The Inner Speech of a Cognitive Architecture

Antonio Chella
Department of Engineering
University of Palermo, Italy
antonio.chella@unipa.it
Joint work

Antonio Chella
Arianna Pipitone
Valeria Seidita
Antonella D’Amico
Alessandro Geraci
Researchers at the Università degli Studi di Palermo have designed a robot that ‘thinks out loud’ so that users can hear its thought process and better understand the robot’s motivations and decisions.

"If you were able to hear what the robots are thinking, then the robot might be more trustworthy. The robots will be easier to understand for laypeople, and you don’t need to be a technician or engineer. In a sense, we can communicate and collaborate with the robot better." - Professor Antonio Chella

https://lnkd.in/dFr3Jqx

The research is supported by the AFRL/AFOSR Life Sciences and Human Performance program, managed by Dr. Nandini Iyer.

Air Force Research Laboratory #AFOSR_QO #AFOSRBoldResearch #HighRiskHighReward #HumanMachineTeaming #EOARD
Why inner speech for robot?

Inner speech:

- is linked to self-consciousness [Morin, 2005]:
  - comes to play important roles in self-regulation and planning
  - focusing attention and self-attention

it enables high-level cognition (inferences, planning...)

it makes the robot operations more transparent

it allows to internalize human’s explanation (task learning)
Models of inner speech

• Morin
  inner speech is linked to self-consciousness

• Baddeley
  inner speech as a component of the phonological loop of the working memory

• Vygotskij
  inner speech is the result of a developmental internalization process

• Martínez-Manrique and Vicente
  vehicle or activity forms of inner speech

...and others
Baddley

Baddley and the Common Model of Cognition

Integrating Baddley’s model of inner speech and the Common Model of Cognition

A cognitive architecture for inner speech
Implementation of the architecture
ACT-R inner speech model

Modeling perception:

(chunk-type perception channel objects self-reflect)

Modeling inner dialogue turn:

(chunk-type link-turn turn1 turn2)

Turns for [Morin, 2005]:
- Social milieu («What do I see in a video?», «What do the other think about me?»)
- Self-direct question («What am I doing?»)
- Self-control («I move my hand»)
- Self-focus («My arm hurts me»)
Operation of the architecture

Perceptual motor layer

«apple»
«green»
The covert articulator

"Apple is a fruit"
Searching for an orange...

Perceptual motor layer

«apple»
«green»
«is»
«fruit»
«orange»

«Orange is a fruit»
The inner speech implementation in Pepper
Inner speech for transparency

- The robot talks about its inner reasoning process, making the interaction more transparent.
  - Self-direct question («What am I doing?» «Did I pick the object before?»)
  - Self-control («I move my hand»)
  - Self-focus («My arm is overheated»)
How inner speech improves interaction

• By inner speech, facts and information are retrieved from the knowledge base.

• Inner speech improves the interaction with human, enabling the partner to focus on problems otherwise not considered.

Attention! The knife cuts!
Inner speech for conflict resolution

• What to do when human requires to break a rule? It is a dilemma to solve for robot.

• To place object in a location which is different from those in the etiquette schema, falls in conflict situation.

• inner speech and conflicts (Tappan, 1962)
Mirror Self-Recognition Test

- Mirror self-recognition (MSR) test
- ...by using inner speech [Morin, 1989]

Passing test

- Social milieu («What do I see in a video?», «What do the other think about me?»)
- Self-direct question («What am I doing?»)
- Self-control («I move my hand»)
- Self-focus («My arm hurts me»)

Morin, Deblois
Thank you for your attention!
Developing Self-Awareness in Robots via Inner Speech

Antonio Chella1,2*, Arianna Pipitone1, Alain Morin3 and Famira Racy4

1 RoboticsLab, Dipartimento di Ingegneria, Università degli Studi di Palermo, Palermo, Italy
2 Istituto di Calcolo e Reti ad Alte Prestazioni (ICAR), Consiglio Nazionale delle Ricerche, Palermo, Italy
3 Department of Psychology, Mount Royal University, Calgary, AB, Canada
4 Researcher, Mount Royal University, Calgary, AB, Canada

The experience of inner speech is a common one. Such a dialogue accompanies the introspection of mental life and fulfills essential roles in human behavior, such as self-restructuring, self-regulation, and re-focusing on attentional resources. Although the underpinning of inner speech is mostly investigated in psychological and philosophical fields, the research in robotics generally does not address such a form of self-aware behavior. Existing models of inner speech inspire computational tools to provide a robot with this form of self-awareness. Here, the widespread psychological models of inner speech are reviewed, and a cognitive architecture for a robot implementing such a capability is outlined in a simplified setup.

Keywords: inner speech, self-awareness, robot, human-robot interaction, cognitive cycle

INTRODUCTION

The idea of implementing self-awareness in robots has been popular in science-fiction literature and movies for a long time. This quest is also becoming increasingly prevalent in scientific research, with articles, special topics, books, workshops, and conferences dedicated to it.

It is widely assumed that there are two dimensions of awareness (see Dehaene et al., 2017), and namely, awareness as experience and awareness as self-monitoring, i.e., self-awareness. In essence, awareness as experience occurs when an agent perceives the environment and experiences it from within in the form of images, sensations, thoughts, and so on (see Block, et al., 2010); as such, awareness (or consciousness) exists when an organism can focus attention outward toward the environment (Dennett and Rausch, 1972). Instead, self-awareness takes place when the agent focuses attention inward and apprehends the self in its diverse manifestations, like emotions, thoughts, attitudes, sensations, motives, physical attributes, which frequently involves a verbal narration of inner experiences (Morin, 2011).

Models of awareness and self-awareness are being proposed, each with idiosyncratic views of what the aforementioned concepts constitute, as well as different suggestions on how to implement them in artificial agents (see among others, Tononi and Edelman, 1998; Gray et al., 2007; Seth, 2010; Oizumi et al., 2014; Tononi et al., 2016; Juel et al., 2019). For reviews, see Reggia (2013) and Chella et al. (2019).

The proposed approach focuses on implementing a form of robot self-awareness by developing inner speech in the robot. Inner speech is known to importantly participate in the development and maintenance of human self-awareness (Morin, 2018); thus, self-talk in robots is an essential behavioral capability of robot self-awareness. More in detail, the paper discusses a computational model of inner speech. The proposed model is based on the cognitive architecture described by Laird et al. (2017). Therefore, the approach

Keywords: inner speech, self-awareness, robot, human-robot interaction, cognitive cycle
In the past years, robots and automation development and implementation have increased exponentially in every context, leading to growing interactions with humans (Merritt and Ilgen, 2004). Robots are now used in different contexts, such as military, security, medical, domestic, and entertainment (Li et al., 2010). Robots, compared with other types of automation (e.g., machines, calculators), are designed to be self-governed to some extent to respond to situations that are not preplanned (Leev et al., 2018). Therefore, the greater the complexity of robots, the higher the importance to focus on factors that influence human–automation interaction (HAI) as their collaboration increases over time (Lee and See, 2004; Schaefer et al., 2016). In this paper, we aim to discuss the possible role of inner speech in influencing trust in human–automation interaction. Inner speech is an everyday covert inner monologue or dialog with oneself, which is essential for human psychological life and functioning as it is linked to self-regulation and self-awareness. Recently, in the field of machine consciousness, computational models using different forms of robot speech have been developed that make it possible to implement inner speech in robots. As is discussed, robot inner speech could be a new feature affecting human trust by increasing robot transparency and anthropomorphism.

Keywords: inner speech, trust, anthropomorphism, human-automation interaction, human-robot interaction, robot, automation

OPEN ACCESS

HYPOTHESIS AND THEORY

INTRODUCTION

In the past years, robots and automation development and implementation have increased exponentially in every context, leading to growing interactions with humans (Merritt and Ilgen, 2004). Robots are now used in different contexts, such as military, security, medical, domestic, and entertainment (Li et al., 2010). Robots, compared with other types of automation (e.g., machines, computers), are designed to be self-governed to some extent to respond to situations that are not preplanned (Leev et al., 2018). Therefore, the greater the complexity of robots, the higher the importance to focus on factors that influence human–automation interaction (HAI) as their collaboration increases over time (Lee and See, 2004; Schaefer et al., 2016). In this paper, we aim to start the exploration of the role of inner speech in HAI and, in particular, on its role in improving human trust toward automation. For this purpose, we first focus on the concept of inner speech in psychological literature, also examining the first results of its implementation in automation. Then, we discuss the possible role of inner speech as one of the anthropomorphic automation features that may affect human trust in HAI.

INNER SPEECH

Inner speech is an everyday covert inner monologue or dialog with oneself, which is essential for human psychological life and functioning because it is linked to reasoning, self-regulation, and self-awareness (Merrit, 2012).

A cognitive architecture for inner speech

Antonio Chella1,⇑, Arianna Pipitone2

1 Dipartimento di Informatica – Università degli Studi di Palermo, Italy
2 C.N.R., Istituto di Calcolo e Reti ad Alte Prestazioni, Palermo, Italy

Abstract

A cognitive architecture for inner speech is presented. It is based on the Standard Model of Mind, integrated with modules for self-reflecting. Briefly, the working memory of the proposed architecture includes the phonological loop as a component which manages the exchange information between the phonological store and the articulatory control system. The inner dialogue is modeled as a loop where the phonological store hears the inner voice produced by the hidden articulator process. A central executive module drives the whole system, and contributes to the generation of conscious thoughts by retrieving information from long-term memory. The surface form of thoughts thus emerges by the phonological loop. Once a conscious thought is elicited by inner speech, the perception of new content takes place and then repeating the cognitive loop. A preliminary formalization of some of the described processes by event calculus, and early results of their implementation on the humanoid robot Pepper by SoftBank Robotics are discussed.

© 2019 Elsevier B.V. All rights reserved.

Keywords: inner speech, cognitive architecture