

“© 2017 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.”

# Evolving Artificial Pain from Fault Detection through Pattern Data Analysis

Muh Anshar<sup>1</sup> and Mary-Anne Williams<sup>2</sup>

**Abstract**—Fault detection is a classical area of study in robotics and extensive research works have been dedicated to investigate its broad applications. As the breath of robots applications requiring human interaction grow, it is important for robots to acquire sophisticated social skills such as empathy towards pain. However, it turns out that this is difficult to achieve without having an appropriate concept of pain that relies on robots being aware of their own body machinery aspects. This paper introduces the concept of pain, based on the ability to develop a state of awareness of robots own body and the use of the fault detection approach to generate artificial robot pain. Faults provide the stimulus and defines a classified magnitude value, which constitutes artificial pain generation, comprised of synthetic pain classes. Our experiment evaluates some of synthetic pain classes and the results show that the robot gains awareness of its internal state through its ability to predict its joint motion and generate appropriate artificial pain. The robot is also capable of alerting humans whenever a task will generate artificial pain, or whenever humans fails to acknowledge the alert, the robot can take a considerable preventive actions through joint stiffness adjustment.

## I. INTRODUCTION

Fault detection is a classical area of study in robotics which mainly considered as a stimulus in robot motion planning. Earlier studies, reported in [1], [2], [3] provide the foundation for the importance of incorporating failure detection into robot planning mechanisms. Various aspects of robot motion planning are investigated in [4], [5], [6], [7], [8] and extensions the scope to multiple robot planning [9], [10]. All of these studies assume that the robot is fully functional. In practice, however, robots fail and their failure can affect not only the plans but also put resources and people at risk. As the use of robots grow, such as in human-robot interaction or robot-to robot interaction, a new and growing field of research, robots are required to develop more sophisticated social skills. Understanding the concept of pain in humans themselves is critical for planning and tasks that require human-robot interaction. However, this raises an issue on how a robot can develop a proper concept of pain which relies on the robot being aware of its own body machinery aspects. [11] proposes a damage recovery approach where robots are aware of their body hardware failure, such as one or few of robot joints suffer from malfunctions. The study reports that the robot successfully discovers a new qualitative behavior of hexapod gaits. However, when this

type of fault is detected, it merely functions as a stimulus to activate the new motion plan or new motion behaviour generation. Unlike in robots, in human mechanism, any machinery of body failure will generate internal states where humans experience what is called 'pain' [12]. In fact, if the faults in robots themselves are associated with not only stimulus but also specific meaningful magnitude, it will be beneficial for robots to incorporate them as part of their experience. This paper intends to derive machinery faults of the robots, detected from robot proprioceptive sensors, into an appropriate representation of pain by introducing an artificial pain concept containing synthetic pain classification. This paper provides experimental results by presenting different scenarios to evaluate each of synthetic pain classification proposals.

The reminder of the paper proceeds as follows: section II gives an overview of proposed artificial pain concept, including description of the proposed synthetic pain and the integration process into the robot mechanism. Section III presents the raw data analysis and synthetic pain classes followed by section IV which covers the experiment stages. Results and discussions are presented in section V followed by section VI which focuses on the overall achievement and possible future developments.

## II. APPROACH OVERVIEW

Our approach contains three major steps of development, firstly, redefining the artificial pain, and then integrating this artificial pain concept into real robot mechanism. Finally, experiments, which involve human-and-robot interaction through a hand pushing task, are carried out.

### A. Artificial Pain

Our artificial pain for robots contains three classifications of synthetic pain de-rived from the report [13], described as follows:

- Category 1 :Proprioceptive pain which is associated with potential hardware damage, functions as an alert signal.
- Category 2 :Inflammatory reduction pain associated with the change of robot behaviour as the result of:
  - 2.1 Predicted robot hardware damage.
  - 2.2 Real robot hardware damage.
- Category 3 :Sensory malfunction pain, associated with:
  - 3.1 Abnormal function of internal sensors.
  - 3.2 Damage internal sensors.

<sup>1</sup>Muh Anshar is with Faculty of Electrical Engineering, University of Hasanuddin UNHAS, Makassar, Indonesia [muh.anshar@gmail.com](mailto:muh.anshar@gmail.com)/[anshar@unhas.ac.id](mailto:anshar@unhas.ac.id)

<sup>2</sup>Mary-Anne Williams with the Faculty of Engineering and Information Technology Sydney UTS, Sydney, Australia [mary-anne.williams@uts.edu.au](mailto:mary-anne.williams@uts.edu.au)

The region in which each body part motion occurs, whether the motion is lateral or rotational, determines the pain level. The pain level is assigned by measuring the distance between the position of the respective body part in the region and the threshold values assigned by the robot awareness framework (see Figure 1). The physical motions associated with the joint movements of the robot hardware are actively monitored by the sensory mechanisms, which contain proprioceptive and exteroceptive sensors. The further the distance from the threshold value, the higher the pain level to be assigned. The threshold values can be manually designed by the human user and placed in the database as a reference (static threshold), or they can be generated and configured autonomously by the robot framework itself (self-generated).

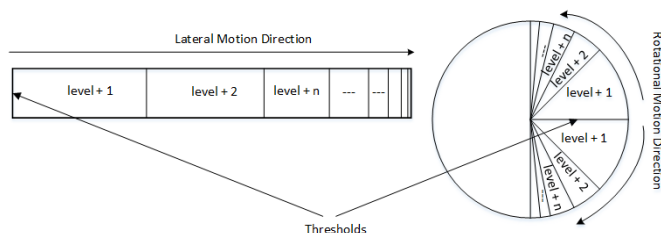


Fig. 1. Pain Region Assignment

### B. Integration

Artificial pain definition is integrated into a NAO robot mechanism, which utilise our novel Adaptive Self-awareness framework for robots (ASAF) [13]. The ASAF framework consists of several elements as shown in Fig. 2 and we will only discuss briefly the key elements of the framework.

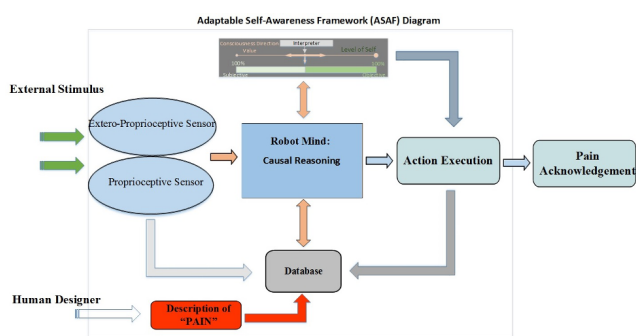


Fig. 2. Adaptive Self-Awareness Framework for Robot - ASAF

There are two predominant factors in directing robot attention:

- (i) the ability to focus attention on a specified physical aspect of the machinery of robot body.
- (ii) the ability to foresee, and at the same time, to be aware of the consequences of predicted actions.

The ASAF addresses these two aspects and integrate them into the framework so that the detection of synthetic pain can be acknowledged and responded to in an appropriate way.

The framework maps the robot consciousness region into a discrete range 1 - 3 for subjective and 4 - 6 for objective elements. This region values are called Consciousness Direction Values (CDV) as shown in Fig. 3.

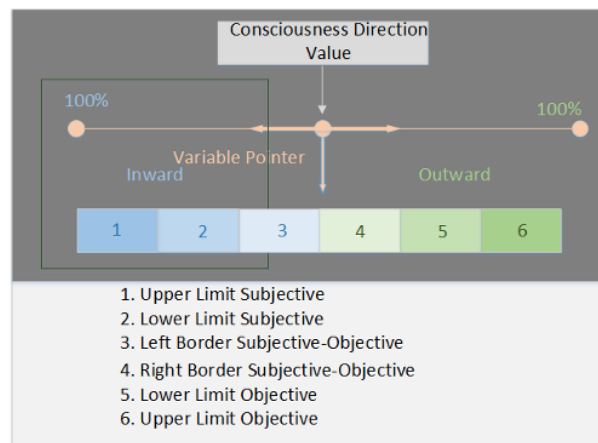


Fig. 3. CDV Region Distribution

In order to generate synthetic pain on the robot, we set the robot joint restriction region that should be avoided. Synthetic pain can then be generated when the robot joint moves into this region. Pattern data analyses takes place inside the ASAF where the raw data is obtained from the robot joint motion through the joint proprioceptive sensor. At early implementation, the ASAF framework utilise the sequential pattern prediction [14], [15] in order to capture the behaviour of the observed real data and then use them to predict the future possible conditions. The inter-correlation among data is analysed using the associative theories [16], which utilises covariance information obtained from sequence data. The ASAF analyses the relationship among data covariance by making predictions of sequence data patterns obtained from robot's proprioceptive sensor (joint position sensor). The prediction process only takes place after several sequences of data so as to reduce biased analyses. Any result made from previous sequence predictions are reassessed with the current incoming data, and the results either kept for future prediction or amendment actions take place before execution proceeds. This cycle repeats only if current data and predicted values are not classified in restricted region which refers to the synthetic pain classifications. When joint values falls into any region, the robot's action will be either send an alert which signifies proprioceptive pain (Category 1) or modify motor stiffness values on the robot's joint signifying that the robot is experiencing Inflammatory reduction pain (Category 2.1). By increasing motor stiffness, the robot joint will resist any force generated from the physical pushing interaction, and as a result, the robot will be prevented from experiencing higher degree of synthetic pain. For Sensory malfunction pain (Category 3.1), the joint sensor of the robot being used at times produces abnormal data reading. When this takes place, robots will analyse it, and the data sequence is amended in order to preserve uniformity of the overall

data.

### III. DATA ANALYSIS

Raw data from the sensor is collected, and arranged according to retrieval time and are further used to determine the covariance data. By substituting the covariance data into the latest obtained raw data, predicted data can be obtained. This process is further discussed in the following subsections.

#### A. Raw Proprioceptive Data

The interaction occurs within a specified constant time span,  $T$ . The representation of collected data obtained at a specified time,  $t_i$  is:

$$\prod_{t_i, i=0}^{i < T} d_{t_i}$$

where:

$d_{t_i}$  represents a joint value at time  $t_i$ , and the value of  $t_i$  is defined by:

$$t_i = \begin{cases} i = 0, \text{ initiating experiment} \\ i < T, \text{ time span of experiment} \end{cases}$$

#### B. Raw Proprioceptive Data

Data covariance is derived from the difference of the last obtained joint values and previous one, as depicted in Eq. 1:

$$\Delta int = d_{t_T} - d_{t_{T-1}} \quad (1)$$

#### C. Prediction Data

Data covariance is used during the analyses process in formulating predicted sequence data, allowing the system to reproduce a sequence predicted data. By substituting Eq. 1 into the obtained data,  $d_{t_i}$ , we can reproduce the sequence of predicted data shown in Eq. 2.

$$\prod_{t_i, i=m}^{i < \bar{T}} \bar{d}_{t_i} = \prod_{t_i, i=m}^{i < \bar{T}} (d_{t_i} + \Delta int) \quad (2)$$

$\bar{d}_{t_i}$  represents predicted data at sequence time  $t_i$ , where the values of  $t_i$  are determined by:

$$t_i = \begin{cases} i = m, \text{ data at time m analysing process initiates} \\ i < \bar{T}, \text{ time discrete of prediction} \end{cases}$$

where the value of m must satisfy the following conditions (Criteria 1):

$$t_i = \begin{cases} c_s > 0, \text{ total similarity of joint values reference} \\ c_d > 0, \text{ total difference of joint values reference} \\ c_u \gg c_d; c_u \gg c_s, \text{ unique data} \end{cases}$$

#### D. Synthetic Pain Classification

Three joint values which are set to justify the range of synthetic pain detection depicted in the Table I.

TABLE I  
JOINT VALUES AS FAULT INDICATOR AND SYNTHETIC PAIN ASSOCIATION

Joint Values	Synthetic Pain Classification		Robot Responses
	Category	Definition	
1.5421	1	Proprioceptive pain	Alert Human
1.5521	2.1	Inflammatory reduction pain	Increase Joint Stiffness
Fluctuate	3.1	Sensory malfunction pain	Disregard/Sequence Values Amendment

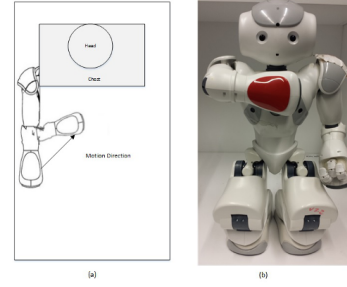


Fig. 4. Robot Posture (a) Motion Direction (b) Initial Posture

### IV. EXPERIMENT OVERVIEW

The purpose of the experiment is threefold:

- (1) to conceptualise synthetic pain through robot faulty joint settings and the response to pain times
- (2) to analyse robot joint motion through sequential data analyses
- (3) to verify time responses of a robot through proportionate joint stiffness settings that change with respect to the force provided by the human to the robot's arm, e.g. a strong human force is met with stronger joint stiffness

Besides that, we carry out an additional experiment where there is no physical interaction involved (human partner does not touch the robot hand). This experiment is dedicated solely to measure the accuracy of the prediction output of the pattern data analyser.

#### A. Experimental Set Up

The overview of the experimental set up is described as follows: It involves the NAO robot and a human partner, which is separated by a distance of 0.2 - 0.3 meter from the robot feet. The robot raises its right arm until it is perpendicular relative to its chest. Fig. 3 shows the illustration of motion direction and the robot posture.

#### B. Scenario

The right elbow roll joint of the robot sets to minimum stiffness, allowing the human to push the hand towards robot's chest. Then, data is collected the associated joint sensor which acting as proprioceptive sensor, and ordered according to their reading sequences, unless it is stated differently. This experiment scenario repeats three times in order to have a considerable reliable data analyses.

TABLE II  
JOINT VALUES DURING THREE TRIALS

Trial	1		2		3	
Sequence	Joint Data	Time	Joint Data	Time	Joint Data	Time
1	0.22247	38.88	0.02765	776.46	0.02765	267.12
2	0.26696	39.40	0.02765	776.98	0.02765	267.64
3	0.37127	39.91	0.02919	777.49	0.02765	268.15
4	0.49246	40.44	0.21940	778.01	0.06907	268.68
5	0.63205	41.00	0.39735	778.54	0.29917	269.20
6	0.78852	41.54	0.68421	779.11	0.52774	269.73
7	0.95572	42.10	1.30548	781.36	0.71642	270.28
8	1.32695	44.37			0.87902	270.84
9					1.04470	271.39
10					1.41132	273.68

## V. RESULTS AND ANALYSIS

The results of three trials is shown in Table II.

It can be seen that most of trials end up within different time sequence this is the result from the values of  $m$  which is determined by  $c_s = c_d = 3$  and  $c_u = 5$  and  $\bar{T} = 4$ . Setting the values of  $c_s$  and  $c_d$  equal to 3 will allow the sequence pattern analyser to collect sufficient sensor data before commencing prediction. The value of  $c_u$  is set to 5 as the incoming sensor data may not perform a uniform sequence values. As the hand movement lasts for a short time, the system is restricted to produce 4 interval data prediction. However, the length of data prediction is a subject of adjustment whenever the the workspace of the robot arm changes.

Using Eq. 1 and Eq. 2, the robot system analyses the incoming joint data and produced the resulting sequence data prediction as Table III.

TABLE III  
JOINT VALUES PREDICTION

Trial	Sequence	Joint Values	Prediction Cycles				
			1st	2nd	3rd	4th	5th
1	1	0.22247					
	2	0.26696					
	3	0.37127					
	4	0.49246	0.49246				
	5	0.63205	0.61365	0.63205			
	6	0.78852	0.73484	0.77164	0.78852		
	7	0.95572	0.85603	0.91123	0.94499	0.95572	
	8	1.32695	0.97722	1.05082	1.10146	1.12292	1.32695
	9						1.69818
2	1	0.02765					
	2	0.02765					
	3	0.02919					
	4	0.21940					
	5	0.39735	0.39735				
	6	0.68421	0.57530	0.68421			
	7	1.30548	0.75325	0.97107	1.30548		
	8				1.92675		
3	1	0.02765					
	2	0.02765					
	3	0.02765					
	4	0.06907					
	5	0.29917					
	6	0.52774	0.52774				
	7	0.71642	0.75631	0.71642			
	8	0.87902	0.98488	0.90510	0.87902		
	9	1.04470	1.21345	1.09378	1.04162	1.04470	
	10	1.41132	1.44202	1.28246	1.20422	1.21038	1.41132
	11						1.77794

### A. Analysis

By using Criteria 1, the robot determines when the prediction starts and it can be seen from Table III that Trial 1 and Trial 2, the robot commences its analyses at sequence data 4 and Trial 3, takes place at sequence 6.

This shows the incoming data from the robot joint elbow sensor is not necessarily the same for each trial. Human hand physical forces during interaction with the robot arm is not the same in all trials. During Trial 1, the highest margin of error takes place at prediction cycle 1, which is the fourth interval, 34.97%. This trend remains the same for the rest of the prediction cycles. Similar situation occurs in Trial 2 which produces the highest margin of error at fourth interval of prediction cycle 1, 55.22%. In contrast, Trial 3 produces lower margin of error at fourth interval of prediction cycle 1, 3.07%. The lowest margin of error takes place at second interval of prediction cycle 3, 0.31%. Over the three trials, the highest margin of error occurs in Trial 2, particularly at prediction cycle 1, second interval, about 55.22%. The average standard deviation of Trial 1 and Trial 2 tend to increase, from the lowest  $\sigma$  value in Trial 1 equals to 0.01 to 0.28 in Trial 2. The margin of  $\sigma$  values in Trial 3 is likely twice of in Trial 1 (see Table IV).

TABLE IV  
PREDICTION ERROR

Trial	Sequence	Prediction Cycles						Std D ( $\sigma$ )	
		1	2	3	4	5	6		
1	4	0							
	5	1.84%	0.00%					0.01	
	6	5.37%	1.69%	0.00%				0.03	
	7	9.97%	4.45%	1.07%	0.00%			0.04	
	8	34.97%	27.61%	22.55%	20.40%	0.00%		0.13	
	9								
	2	5	0.00%						
		6	10.89%	0.00%					0.08
		7	55.22%	33.44%	0.00%				0.28
8									
3	6	0.00%							
	7	3.99%	0.00%					0.03	
	8	10.59%	2.61%	0.00%				0.06	
	9	16.87%	4.91%	0.31%	0.00%			0.08	
	10	3.07%	12.89%	20.71%	20.09%	0.00%		0.11	
	11								

With the additional experiment, we obtained a comparison data as shown in Table V with the standard deviation of accuracy of prediction is 0.00.

TABLE V  
COMPARISON EXPERIMENT

Sequence	Joint Values	Prediction Cycles			Std D( $\sigma$ )
		1st	2nd	3rd	
1	0.02765				
2	0.02765				
3	0.02765				
4	0.02765	0.02765			
5	0.02765	0.02765	0.02765		0.00
6	0.02765	0.02765	0.02765	0.02765	0.00
7		0.02765	0.02765	0.02765	

It can be seen that in Trial 1 and Trial 2, the interaction only lasts for a short time which is caused by the speed of human when pushing the robot hand, while at final trial, human partner reduce the pushing speed allowing the robot to obtained more joint values for analyses process. Table VI below depicts final prediction for each trial produces joint values which are higher than joint values listed in Table I. For Trial 1, sequence 8, the robot elbow experiences

Category 1 of Synthetic Pain and as the consequence, the robot sending an alert to human partner. As the human is not aware of the alert sending by the robot, the robot suffers from Synthetic Pain Category 2.1 which forces the robot to react with resistance by increasing joint motor stiffness. This situation causes robot arm to resist any human physical forces applied on the robot hand.

TABLE VI  
ARTIFICIAL PAIN AND ROBOT RESPONSES

Trial	Seq	Prediction Cycles					Synthetic Pain		Robot Responses	
		1	2	3	4	5	6	Category		Definition
1	4									
	5									
	6									
	7									
	8						1.32695	1	Proprioceptive	Alert Human
2	9						1.69818	2.1	Inflammatory reduction	Joint Stiff
	5									
	6									
	7			1.30548				1	Proprioceptive	Alert Human
	8			1.92675				2.1	Inflammatory reduction	Joint Stiff
3	6									
	7									
	8									
	9									
	10						1.41132	1	Proprioceptive	Alert Human
11						1.77794	2.1	Inflammatory reduction	Joint Stiff	

For instance in Trial 3, as the robot is experiencing the kind of proprioceptive pain, Category 1 at sequence 10, it is not subsequently increasing its elbow stiffness. The mechanism still allows the joint to be moveable as previous average prediction is still below the limit joint values. On the opposite, the robot may possibly alerting human and subsequently maximizing its elbow joint stiffness. It can be seen that the robot's mind generates proprioceptive pain, which as per our definition, functions as an alert signal. Apart from sending a voice alert to human partner, the robot also takes a preventive action by increasing its elbow joint stiffness to a degree that the human partner detects resistance.

Overall, robot is capable of predicting in advance the kind of synthetic pain that the robot will experience whenever an interaction proceeds.

## VI. CONCLUSIONS

As the robot utilises our ASAF framework, we are able to demonstrate that a robot can develop an accurate pain acknowledgement and appropriate response. The causal reasoning through sequential pattern prediction enables the robot's decision making to embrace the past, the current and the future considerations as it builds its expectation during interaction, which leads to effective and accurate decisions.

The experiment has demonstrated two kinds of Synthetic Pain, which are Category 1, Proprioceptive pain and Category 2.1, Inflammatory reduction pain, and the robot is able to develop associated pain based on prediction of its elbow joint values. The robot's capability to predict its joint value as the consequence of interaction with human, has allow the robot develop artificial pain through appropriate synthetic pain classification, and at the same time, takes proper counter responses in a timely fashion. This capability allows the robot to expose its own internal state to its human partner whenever the task would lead to physical damage. At the same time, robot can take considerable preventive actions whenever

human fail to acknowledge the alert. Besides, response time which becomes an important factor to be considered from the detection of pain till the counter reactions taken by the robot, has been successfully enabled the robot whether to follow up with subsequent actions or not.

Building on this implementation and proof-of-concept work, future research will extend the pain acknowledgement and responses further by integrating sensor data across more than one sensor using more sophisticated data integration. Furthermore, future works will include carefully designed experiments that evaluate the other categories, which are Inflammatory reduction pain, Category 2.2 (Real robot hardware damage) and Sensory malfunction pain, Category 3.1 (Abnormal function of internal sensors) and Category 3.1 (Damage internal sensors).

## ACKNOWLEDGMENT

This work expressed the appreciation to all members of The Advanced Artificial Intelligence Research Group - Cognitive and Social Robotics (established in 2009) Electrical Engineering Department, University of Hasanuddin UNHAS.

## REFERENCES

- [1] S. Elliott Fahlman, "A planning system for robot construction tasks," *Artificial Intelligence*, vol. 5, no. 1, pp. 1-49, 1974.
- [2] R. J. Firby, "An investigation into reactive planning in complex domains," in *Proceedings of the sixth National conference on Artificial intelligence*, vol. 1. Seattle, WA, 1987, Conference Proceedings, pp. 202-206.
- [3] D. E. Koditschek, "Task encoding: Toward a scientific paradigm for robot planning and control," *Robotics and autonomous systems*, vol. 9, no. 1, pp. 5-39, 1992.
- [4] M. Anshar and M.-A. Williams, "Extended evolutionary fast learn-to-walk approach for four-legged robots," *Journal of Bionic Engineering*, vol. 4, no. 4, pp. 255-263, 2007. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1672652907600390>
- [5] C. J. J. Paredis and P. K. Khosla, "Global trajectory planning for fault tolerant manipulators," in *Human Robot Interaction and Cooperative Robots. 'Proceedings. 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems 95*, vol. 2, 1995, Conference Proceedings, pp. 428-434 vol.2.
- [6] S. K. Ralph and D. K. Pai, "Fault tolerant locomotion for walking robots," in *Proceedings., 1997 IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA'97)*, 1997, Conference Proceedings, pp. 130-137.
- [7] M. Soika, "A sensor failure detection framework for autonomous mobile robots," in *Proceedings of the 1997 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '97)*, vol. 3, 1997, Conference Proceedings, pp. 1735-1740 vol.3.
- [8] S. Tosunoglu, "Fault-tolerant control of mechanical systems," in *Proceedings of the 21st International Conference on Industrial Electronics, Control, and Instrumentation (IEEE IECON)*, vol. 1, 1995, Conference Proceedings, pp. 127-132 vol.1.
- [9] A. Fagiolini, G. Valenti, L. Pallottino, G. Dini, and A. Bicchi, "Decentralized intrusion detection for secure cooperative multi-agent systems," in *46th IEEE Conference on Decision and Control*, 2007, Conference Proceedings, pp. 1553-1558.
- [10] G. Kaminka and M. Tambe, *Social comparison for failure detection and recovery*, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 1998, vol. 1365, book section 11, pp. 127-141. [Online]. Available: <http://dx.doi.org/10.1007/BFb0026755>
- [11] S. Koos, A. Cully, and J.-B. Mouret, "Fast damage recovery in robotics with the t-resilience algorithm," 2013. [Online]. Available: <http://arxiv.org/abs/1302.0386>
- [12] C. J. Woolf and Q. Ma, "Nociceptorsnoxious stimulus detectors," *Neuron*, vol. 55, no. 3, pp. 353-364, 2007. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0896627307005375>

- [13] M. Anshar and M.-A. Williams, "Evolving synthetic pain into an adaptive self-awareness framework for robots," *Biologically Inspired Cognitive Architectures*, vol. 16, pp. 8–18, 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S2212683X15000626>
- [14] R. Agrawal and R. Srikant, "Mining sequential patterns," in *Proceedings of the Eleventh International Conference on Data Engineering*. IEEE, 1995, Conference Proceedings, pp. 3–14.
- [15] P. Laird, "Identifying and using patterns in sequential data," in *Algorithmic learning theory*. Springer, 1993, Conference Proceedings, pp. 1–18.
- [16] C. Schwind, *A Theory for Causal Reasoning*, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 1999, vol. 1701, book section 24, pp. 285–288. [Online]. Available: [http://dx.doi.org/10.1007/3-540-48238-5\\_24](http://dx.doi.org/10.1007/3-540-48238-5_24)