

A Multi-Level Cognitive Architecture for Self-Referencing, Self-Awareness and Self-Interpretation

Jan Treur¹ and Gerrit Glas^{2,3}

¹Social AI Group, Vrije Universiteit Amsterdam

²Epistemology and Metaphysics Group, Vrije Universiteit Amsterdam

³Anatomy and Neurosciences Department, Amsterdam University Medical Center



Jan Treur

Gerrit Glas

ABSTRACT

In this paper, a multilevel cognitive architecture is introduced that can be used to model mental processes in **clients of psychotherapeutic sessions**. The architecture does not only cover base level mental processes but also mental processes involving self-referencing, self-awareness and self-interpretation. To this end, the cognitive architecture was designed according to **four levels**, where (part of) the structure of each level is represented by an **explicit self-model** of it at the next-higher level of the architecture. At that next-higher level, states reify part of the structure of the level below; these states have a referencing relation to it. In this way the overall architecture includes its own overall self-model. The cognitive architecture was evaluated for a case study of a realistic type of **therapeutic session from clinical practice**.

INTRODUCTION

During therapy, often the client's knowledge and awareness of her or himself is enhanced and can become useful to his or her benefit:

- Getting familiar with one's **own personal characteristics**
- Being **aware** of important aspects of oneself
- Being able to **interpret** one's behaviour in relation to other aspects of oneself
- Based on such insights being able to **manage oneself** more effectively

Three important notions for a therapeutic context:

Self-referentiality

- behaviors by themselves 'say' (signify, indicate) something about the person having them

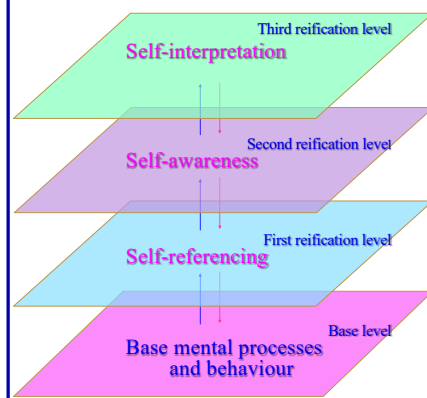
Self-awareness

- awareness of an aspect of oneself or of oneself as a whole

Self-interpretation

- the way people understand (perceive, value) themselves

MULTILEVEL COGNITIVE ARCHITECTURE: OVERALL VIEW



NETWORK STATES VERSUS NETWORK STRUCTURE

Network nodes or states

- Informational view on the network
- For example, sensor, sensory representation preparation, emotion states
- Have activation levels that change over time within a given network structure

Network structure characteristics

- Structural view on the network
- For example, connection weights and excitability thresholds
- Have values and determine the activation levels of the states over time
- For an adaptive network they can change

NETWORK REIFICATION: INCLUDING A SELF-MODEL FOR A NETWORK

Reification states are added that **represent or refer** to the network's structure characteristics; for example:

- **Self-model for its connections:** states $W_{X,Y}$ that refer to the **weight** of the connection from state X to state Y
- **Self-model for its excitability:** states T_Y referring to the **excitability threshold** of state Y

USES OF A NETWORK'S SELF-MODEL

Monitoring and analysis of network evolution

by a dynamically updated self-model over time:

- adding connections from base states to self-model states
- adding connections from self-model to other analysis states

Network adaptation

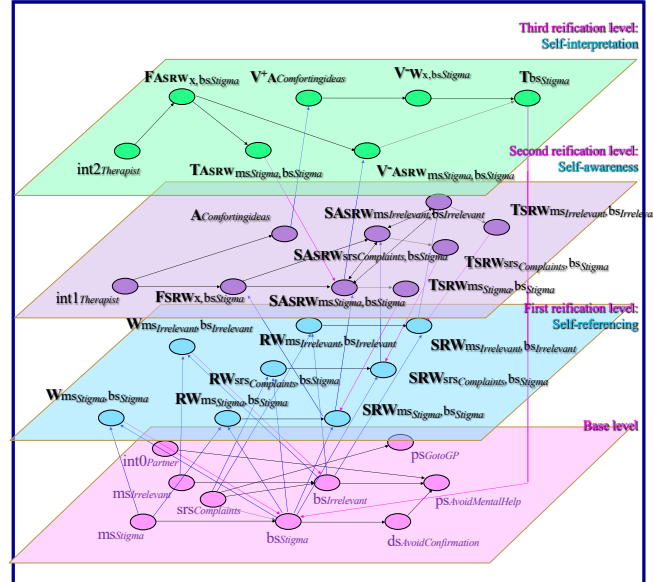
by a dynamic self-model determining the base network characteristics used:

- adding connections from self-model states to lower level states

For example

- A network with **adaptive connection weights**: a specific downward connection from $W_{X,Y}$ to Y for using the value of $W_{X,Y}$ as actual connection weight for the connection from X to Y
- A network with **adaptive excitability thresholds**: a specific downward connection from T_Y to Y for using the value of T_Y as actual excitability threshold for Y

DETAILED ARCHITECTURE FOR SELF-REFERENCING, SELF-AWARENESS AND SELF-INTERPRETATION



EXAMPLE SIMULATIONS

$srs_{Complaints} \rightarrow ps_{GotoGP}$

complaints make preparing to go to GP

$srs_{Complaints} ms_{Stigma} \rightarrow bs_{Stigma}$

complaints and memory about stigma lead to the belief that confirming a mental problem provides a stigma

$bs_{Stigma} \rightarrow ds_{AvoidConfirmation}$

belief bs_{Stigma} leads to desire $ds_{AvoidConfirmation}$ to avoid confirming mental problems

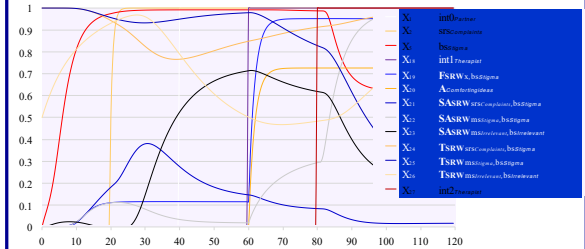
$srs_{Complaints} \rightarrow bs_{Irrelevant}$

complaints lead to the belief $bs_{Irrelevant}$ that after a while mental problems go away by themselves

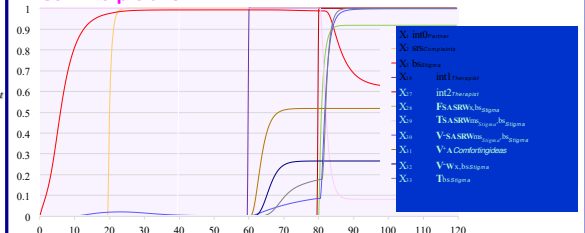
$bs_{Irrelevant} ds_{AvoidConfirmation} \rightarrow ps_{AvoidMentalHelp}$

the desire to avoid confirming mental problems, and belief $bs_{Irrelevant}$ make preparing to avoid mental help $ps_{AvoidMentalHelp}$

Self-Awareness



Self-Interpretation



REFERENCES

- Glas, G. (2017). Dimensions of the self in emotion and psychopathology. *Philosophy, Psychiatry, & Psychology*, 24(2), 143–155.
- Glas, G. (2019). *Person-centered care in psychiatry*. Selfrelational, contextual, and normative perspectives. Abingdon/London: Routledge.
- Treur, J. (2020). *Network-Oriented Modeling for Adaptive Networks: Designing Higher-Order Adaptive Biological Mental and Social Network Models*. Cham, Switzerland: Springer Nature Publishers