Affect-grounded Language Learning in a Robot

Zakaria Lemhaouri, Laura Cohen and Lola Canamero

I. Introduction

Most of the computational models of language development adopt a passive-learner view on language learning, and disregard the important role that affect plays in the development of communication, intersubjectivity, and the (co-)construction and sharing of meaning. Typical solutions propose teaching the artificial agent (infant) an association between a sensory perception of an object (e.g., an image) and a label (i.e., the name of the object) given by a knowledgeable caregiver [11], [10]. In this view, commonly adopted in AI, learning language is a goal in itself during the course of development. We propose to adopt an alternative view: that communication has an extrinsic functionality, i.e., a goal to achieve in the world that lies outside of language itself and can be better described as a means to reach this goal. This idea is in line with typical observations of infants' development, who can convey functional meanings before they master the adults' language [6]. For example, communication can be a way to obtain a desired object by requesting it from an adult, or a means to strengthen a social bond. Halliday argues that children initially develop "meaning potentials" to serve some functions that he identifies as instrumental, regulatory, interactional/social, personal, heuristic, imaginative and informative. We posit that endowing robots with the ability to learn language in this functional way is key towards bridging the gap between language and meaning in artificial agents, which remains one of the big challenges in artificial intelligence.

Previous work [3] uses a reinforcement learning paradigm [12] to model the learning of verbal and gestural communication skills through the interaction with a caregiver. It shows that a robot can learn both symbolic words and gestures to request objects by interacting with a caregiver. This corresponds to the instrumental function of language. We propose to extend this work to the other early functions of language described by Halliday, namely the regulatory, social and personal functions. For this purpose, we propose to include a model of affect [2], [9] as a prerequisite to motivate the acquisition of these functions of language [13], to ground the development of "meaning potentials" in emotional and affective internal states, beyond the basic biological needs. In this workshop, we will introduce the proposed framework and set the basis of our approach to a functionalist and embodied model of language learning by a robot, and we will present some preliminary results.

II. Method

In the present work, we propose to extend the RL approach proposed by [3] dedicated to learning associations between internal needs and words in a robot. The goal is to endow the robot with the ability to express its internal states by requesting an object from a human caregiver. However, the internal states of the robot remain limited as they are modeled by binary variables and do not evolve with time. Furthermore, the robot is not able to modulate these internal states depending on the visual perception of an object, as it does not have a vision module. In the present work, we propose to include more realistic affective internal states and a visual perception module to overcome these limitations. The proposed architecture is shown on Fig.1. The formalism is based on the sensory-motor PerAc neural architecture [5].

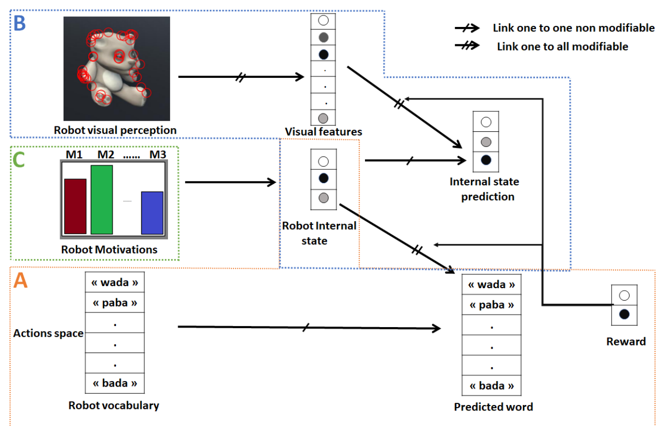

Fig. 1. The overall architecture of the model. A: corresponds to the RL block which allows the robot to create the association internal needs/words. B: The visual perception module for the association object/need. C: The motivation module for the affective internal states modeling.

A. Words / internal needs association

The association between the robot vocabulary and internal state (fig.1.A) follows the reinforcement learning approach proposed in [3]. In this context, the robot has three internal needs that can be satisfied with specific objects. The robot vocabulary is composed of words formed of two syllables (that corresponds to 10 of the most frequent syllables of an 8-month-old infant [8]). The robot starts by randomly producing a word when one need outweighs the others, the caregiver/human partner does not have access to the internal need of the robot. He/she reacts to the robot's vocalization by selecting an object and handing it to the robot. If the given object satisfies the robot's need, the motivation related to this need decreases, a reward of +1 is given to the robot which expresses its satisfaction with a happy gesture, otherwise the word receives a reward of -1, which decrease the probability of reusing the word in this context, and the robot expresses its dissatisfaction.

In RL, this problem is formulated as a contextual multi-armed bandit problem, in which the action space consists of the words that the robot can vocalize and the contexts correspond to the internal needs. In each context, the Q-value of each action a is estimated with the equations:

$$Q_{n+1}(a) = \frac{h-1}{h} Q_n(a) + R_n \tag{1}$$

with h, a parameter used to avoid the Q-value divergence, and R_n the reward received at time step n.

The robot uses a greedy approach to select a word according to its internal needs:

$$A_n = \underset{a}{\operatorname{argmax}}\, Q_n(a) \tag{2}$$

B. Object's recognition

To learn the association between the visual appearance of an object and the satisfaction of a need, we propose to include a visual perception module (fig.1.B) to our architecture, we used an online incremental learning method, similar to Kohonen's map[7], called SAW (Self Adaptive Winner). In this method, we define a matrix of visual features VF as:

$$VF_j = net_j.H_{max(\gamma, \overline{net}+\sigma_{net})}(net_j) \tag{3}$$

with :

$$net_j = 1 - \frac{1}{N}.\sum_{i=1}^{N} |W_{ij} - I_i| \tag{4}$$

net_j : is a measure of the similarity between the features bank VF and the new visual input I, this latter is a size N descriptor calculated around an extracted key points (fig.2).
W : the synaptic weight of the connection between the visual input I and the visual features bank VF.
H : the Heaviside function that allow the recruitment of a new neuron,i.e. adding the new visual input to VF, when the similarity is below the threshold of recognition γ:

$$H_\theta(x) = \begin{cases} 1 \text{ if } \theta < x \\ 0 \text{ otherwise} \end{cases} \tag{5}$$

The modification of the weights W is computed as:

$$\Delta W_{ij} = a_j(t) I_i + \mu(1 - \delta_j^k).(I_i - W_{ij})(1 - VF_j) \tag{6}$$

with:

$$k = argMax(a_j) \tag{7}$$

$$a_j = \begin{cases} 1 \text{ if a new neuron is recruited} \\ 0 \text{ otherwise} \end{cases} \tag{8}$$

δ_j^k Kronecker symbol :

$$\delta_j^k = \begin{cases} 1 \text{ if j=k} \\ 0 \text{ otherwise} \end{cases} \tag{9}$$

The VF bank is then used for training a neural network to predict the internal need that can be satisfied with the detected object. The synaptic weights update of this neural network follows the least mean square rule:

$$\Delta\omega_{ij} = \epsilon VF_i(RIS_j - ISP_j) \tag{10}$$

with : ϵ the learning rate, RIS the robot internal state and ISP internal state prediction.

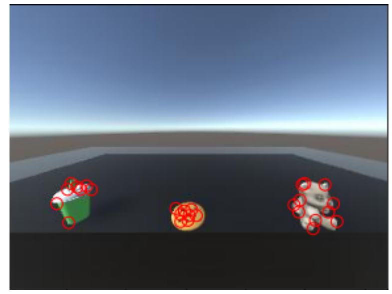

Fig. 2. Robot visual perception and the extracted key-points used to recognize objects.

C. Internal motivation

To endow the robot with a more realistic intrinsic motivation (fig.1.C), each internal need is modeled by a homeostatic variable $h(t)$ [4] that decreases over time and increases when the need is fulfilled :

$$h_i(t) = h_0 e^{-t/\tau_i} + \sum_k \alpha_{ki} \sum_{j_k} u(t - t_{jk}) e^{-(t-t_{jk})/\tau_i} \tag{11}$$

τ_i is the variable decay and u(t) is the Heaviside-step function ,α_{ki} indicates the amount by which the homeostatic variable increases when the need is fulfilled by an object k at time t_{jk}

The robot drive, representing the urge to act, is defined as the difference between the current homeostatic variable and its optimal value h_{op}:

$$d_i(t) = h_{op} - h_i(t) \tag{12}$$

The robot's motivation to satisfy a need depends on the related drive and the intensity of the stimulus / object that can satisfy this need [1]:

$$m = d + d.s \tag{13}$$

The intensity of the stimulus s is estimated by the object recognition module, for this purpose, we used the estimated activation of each class as the intensity of the stimulus corresponding to each object.

III. Experimental setup

To test our model, we used the social robot Reachy with the Unity simulation environment, the robot has three internal needs: hunger, thirst and curiosity, each need is satisfied by one of the three objects present in scene. These objects can be given to the robot by a human caregiver when the robot expresses its need. The vocabulary of the robot consists in 10 words, when a need is triggered, the robot says a word and the correspondent text is displayed (fig.3).

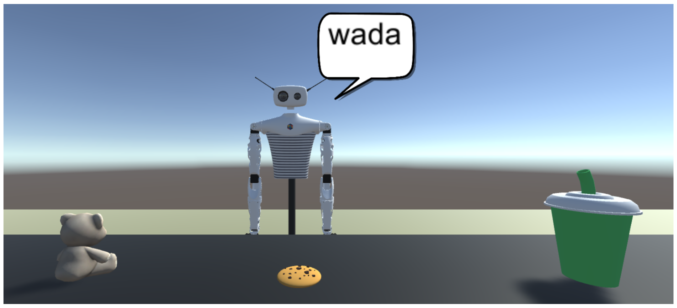

Fig. 3. Experimental setup. The Reachy robot is placed in front of a table with 3 objects that can satisfy each of his 3 internal states (i.e., curiosity, hunger and thirst). In a given internal state, he says a word from his word repertoire to get the corresponding object. The human partner tries to guess which object is desired and clicks on it. The robot then shows an expressive attitude as a feedback on his satisfaction or frustration.

IV. Results

After the learning phase, the results show the convergence of the moving average of the reward (fig.4). Table I represents an example of the Q-table of the association between the robot's vocabulary and the internal needs, after learning, each need has only one word with a max Q-value which confirms the convergence.

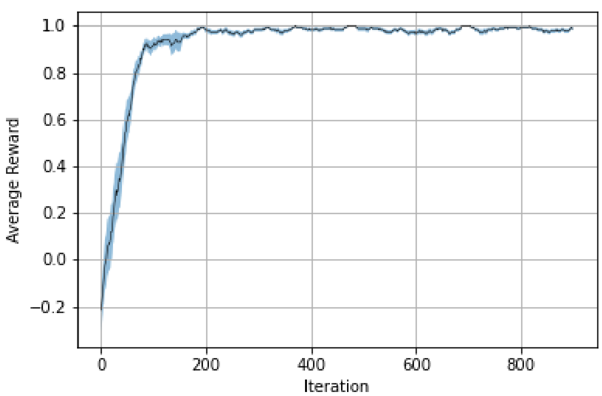

Fig. 4. Evolution of Average Reward.

	"wada"	"naba"	"maba"	"daba"	"paba"	"bada"	"bama"	"babe"	"waba"	"wama"
"Hunger"	-1.5	2	-1.5	-1.5	-1.5	-1.5	-1.5	-1.5	-1.5	-1.5
"Thirst"	-1.375	-1.5	-1.25	-1.15	-1.5	-1.5	-1.5	-1.5	-1.5	2
"Curiosity"	-1.69	-1.75	2	-1.83	-1.79	-1.6	-1.76	-1.71	-1.83	-1.75

TABLE I

Q-TABLE OF THE ASSOCIATION BETWEEN THE ROBOT VOCABULARY AND ITS INTERNAL STATES.

As a consequence of this words/needs mapping obtained by reinforcement learning and objects/needs correspondence made possible by the visual perception module, the robot was able to create a general association of word-object-need.

V. Conclusion

We have presented an improved model of active language learning. As in the previous work [3], the robot was able to create a mapping between its vocabulary and the goals it can achieve with words. The visual perception module allowed the robot to gain more autonomy and become able of recognizing objects to associate them with its needs. This association was made with respect to the developmental approach, since we used only online learning methods without prior information. The motivation module made the language learning more realistic and more similar to child's language acquisition. The expansion of this motivational module with new emotional and affective internal states will increase the number of meaningful words in the robot's vocabulary and will allow the computational modeling of the more advanced functions of language.

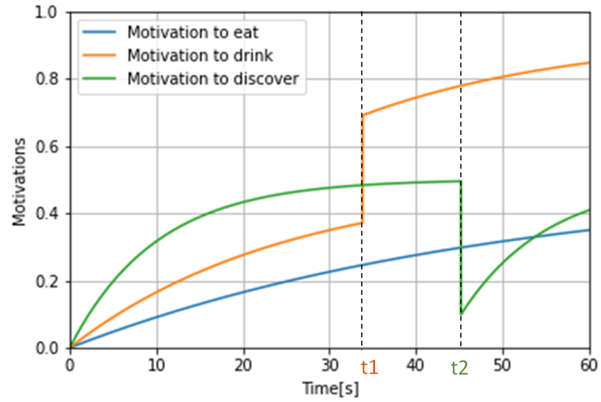

Fig. 5. The evolution of the robot's motivations. In this time interval, a drink was presented in the environment of the robot at t1, which increased the motivation to drink. At time t2, the caregiver gave the robot a toy, this act reduced the motivation to discover.

References

[1] Orlando Avila-Garcıa and Lola Cañamero. Using hormonal feedback to modulate action selection in a competitive scenario. In *From Animals to Animats 8: Proceedings of the 8th International Conference on Simulation of Adaptive Behaviour, MIT Press, Cambridge, MA*, pages 243–52. Citeseer, 2004.

[2] Dolores Canamero. Modeling motivations and emotions as a basis for intelligent behavior. In *Proceedings of the first international conference on Autonomous agents*, pages 148–155, 1997.

[3] Laura Cohen and Aude Billard. Social babbling: The emergence of symbolic gestures and words. *Neural Networks*, 106:194–204, 2018.

[4] Ignasi Cos, Lola Canamero, Gillian M Hayes, and Andrew Gillies. Hedonic value: Enhancing adaptation for motivated agents. *Adaptive Behavior*, 21(6):465–483, 2013.

[5] Philippe Gaussier and Stéphane Zrehen. Perac: A neural architecture to control artificial animals. *Robotics and Autonomous Systems*, 16(2-4):291–320, 1995.

[6] Michael Alexander Kirkwood Halliday. *Language of early childhood*, volume 4. A&C Black, 2006.

[7] Teuvo Kohonen. *Self-organization and associative memory*, volume 8. Springer Science & Business Media, 2012.

[8] Andrea G Levitt and Jennifer G Aydelott Utman. From babbling towards the sound systems of english and french: A longitudinal two-case study. *Journal of child language*, 19(1):19–49, 1992.

[9] Matthew Lewis and Lola Canamero. Hedonic quality or reward? a study of basic pleasure in homeostasis and decision making of a motivated autonomous robot. *Adaptive Behavior*, 24(5):267–291, 2016.

[10] Anthony F Morse and Angelo Cangelosi. Why are there developmental stages in language learning? a developmental robotics model of language development. *Cognitive Science*, 41:32–51, 2017.

[11] Luc Steels and Frederic Kaplan. Aibo's first words: The social learning of language and meaning. *Evolution of communication*, 4(1):3–32, 2000.

[12] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*, pages 25–45. MIT press, 2018.

[13] Colwyn Trevarthen. Making sense of infants making sense. *Intellectica*, 34(1):161–188, 2002.