Biologically Inspired Cognitive Architectures (2016) **16**, 8–18

Available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/bica

RESEARCH ARTICLE

Evolving synthetic pain into an adaptive self-awareness framework for robots

Muh Anshar [a,b,*], **Mary-Anne Williams** [a]

[a] *University of Technology Sydney, UTS, Australia*
[b] *University of Hasanuddin, UNHAS, Makassar, Indonesia*

Received 15 September 2015; received in revised form 30 September 2015; accepted 25 October 2015

KEYWORDS
Self-Awareness;
Consciousness;
Synthetic pain;
Sensory perceptions;
Reasoning;
Prediction

Abstract
In human–robot interaction, physical contact is the most common medium to be used, and the more physical interaction occurs, at certain times, the higher possibilities of causing humans to experience pain. Humans, at times, send this message out through social cues, such as verbal and facial expressions in which requires robots to have the skill to capture and translate these cues into useful information. The task of understanding human pain concept and its implementation on robots plays a dominant factor in allowing robots to acquire this social skill. However, it is reported that the concept of human pain is strongly related to the concept of human self-awareness concept and cognitive aspects with complex nerve mechanisms, hence, it is crucial to evolving appropriate self-awareness and pain concepts for robots. This paper focuses on imitating the concept of pain into a synthetic pain model, utilised in justifying the integration and implementation an adaptive self-awareness into a real robot design framework, named ASAF. The framework develops an appropriate robot cognitive system-"self-consciousness" that includes two primary levels of self-concept, namely subjective and objective. Novel experiments designated to measure whether a robot is capable of generating appropriate synthetic pain; whether the framework's reasoning skills support an accurate synthetic pain acknowledgement, and at the same time, develop appropriate counter responses. We find that the proposed framework enhances the awareness of robot's body parts and prevent further catastrophic impact on robot hardware and possible harm to human peers.

Introduction

As the number of robots application in various areas of human life grows, it is inevitable to see more collaboration

* Corresponding author at: University of Technology Sydney, UTS, Australia.
E-mail addresses: muh.anshar@gmail.com, anshar@unhas.ac.id (M. Anshar), mary-anne.williams@uts.edu.au (M.-A. Williams).

http://dx.doi.org/10.1016/j.bica.2015.10.002
2212-683X/

tasks take place. During the interaction, humans and robots commonly utilise physical medium to engage, and at a certain situation, this may cause humans to experience pain. At times, humans deliver this pain message through social cues in which requires robots to have the ability to capture and translate these cues into useful information. It is primarily important to have an adequate understanding of human pain concept before addressing the issue on how this concept can be integrated into robot framework.

Studies suggest that human pain is strongly related to the concept of self and it is reported that the degree of one's self-awareness is related to the ability to feel the pain Hsu, Schubiner, Stracks, and Clauw (2010), and the kinds of counter responses humans might generate Steen and Haugli (2001). Investigating pain as a means to develop robot self-awareness raises important questions of how robots can conceptualise, detect and respond to the severity of the pain levels. This pain investigation requires an appropriate model of the pain to be developed so as robots can have an understanding of pain concept. An important question raises about how this concept of pain can be generated as compared to humans; pain is strongly related to human self-awareness concept and cognitive aspects of complex nerve mechanisms. Hence, having an appropriate concept of self-awareness are important as Gorbenko, Popov, and Sheka (2012) mentions that robots with self-awareness have the ability to behave more effectively in novel situations compared to those without it. On top of that, evolving pain concept dedicated for robots becomes priority by which can be used by robots to understand and further to interpret human pain captured through social cues.

This paper addresses the challenge of developing self-aware robots by conceptualising pain and implementing it in a synthetic pain model within a new robot self-awareness framework that provides a novel approach towards the detection of and response to synthetic pain for a robot. Our approach implements the concept of the self divided into subjective and objective self-awareness Lewis (1991). It is important to point out that the terminology of "consciousness" for our robot is to signify the robot's focus of attention, and it should not be understood as to mean human consciousness. By changing the nature of robot consciousness, from objective to subjective or vice versa, a robot has the ability to categorise and recognise internal events. Such a robot would be able to develop an appropriate behaviour in response to the synthetic pain that it is experiencing, e.g., alerting a person for help or changing its behaviour as a preventive action.

Integrating the pain concept into the robot's self-awareness framework enables the robot to perform assigned tasks whilst at the same time, being conscious/aware of its internal state of well-being. This is particularly important in human—robot interaction tasks because if a robot is damaged then that may put people at risk of injury. The remainder of the paper proceeds as follows: related works and highlights of the state-of-the-art in self-awareness and pain studies are presented followed by the overview of proposed robot self-awareness framework. Experiment stages including environmental set up are briefly described followed by the results, evaluation and discussions. Finally, the overall achievement and possible future developments conclude the paper.

Related studies

Self-aware robots and self concept

Studies on the notion of robots being self-aware which are first reported in Michel, Gold, and Scassellati (2004) and Scassellati (2002) do not include how to develop the concept of self. The study reported in Michel et al. (2004) presents the NICO robot that possesses the capability of recognising itself in the mirror. Since then, studies of self-aware robots have continued to grow as reported in Birlo and Tapus (2011) and Bongard, Zykov, and Lipson (2006). In other areas, health services, for instance, the implementation of self-awareness has been highlighted in Agha-Mohammad, Ure, How, and Vian (2014) where systems need to diagnose their state health. The key element of the study captures the uncertainties needed to develop the appropriate control and planning scheme as described in the earlier paper by Marier, Rabbath, and Lechevin (2013).

Much of the literature also identifies the lack of the concept of "self". This paper uses the definition of self-awareness in Lewis (1991) which divided the concept of self into two levels, subjective awareness and objective awareness. Subjective self-awareness concerns the machinery of body and objective self-awareness concerns focusing the attention capability towards one's self, thought, actions and feelings. The author shows that human adults have the ability to function at both levels. Under certain conditions, human adults utilise one level of self-awareness at a time. However, it can be inferred that these two primary levels of self-awareness coexist and that human adults utilise them by switching between them. We utilises this insight, particularly the ability to switch between subjective and objective awareness, and through this action, a new framework can be used to change the robot awareness from subjective to objective, and vice versa. In this framework, we refer to the subjective element as the physical parts of a robot, such as robot motors and joints; and the objective elements as the metaphysical aspects of the robot, such as robot's representation of its position towards an external object.

Human pain concept and self-awareness

Few studies have correlated pain with the self-awareness concept in humans. The earliest reported study conducted by Steen and Haugli (2001) investigates the correlation between musculoskeletal pain and the increase of self-awareness in humans. The study suggests that having an awareness of the internal relationship between body, mind and emotions enable a person to address and understand messages communicated by the pain. Another study carried out by Hsu et al. (2010) proposes that the development of affective self-awareness is associated with the pain's severity level based on reports of people who suffer from fibromyalgia.[1] The study by Steen and Haugli (2001) is based on the assumption that humans attention is directed from

[1] A widespread pain and tenderness in the human body and sometimes accompanied by fatigue, cognitive disturbance and emotional distress Channel (2014).

Table 1 Artificial synthetic pain for robots.

Category	Synthetic pain	Description	Definition	Intensity level
1	Proprioceptive pain	1.0	Potential hardware damage, as an alert signal	"None", "Slight"
2	Inflammatory Reduction Pain	2.1	Predicted robot hardware damage	"None", "Slight"
		2.2	Real robot hardware damage	"Moderate", "Severe"
3	Sensory Malfunction Pain	3.1	Abnormal function of internal sensors	"None", "Slight"
		3.2	Damage internal sensors	"Moderate", "Severe"

pain acknowledgement to self-awareness, in contrast for Hsu et al. (2010) attention is directed from affective self-awareness to accurate pain acknowledgement. In terms of the study of self-awareness in robotics in relation to pain, Koos, Cully, and Mouret (2013) was first to refer to the term "pain" in their report. The paper mentions that when humans are injured, the cause of the "pain" is not fully understood and for some painful movements, humans prefer to learn and avoid those movements. However, the paper does not explicitly conceptualise "pain" into their approach.

In the health area, Prkachin (2009) and Sikka et al. (2015) reported related findings that focus on pain acknowledgement through human facial expression. The studies demonstrate that their approaches allow robots to recognise human pain through facial expression. However, their concepts of pain do not derive from robot self-awareness concept. A study reported in Koos et al. (2013) also utilises the term "pain" to represent when a robot suffered from joint failure, however, there is no a clear concept of how this kind of pain is derived.

Artificial pain concept

To fill this research gap, we are inspired by the definition of pain proposed by Woolf (2010) and our artificial pain classes proposal are developed accordingly. There are three classifications of synthetic pain derived from the report and for each class, the intensity level of each pain derived from Zhu (2014). Descriptions of the proposal are described in the following Table 1.

Adaptive self-awareness framework

Our adaptive self-awareness framework for robots or known as ASAF is comprised of several elements as shown in Fig. 1. Important elements of the ASAF, Consciousness Direction, Synthetic Pain Description, Robot Mind, Action Execution and Database are discussed briefly in the following subsections.

Consciousness Direction

The ability to redirect attention between the two levels of awareness refers to as consciousness by Lewis (1991). We utilise this concept in this paper, therefore, our robot consciousness refers to the cognitive aspect of the robot that is used to specifically signify the focus of robot's attention. There are two predominant factors in directing robot consciousness: (i) the ability to focus attention on a specified physical aspect of self and (ii) the ability to foresee, and at the same time, to be aware of the consequences of predicted actions. Our proposal formulates how to address these two aspects so that they can be developed and built into a robot self-awareness framework so that the detection of synthetic pain can be acknowledged and responded to in an appropriate way. Robot awareness is mapped to a discrete range 1–3 for subjective and 4–6 for objective elements. Changing the value of Consciousness Direction (CDV) allows the exploration of these regions, and at the same time, changes the focus of robot attention. It is important to keep in mind, as mentioned in section 'Self-aware robots and self concept', that our subjective element specifies the physical parts of a robot, such as robot motors and joints; and the objective elements signifies the metaphysical aspects of the robot, such as robot's representation of its position towards an external object. The Robot Mind sets the CDV and determines the conditions of exploration of robot awareness regions, either constrained or unconstrained conditions. The structure of robot awareness regions and CDV are shown in Fig. 2.

Synthetic Pain Description

In order to generate synthetic pain on the robot, we set the robot joint restriction region that should be avoided. These joint restriction regions contain joint robot values that are considered as faulty joint values. Synthetic pain can then be generated when the robot joint moves into this region, mimicking the effect of musculoskeletal pain in human, mentioned in previous section 'Self-aware robots and self concept'. Joint movement is monitored by the proprioceptive perception of the robot, which can subsequently be used by the Robot Mind to reason upon.

Robot Mind

The Robot Mind can utilise causal reasoning as reported in Morgenstern and Stein (1988), Schwind (1999, chapter 24), and Stein and Morgenstern (1994) to draw conclusions from its perceptions. Our idea of reasoning is derived from human cognitive competencies that integrate the cause and effect relationship Reisberg (2013). This allows our framework to allow robots to adapt to the world by predicting their own future states through reasoning about the perceived/detected facts. We integrate our approach with sequential pattern prediction Laird (1993) and Agrawal and Srikant

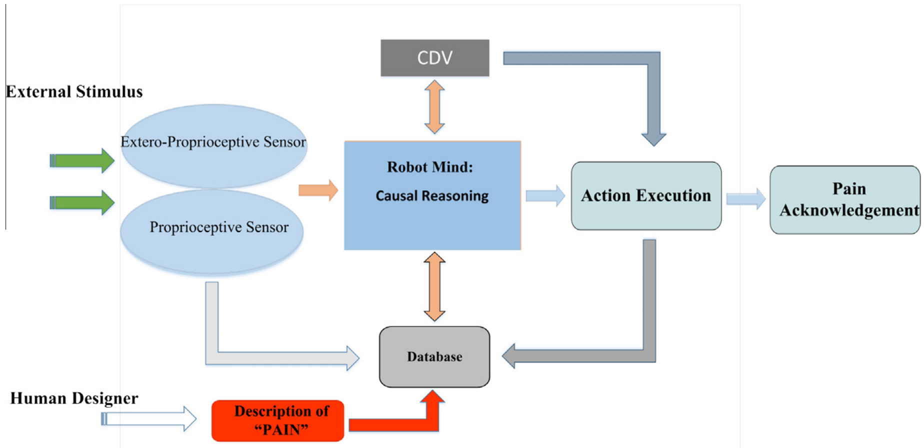

Fig. 1 Adaptive robot self-awareness framework (ASAF).

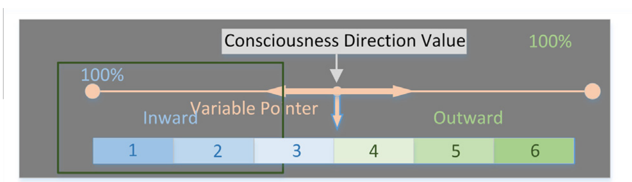

Fig. 2 Robot awareness region and CDV: (1) Upper Limit Subjective, (2) Lower Limit Subjective, (3) Left Border Subjective—Objective, (4) Right Border Subjective—Objective, (5) Lower Limit Objective, and (6) Upper Limit Objective.

(1995) to capture the behaviour of the observed facts and then use them to predict the possible future conditions. In ASAF, a robot's decision making is built on the associative theories Schwind (1999, chapter 24), which utilises covariance information obtained from sequence data so as to facilitate the causal reasoning process. The Robot Mind analyses the relationship amongst 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 decisions made from previous sequence predictions are reassessed with the current state, and the results are either kept as history for future prediction or amendment actions take place before placement proceeds. This cycle repeats only if current data and predicted values are not classified in the restricted region that refers to the painful joint settings. Once the reasoning of Robot Mind indicates that the joint moves are heading towards or falls into these joint restricted regions, the Robot Mind performs three consecutive actions:

- Setting the robot awareness into constrained condition.
- Modifying the CDV, which will shift robot's focus of attention to the subjective element of its awareness.
- Providing counter response actions, by collecting available pre-defined sets of counter response actions (Event—Goal Pair stored in the Database), such as alerting human peers through verbal expressions and increasing robot joint stiffness.

The components and pathways of the overall reasoning of the Robot Mind are illustrated in Fig. 3. The values of faulty joint settings and the limit of the Consciousness Region areas are defined and placed in the Database. Once the collaboration task, that involves human and robot, takes place, the Robot Mind specifies the awareness to random style. This style means that the attention may focus in one of amongst the sixth regions by randomly selecting the CDV. Once selected, the Robot Mind is set in an unconstrained condition, allowing task execution and collaboration to proceed. Whilst the awareness is on the afore-selected region, the Robot Mind at the same time, monitors its proprioceptive sensor, joint arm sensors which physically involved in interaction with a human peer. The change of joint sensor readings produces the change in the pattern, and this situation is captured and used as the element of reasoning of the Robot Mind. As the joint moves, the robot *Belief*, *Desire* and *Intention* are subjects to changes and the Action Executions transformed the results into primitive actions for further execution. For every prediction that may introduce higher risk of arm joint to experience the faulty joint settings, the Robot Mind alters the CDV, causing the awareness to focus on the robot arm (subjective awareness); and at the same time, sets the robot internal state into constrained. Once this situation is reached, the robot joint stiffness is set to a maximum value and alerting the human peers via verbal notification.

The decision-making process of a robot uses the ASAF based BDI architecture as illustrated in Fig. 4.

Action Execution

The Action Execution module is responsible for translating each decision into one of three intentions: (i) Send alert; (ii) shift the awareness level through CDV; or (iii) modify joint stiffness values in the robot's body. If the decision is to maximise the joint stiffness, the robot will disregard any external physical interaction, e.g., interaction with a human. By increasing stiffness, the robot joint will resist any force generated from the physical interaction, and as a result, the robot will be prevented from experiencing the faulty joint settings. In addition to that, having a

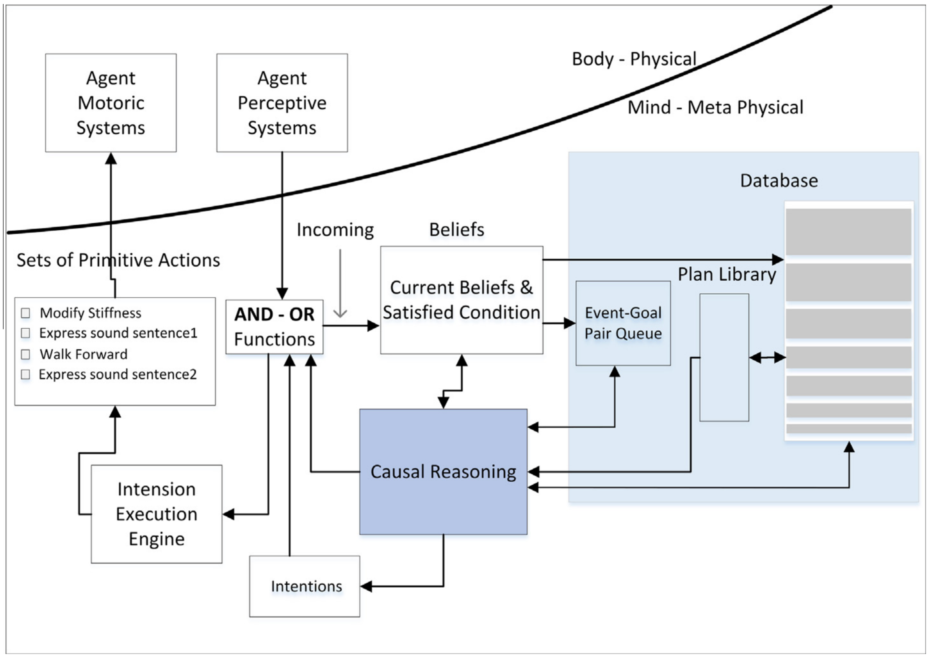

Fig. 3 Robot Mind structure.

resistance of the robot joint, the human, on the other hand, will realise that the robot is no longer willing to involve in the interaction.

Database

The Robot Database contains of a set of predefined Consciousness Regions, a set of faulty joint setting constitutes as joint pain areas, pre-recorded sequence of arm joint movements (explained in section 'Experiment scenario'), Event-Goal pairs and temporary arm joint position readings.

Experiments

The implementation of the ASAF for the experiments involved four key issues:

1. The realisation of "self-consciousness" has two elements: (i) the ability to focus attention on a specified physical aspect of self, and (ii) the ability to foresee, and at the same time, be aware of the consequences of predicted joint positions.
2. The position of the right elbow joint and the time corresponds to the collection time are obtained from the proprioceptive sensor, namely, joint data and time data.
3. The reasoning process produces response times by predicting time data for specific elbow motion joint data.
4. "Robot consciousness" states are divided using two state conditions: (i) constrained where the Robot Mind is allowed to explore its entire "consciousness" region, e.g., Region 1 to Region 6, and (ii) unconstrained where the Robot Mind is limited to the highest level of its subjectivity, Region 1.

The purpose of the experiment is threefold: (1) to conceptualise "robot self-awareness" and use it to learn and predict joint elbow parameters; (2) to conceptualise synthetic pain using robot faulty joint settings and the response to pain times; and (3) to verify proper responses of a robot through proportionate joint stiffness settings that change with respect to the force provided by the human to the

```
1: database ← (on real time or pre-recorded)
2: CDV region ← random
3: Consciousness ← (CDV)
4: Input ← joint data; Input ← sensing time
5: While Assessing quantity of data retrieve
      If Similarity Count == Limit set:
         same_count ← Limit
      If Different Count == Limit set:
         differ_count ← Limit
      Else:
         uni_count ← Limit
6: for each incoming data, DO reasoning
    data_Interval ← same_count OR differ_count OR uni_count
    update (current_data, current Belief, current state)
      ψ ← update
    start predicting ← (current, past, future)
    faulty_joint assessment limit:
       faulty_prediction ← update_prediction(Belief)
       update(Belief)
    P ← update
    database: update( prediction)
    end for
7: get event-goal pair
    e ← database
8: execute Intention
    Consciousness ← update
    Alert (activation)
    Stiffness (setting)
9: return Input
```

Fig. 4 Robot Mind reasoning process.

robot's arm, e.g. a strong human force is met with stronger stiffness.

Environmental set up

We implement the Adaptable Robot Self-Awareness Framework (ASAF), test and evaluate it using several human—robot interaction experiments. The overview of the experimental set-up is described as follows: It involves two agents, the NAO humanoid robot and a human partner. The robot and the human are working together to perform a simplified task. The robot is situated in a static environment where it is required to walk towards the human partner. When it reaches a certain distance, the robot stops and then raises its right hand to point to the human partner. This position initiates the interaction in which the human pushes the robot's right arm. The robot uses verbal communication to direct the human to undertake actions.

Human interaction scenario

The experiment involved two interacting agents, the NAO as Agent 1 and a human partner as Agent 2. The agents are initially separated by a distance of 1 m. When the experiment commenced, Agent 1 (robot) walks towards Agent 2 (human partner) and stops at about 0.3—0.4 m from Agent 2's feet. Agent 1 raises its right arm until it is perpendicular about its chest.

Experiment scenario

Experiments are divided into offline and online scenarios. In offline scenarios, the experiment contained two stages. In stage one, the interaction between Agent 1 and Agent 2 takes place with the exception that Agent 1 does not pose the awareness framework. The purpose for the stage one experiment is to collect a set of joint elbow data — Joint Data and Time Data, and places them into Agent 1 database. Two types of interaction are used to collect the data sets, (i) without physical interaction (phase 1) and (ii) with physical interaction (phase 2). Physical interaction means that Agent 2 reacted by pushing the arm of Agent 1, whilst without physical interaction means that Agent 2 remains standing in front of Agent 1 without performing a pushing action. Each phase contains three trials that made up a set of six datasets in total. In the next stage, only Agent 1 with an activated awareness framework performed, without Agent 2 involvement. The experiment is simulated on Agent 1's mind, and the data for interaction is retrieved from the datasets obtained from the previous stage stored in the agent database. This experiment produces an additional set of six datasets, containing data predictions. This stage is designed, first to evaluate the mind simulation of Agent 1's reasoning performances through its capability to shift its "consciousness" using pre-recorded elbow joint datasets. Second, to measure the accuracy of agent's reasoning skills through its capability to predict and generate an accurate pain acknowledgement, and the counter-responses carried out by the intention execution engine. In the online scenario, Agent 1 and Agent 2 perform an interaction. However, this time, Agent 1 performed with an activated self-awareness framework. Hence, interaction with Agent 2 would provide the joint data straight away for further processing. This experiment is divided into phase one without physical interaction and phase two with physical interaction. The objectives of this experiment scenario are to measure the overall performances of the agent with self-awareness framework embedded in its mechanism, including how robust the framework is working within real world environment. All data collected on these two scenarios were ordered according to their reading sequences unless it is stated differently.

Predefined parameters

There are two parameters that are predefined and set uniform throughout the experiment: (1) Consciousness Region and (2) Synthetic Pain.

Consciousness Region

There are six Consciousness Regions with the width of each region is evenly set (Table 2). Visitation of any Consciousness Region under unconstrained condition is set randomly and this condition remains the same through out the experiment. This Consciousness Region is changed to subjective awareness and the condition is set to constrained, however, if the Robot Mind reasoning confirms that the next arm movement falls into the faulty joint settings. Our experiments are set to follow this setting as to demonstrate the ability of the ASAF to shift the robot consciousness whenever the interaction with the human peers increases the possibility of synthetic pain generation. The overall situation of robot awareness during experiments are shown in Table 3.

Synthetic Pain Region

The robot synthetic pain is generated from three predefined of robot's right arm joint positions and each position constitutes to the level of synthetic pain magnitude. The closest join position constitutes to the highest synthetic pain level as shown in Table 4.

Experiment results

Sequence of joint positions data that are used in the offline scenario, without physical interaction and with physical interaction, are shown in Tables 5 and 6 respectively; and for the online scenario, without physical interaction and with physical interaction, are shown in Tables 7 and 8 respectively.

Analyses and discussions

During the offline scenario (without physical interaction), average joint prediction and reasoning reached the margin of error at zero percent with a standard deviation (σ) equal to zero and the standard deviation for time prediction reaches the maximum deviation at 0.05 (see Table 9). Agent 1 commences its prediction at the fourth sequence and the reasoning is capable of maintaining the accuracy of prediction which allows the agent to identify the scenario.

Table 2 Consciousness Region divisions.

Subjective awareness		Subjective—objective awareness		Objective awareness	
Upper limit	Lower limit	Left limit	Right limit	Lower limit	Upper limit
1	50	51	100	101	150

Table 3 Consciousness Region state.

Consciouness Region		Robot action during visitation	
		Unconstrained	Constrained
Subjective awareness	Upper limit	Low stiffness on arm joint	Increased stiffness and alert human peer
	Lower limit	Not modelled	Not available
Subjective—objective awareness	Left limit	Not modelled	Not available
	Right limit	Not modelled	Not available
Objective awareness	Lower limit	Not modelled	Not available
	Upper limit	Not modelled	Not available

Table 4 Faulty joint settings.

High stage	Medium stage	Low stage
1.5621	1.5521	1.5421

During physical interaction experiments, on average, Agent 1 deciphered the incoming interaction data at the fourth sequence with standard deviation relatively low; from 0.03 to 0.11, and the standard deviation for the time prediction increases up to 0.54 (see Table 10). The final sequence, trail data 10, serves as the primary source of determining the Agent 1's decision, whether to proceed to the next sequence or to constrain the robot awareness through joint stiffness, resulting in resistance within the Agent 1 elbow. With prediction error of 16.83%, Agent 1 is still able to deliver accurate decisions in the given situation and predicted the consequence of the data in a timely manner.

In contrast to the offline scenario, the cycles of experiment in online scenario experiments are much shorter. For

Table 5 Offline pre-recorded without physical interaction.

No.	Trial 1		Trial 2		Trial 3	
	Elbow data	Time	Elbow data	Time	Elbow data	Time
1	0.00873	221.72	0.00873	316.95	0.00873	385.60
2	0.00873	222.24	0.00873	317.47	0.00873	386.12
3	0.00873	222.75	0.00873	317.98	0.00873	386.64
4	0.00873	223.28	0.00873	318.52	0.00873	387.16
5	0.00873	223.84	0.00873	319.07	0.00873	387.72
6	0.00873	224.40	0.00873	319.63	0.00873	388.30
7	0.00873	224.96	0.00873	320.19	0.00873	388.85
8	0.00873	225.51	0.00873	320.75	0.00873	389.41
9	0.00873	226.07	0.00873	321.31	0.00873	389.97
10	0.00873	226.63	0.00873	321.86	0.00873	390.54
11	0.00873	227.19	0.00873	322.42	0.00873	391.10
12	0.00873	227.76	0.00873	322.98	0.00873	391.66
13	0.00873	228.32	0.00873	323.54	0.00873	392.22
14	0.00873	228.88	0.00873	324.11	0.00873	392.78
15	0.00873	229.44	0.00873	324.67	0.00873	393.35
16	0.00873	230.01	0.00873	325.24	0.00873	393.92
17	0.00873	230.58	0.00873	325.80	0.00873	394.49
18	0.00873	231.14	0.00873	326.37	0.00873	395.05
19	0.00873	231.71	0.00873	326.95	0.00873	395.62
20	0.00873	232.28	0.00873	327.52	0.00873	396.19
21	0.00873	232.85	0.00873	328.10	0.00873	396.76

Table 6 Offline pre-recorded with physical interaction.

No.	Trial 1		Trial 2		Trial 3	
	Elbow data	Time	Elbow data	Time	Elbow data	Time
1	0.00873	583.66	0.02919	644.54	0.02919	726.37
2	0.02765	584.17	0.02919	645.05	0.02919	726.89
3	0.14884	584.69	0.03072	645.56	0.04606	727.40
4	0.34519	585.22	0.03072	646.09	0.22861	727.93
5	0.57836	585.76	0.13043	646.62	0.40348	728.46
6	0.78238	586.31	0.45871	647.14	0.60444	728.98
7	1.02782	586.85	0.73023	647.66	0.88669	729.50
8	1.30701	587.40	0.94959	648.19	1.08765	730.03
9	1.51870	587.94	1.14441	648.72	1.25485	730.55
10	1.56207	588.48	1.37297	649.25	1.42359	731.08
11	1.56207	589.03	1.53251	649.77	1.56012	731.60
12	1.56207	589.56	1.56207	650.30	1.56207	732.14
13	1.56207	590.08	1.56207	650.83	1.56207	732.66
14	1.56207	590.63	1.56207	651.38	1.56207	733.19
15	1.56207	591.18	1.56207	651.93	1.56207	733.74
16	1.56207	591.74	1.56207	652.48	1.56207	734.29
17	1.56207	592.30	1.56207	653.04	1.56207	734.85
18	1.56207	592.85	1.56207	653.59	1.56207	735.40
19	1.56207	593.40	1.56207	654.14	1.56207	735.96
20	1.56207	593.95	1.56207	654.70	1.56207	736.51

Table 7 Online without physical interaction.

No.	Trial 1		Trial 2		Trial 3	
	Elbow data	Time	Elbow data	Time	Elbow data	Time
1	0.02765	367.27	0.02765	551.84	0.02765	793.73
2	0.02765	367.79	0.02765	552.36	0.02765	794.25
3	0.02765	368.30	0.02612	552.87	0.02765	794.76
4	0.02765	368.83	0.02765	553.4	0.02765	795.28
5	0.02765	369.38	0.02765	553.93	0.02612	795.84
6	0.02765	369.94	0.02765	554.46	0.02765	796.38
7					0.02612	796.94

Table 8 Online with Physical Interaction.

No.	Trial 1		Trial 2		Trial 3	
	Elbow data	Time	Elbow data	Time	Elbow data	Time
1	0.22247	38.88	0.02765	776.5	0.02765	267.12
2	0.26696	39.40	0.02765	777	0.02765	267.64
3	0.37127	39.91	0.02919	777.5	0.02765	268.15
4	0.49246	40.44	0.21940	778	0.06907	268.68
5	0.63205	41.00	0.39735	778.5	0.29917	269.20
6	0.78852	41.54	0.68421	779.1	0.52774	269.73
7	0.95572	42.10	1.30548	781.4	0.71642	270.28
8	1.32695	44.37			0.87902	270.84
9					1.04470	271.39
10					1.41132	273.68

example, in two trials (without physical interaction), trial 1 and trial 3, the data sequence resulting from the hand pushing interaction with Agent 2, Agent 1's arm takes place in a strikingly short time. Decisions made by the Agent 1 were reliable with a standard deviation of prediction considerably very low, 0.00146, with the highest prediction error 0.31%

Table 9 Prediction error — offline no interaction.

Data	Prediction cycles						Std D (σ)	Time
	1 (%)	2 (%)	3 (%)	4 (%)	5 (%)	6 (%)		Std D (σ)
4	0.00							
5	0.00	0.00					0.00	0.02
6	0.00	0.00	0.00				0.00	0.03
7	0.00	0.00	0.00	0.00			0.00	0.05
8	0.00	0.00	0.00	0.00	0.00		0.00	0.05
9		0.00	0.00	0.00	0.00	0.00	0.00	0.01
10			0.00	0.00	0.00	0.00	0.00	0.01
11				0.00	0.00	0.00	0.00	0.02

Table 10 Prediction error — offline physical interaction.

Data	Prediction Cycles						Std D (σ)	Time
	1 (%)	2 (%)	3 (%)	4 (%)	5 (%)	6 (%)		Std D (σ)
4	0.00							
5	3.68	0.00					0.03	0.24
6	4.45	2.92	0.00				0.04	0.35
7	9.36	1.69	4.14	0.00			0.05	0.44
8	17.64	2.91	11.66	3.37	0.00		0.07	0.54
9		0.77	12.43	0.00	6.75	0.00	0.07	0.52
10			3.64	20.21	30.33	16.83	0.11	0.45

Table 11 Prediction error — online without physical interaction.

Data	Prediction cycles			Std D (σ)	Time
	1 (%)	2 (%)	3 (%)		Std D (σ)
4	0.00				0.01
5	0.00	0.00		0.00	0.03
6	0.00	0.00	0.00	0.00	0.04

Table 12 Prediction error — online with physical interaction.

Data	Prediction cycles						Std D (σ)	Time
	1 (%)	2 (%)	3 (%)	4 (%)	5 (%)	6		Std D (σ)
4	0.00							
5	1.84	0.00					0.01	0.31
6	5.37	1.69	0.00				0.03	0.43
7	9.97	4.45	1.07	0.00			0.04	0.56
8	34.97	27.61	22.55	20.40	0.00		0.13	1.33

and the deviation level of the time prediction is considerably low as well, 0.01 (see Table 11). In trial 2, on the contrary, the joint sensor readings are volatile even though it supposed to remain steady for each specific data reading. Agent 1 is capable of identifying this anomalous situation and correctly maps the behaviour of the Agent 1's arm. It can be seen by the standard deviation for joint elbow prediction remains very low, 0.00146 by the end of the experiment.

The reasoning behaviour of Agent 1 for the three experiments in the online scenario with physical interaction has a relatively similar pattern, and it manages to accurately predict the elbow data from its arm movement pattern. The prediction error made at the first cycle is relatively high,

Table 13 State of awareness.

Feeding data	CDV	Experiment	Awareness		Internal states
			Region	Early type	
Offline	131	No physical interaction	6	High priority objective	1
	33		2	Low priority subjective	2
	17		1	High priority subjective	3
	80	Pushing arm	4	Right border subjective—objective	4
	110		5	Low priority objective	5
	3		1	High priority subjective	6
Online	14	No physical interaction	1	High priority subjective	7
	62		3	Left border subjective—objective	8
	116		5	Low priority objective	9
	6	Pushing arm	1	High priority subjective	10
	126		6	High priority objective	11
	138		6	High priority objective	12

Table 14 Internal states after reasoning process.

Internal state	Awareness		Synthetic pain	
	Final type	Level	Categories	Intensity
After reasoning				
1	—	Unconstrained	No pain	—
2	—			
3	High priority subjective			
4	—	Unconstrained	No pain	—
5	—			
6	High priority subjective	Constrained	1:0 Proprioceptive	"Slight"
7	—	Unconstrained	No pain	—
8	—			
9	Low priority objective			
10	—	Unconstrained	No pain	—
11	—			
12	High priority subjective	Constrained	1:0 Proprioceptive	"Slight"

34.97%. The standard deviation prediction for the eight sequences of data reading, however, is still relatively small, 0.13 and at the same time, prediction reaches the highest standard deviation value at 1.33. Overall, the agent is still capable of identifying future consequence and preventing the agent from experiencing the synthetic pain. Table 12 shows the overall prediction cycles of the agent. The awareness during the early stage is freely explored, however, at the Robot Mind recommendation, the attention of awareness is subject to change. The overall states of the Agent 1 awareness during the experiment are shown in Table 13 below. 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 joint elbow stiffness to a degree that the human partner can detect resistance. Overall, the Agent 1 is still capable of identifying future consequences thus preventing it from entering the faulty join region. The robot internal states after the reasoning process shown in Table 14. The internal state contain several states, the offline experiments formed state 1 to state 6, and the online ones constitutes to state 7 to state 12. It can be seen that at early stages, the Robot Mind is at Unconstrained level which indicates that no synthetic pain is generated. In any case, the awareness type is forced to switch to a High Priority Subjective type once the reasoning process predicts that the arm joint moves to the faulty joint regions. In addition, the awareness type of High Priority Subjective may be revisited during an Unconstrained level without generating any synthetic pain.

Conclusions and future works

Using the Adaptive Robot Self-Awareness Framework, we are able to demonstrate that a robot can use it to develop an accurate pain acknowledgement and appropriate response. The causal reasoning through sequential pattern prediction enabled the robot's decision making to embrace the past, the current and the future considerations as it built its expectations during interaction, which leads to effective and accurate decisions using self-awareness and a concept of synthetic pain.

The experiments have shown that the robot was able to become aware of its body parts by demonstrating the ability to foresee its future state from the current state. Skills of predicting the consequence of its body behaviour and classifying it into appropriate synthetic pain categories, and at the same time, taking proper counter-responses in a timely fashion, have been used to prevent the robot from experiencing other kinds of pain, such as inflammatory reduction or Sensory miss functional pain, which may lead to significant hardware damage to the robot. Also, the innovative ASAF successfully disregarded noisy sensor data from the robot's physical body, by proposing amended data sequences derived from previous data sequence prediction. This allows the robot to generate robustly accurate alternatives to its decisions based on self-awareness. Our findings also suggest that during the High Priority Subjective Awareness type, the Robot Mind does not necessarily to be in Constrained level. This finding demonstrates that ASAF framework is capable of exploring any of its Consciousness Regions, including focusing attention to its body parts.

Pain acknowledgement is crucial in human—robot interactions and developing the concept of pain within an adaptive awareness framework for robots requires reasoning skills embedded into the decision-making processes of the robot. 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. Also, the reasoning mechanism will be further developed, utilising and comparing the performance with different learning approaches. Future works will also include carefully designed experiments that evaluate the other two classes of synthetic pain proposed in this paper. Furthermore, experiments should be designed to accommodate the modelling of robot actions during visitation of any other regions than subjective awareness (upper limit) as described in section 'Consciousness Region', particularly in Table 3. Our future works will also look into a further implementation of the ASAF framework in addressing robot empathy towards synthetic pain.

Acknowledgments

The authors would like to thank the Innovation and Enterprise Research Lab, Magic Lab, UTS, in development process of our ASAF framework and conducting the experiments. Also, this work is part of the collaboration research with the Advanced Artificial Research Community, Electrical Engineering Department, University of Hasanuddin, Makassar, Indonesia.

References

Agha-Mohammad, A.-A., Ure, N. K., How, J. P., & Vian, J. (2014). Health-aware stochastic planning for persistent package delivery missions using quadrotors. In *Proceedings of the IEEE international conference on intelligence robotics systems (IROS)*.

Agrawal, R., & Srikant, R. (1995). Mining sequential patterns. In *Proceedings of the eleventh international conference on data engineering, 1995* (pp. 3—14). IEEE.

Birlo, M. & Tapus, A. (2011). The crucial role of robot self-awareness in HRI. In *6th ACM/IEEE international conference on human—robot interaction (HRI), 2011* (pp. 115—116).

Bongard, J., Zykov, V., & Lipson, H. (2006). Resilient machines through continuous self-modeling. *Science, 314*(5802), 1118—1121.

Channel, B. H. (2014). Fibromyalgia.

Gorbenko, A., Popov, V., & Sheka, A. (2012). Robot self-awareness: Exploration of internal states. *Applied Mathematical Sciences, 6* (14), 675—688.

Hsu, M. C., Schubiner, H., Stracks, J. S., & Clauw, D. J. (2010). Sustained pain reduction through affective self-awareness in fibromyalgia: A randomized controlled trial. *Journal of General Internal Medicine, 25*(10), 1064—1070.

Koos, S., Cully, A., & Mouret, J.-B. (2013). Fast damage recovery in robotics with the t-resilience algorithm.

Laird, P. (1993). Identifying and using patterns in sequential data. In *Algorithmic learning theory* (pp. 1—18). Springer.

Lewis, M. (1991). Ways of knowing: Objective self-awareness or consciousness. *Developmental Review, 11*(3), 231—243.

Marier, J. S., Rabbath, C. A., & Lechevin, N. (2013). Health-aware coverage control with application to a team of small uavs. *IEEE Transactions on Control Systems Technology, 21*(5), 1719—1730.

Michel, P., Gold, K., & Scassellati, B. (2004). Motion-based robotic self-recognition. In *Proceedings of the 2004 IEEE/RSJ international conference on intelligent robots and systems, 2004 (IROS 2004)* (Vol. 3, pp. 2763—2768).

Morgenstern, L. & Stein, L.A. (1988). Why things go wrong: A formal theory of causal reasoning. In *AAAI* (pp. 518—523).

Prkachin, K. M. (2009). Assessing pain by facial expression: Facial expression as nexus. *Pain Research & Management: The Journal of the Canadian Pain Society, 14*(1), 53—58.

Reisberg, D. (2013). *The Oxford handbook of cognitive psychology*. Oxford University Press.

Scassellati, B. (2002). Theory of mind for a humanoid robot. *Autonomous Robots, 12*(1), 13—24.

Schwind, C. (1999). *A theory for causal reasoning*. *Lecture Notes in Computer Science* (Vol. 1701). Berlin Heidelberg: Springer, pp. 285—288.

Sikka, K., Ahmed, A. A., Diaz, D., Goodwin, M. S., Craig, K. D., Bartlett, M. S., et al (2015). Automated assessment of children's postoperative pain using computer vision. *Pediatrics, 136*(1), 1—8.

Steen, E., & Haugli, L. (2001). From pain to self-awareness a qualitative analysis of the significance of group participation for persons with chronic musculoskeletal pain. *Patient Education and Counseling, 42*(1), 35—46.

Stein, L. A., & Morgenstern, L. (1994). Motivated action theory: A formal theory of causal reasoning. *Artificial Intelligence, 71*(1), 1—42.

Woolf, C. J. (2010). What is this thing called pain? *The Journal of Clinical Investigation, 120*(11), 3742.

Zhu, S. (2014). Pain expression recognition based on PLSA model. *The Scientific World Journal, 2014*.