20:1–20:17.

Ereteo, G., Buffa, M., Gandon, F., Corby, O., 2009. Analysis of a real online social network using semantic web frameworks, in: Lecture Notes in Computer Science.

Mika, P., 2005. Flink: Semantic Web technology for the extraction and analysis of social networks. Journal of Web Semantics 3, 211 – 223.

Barbieri, D.F., Braga, D., Ceri, S., Della Valle, E., Grossniklaus, M., 2010. C-SPARQL: a continuous query language for RDF data streams. International Journal of Semantic Computing 4, 3–25.

Department of the Environment and Heritag, 2005. National standards for criteria air pollutants 1 in Australia.

Giustozzi, F., Saunier, J., & Zanni-Merk, C. (2019). Abnormal Situations Interpretation in Industry 4.0 using Stream Reasoning. Procedia Computer Science, 159, 620-629.