ELSEVIER

Robotics and Autonomous Systems 43 (2003) 97–108

Robotics and Autonomous Systems

www.elsevier.com/locate/robot

Anchoring in a grounded layered architecture with integrated reasoning

Stuart C. Shapiro [a,*], Haythem O. Ismail [b]

[a] Department of Computer Science and Engineering, and Center for Cognitive Science, University at Buffalo,
The State University of New York, 201 Bell Hall, Buffalo, NY 14260-2000, USA
[b] Department of Engineering Mathematics and Physics, Cairo University, Giza, Egypt

Abstract

The GLAIR grounded layered architecture with integrated reasoning for cognitive robots and intelligent autonomous agents has been used in a series of projects in which Cassie, the SNePS cognitive agent, has been incorporated into hardware- or software-simulated cognitive robots. In this paper, we present an informal, but coherent, overview of the GLAIR approach to anchoring the abstract symbolic terms that denote an agent's mental entities in the lower-level structures used by the embodied agent to operate in the real (or simulated) world. We discuss anchoring in the domains of: perceivable entities and properties, actions, time, and language.

Keywords: Anchoring; Symbol grounding; Autonomous agents; Cognitive robotics

1. Introduction

GLAIR (grounded layered architecture with integrated reasoning) is a three-level architecture for cognitive robots and intelligent autonomous agents [15,16]. GLAIR has been used in the design and implementation of Cassie, a cognitive robot [20–23,39,41, 44,46–48,50], which has been implemented as a hardware robot and in various software-simulated versions. The capabilities of the embodied Cassie have included: input and output in fragments of English, reasoning, performance of primitive and composite acts, motion, and vision.

Previous papers have described various aspects of GLAIR and Cassie. In this paper, we present, for the first time, a coherent, unified, overview of the GLAIR approach to anchoring the abstract symbolic terms that denote an agent's mental entities in the lower-level structures used by the embodied agent to operate in the real (or simulated) world.

In Section 2 we give an overview of the three levels of the GLAIR architecture. In Section 3 we discuss a hardware implementation of Cassie. In Section 4, we discuss anchoring in the domains of: perceivable entities and properties, actions, time, and language. In Section 5, we discuss some related work, and in Section 6, we summarize the paper. This paper has deliberately been kept as an informal, but coherent, overview of our approach. For more details, and more formal presentations, of particular aspects of our approach, see the papers cited herein.

2. GLAIR

GLAIR (grounded layered architecture with integrated reasoning) consists of three levels: the

---

[*] Corresponding author. Fax: +1-716-645-3464.
E-mail address: shapiro@cse.buffalo.edu (S.C. Shapiro).

0921-8890/03/$ – see front matter © 2003 Elsevier Science B.V. All rights reserved.
PII: S0921-8890(02)00352-4

knowledge level, the perceptuo-motor level, and the sensori-actuator level.

The knowledge level (KL) is the level at which conscious reasoning takes place. The KL is implemented by the SNePS knowledge representation and reasoning system [46,48,50], and its subsystem SNeRE (the SNePS rational engine) (see [28–31] and [53, Chapter 4]), which is used for scheduling and initiating the execution of intentional acts.

We refer to the KL as the "conscious" level, since that is the locus of symbols accessible to reasoning and to natural language interaction. It is the level containing the "abstract-level representations of objects" [5,6]. Similarly, the KL-level acts are "intentional" in the sense that they are scheduled as a result of natural language understanding and reasoning.

Atomic symbols in the KL are terms of the SNePS logic [42]. Symbol structures in the KL are functional terms in the same logic [40,42]. All terms denote mental entities [31,46]. For example, if Cassie is asked to "*Find a green thing*", she conceives of an entity whose only properties are being green and being a thing, by creating a KL term denoting that entity, and KL terms denoting propositions that the entity is green and that the entity is a thing, even though no such object, without further properties, exists in the world. When, in response to this request, Cassie find a particular green robot she recognizes (re-cognizes), by already having a KL term for it, she adds a KL term for the proposition that the two entities have the same extension. (Compare Frege's example that "The Morning Star is the Evening Star" [10].) This approach is in general accord with what Jackendoff calls "conceptualist semantics" [25,26]. We will consistently use "entity" for such a mental entity—the denotation of a KL term, and "object" for an object in the real (or simulated) world.

SNePS (and hence the KL) is implemented in Common Lisp.

The perceptuo-motor level (PML) is the level containing the "physical-level representations of objects" [5,6] consisting of object characteristics such as size, weight, texture, color, and shape. At this level objects are not characterized by KL terms such as categories (box, robot, person, etc.) or properties (green, tall, etc.). The PML also contains routines for well-practiced behaviors, including those that are primitive acts at the KL, and other subconscious activities that ground Cassie's consciousness of its body and surroundings.

The PML has been implemented in three sub-levels:

1. The highest sub-level (which we will refer to as PMLa) has been implemented in Common Lisp, and contains the definitions of the functions that implement the activity represented by KL primitive acts.
2. The middle sub-level (henceforth PMLw) contains a set of Common Lisp symbols and functions defined in the `World` package which use Common Lisp's foreign function facility to link to the lowest sub-level.
3. The lowest sub-level (henceforth PMLc) has been a C implementation of "behavioral networks" [17,18].

The sensori-actuator level (SAL) is the level controlling the operation of sensors and actuators (being either hardware or simulated). The SAL has been implemented in C and other languages, depending on the implementation of the hardware or software-simulated robot.

The Common Lisp programs, PMLc, and the SAL run on different processes, and, in some circumstances, on different machines.

The topic of this paper is our approach to anchoring the KL terms that denote Cassie's (or any GLAIR-based agent's) mental entities in the PML structures used by embodied Cassie to operate in the real world. Briefly, our theoretical stance is that a KL term (symbol) serves as a *pivot*, supporting and coordinating various modalities. Anchoring is achieved by associating (we use the term "aligning") a KL term with one or more PML structures—more than one, if different PML structures are used by different modalities. Some PML structures are accessible to sensors, some to effectors. Others are accessible to natural language interaction. KL terms, but not PML structures, are accessible to reasoning. Cassie's ability to understand a natural language description, and then visually locate an object in the world satisfying that description depends on going from PML structures supporting natural language perception to KL symbol structures, possibly clarified and enhanced by reasoning, to PML structures supporting visual perception. Her ability to describe in natural language an object she is seeing in the world depends on the following that same path in the other direction.

3. The FEVAHR

Cassie in the role of a FEVAHR (foveal extravehicular activity helper-retriever) [2,14,41] was implemented, as a joint project of researchers at the University at Buffalo and researchers at Amherst Systems, Inc., on a Nomad 200 mobile robot, including sonar, bumpers, and wheels, enhanced with an hierarchical foveal vision system [1] consisting of a pair of cameras with associated hardware and software [7]. Henceforth, we will refer to Cassie in the role of a FEVAHR as Cassie$_\mathrm{F}$ (in [14], Cassie$_\mathrm{F}$ is referred to as Freddy).

Cassie$_\mathrm{F}$ operates in a 17 × 17 ft. room containing: Cassie$_\mathrm{F}$; Stu, a human supervisor; Bill, another human; a green robot; three indistinguishable red robots. In the actual room in which the Nomad robot operated, “Stu” was a yellow cube, “Bill” was a blue cube, the green robot was a green ball, and the red robots were red balls. Cassie$_\mathrm{F}$ is always talking to either Stu or Bill. That person addresses Cassie$_\mathrm{F}$ when he talks, and Cassie$_\mathrm{F}$ always addresses that person when she talks. Cassie$_\mathrm{F}$ can be told to talk to the other person, to find, look at, go to, or follow any of the people or other robots in the room, to wander, or to stop. Cassie$_\mathrm{F}$ can also engage in conversations on a limited number of other topics in a fragment of English, similar to some of the conversations in [39]. While Cassie$_\mathrm{F}$ is moving, she avoids obstacles.

Cassie$_\mathrm{F}$’s SAL was designed and implemented by the researchers at Amherst Systems, Inc. Its hierarchical foveal vision system [1,2,7] was implemented and trained to recognize the several colors and shapes of the objects in the room.

Cassie$_\mathrm{F}$’s KL and PML were designed and implemented by the researchers at the University at Buffalo, including the senior author of this paper. During development of the KL, and subsequently, we used several simulations of the robot and of the world it operates in:

The Nomad simulator uses the commercial simulator that was included with the Nomad robot, enhanced by a simulation of Cassie$_\mathrm{F}$’s world and its vision system.

The VRML simulation simulates Cassie$_\mathrm{F}$ and her world by VRML (virtual reality modeling language [3]) objects visible through a world-wide web browser.

The Garnet simulation simulates Cassie$_\mathrm{F}$ and her world by Garnet [11] objects in a Garnet window.

The ASCII simulation, used to create examples for Section 4, implements the PMLw, PMLc, and SAL as sets of Common Lisp functions which print indications of what Cassie$_\mathrm{F}$ would do.

No code at the KL or PMLa levels need be changed when switching among the hardware robot and these four different simulations. All that is required is a different PMLw file of functions that just print messages, or make calls to the appropriate PMLc sub-level.

4. Anchoring in GLAIR

4.1. Perceivable entities

There are KL terms for every mental entity Cassie has conceived of, including individual entities, categories of entities, colors, shapes, and other properties of entities.

There are PML structures (at the PMLw and PMLc sub-levels) for features of the perceivable world that Cassie’s perceptual apparatus can detect and distinguish. For example, in the hardware and Nomad simulator versions of Cassie$_\mathrm{F}$, each distinguishable color and each distinguishable shape is represented by a single integer, while in the VRML simulation, each is represented by a string, and in the Garnet and ASCII simulations, each is represented by a Lisp symbol. Each particular perceived object is represented at this level by an n-tuple of such structures, $\langle v_1, \ldots, v_n \rangle$, where each component, v_i, is a possible value of some perceptual feature domain, D_i. What domains are used and what values exist in each domain depend on the perceptual apparatus of the robot. We will call the n-tuples of feature values “PML-descriptions”.

Our approach to grounding KL terms for perceivable entities, categories, and properties is to align a KL term with a PML-description, possibly with unfilled (null) components. For example, Cassie$_\mathrm{F}$ used two-component PML-descriptions in which the domains were color and shape. In the hardware and Nomad simulator versions, the KL term denoting Cassie$_F$'s idea of blue was aligned with a PML-description whose color component was the PML structure the vision system used when it detected blue in the visual field, but whose shape component was null. The KL term denoting people was aligned with a PML-description whose shape component was the PML structure the vision system used when it detected a cube in the visual field, but whose color component was null. We have implemented alignment in various ways, including association lists, hash tables, and property lists.

Call a PML-description with some null components an "incomplete PML-description", and one with no null components a "complete PML-description". KL terms denoting perceivable properties and KL terms denoting recognizable categories of entities are aligned with incomplete PML-descriptions. Examples include the terms for blue and for people mentioned above, and may also include terms for the properties tall, fat, and bearded, and the categories man and woman. The words for these terms may be combined into verbal descriptions, such as "a tall, fat, bearded man", whose incomplete PML-descriptions may be used to perceptually recognize the object corresponding to the entity so described.

In this paper, we will use "description" (unqualified by "PML") only to mean a verbal description that can be used for perceptual recognition, such as "a tall, fat, bearded man", and not to mean a verbal description that cannot be used for perceptual recognition, such as "a college-educated businessman who lives in Amherst, NY". Cassie might have a KL term for an entity about which she knows no descriptive terms. For example, all she might believe about Fred is that he is a college-educated businessman who lives in Amherst, NY. Thus, she would be incapable of describing Fred (the way we are using "describe"). Nevertheless, it might be the case that Cassie's term denoting Fred is aligned with a complete PML-description. In this case, Cassie would be able to recognize Fred, though not describe him verbally. We call such a PML-description aligned with an entity-denoting term, the entity's PML-description.

A complete PML-description may be assembled for an entity by unifying the incomplete PML-descriptions of its known (conceived of) properties and categories. For example, if Cassie knows nothing about Harry, and we tell her that Harry is a tall, fat, bearded man, she would be able to assemble a PML-description of Harry and recognize him on the street (assuming that Cassie's terms for tall, fat, bearded, and man are aligned with incomplete PML-descriptions). In some cases, this might result in a set of several complete PML-descriptions. For example, the PML-descriptions of some, but not a particular, red chair might include PML-descriptions with different shape components. Once a PML-description is assembled for an entity, it can be cached by aligning the term denoting the entity directly with it. Afterwards, Cassie could recognize the entity without thinking about its description.

To find (come to be looking at) an entity, Cassie finds a PML-description of the entity that is as complete as possible, and directs her perceptual apparatus (via the SAL) to do what is necessary to cause an object satisfying it to be in her visual field. For example, in the Nomad version of Cassie$_F$, the PML-description of Bill is the 2-tuple $\langle 13, 21\rangle$, which is passed to the appropriate SAL routines, which move the cameras until a blue cube is in their field-of-view (see the section on actions, for a description of how actions are grounded).

If Cassie is looking at some object, she can recognize it if its PML-description is the PML-description of some entity she has already conceived of. If there is no such entity, Cassie can create a new KL term to denote this new entity, align it with the PML-description, and believe of it that it has those properties and is a member of those categories whose incomplete PML-descriptions unify with the PML-description of the new entity.

If there are multiple entities whose PML-descriptions match the object's PML-description, disambiguation is needed, or Cassie might simply not know which one of the entities she is looking at.

We are currently investigating the issue of when Cassie might decide that the object she is looking at is new, even though it looks exactly like another she has already conceived of (see [36]).

We have not worked on the problem of recognizing an entity by context. For example, a store clerk might be recognized as any person standing behind a cash register.[1] We speculate that this problem requires

---

[1] This example was suggested by one of the anonymous reviewers of Shapiro and Ismail [43].

Table 1
Objects and descriptions of Cassie$_{\text{F}}$'s world

Object	Color	Shape
`World:Bill`	`World:blue`	`World:square`
`World:Stu`	`World:yellow`	`World:square`
`World:Cassie`	`World:cyan`	`World:circle`
`World:Greenie`	`World:green`	`World:circle`
`World:Redrob-1`	`World:red`	`World:circle`
`World:Redrob-2`	`World:red`	`World:circle`
`World:Redrob-3`	`World:red`	`World:circle`

Table 2
Some of Cassie$_{\text{F}}$'s KL terms and their PML-descriptions

KL term	⟨Color, Shape⟩
`b1`	⟨`World:cyan`, `World:circle`⟩
`b5`	⟨`World:yellow`, `World:square`⟩
`b6`	⟨`World:blue`, `World:square`⟩
`m21`	⟨`World:green`, `nil`⟩
`m25`	⟨`World:red`, `nil`⟩
`m19`	⟨`nil`, `World:square`⟩
`m22`	⟨`nil`, `World:circle`⟩

Table 3
Some of Cassie$_{\text{F}}$'s beliefs

b1's name is Cassie
b5's name is Stu
b6's name is Bill
Cassie is a FEVAHR
FEVAHRs are robots
Bill and Stu are people
Robbie is a green robot
b8, b9, and b10 are red robots
People and robots are things

a combination of KL knowledge and KL–PML alignment. Knowing that a person standing behind a cash register is a clerk is KL knowledge. Recognizing a person, a cash register, and the "behind" relation requires KL–PML alignment.

Consider an example interaction with the ASCII version of Cassie$_{\text{F}}$. In this simulation, created so that interactions can be shown in print, the entire PML and the simulated world are implemented in Common Lisp. The PML-descriptions have two domains, called "color" and "shape". There are seven objects in the simulated world. The Common Lisp symbols that represent these objects and their PML-descriptions are shown in Table 1.[2] Recall that Lisp symbols of the PMLw are in the `World` package, so Lisp prints them preceded by "`World:`".

The KL terms that are aligned with PML-descriptions are shown in Table 2. Notice that `b1`, `b5`, and `b6` are aligned with complete PML-descriptions, while `m21`, `m25`, `m19`, and `m22` are aligned with incomplete PML-descriptions. `b1`, `b5`, and `b6` denote individuals. `m21` and `m25` denote the properties green and red, respectively. `m19` and `m22` denote the categories of people and robots, respectively.

Cassie$_{\text{F}}$'s relevant beliefs about the entities denoted by these terms may be glossed as shown in Table 3. The only descriptive terms Cassie$_{\text{F}}$ has for Bill and Stu are that they are people, and the only descriptive term she has for herself is that she is a robot. Nevertheless, Bill, Stu, and Cassie are aligned with complete PML-descriptions, so she can recognize them. On the other hand, neither Robbie, `b8`, `b9`, nor `b10` are aligned with PML-descriptions, although PML-descriptions can be assembled for them from their properties and categories.

Following is an interaction with Cassie$_{\text{F}}$ about these entities. Sentences preceded by "`:`" are human inputs. Sentences preceded by "`PML:`" and "`SAL:`" are reports of behaviors and simulated actions and perceptions by the ASCII version of Cassie$_{\text{F}}$ at the respective levels, and are not output by the other four versions. Notice that the PML deals with PML-descriptions, and only the SAL deals with (simulated) objects in the world. Sentences beginning with "`I`" are generated by Cassie$_{\text{F}}$. At the beginning of the interaction, Cassie$_{\text{F}}$ is looking at, listening to, and talking to Stu. (See next page).

4.2. Deictic registers

An important aspect of being embodied is being situated in the world and having direct access to components of that situatedness. This is modeled in GLAIR via a set of PML registers (variables), each of which can hold one or more KL terms or PML structures. Some of these registers derive from the theory of the Deictic Center [8], and include: `I`, the register that holds the KL term denoting the agent itself; `YOU`, the register that holds the KL term denoting the individual the agent is talking with; and `NOW`, the register that holds the KL term denoting the current time.

---

[2] The examples in this paper were created using SNePS 2.6 [50] running under Franz, Inc.'s Allegro CL 6.2 [9].

```
 : Find a robot.
PML: The FEVAHR is looking at (World:yellow World:square)
PML: The FEVAHR is looking for something that's (nil World:circle)
SAL: The FEVAHR found World:RedRob-1
PML: The FEVAHR found (World:red World:circle)
 I found a red robot.
PML: The FEVAHR is looking at (World:red World:circle)
 I am looking at a red robot.

 : Find a person.
PML: The FEVAHR is looking at (World:red World:circle)
PML: The FEVAHR is looking for something that's (nil World:square)
SAL: The FEVAHR found World:Stu
PML: The FEVAHR found (World:yellow World:square)
 I found you, Stu.
PML: The FEVAHR is looking at (World:yellow World:square)
 I am looking at you.

 : Find a green thing.
PML: The FEVAHR is looking at (World:yellow World:square)
PML: The FEVAHR is looking for something that's (World:green nil)
SAL: The FEVAHR found World:Greenie
PML: The FEVAHR found (World:green World:circle)
 I found Robbie.
PML: The FEVAHR is looking at (World:green World:circle)
 I am looking at Robbie.

 : Find Bill.
PML: The FEVAHR is looking at (World:green World:circle)
PML: The FEVAHR is looking for something that's (World:blue World:square)
SAL: The FEVAHR found World:Bill
PML: The FEVAHR found (World:blue World:square)
 I found Bill.
PML: The FEVAHR is looking at (World:blue World:square)
 I am looking at Bill.
```

It was by use of these registers that, in the example interaction shown in Section 4.1, Cassie used "I" to refer to the individual denoted by `b1` (herself), "you" to refer to the individual denoted by `b5` (Stu), and the appropriate tense in all the sentences she generated. The use of `NOW` is discussed further in Section 4.5, and language is discussed further in Section 4.6.

Embodiment is further modeled in GLAIR via a set of modality registers.

4.3. *Modality registers*

How does an agent know what it is doing? A standard technique in the Artificial Intelligence literature amounts to the following steps:

(1) I started doing *a* at some previous time or in some previous situation.
(2) I have not done anything since then to stop me from doing *a*.
(3) Therefore, I am still doing *a*.

However, we human’s do not have to follow these steps to know what we are doing, because we have direct access to our bodies.

GLAIR agents know what they are doing via direct access to a set of PML registers termed “modality registers”. For example, if one of Cassie’s modalities were speech, and she were currently talking to Stu, her `SPEECH` register would contain the KL term denoting the state of Cassie’s talking to Stu (and the term denoting Stu would be in the `YOU` register). In many cases, a single modality of an agent can be occupied by only one activity at a time. In that case the register for that modality would be constrained to contain only one term at a time.

One of the modality registers we have used is for keeping track of what Cassie is looking at. When she recognizes an object in her visual field, the KL term denoting the state of looking at the recognized entity is placed in the register, and is removed when the object is no longer in the visual field. If one assumed that Cassie could be looking at several objects at once, this register would be allowed to contain several terms. If asked to look at or find something that is already in her visual field, Cassie recognizes that fact, and doesn’t need to do anything. The following interaction with Cassie$_{\text{F}}$ continues from the previous one:

```
 : Look at Robbie.
PML: The FEVAHR is looking at (World:blue World:square)
PML: The FEVAHR is looking for something that’s (World:green World:circle)
SAL: The FEVAHR found World:Greenie
PML: The FEVAHR found (World:green World:circle)
 I found Robbie.
PML: The FEVAHR is looking at (World:green World:circle)
 I am looking at Robbie.

 : Find a robot.
PML: The FEVAHR is looking at (World:green World:circle)
 I am looking at Robbie.
```

Comparing Cassie’s response to the second request with her response to the previous requests, one can see that she realized that she was already looking at a robot, and so did not need to do anything to find one.

4.4. Actions

Some KL terms denote primitive actions that the GLAIR agent can perform. We call a structure consisting of an action and the entity or entities it is performed on, an “act”. For example, the act of going to Bill consists of the action of going and the object Bill. Acts are denoted by KL functional terms.

Each KL action term that denotes a primitive action is aligned with a procedure in the PML. The procedure takes as arguments the KL terms for the arguments of the act to be performed. For example, when Cassie is asked to perform the act of going to Bill, the PML going-procedure is called on the KL Bill-term. It then finds the PML-description of Bill, and (via the SAL) causes the robot hardware to go to an object in the world that satisfies that description (or causes the robot simulation to simulate that behavior). The PML going-procedure also inserts the KL term denoting the state of Cassie’s going to Bill into the relevant modality register(s), which when `NOW` moves (see Section 4.5), causes an appropriate proposition to be inserted into Cassie’s belief space.

Acts whose actions are primitive are considered to be primitive acts. Composite acts are composed of primitive “control actions” and their arguments, which, themselves are primitive or composite acts. Control actions include sequence, selection, iteration, and non-deterministic choice [21,27–30,50]. There are also propositions for act preconditions, goals, effects, and for plans (what some call recipes) for carrying out non-primitive acts.

In the interactions shown above, sentences preceded by "`SAL:`" were printed by the simulated action function, which was called by the PML procedure aligned with the KL term for finding something. When Cassie was asked to look at Robbie, she did so by finding Robbie, because there is a KL belief that the plan for carrying out the non-primitive act of looking at something is to find that thing.

4.5. Time

As mentioned above, the `NOW` register always contains the KL term denoting the current time [20,23,24,41]. Actually, since "now" is vague (it could mean this minute, this day, this year, this century, etc.), `NOW` is considered to include the entire semi-lattice of times that include the smallest current now-interval Cassie has conceived of, as well as all other times containing that interval.

`NOW` moves whenever Cassie becomes aware of a new state. Some of the circumstances that cause her to become aware of a new state are: she acts, she observes a state holding, she is informed of a state that holds. `NOW` moves by Cassie's conceiving of a new smallest current now-interval (a new KL term is introduced with that denotation), and `NOW` is changed to contain that time. The other times in the old `NOW` are defeasibly extended into the new one by adding propositions asserting that the new `NOW` is a subinterval of them.

Whenever Cassie acts, the modality registers change (see above), and `NOW` moves. The times of the state(s) newly added to the modality registers are included in the new `NOW` semi-lattice, and the times of the state(s) deleted from the modality registers are placed into the past by adding propositions that assert that they precede the new `NOW`.

The following interaction, following the ones shown above, shows an action of Cassie's first being in the present, and then being in the past:

```
 : Who have you talked to?
 I am talking to you.

 : Talk to Bill.
PML: The FEVAHR is starting to talk
 to b6
 I am talking to you, Bill.
 : Who have you talked to?
 I talked to Stu
and I am talking to you.
```

The term denoting the state of Cassie's talking to Stu did not change between the first of these interactions and the third. What did change were: the state of Cassie's talking to Stu was replaced in the `SPEECH` register by the state of Cassie's talking to Bill; a propositional term was added to the KL that the time of talking to Stu was before the time of talking to Bill; and the `NOW` register was changed to include the time of talking to Bill and the times that include it.

To give GLAIR agents a "feel" for the amount of time that has passed, the PML has a `COUNT` register acting as an internal pacemaker [20,24]. The value of `COUNT` is a non-negative integer, incremented at regular intervals. Whenever `NOW` moves, the following happens:

1. The old now-interval t_0 is aligned with the current value of `COUNT`, grounding it in a PML-measure of its duration.
2. The value of `COUNT` is quantized into a value δ which is the nearest half-order of magnitude [19] to `COUNT`, providing an equivalence class of PML-measures that are not noticeably different.
3. A KL term d, aligned with δ, is found or created, providing a mental entity denoting each class of durations.
4. A belief is introduced into the KL that the duration of t_0 is d, so that the agent can have beliefs that two different states occurred for about the same length of time.
5. `COUNT` is reset to 0, to prepare for measuring the new now-interval.

4.6. Language

Cassie interacts with humans in a fragment of English. Although it is possible to represent the linguistic knowledge of GLAIR agents in the KL, use reasoning to analyze input utterances [32–34,45], and use the acting system to generate utterances [12,13], we do not currently do this. Instead, the parsing and generation grammars, as well as the lexicon, are at the PML (see, e.g. [35,38,49]). There are KL terms for lexemes, and these are aligned with lexemes in the PML lexicon. We most frequently use a KL unary functional term to denote the concept expressed by a given lexeme, but this does not allow for polysemy, so we have occasionally used binary propositions that assert that some concept may be expressed by some lexeme. There may also be KL terms for inflected words, strings of words, and sentences. This allows one to discuss sentences and other language constructs with GLAIR agents.

This facility was used for Cassie to understand the human inputs shown in the example interactions in this paper, and for her to generate her responses (the sentences beginning with "`I`"). We can also use the low level `surface` function to see the NL expression Cassie would use to express the denotation of various SNePS terms (the prompt for this Lispish interaction level is "∗"):

```
∗ (surface b1)
 me
∗ (surface b5)
 Stu
∗ (surface b6)
 you
∗ (surface m21)
 green
∗ (surface m115)
 I found a red robot.
∗ (surface m332)
 I am looking at Robbie.
```

(Remember, Cassie is currently looking at Robbie and talking to Bill.)

5. Related work

Coradeschi and Saffiotti [4,6] present a model of anchoring in an agent with a symbol system, which includes object symbols and unary predicate symbols, and a perceptual system, which includes attributes and percepts. Their grounding relation relates predicate symbols, attributes, and attribute values. Their anchor is a partial function from time to quadruples of: object symbols; percepts; partial functions from attributes to attribute values; and sets of predicate symbols. Their anchor is "reified in an internal data structure" [7, p. 408]. Their symbol system corresponds to our KL, and their perceptual system to a combination of our PML and SAL. While their anchor is a data structure that cuts across their symbol and perceptual systems, our KL and PML communicate by passing PML-descriptions from one to the other, sometimes by socket connections between different computers. Their discussion of "perceptual anchoring of symbols for action" [6] concerns the anchoring of object symbols of objects the actions are performed on. We also discussed the anchoring of action symbols to the PML procedures that carry them out.

Santos and Shanahan [37] discuss anchoring as the "process of assigning abstract symbols to real sensor data" and develop a theory whose "universe of discourse includes sorts for time points, depth, size, peaks, physical bodies and viewpoints. *Time points*, *depth* and *size* are variables that range over positive real numbers (R^+), *peaks* are variables for depth peaks, *physical bodies* are variables for objects of the world, *viewpoints* are points in R^3" [pp. 39–40, italics in the original]. We consider data such as these to belong at the PML, as not being the sort of entities people reason and talk about, and therefore, not the sort of entities cognitive robots should have at the KL. We view anchoring as the aligning of physical-level representations such as these to the KL terms used for reasoning.

Jackendoff [26] explicates a theory in which "the character of a consciously experienced entity is functionally determined by a cognitive structure that contains the following feature types: an indexical feature to which descriptive features can be attached; one or more modalities in which descriptive features are present; the actual descriptive features in the available modalities" [27, p. 313]. His indexical features correspond with our KL term, and his descriptive features correspond with our PML-descriptions. His suggestion that "we think of the descriptive features as being linked to a common indexical feature" [27, pp. 311–312] parallels our suggestion in Section 2 of KL terms as pivots.

6. Summary

We have given an informal, but coherent, unified, overview of our approach to connecting the abstract-level representations to the physical-level representations in GLAIR, an architecture for cognitive robots and intelligent autonomous agents. The abstract-level representations are terms of SNePS logic contained in the knowledge level (KL) of the GLAIR architecture, while the physical-level representations are n-tuples of perceptual features, procedures, and other symbol structures contained at the perceptuo-motor level (PML) of the architecture.

KL terms denoting perceivable entities, perceivable properties, and recognizable categories are aligned with PML-descriptions. Primitive actions are aligned with PML procedures. Deictic and modality registers hold KL terms for individuals and states that the agent is currently aware of, including states of its own body. They are updated by the PML procedures. The `NOW` register is used to give the agent a personal sense of time, including keeping track of current and past states. KL terms denoting times and temporal durations are aligned with PML numeric measures of durations created by the PML pacemaker. Lexemes are represented by KL terms that are aligned with PML lexicon entries used by the parsing and generation grammars, which, like PML procedures, mediate between the agent and the outside world, in this case, humans with which she communicates.

Acknowledgements

Development of GLAIR and the FEVAHR were supported in part by NASA under contracts NAS 9-19004 and NAS 9-19335. SNePS, GLAIR, and Cassie are products of SNeRG, the SNePS Research Group, Department of Computer Science and Engineering, University at Buffalo. Many group members, past and present, have contributed to these efforts. This paper is a revised version of Shapiro and Ismail [43]. We appreciate the valuable comments on previous drafts made by William J. Rapaport and other members of SNeRG, and by Venugopal Govindaraju, Peter D. Scott, and the anonymous reviewers. We also appreciate the efforts and discussions of Silvia Coradeschi, Alessandro Saffiotti, and other organizers and attendees of the 2001 AAAI Fall Symposium on Anchoring Symbols to Sensor Data in Single and Multiple Robot Systems.

References

[1] C. Bandera, P. Scott, Foveal machine vision systems, in: Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, November 1989, pp. 596–599.

[2] C. Bandera, S. Shapiro, H. Hexmoor, Foveal machine vision for robots using agent based gaze control, Final Technical Report No. 613-9160001, Amherst Systems, Inc., Buffalo, NY, September 1994.

[3] R. Carey, G. Bell, The Annotated VRML 2.0 Reference Manual, Addison-Wesley Developers Press, Reading, MA, 1997.

[4] S. Coradeschi, A. Saffiotti, Anchoring symbols to sensor data: Preliminary report, in: Proceedings of the 17th National Conference on Artificial Intelligence (AAAI-2000) and 12th Innovative Applications of Artificial Intelligence Conference (IAAI-2000), AAAI Press/The MIT Press, Menlo Park, CA, 2000, pp. 129–135.

[5] S. Coradeschi, A. Saffiotti, Forward, in: S. Coradeschi, A. Saffiotti (Eds.), Anchoring Symbols to Sensor Data in Single and Multiple Robot Systems: Papers from the 2001 AAAI Fall Symposium, Technical Report No. FS-01-01, AAAI Press, Menlo Park, CA, 2001, p. viii.

[6] S. Coradeschi, A. Saffiotti, Perceptual anchoring of symbols for action, in: Proceedings of the 17th International Joint Conference on Artificial Intelligence (IJCAI-01), Morgan Kaufmann, San Francisco, CA, 2001, pp. 407–412.

[7] F. Du, A. Izatt, C. Bandera, An MIMD computing platform for a hierarchical foveal machine vision system, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'96), IEEE Computer Society Press, Silver Spring, MD, June 1996, pp. 720–725.

[8] J.F. Duchan, G.A. Bruder, L.E. Hewitt (Eds.), Deixis in Narrative: A Cognitive Science Perspective, Lawrence Erlbaum Associates, Inc., Hillsdale, NJ, 1995.

[9] Franz, Inc., Oakland, CA, Allegro CL 6.2 Documentation, 2002. http://www.franz.com/support/documentation/6.2/doc/introduction.htm.

[10] G. Frege, On sense and reference, in: P. Geach, M. Black (Eds.), Translations from the Philosophical Writings of Gottlob Frege, Blackwell Scientific Publications, Oxford, 1970, pp. 56–78 (original work published in 1892).

[11] Garnet Group, Garnet Reference Manual, Version 2.2, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, 1993.

[12] S. Haller, Planning text about plans interactively, International Journal of Expert Systems 9 (1) (1996) 85–112.

[13] S. Haller, An introduction to interactive discourse processing from the perspective of plan recognition and text planning, Artificial Intelligence Review 13 (4) (1999) 259–333.

[14] H. Hexmoor, C. Bandera, Architectural issues for integration of sensing and acting modalities, in: Proceedings of the IEEE International Symposium on Intelligent Control: ISIC/CIRA/ISIS Joint Conference, IEEE Press, New York, 1998, pp. 319–324.

[15] H. Hexmoor, J. Lammens, S.C. Shapiro, Embodiment in GLAIR: a grounded layered architecture with integrated reasoning for autonomous agents, in: D.D. Dankel II, J. Stewman (Eds.), Proceedings of the Sixth Florida AI Research Symposium (FLAIRS'93), The Florida AI Research Society, April 1993, pp. 325–329.

[16] H. Hexmoor, S.C. Shapiro, Integrating skill and knowledge in expert agents, in: P.J. Feltovich, K.M. Ford, R.R. Hoffman (Eds.), Expertise in Context, AAAI Press/MIT Press, Cambridge, MA, 1997, pp. 383–404.

[17] H.H. Hexmoor, Representing and learning routine activities, Ph.D. Dissertation, Technical Report No. 98-04, Department of Computer Science, State University of New York at Buffalo, Buffalo, NY, December 1995.

[18] H.H. Hexmoor, Learning routines, in: M.J. Wooldrige, J.P. Muller, M. Tambe (Eds.), Intelligent Agents. II. Agent Theories, Architectures and Languages, Lecture Notes in Artificial Intelligence, vol. 1037, Springer, Berlin, 1996, pp. 97–110.

[19] J.R. Hobbs, Half orders of magnitude, in: L. Obrst, I. Mani (Eds.), Papers from the Workshop on Semantic Approximation, Granularity, and Vagueness, Proceedings of the Workshop of the Seventh International Conference on Principles of Knowledge Representation and Reasoning, Breckenridge, CO, 2000, pp. 28–38.

[20] H.O. Ismail, Reasoning and acting in time, Ph.D. Dissertation, Technical Report No. 2001-11, University at Buffalo, The State University of New York, Buffalo, NY, August 2001.

[21] H.O. Ismail, S.C. Shapiro, Cascaded acts: Conscious sequential acting for embodied agents, Technical Report No. 99-10, Department of Computer Science and Engineering, University at Buffalo, Buffalo, NY, November 1999.

[22] H.O. Ismail, S.C. Shapiro, Conscious error recovery and interrupt handling, in: H.R. Arabnia (Ed.), Proceedings of the International Conference on Artificial Intelligence (IC-AI'2000), CSREA Press, Las Vegas, NV, 2000, pp. 633–639.

[23] H.O. Ismail, S.C. Shapiro, Two problems with reasoning and acting in time, in: A.G. Cohn, F. Giunchiglia, B. Selman (Eds.), Proceedings of the Seventh International Conference on Principles of Knowledge Representation and Reasoning (KR'2000), Morgan Kaufmann, San Francisco, 2000, pp. 355–365.

[24] H.O. Ismail, S.C. Shapiro, The cognitive clock: A formal investigation of the epistemology of time, Technical Report No. 2001–08, Department of Computer Science and Engineering, University at Buffalo, Buffalo, NY, August 2001.

[25] R. Jackendoff, Semantic Structures, MIT Press, Cambridge, MA, 1990.

[26] R. Jackendoff, Foundations of Language: Brain, Meaning, Grammar, Evolution, Oxford University Press, Oxford, UK, 2002.

[27] D. Kumar, From beliefs and goals to intentions and actions: an amalgamated model of inference and acting, Ph.D. Dissertation, Technical Report No. 94-04, State University of New York at Buffalo, Buffalo, NY, 1994.

[28] D. Kumar, The SNePS BDI architecture, Decision Support Systems 16 (1) (1996) 3–19.

[29] D. Kumar, S.C. Shapiro, Acting in service of inference (and vice versa), in: D.D. Dankel II (Ed.), Proceedings of the Seventh Florida AI Research Symposium (FLAIRS'94), The Florida AI Research Society, May 1994, pp. 207–211.

[30] D. Kumar, S.C. Shapiro, The OK BDI architecture, International Journal on Artificial Intelligence Tools 3 (3) (1994) 349–366.

[31] A.S. Maida, S.C. Shapiro, Intensional concepts in propositional semantic networks, Cognitive Science 6 (4) (1982) 291–330. Reprinted in: R.J. Brachman, H.J. Levesque (Eds.), Readings in Knowledge Representation, Morgan Kaufmann, San Mateo, CA, 1985, pp. 170–189.

[32] J.G. Neal, S.C. Shapiro, Parsing as a form of inference in a multiprocessing environment, in: Proceedings of the Conference on Intelligent Systems and Machines, Oakland University, Rochester, MI, 1985, pp. 19–24.

[33] J.G. Neal, S.C. Shapiro, Knowledge-based parsing, in: L. Bolc (Ed.), Natural Language Parsing Systems, Springer, Berlin, 1987, pp. 49–92.

[34] J.G. Neal, S.C. Shapiro, Knowledge representation for reasoning about language, in: J.C. Boudreaux, B.W. Hamill, R. Jernigan (Eds.), The Role of Language in Problem Solving, vol. 2, Elsevier, Amsterdam, 1987, pp. 27–46.

[35] W.J. Rapaport, S.C. Shapiro, J.M. Wiebe, Quasi-indexicals and knowledge reports, Cognitive Science 21 (1) (1997) 63–107. Reprinted in: F. Orilia, W.J. Rapaport (Eds.), Thought, Language, and Ontology: Essays in Memory of Hector–Neri Castañeda, Kluwer Academic Publishers, Dordrecht, 1998, pp. 235–294.

[36] J.F. Santore, S.C. Shapiro, Identifying perceptually indistinguishable objects: Is that the same one you saw before? in: C. Baral, S. McIlraith (Eds.), Cognitive Robotics (CogRob2002), Papers from the AAAI Workshop, Technical Report No. WS-02-05, AAAI Press, Menlo Park, CA, 2002, pp. 96–102.

[37] P.E. Santos, M.P. Shanahan, From stereoscopic vision to symbolic representation, in: S. Coradeschi, A. Saffiotti (Eds.), Anchoring Symbols to Sensor Data in Single and Multiple Robot Systems: Papers from the 2001 AAAI Fall Symposium, Technical Report No. FS-01-01, AAAI Press, Menlo Park, CA, 2001, pp. 37–43.

[38] S.C. Shapiro, Generalized augmented transition network grammars for generation from semantic networks, The American Journal of Computational Linguistics 8 (1) (1982) 12–25.

[39] S.C. Shapiro, The CASSIE projects: An approach to natural language competence, in: J.P. Martins, E.M. Morgado (Eds.), Proceedings of the Fourth Portugese Conference on Artificial Intelligence (EPIA'89), Lecture Notes in Artificial Intelligence, vol. 390, Springer, Berlin, 1989, pp. 362–380.

[40] S.C. Shapiro, Belief spaces as sets of propositions, Journal of Experimental and Theoretical Artificial Intelligence (JETAI) 5 (2–3) (1993) 225–235.

[41] S.C. Shapiro, Embodied Cassie, in: Cognitive Robotics: Papers from the 1998 AAAI Fall Symposium, Technical Report No. FS-98-02, AAAI Press, Menlo Park, CA, October 1998, pp. 136–143.

[42] S.C. Shapiro, SNePS: A logic for natural language understanding and commonsense reasoning, in: Ł. Iwańska, S.C. Shapiro (Eds.), Natural Language Processing and Knowledge Representation: Language for Knowledge and

Knowledge for Language, AAAI Press/The MIT Press, Menlo Park, CA, 2000, pp. 175–195.

[43] S.C. Shapiro, H.O. Ismail, Symbol-anchoring in Cassie, in: S. Coradeschi, A. Saffiotti (Eds.), Anchoring Symbols to Sensor Data in Single and Multiple Robot Systems: Papers from the 2001 AAAI Fall Symposium, Technical Report No. FS-01-01, AAAI Press, Menlo Park, CA, 2001, pp. 2–8.

[44] S.C. Shapiro, H.O. Ismail, J.F. Santore, Our dinner with Cassie, in: Working Notes for the AAAI 2000 Spring Symposium on Natural Dialogues with Practical Robotic Devices, AAAI Press, Menlo Park, CA, 2000, pp. 57–61.

[45] S.C. Shapiro, J.G. Neal, A knowledge engineering approach to natural language understanding, in: Proceedings of the 20th Annual Meeting of the Association for Computational Linguistics, ACL, Menlo Park, CA, 1982, pp. 136–144.

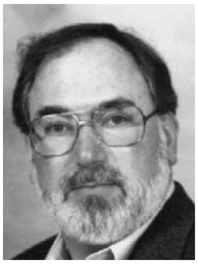
[46] S.C. Shapiro, W.J. Rapaport, SNePS considered as a fully intensional propositional semantic network, in: N. Cercone, G. McCalla (Eds.), The Knowledge Frontier, Springer, New York, 1987, pp. 263–315.

[47] S.C. Shapiro, W.J. Rapaport, Models and minds: Knowledge representation for natural-language competence, in: R. Cummins, J. Pollock (Eds.), Philosophy and AI: Essays at the Interface, MIT Press, Cambridge, MA, 1991, pp. 215–259.

[48] S.C. Shapiro, W.J. Rapaport, The SNePS family, Computers and Mathematics with Applications 23 (2–5) (1992) 243–275. Reprinted in: F. Lehmann (Ed.), Semantic Networks in Artificial Intelligence, Pergamon Press, Oxford, 1992, pp. 243–275.

[49] S.C. Shapiro, W.J. Rapaport, An introduction to a computational reader of narratives, in: J.F. Duchan, G.A. Bruder, L.E. Hewitt (Eds.), Deixis in Narrative: A Cognitive Science Perspective, Lawrence Erlbaum Associates, Inc., Hillsdale, NJ, 1995, pp. 79–105.

[50] S.C. Shapiro, The SNePS Implementation Group, SNePS 2.6 User's Manual, Department of Computer Science and Engineering, University at Buffalo, The State University of New York, Buffalo, NY, 2002.



Stuart C. Shapiro received his S.B. degree in mathematics from the Massachusetts Institute of Technology in 1966, and the M.S. and Ph.D. degrees in computer sciences from the University of Wisconsin–Madison in 1968 and 1971. He is currently Professor of Computer Science and Engineering at the University at Buffalo, The State University of New York. where he was Department Chair from 1984 to 1990, and from 1996 to 1999.

Prof. Shapiro is editor-in-chief of *The Encyclopedia of Artificial Intelligence* (Wiley, First Edition, 1987; Second Edition, 1992), co-editor of *Natural Language Processing and Knowledge Representation: Language for Knowledge and Knowledge for Language* (AAAI Press/The MIT Press, 2000), author of *Techniques of Artificial Intelligence* (Van Nostrand, 1979), *LISP: An Interactive Approach* (Computer Science Press, 1986), *Common Lisp: An Interactive Approach* (Computer Science Press, 1992), and author or co-author of over 190 technical articles and reports. He has served on the editorial board of the *American Journal of Computational Linguistics*, and as Guest Editor of special issues of *Minds and Machines*, and of the *International Journal of Expert Systems*. Prof. Shapiro is a member of the ACM, the ACL, the Cognitive Science Society, and Sigma Xi, a senior member of the IEEE, and a fellow of the American Association for Artificial Intelligence. He has served as Chair of ACM/SIGART and as President of Knowledge Representation and Reasoning, Inc.

Haythem O. Ismail received his B.Sc. degree in communications and electronics in 1992, Diploma in mathematics in 1993, and M.Sc. in computer science (Natural Language Processing) in 1996, all from Cairo University, and his Ph.D. in computer science and engineering from the University at Buffalo, The State University of New York, in 2001. He is currently Assistant Professor of Engineering Mathematics and Physics at Cairo University.