Grounding a Descriptive Language in Cognitive Agents Using Consensus Methods

Agnieszka Pieczynska-Kuchtiak

Institute of Information Science and Engineering, Wroclaw University of Technology,
Wybrzeze Wyspianskiego 27, 50-370 Wroclaw, Poland
agnieszka.pieczynska-kuchtiak@pwr.wroc.pl

Abstract. In biological systems the process of symbol grounding is realised at a neural level. But in case of artificial subjects it must be done by certain organisations of private data structures and data collections, and by internal data processing and recognition algorithms. In this work a possible way of implementation of the well - known problem of symbol grounding is presented. The idea of an algorithm of grounding formulas in a private agent's data structure is given. It is assumed that from the agent's point of view stored perceptions are the original source of any meaning accessible for cognitive processes of symbol grounding. This algorithm is used when the state of a particular object from external world is not known. As a results the logic formulas with modal operators and descriptive connectives are generated.

1 Introduction

The problem of symbol grounding has been widely discussed in artificial intelligence, robotics, cognitive science and linguistics [2], but it still fails to practical solution applied by artificial systems [11]. Complete study of the grounding problem is very difficult task and practical solutions of this problem seem to be a big challenge. In biological systems the process of grounding is realised at a neural level. But in case of artificial subjects it must be done by certain organisations of private data structures and data collections, and by internal data processing and recognition algorithms [1].

The symbol grounding is strictly related to the problem of generating statements about the external world. According to the cognitive linguistics and the phenomenology of knowledge each external formula generated by the cognitive agent needs to be grounded in relevant knowledge structures. These relevant structures are called the grounding experience of the related formula [9]. In this case a descriptive language consisted of logic formulas with modal operator of belief or possibility one is taken into account. External formulas are the reflection of agent grounded opinion about the current state of external world. They describe the states of object from external world. Such an approach to define the epistemic concept of belief is an alternative way to definition of formulas' semantics, because each formula is treated by the cognitive agent as true, if and only if this formula is satisfied by the overall state of agent-encapsulated knowledge [4], [5]. In this work a possible way of

grounding descriptive language, represented by the algorithm for grounding formulas in the agent's knowledge structures, is proposed. This algorithm uses consensus in the agent's knowledge structures, is proposed. This algorithm uses consensus methods [3]. The process of grounding is carried out when the state of particular object from external world is not known for an agent and the empirical verification is not possible.

2 Assumptions

It is assumed that the agent is independent entity equipped with sensors and situated in some world. This world is consisted of the atomic objects $O=\{o_1, o_2, \dots, o_Z\}$. Each object $o_z \in O$ at the particular time point $t_n \in T$, $T=\{t_0, t_1, t_2, \dots\}$ possesses or doesn't posses the property $P_i \in P$, $P=\{P_1, P_2, \dots, P_K\}$. The states of the objects from O are a target of agent's cognitive processes.

2.1 Agent's Internal Organization

The states of the objects from external world at the particular time point t_k are remembered by the agent in the structure called base profile $BP(t_n)$.

Definition. Let the base profile be given as the following system of mathematical relations:

$$BP(t_n) = \langle O, P^+_1(t_n), P^-_1(t_n), P^\pm_1(t_n), \dots, P^+_K(t_n), P^-_K(t_n), P^\pm_K(t_n) \rangle, \quad (1)$$

It is assumed that:

1. $O=\{o_1, o_2, \dots, o_Z\}$, where each o denotes a unique cognitive representation of a particular object of the world W .
2. For each $i=1, 2, \dots, K$, relations $P^+_i(t_n) \subseteq O$, $P^-_i(t_n) \subseteq O$, $P^\pm_i(t_n) \subseteq O$ hold. For each $o_z \in O$ the relation $o_z \in P^+_i(t_n)$ holds if and only if the agent has perceived that this object o_z possesses atomic property P_i . For each $o_z \in O$ the relation $o_z \in P^-_i(t_n)$ holds if and only if the agent has perceived that this object o_z does not posses the atomic property P_i . For each $o_z \in O$ the relation $o_z \in P^\pm_i(t_n)$ holds if and only if the agent doesn't known the state of an object o_z in relation to the property P_i .

Obviously, for each $i=1, 2, \dots, K$, the conditions $P^+_i(t_n) \cap P^-_i(t_n) = \emptyset$, $P^+_i(t_n) \cap P^\pm_i(t_n) = \emptyset$, $P^-_i(t_n) \cap P^\pm_i(t_n) = \emptyset$ hold.

All base profiles are ordered and stored in agent's private temporal database $KS(t_c)$.

Definition. The overall state of stored perceptions is given as a temporal data base:

$$KS(t_c) = \{ BP(t_n): t_n \in T \text{ and } t_n \leq t_c \}. \quad (2)$$

2.2 Agent's Language of Communication

The agent is provided with the descriptive language of communication. This language makes it possible for an agent to generate the external messages (logic formulas).

Definition. The agent's language is defined as a pair L:

$$L = \{A, F\}, \quad (3)$$

where A denotes the alphabet consisted of terms $O = \{o_1, o_2, \dots, o_z\}$, descriptive connectives: negation ($\neg$), conjunction ($\wedge$), exclusive alternative ($\vee$) and alternative ($\vee$); modal operators of belief (*Bel*) and possibility (*Pos*). F denotes the set of logic formulas:

$$F = \Phi \cup \Pi, \quad (4)$$

where Φ is a set of formulas with belief (*Bel*) operator:

$$\Phi = \{ \text{Bel}(P_i(o_z) \wedge P_j(o_z)), \text{Bel}(P_i(o_z) \wedge \neg P_j(o_z)), \text{Bel}(\neg P_i(o_z) \wedge P_j(o_z)), \text{Bel}(\neg P_i(o_z) \wedge \neg P_j(o_z)), \text{Bel}(P_i(o_z) \underline{\vee} P_j(o_z)), \text{Bel}(P_i(o_z) \vee P_j(o_z)) \} \quad (5)$$

For example if the agent at the time point t_c generates the message $\varphi_c = \text{Bel}(P_i(o_z) \wedge P_j(o_z))$ it is equivalent to the spoken language: *I belief that the object o_z at the time point t_c possesses the property P_i and property P_j .*

Π is a set of logic formulas with possibility operators:

$$\Pi = \{ \text{Pos}(P_i(o_z) \wedge P_j(o_z)), \text{Pos}(P_i(o_z) \wedge \neg P_j(o_z)), \text{Pos}(\neg P_i(o_z) \wedge P_j(o_z)), \text{Pos}(\neg P_i(o_z) \wedge \neg P_j(o_z)) \} \quad (6)$$

For example if the agent at the time point t_c generates the message $\varphi_c = \text{Pos}(P_i(o_z) \wedge P_j(o_z))$ it is equivalent to the spoken language: *It is possible that object o_z at the time point t_c possesses both the property P_i and P_j .* Each formula is interpreted from the agent's point of view as external, logical representation of beliefs on the current state of the object $o_z \in O$ [4], [5]. It is assumed that at the same time t_c it is possible to generate the only one message with belief operator and few with the possibility one.

3 The Idea of an Algorithm for Grounding Formulas

In this section the idea of the algorithm for grounding formulas will be given. If the state of particular object $o_z \in O$ is not known, then the grounding algorithm is used.

This algorithm is consisted of six steps [6], [7], [8]:

Step 1. Elimination of perceptions

All the base profiles $BP(t_n)$, in which the state of an object o_z in relation to the at least one property P_i or P_j is not known, are rejected. In this way the set $KS'(t_c)$ is created.

Definition. The set $SW'(t_c)$ of base profile is defined as:

$$SW'(t_c) = \{ BP(t_k) \mid ((o_z \in P_i^+(t_k) \text{ and } o_z \in P_j^+(t_k)) \text{ or } (o_z \in P_i^+(t_k) \text{ and } o_z \in P_j^-(t_k)) \text{ or } (o_z \in P_i^-(t_k) \text{ and } o_z \in P_j^+(t_k)) \text{ or } (o_z \in P_i^-(t_k) \text{ and } o_z \in P_j^-(t_k))) \text{ and } k \leq c, i, j \in K, i \neq j \}. \quad (7)$$

According to (7) $SW'(t_c)$ is the set of all base profiles in which the state of an object o_z was known for an agent in relation to the both properties P_i and P_j .

Step 2. Classification of perceptions

All the base profiles from $KS'(t_c)$ are grouped into four classes $C^m(t_c)$, $m=1,2,3,4$ of observations. The criterion of the membership of particular base profile $BP(t_n)$ to the proper class $C^m(t_c)$ is the state of the object o_z in relation to the properties P_i and P_j .

In particular, the following classes of base profiles are taken into account:

a) $C^1(t_c) = \{BP(t_n) :$

$t_n \leq t_c$ and $BP(t_n) \in KS'(t_c)$ and both $o_z \in P_i^+(t_n)$ and $o_z \in P_j^+(t_n)$ hold for $BP(t_n)\}$

Class $C^1(t_c)$ of data consists of all base profiles $BP(t_n)$ stored up to the time point t_c , in which the object o_z has been found as having both properties P_i and P_j .

b) $C^2(t_c) = \{BP(t_n) :$

$t_n \leq t_c$ and $BP(t_n) \in KS'(t_c)$ and both $o_z \in P_i^+(t_n)$ and $o_z \in P_j^-(t_n)$ hold for $BP(t_n)\}$

c) $C^3(t_c) = \{BP(t_n) :$

$t_n \leq t_c$ and $BP(t_n) \in KS'(t_c)$ and both $o_z \in P_i^-(t_n)$ and $o_z \in P_j^+(t_n)$ hold for $BP(t_n)\}$

d) $C^4(t_c) = \{BP(t_n) :$

$t_n \leq t_c$ and $BP(t_n) \in KS'(t_c)$ and both $o_z \in P_i^-(t_n)$ and $o_z \in P_j^-(t_n)$ hold for $BP(t_n)\}$

Interpretations for $C^2(t_c)$, $C^3(t_c)$ and $C^4(t_c)$ are similar to $C^1(t_c)$.

Each class $C^m(t_c)$ is semantically correlated with one logic formula without modal operator in the following way:

$$\begin{aligned}
 P_i(o_z) \wedge P_j(o_z) & \text{ is related to the content of } C^1(t_c) \\
 P_i(o_z) \wedge \neg P_j(o_z) & \text{ is related to the content of } C^2(t_c) \\
 P_i(o_z) \wedge \neg P_j(o_z) & \text{ is related to the content of } C^2(t_c) \\
 \neg P_i(o_z) \wedge P_j(o_z) & \text{ is related to the content of } C^3(t_c) \\
 \neg P_i(o_z) \wedge \neg P_j(o_z) & \text{ is related to the content of } C^4(t_c) \\
 P_i(o_z) \supseteq P_j(o_z) & \text{ is related to the content of } C^5(t_c) = C^2(t_c) \cup C^3(t_c) \\
 P_i(o_z) \vee P_j(o_z) & \text{ is related to the content of } C^6(t_c) = C^1(t_c) \cup C^2(t_c) \cup C^3(t_c)
 \end{aligned} \tag{8}$$

Step 3. Consensus profiles computing

For each class $C^m(t_c)$, $m=1,2,3,4$ an agreement of the knowledge called consensus profile $CBP^m(t_c)$, $m=1,2,3,4$ is computed. In order to determine consensus profiles some postulates need to be applied. See Section 4.

Step 4. Distance between each consensus profile and current base profile computing

The agent computes the distance d_m , $m=1,2,3,4$ between each consensus profile $CBP^m(t_c)$ and current base profile $BP(t_c)$. This distance reflects numerical similarity between each consensus profile and current base profile. One of the possible ways for determining such distance is computing of the total cost of transformation one object into another. In this case transformation $CBP^m(t_c)$ into $BP(t_c)$ is based on the objects movements between the sets $cP_{im}^+(t_c)$, $cP_{im}^-(t_c)$ and $cP_{im}^\pm(t_c)$ in order to obtain the sets $P_i^+(t_c)$, $P_i^-(t_c)$ and $P_i^\pm(t_c)$.

Step 5. Choice function value computing

The agent computes the values of the choice function $V_w(X)$, $w=1,2,\dots,6$. Each of these value is derived from a subset of $\{d_1, d_2, d_3, d_4\}$:

$$\begin{aligned}
V_1(X) & \text{ is derived for } P_i(o_z) \wedge P_j(o_z) \text{ from the set of distances } X=\{d_1\} \\
V_2(X) & \text{ is derived for } P_i(o_z) \wedge \neg P_j(o_z) \text{ from the set of distances } X=\{d_2\} \\
V_3(X) & \text{ is derived for } \neg P_i(o_z) \wedge P_j(o_z) \text{ from the set of distances } X=\{d_3\} \\
V_4(X) & \text{ is derived for } \neg P_i(o_z) \wedge \neg P_j(o_z) \text{ from the set of distances } X=\{d_4\} \\
V_5(X) & \text{ is derived for } P_i(o_z) \vee P_j(o_z) \text{ from the set of distances } X=\{d_2, d_3\} \\
V_6(X) & \text{ is derived for } P_i(o_z) \vee P_j(o_z) \text{ from the set of distances } X=\{d_1, d_2, d_3\}
\end{aligned} \tag{9}$$

The aim of this step is to create the set of logic formulas without modal operators that are considered as candidates for external messages. Smaller the average value of distances of the set X and the variance of distance values of the set X bigger choice function value.

Step 6. Decision procedure

The formula with belief operator is chosen as external, for which the correlated choice function value is maximal. If there are more than one maximal choice function, value the cardinality of the related class of perceptions $C^m(t_c)$ are taken into account. If these cardinalities are mutually closed then all formulas with possibility operator are chosen as external.

4 Consensus Profiles Computing

In the algorithm for grounding formulas some conceptions from the theory of consensus have been applied [3], [10]. In the third step of this algorithm for each class of perceptions the consensus profile $CBP^m(t_c)$, $m=1,2,3,4$ is computed. Consensus profile is an agreement of the knowledge stored in each class $C^m(t_c)$. On this stage of an algorithm the knowledge that is represented by the set of base profiles into single profile is transformed.

Definition. Let the consensus profile be given as:

$$CBP^m(t_n) = \langle O, cP_{1m}^+(t_n), cP_{1m}^-(t_n), cP_{1m}^\pm(t_n), \dots, cP_{Km}^+(t_n), cP_{Km}^-(t_n), cP_{Km}^\pm(t_n) \rangle, \tag{10}$$

It is assumed that for $y \in \{1, 2, \dots, 7\}$:

1. $o_z \in cP_{im}^+(t_c)$ if and only if the relation is consistent with at least one postulate PS_y
2. $o_z \in cP_{im}^-(t_c)$ if and only if the relation is consistent with at least one postulate PS_y
3. $o_z \in cP_{im}^\pm(t_c)$ if and only if the relation is consistent with at least one postulate PS_y
4. for each $i=1,2,\dots,K$, the conditions $cP_{im}^+(t_n) \cap cP_{im}^-(t_n) = \emptyset$, $cP_{im}^+(t_n) \cap cP_{im}^\pm(t_n) = \emptyset$, $cP_{im}^-(t_n) \cap cP_{im}^\pm(t_n) = \emptyset$ hold.

Each postulate PS_y is divided into two parts: in the first parts the set of premises is specified and in the second – the conclusions. The premises represent some conditions that should be fulfilled in order to execute the action from conclusions.

Definition. (postulate's fulfillment)

The postulate PS_y , $y \in \{1, 2, \dots, 7\}$ is fulfilled by the structure S if and only if at least one of the following condition holds:

1. All the premises $\delta = \{\delta_1, \dots, \delta_x\}$ and all the conclusions $\phi = \{\phi_1, \dots, \phi_z\}$ of the postulate PS_y : $\delta = \{\delta_1, \dots, \delta_x\}$ are true:

$$S \models PS_y \text{ if and only if } (\neg \exists \delta_i \in \delta \mid \delta_i =^{\text{bool}} \text{false}) \text{ and } (\neg \exists \phi_i \in \phi \mid \phi_i =^{\text{bool}} \text{false}), \quad (11)$$

2. At least one premise of the postulate PS_y , $y = \{1, 2, \dots, 7\}$ is false.

$$S \models PS_y \text{ if and only if } (\exists \delta_i \in \delta \mid \delta_i =^{\text{bool}} \text{false})$$

where:

$i, x, z \in \mathbb{N}$

$S \in U_S = U_C \times U_{CBP}$

U_C is universe of the all class of perceptions $C^m(t_c)$

U_{CBP} is universe of all consensus profiles

δ - set of all premises of the postulate PS_y

ϕ - set of all conclusions of the postulate PS_y

4.1 Postulates for Consensus

In order to determine consensus profile $CBP^m(t_c)$ some rational requirements (postulates) need to be applied. Following postulates are taken into account:

Postulate PS_1 (unanimity of knowledge)

If object o_z possesses (doesn't possess) the property P_i in each base profile $BP(t_k) \in C^m(t_c)$, $k \leq c$, then in a consensus profile $CBP(t_c)$ possesses (doesn't possess) the property P_i .

Postulate PS_2 (superior of knowledge)

If object o_z possesses (doesn't possess) the property P_i in at least one base profile $BP(t_k) \in C^m(t_c)$, $k \leq c$ and there doesn't exist base profile $BP(t_n) \in C^m(t_c)$, $k \leq c$, $k \neq n$, in which the object o_z doesn't possess (possesses) the property P_i , then in a consensus profile $CBP(t_c)$ the object o_z possesses (doesn't possess) the property P_i .

Postulate PS_3 (the rule of majority)

If the number of base profiles $BP(t_k) \in C^m(t_c)$, $k \leq c$ in which the object o_z possesses (doesn't possess) the property P_i is bigger then the number of base profiles $BP(t_n) \in C^m(t_c)$, $k \leq c$, $k \neq n$, in which the objects o_z doesn't possess (possesses) the property P_i , then in a consensus profile $CBP(t_c)$ the object o_z possesses (doesn't possess) the property P_i .

Postulate PS_4 (the rule of majority of knowledge)

If the number of base profiles $BP(t_k) \in C^m(t_c)$, $k \leq c$ in which the object o_z possesses (doesn't possess) the property P_i is bigger then the number of base profiles $BP(t_n) \in C^m(t_c)$, $k \leq c$, $k \neq n$, in which the objects o_z doesn't possess (possesses) the

property P_i the state of an object o_z is unknown, then in a consensus profile $CBP(t_c)$ the object o_z possesses (doesn't possess) the property P_i .

Postulate PS₅ (restrictive rule of majority of knowledge)

If the number of base profiles $BP(t_k) \in C^m(t_c)$, $k \leq c$ in which the object o_z possesses (doesn't possess) the property P_i is bigger than the sum of the number of base profiles $BP(t_n) \in C^m(t_c)$, $k \leq c$, $k \neq n$, in which the objects o_z doesn't possess (possesses) the property P_i and the base profiles $BP(t_n) \in C^m(t_c)$, $k \leq c$, $k \neq n$, in which the state of an objects o_z is unknown then in a consensus profile $CBP(t_c)$ the object o_z possesses (doesn't possess) the property P_i .

4.2 Algorithm for Determining Consensus Profile

Taking the postulates PS_1 - PS_5 into account the algorithm for the consensus profile computing is given. In consequence the structure S is obtained that fulfills the postulates PS_1 - PS_5 .

Algorithm (consensus profile computing)

Input: The set $C^m(t_c) \subset SW(t_c)$, $m \in \{1, 2, 3, 4\}$

Output: The structure $S = \{C^m(t_c), CBP^m(t_c)\}$, $m \in \{1, 2, 3, 4\}$, $S \in U_s$, U_s is universe of all structures S .

Start.

Let $cP_{im}^+(t_c) = cP_{im}^-(t_c) = cP_{im}^\pm(t_c) = \emptyset$, $P_i^+(t_n) \in BP(t_n)$, $P_i^-(t_n) \in BP(t_n)$, $P_i^\pm(t_n) \in BP(t_n)$, $n \leq c$.

For each property $P_i \in P$

For each object $o_z \in O$

Step 1.

If in each base profile $BP(t_n) \in C^m(t_c)$ $o_z \in P_i^+(t_n)$ then $cP_{im}^+(t_c) := cP_{im}^+(t_c) \cup \{o_z\}$.

Step 2.

If in each base profile $BP(t_n) \in C^m(t_c)$, $o_z \in P_i^-(t_n)$ then $cP_{im}^-(t_c) := cP_{im}^-(t_c) \cup \{o_z\}$.

Step 3.

If there is at least one base profile $BP(t_n) \in C^m(t_c)$, in which $o_z \in P_i^+(t_n)$ and there doesn't exist any profile $BP(t_n) \in C^m(t_c)$, in which $o_z \in P_i^-(t_n)$, then $cP_{im}^+(t_c) := cP_{im}^+(t_c) \cup \{o_z\}$.

Step 4.

If there is at least one base profile $BP(t_n) \in C^m(t_c)$, in which $o_z \in P_i^-(t_n)$ and there doesn't exist base profile $BP(t_n) \in C^m(t_c)$, in which $o_z \in P_i^+(t_n)$, then $cP_{im}^-(t_c) := cP_{im}^-(t_c) \cup \{o_z\}$.

Step 5.

If the number of base profiles $BP(t_n) \in C^m(t_c)$, in which $o_z \in P_i^+(t_n)$ is bigger than the number of base profiles $BP(t_k)$ in which $o_z \in P_i^-(t_k)$, then $cP_{im}^+(t_c) := cP_{im}^+(t_c) \cup \{o_z\}$.

Step 6.

If the number of base profiles $BP(t_n) \in C^m(t_c)$, in which $o_z \in P_i^-(t_k)$ is bigger than the number of base profiles $BP(t_k)$ in which $o_z \in P_i^+(t_n)$, then $cP_{im}^-(t_c) := cP_{im}^-(t_c) \cup \{o_z\}$.

Step 7.

For each base profile $BP(t_n) \in C^m(t_c)$: $cP_{im}^+(t_c) = cP_{im}^+(t_c) \cup P_i^+(t_n) \cup (P_i^+(t_n) \setminus cP_{im}^+) \cup (P_i^-(t_n) \setminus cP_{im}^-)$.

Stop.

5 Conclusions

Although the problem of symbol grounding has been studied for quite a long time, there is still a lack of mathematical models for simplest languages' grounding.

The aim of this paper was to investigate the problem of grounding modal formulas generated individually by the agents after applying dedicated algorithm. Presented algorithm uses consensus methods and represents the process of grounding logic formulas in agent's knowledge structures. These formulas describe the states of objects from external world in relation to some properties. Proposed solution realises also an original way of defining semantics of modal formulas, because logic formulas are related to the internal representations of an object rather than to ontologically existing entity.

References

1. Coradeschi S., Saffiotti A., An Introduction to the Anchoring Problem, *Robotics and Autonomous Systems* 43, (2003), 85-96.
2. Harnad, S.: The Symbol Grounding Problem. *Physica*, 42, 335-236
3. Katarzyniak, R., Nguyen, N.T.: Reconciling Inconsistent Profiles of Agents' Knowledge States in Distributed Multiagent Systems Using Consensus Methods. *Systems Science*, Vol. 26 No. 4, (2000) 93-119.
4. Katarzyniak, R., Pieczynska-Kuchtiak, A.: Formal Modelling of the Semantics for Communication Languages in Systems of Believable Agents. In: *Proc. of ISAT'2001*, Szklarska Poreba, (2001), 174-181
5. Katarzyniak, R., Pieczynska-Kuchtiak, A.: Intentional Semantics for Logic Disjunctions, Alternatives and Cognitive agent's Belief. In: *Proc. of the 14th International Conference on System Science*, Wroclaw, Poland, (2001), 370-382
6. Katarzyniak, R., Pieczynska-Kuchtiak, A.: A Consensus Based Algorithm for Grounding Belief formulas in Internally Stored Perceptions. *Neural Network World*, 5, (2002) 671-682
7. Katarzyniak, R., Pieczynska-Kuchtiak, A.: Distance Measure Between Cognitive Agent's Stored Perceptions. In: *Proc. of IASTED International Conference on Modelling, Identification and Control, MIC'2002*, Innsbruck, Austria (2002) 517-522
8. Katarzyniak, R., Pieczynska-Kuchtiak, A.: Grounding Languages in Cognitive Agents and Robots. In: *Proc. of Sixteenth International Conference on System Engineering*, Coventry (2003) 332-337
9. Katarzyniak, R., Pieczynska-Kuchtiak, A.: Grounding and extracting modal responses in cognitive agents: AND query and states of incomplete knowledge. *International Journal of Applied Mathematics and Computer Science*, 14(2), (2004), 249-263.
10. Nguyen, N.T.: Consensus System for Solving Conflicts in Distributed Systems. *Information Sciences*, 147, (2002) 91-122
11. Vogt P., Anchoring of semiotics symbols, *Robotics and Autonomous Systems*, 43, (2003) 109-120.