The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

Incremental Learning of Integrated Semiotics Based on Linguistic and Behavioral Symbols

Wataru Takano and Yoshihiko Nakamura

Abstract— **This paper describes an novel approach towards linguistic processing for robots through integration of a motion language module and a natural language module. The motion language module represents association between symbolized motion patterns and words. The natural language module models sentences. The motion language module and the natural language module are graphically integrated. The integration allows robots not only to interpret observed motion as a sentence but also to generate motion with a sentence. This paper proposes incremental learning algorithm of association between symbolized motion patterns and words. The incremental learning is required for robot to autonomously develop the linguistic skill. The algorithm can be derived from optimization of the motion language module under stochastic constraints such that the associative probability of a new training pair composed of symbolized motion pattern and sentence becomes larger. Test of interpreting observed motion as sentences demonstrates the validity of the proposed incremental learning algorithm.**

I. Introduction

Language distinguish human from all other species. Human can interpret phenomenon, logically think and communicate with each other through the linguistic symbol system.

Saussure defined a “signe”(sign) as being composed of “signifié”(signified) and “signifiant”(signifier) [1]. The signifié represents concept of the sign. The signifiant is the sound image. According to his theory, the combination of the signifié and the signifiant is not inevitable but arbitrary entity. There is no inherent connection between the signe and an external world. The arbitrary nature provides the versatility and the operability of the symbolic system which leads to “langue”(language).

Peirce also defined three categories of signs : “icon”, “index” and “symbol” [2]. These three signs are distinguished by their referential associations. The icon refers to a thing with its close resemblance. The index is a sign where there is direct link between the sign and things. The symbol is more abstract sign which has no link with things. Deacon explained the origin of the language from Peirce’s semiotics [3].

These theories of semiotics have inspired researches on robot’s imitation learning based on symbolization of motion pattern such as multiple pairs of forward and inverse modules [4], neural networks [5][6][7], stochastic models [8][9][10]. Although these approaches enable robots to recognize observed behavior as symbol, the robot does not have linguistic ability. Language is crucial for robot’s intelligence. Sugita et al. proposed a novel approach to generation of a motion from a sentence by combining two neural networks. One learns mobile robot’s motions, and the other learns sentences [11]. Ogata et al. additionally proposed a method to generate a sentence from a motion by using the same framework as Sugita’s model [12]. In these framework only one sentence is generated from a motion. A motion can be expressed by various kinds of sentences. We also have made research on linguistic processing based on symbolization of motion for humanoid robots [13][14], where motion symbols and natural language are stochastically integrated. Our stochastic approach can not only generate multiple motions from a sentence but also generate various kinds of sentences from a motion.

W. Takano and Y. Nakamura are with Mechano-Informatics, University of Tokyo, 7-3-1 Hongo, Bunkyoku, Tokyo, Japan, {takano, nakamura}@ynl.t.u-tokyo.ac.jp

This paper describes incremental learning of association between motion symbols and natural language. The proposed framework consists of a motion language module and a natural language module. The motion language module associates words with motion symbols. The motion symbols represent motion pattern data as HMMs (Hidden Markov Models). The natural language module stochastically models sequences of words. The motion language module and the natural language module have properties of semantic and syntax respectively. The integration of the motion language module and the natural language module allows robots to both interpret observed motion as a sentence and generate motion from a perceived sentence. The proposed learning algorithm makes it possible for robots to gradually improve the ability of unknown association between the motion symbol and the words. This paper also verifies the validity of the incremental learning algorithm on experiment where a robot makes a sentence corresponding to observed motion pattern.

II. Integration of Motion Language Module and Natural Language Module [13]

The proposed framework is composed of motion language module and natural language module as illustrated by Fig.1. The motion module represents association between motion patterns and words. The natural language module models sequences of words. The integration of the motion language module and the natural language module realize linguistic bidirectional processing : interpretation of motion as a sentence and association of motion with a sentence. These two computations can be achieved by graphical search since both the motion language module and the natural language module are stochastic graphical models.

Motion language module stochastically represents associative structure between motion symbols and words through hidden variables. Note that an HMM representing motion

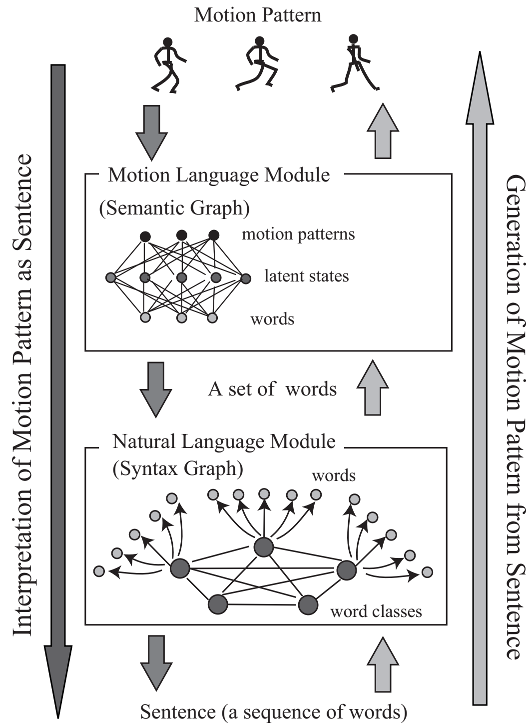

Fig. 1. The framework of linguistic processing is composed of motion language module and natural language module. The motion language module associates words with a motion pattern. The natural language module represents sequences of words. The integration of the motion language module and the natural language module makes it possible for robots not only to interpret motion data as a sentence but also to associate motion data with a sentence.

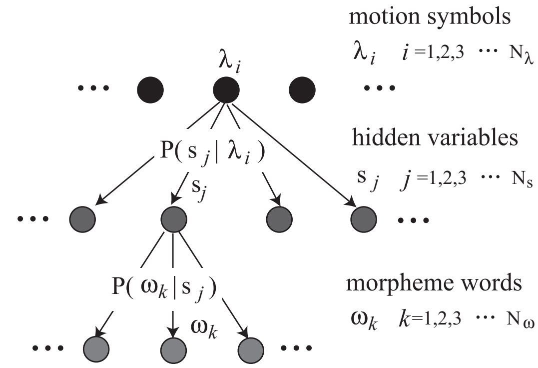

Fig. 2. Motion language module is composed of three layers : motion symbols, hidden variables and words. Association between the motion symbols and words are represented by two kinds of probabilities. One is the probability that a motion symbol generates a hidden variable. Another is the probability that a hidden variable generate a word.

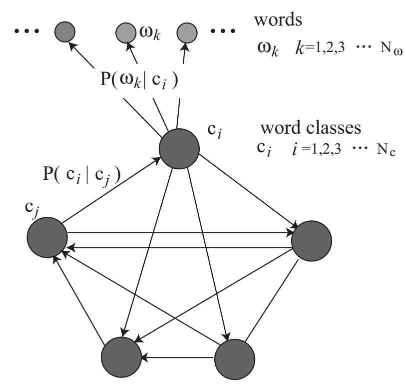

Fig. 3. Natural language module represents sentences (sequences of words). This module expresses sequences of word classes as node transitions and words generated by word classes as output probabilities in each node.

data is called motion symbol. Fig.2 illustrates the motion language module, where association between the motion symbols and the words are represented by using two kinds of probabilities. One is probability $P(s|\lambda)$ that a hidden variable s is associated with a motion symbol λ. Another is probability $P(\omega|s)$ that a hidden variable s generates a word ω. Therefore probability $P(\omega|\lambda)$ that a word ω is associated with a motion symbol λ is calculated as following equation.

$$P(\omega|\lambda) = \sum_{i}^{N_s} P(\omega|s_i)P(s_i|\lambda) \tag{1}$$

where N_s is the number of the hidden variables.

The model parameters, $P(s|\lambda)$ and $P(\omega|s)$, are optimized by EM algorithm given by training pairs composed of a motion symbol and a sentence (a sequence of words). The k-th training pair, $\{\lambda^k; \omega_1^k, \omega_2^k, \cdots, \omega_{n_k}^k\}$, means that the k-th observed motion is recognized as the motion symbol λ^k and that the same motion can be expressed by the sentence $\{\omega_1^k, \omega_2^k, \cdots, \omega_{n_k}^k\}$. Note that n_k is the number of words included in the k-th sentence. The optimized parameters are expressed as follows.

$$P(s|\lambda) = \frac{\sum_{k=1}^{N}\sum_{i=1}^{n_k}\delta(\lambda,\lambda^k)P(s|\lambda^k,\omega_i^k)}{\sum_{j=1}^{N_s}\sum_{k=1}^{N}\sum_{i=1}^{n_k}\delta(\lambda,\lambda^k)P(s_j|\lambda^k,\omega_i^k)} \tag{2}$$

$$P(\omega|s) = \frac{\sum_{k=1}^{N}\sum_{i=1}^{n_k}\delta(\omega,\omega_i^k)P(s|\lambda^k,\omega_i^k)}{\sum_{j=1}^{N_\omega}\sum_{k=1}^{N}\sum_{i=1}^{n_k}\delta(\omega_j,\omega_i^k)P(s|\lambda^k,\omega_i^k)} \tag{3}$$

where δ is kronecker delta. N and N_ω is the number of training pairs and words respectively. The numerators in Eqn.2 and Eqn.3 express expected number of times that hidden variable s is generated from motion symbol λ and expected number of times that hidden variable s is generated from word ω respectively. The denominators in Eqn.2 and Eqn.3 express the number of motion symbol λ in the training pairs and the expected number of times of hidden variable s in the training pairs.

The distribution of the hidden variable is estimated as the following equation.

$$P(s|\lambda^k, \omega_i^k) = \frac{P(\omega_i^k|s, \lambda^k, \theta)P(s|\lambda^k, \theta)}{\sum_{j=1}^{N_s} P(\omega_i^k|s_j, \lambda^k, \theta)P(s_j|\lambda^k, \theta)} \tag{4}$$

where θ is a set of the previously estimated model parameters $P(s|\lambda)$ and $P(\omega|s)$.

The optimized model parameters can be derived by alternately performing E-step (Eqn.4) and M-step (Eqn.2, Eqn.3). However this optimization can be achieved offline only in the case that the training pairs are given in advanced. Incrementally learning association between motion symbols and words are required for robots which develops the linguistic skill through interaction with their partners.

Natural language module also stochastically represents sentences (sequences of words). Fig.3 illustrates the natural language module which is expressed by an HMM [15]. A node in the natural language model corresponds to a word class such as noun verb and so on. A word ω is generated by a node c with probability $P(\omega|c)$. A sequence of word classes are represented by transition probability $P(c_i|c_j)$ from a node c_j to a node c_i. These parameters are can be derived from following equation.

$$\begin{aligned} P(c_i|c_j) &= \frac{N(c_i, c_j)}{\sum_{i=1}^{N_c} N(c_i, c_j)} && (5) \\ P(\omega_k|c_i) &= \frac{N(\omega_k, c_i)}{\sum_{k=1}^{N_\omega} N(\omega_k, c_i)} && (6) \end{aligned}$$

where $N(c_i, c_j)$ is the number of times that the word class c_i follows the word class c_j, $N(\omega_k, c_i)$ is the number of times that the word ω_k is categorized as the word class c_i in given training sentences.

Not only the interpretation of a motion pattern as a sentence but also the association of motion with a sentence can be realized by solving a graphical search by using the motion language module and the natural language module. The searches stated by Eqn.7 and Eqn.8 can be efficiently executed by A^* search.

$$\begin{aligned} \boldsymbol{\omega}^o &= \arg\max_{\forall \boldsymbol{\omega}} P(\boldsymbol{\omega}|\lambda) \\ &\approx \arg\max_{\forall \boldsymbol{\omega}} [\log P(\boldsymbol{\omega}|\lambda) + \log P(\boldsymbol{\omega}|L)] && (7) \\ \lambda^o &= \arg\max_{\forall \lambda} P(\lambda|\boldsymbol{\omega}) \\ &= \arg\max_{\forall \lambda} \sum_{j=1}^{n_*} \log P(\omega_j^*|\lambda) && (8) \end{aligned}$$

where L signifies the natural language module, that is, $P(\boldsymbol{\omega}|L)$ is the probability that the sentence, $\boldsymbol{\omega}$, is generated by the natural language module, L. n_* is the length of the sentence, $\boldsymbol{\omega}$. 1st term and 2nd in Eqn.7 can be separately computed by using the motion language module and the natural language module. Eqn.8 can be computed by using only the motion language module. Eqn.7 and Eqn.8 respectively correspond to the interpretation of a motion pattern as a sentence and the association of motion with a sentence.

III. Incremental Learning of Association between Motion Symbols and Words

Robots are required to gradually improve linguistic skill through interaction with their partners. Incremental learning of motion language module is neccesary.

A new pair of a motion symbol λ^{new} and a sentence $\boldsymbol{\omega}^{new} = \left\{\omega_1^{new}, \omega_2^{new}, \cdots, \omega_{n_{new}}^{new}\right\}$ is given as training data. The evaluation function ϕ for the newly given pair is calulated by logarithm of the probability that the sentence $\boldsymbol{\omega}^{new}$ is generated by the motion symbol λ^{new}.

$$\begin{aligned} \phi &= \log P(\boldsymbol{\omega}^{new}|\lambda^{new}) \\ &= \sum_{i=1}^{n_{new}} \log P(\omega_i^{new}|\lambda^{new}) \\ &= \sum_{i=1}^{n_{new}} \log \sum_{j=1}^{N_s} P(\omega_i^{new}|s_j)P(s_j|\lambda^{new}) && (9) \end{aligned}$$

The parameter changes, $\delta P(\omega|s)$ and $\delta P(s|\lambda)$, must be calculated such that the evaluation function ϕ can become larger; that is, variation of the evaluation function, $\delta\phi(\boldsymbol{\delta P})$, is maximized, where $\delta\phi(\boldsymbol{\delta P}) = \phi(\boldsymbol{P} + \boldsymbol{\delta P}) - \phi(\boldsymbol{P})$, and $\boldsymbol{P}$ is a vector with all the parameter expressed as $[P(\omega_1|s_1), \cdots, P(\omega_{N_\omega}|s_{N_s}), P(s_1|\lambda_1), \cdots, P(s_{N_s}|\lambda_{N_\lambda})]$. Note that N_λ is the number of motion symbols. But the changes of probabilistic parameters are subject to following constraints

$$\sum_{i=1}^{N_\omega} \delta P(\omega_i|s_j) = 0, \quad j = 1, 2, \cdots, N_s \tag{10}$$

$$\sum_{i=1}^{N_s} \delta P(s_i|\lambda_j) = 0, \quad j = 1, 2, \cdots, N_\lambda \tag{11}$$

They imply that constraints of $\sum_i P(\omega_i|s) = 1$ and $\sum_i P(s_i|\lambda) = 1$ are satisfied. We add other following constrains

$$\sum_{i=1}^{N_\omega} \left\{\delta P(\omega_i|s_j)\right\}^2 = \epsilon_j, \quad j = 1, 2, \cdots, N_s \tag{12}$$

$$\sum_{i=1}^{N_s} \left\{\delta P(s_i|\lambda_j)\right\}^2 = \eta_j, \quad j = 1, 2, \cdots, N_\lambda \tag{13}$$

The ϵ and η represent the degree of change. They correspond to forgetting rate of the incremental learning. Adding the Lagrang multipliers α, β, μ, ν, the following function L can

TABLE I

COMPARISON BETWEEN WITHOUT AND WITH INCREMENTAL LEARNING. 5 NEW PAIRS OF A MOTION SYMBOL AND A SENTENCE ARE LEARNED INCREMENTALLY. THIS TABLE SHOWS PROBABILITIES THAT TWO MOTION LANGUAGE MODULES GENERATE THE WORDS FROM THE MOTION PATTERN. ONE MODULE IS COMPOSED OF RANDOMLY SET PARAMETERS, ANOTHER IS OPTIMIZED BY USING THE NEW PAIRS. NOTE THAT THE LOGARITHMS OF THE PROBABILITIES ARE TAKEN.

♯ Motion Symbol	Sentence	Probability without Learning	Probability with Learning
21	A hitter starts running	-18.36	-14.24
22	A runner jumps	-18.28	-14.66
23	A coarch crosses his arms	-18.37	-14.14
24	A player stretches	-18.42	-14.33
25	A hitter swings a bat	-25.87	-21.92

be obtained.

$$\begin{aligned} L &= \delta\phi(\boldsymbol{\delta P}) \\ &+ \sum_{i}^{N_s}\left[\alpha_i \sum_{j=1}^{N_\omega} \delta P(\omega_j|s_i)\right] \\ &+ \sum_{i}^{N_\lambda}\left[\beta_i \sum_{j=1}^{N_s} \delta P(s_j|\lambda_i)\right] \\ &+ \frac{1}{2}\sum_{i}^{N_s}\left[\mu_i\left\{\sum_{j=1}^{N_\omega}\{\delta P(\omega_j|s_i)\}^2 - \epsilon_i\right\}\right] \\ &+ \frac{1}{2}\sum_{i}^{N_\lambda}\left[\nu_i\left\{\sum_{j=1}^{N_s}\{\delta P(s_j|\lambda_i)\}^2 - \eta_i\right\}\right] \end{aligned} \quad (14)$$

Setting the derivative of L with respect to the changes of model parameters, $\delta P(\omega|s)$ and $\delta P(s|\lambda)$, and the Lagrange multipliers, α, β, μ, ν, equal to zero, we can get the following optimum solution.

$$\delta P(\omega_i|s_j) = -\frac{\alpha_j + \dfrac{\partial\delta\phi}{\partial\delta P(\omega_i|s_j)}}{\mu_j} \quad (15)$$

$$\delta P(s_i|\lambda_j) = -\frac{\beta_j + \dfrac{\partial\delta\phi}{\partial\delta P(s_i|\lambda_j)}}{\nu_j} \quad (16)$$

where

$$\alpha_i = -\frac{1}{N_\omega}\sum_{j=1}^{N_\omega}\frac{\partial\delta\phi}{\partial\delta P(\omega_j|s_i)} \quad (17)$$

$$\beta_i = -\frac{1}{N_s}\sum_{j=1}^{N_s}\frac{\partial\delta\phi}{\partial\delta P(s_j|\lambda_i)} \quad (18)$$

$$\mu_i = \begin{cases} -\alpha_i\sqrt{\dfrac{N_\omega(N_\omega-1)}{\epsilon_i}} & if\ \alpha_i \geq 0 \\ \alpha_i\sqrt{\dfrac{N_\omega(N_\omega-1)}{\epsilon_i}} & if\ \alpha_i < 0 \end{cases} \quad (19)$$

$$\nu_i = \begin{cases} -\beta_i\sqrt{\dfrac{N_s(N_s-1)}{\eta_i}} & if\ \alpha_i \geq 0 \\ \beta\sqrt{\dfrac{N_s(N_s-1)}{\eta_i}} & if\ \alpha_i < 0 \end{cases} \quad (20)$$

$$\frac{\partial\delta\phi}{\partial\delta P(\omega|s)} = \sum_{i=1}^{n_{new}}\frac{\delta(\omega,\omega_i^{new})P(s|\lambda^{new})}{P(\omega_i^{new}|\lambda^{new})} \quad (21)$$

$$\frac{\partial\delta\phi}{\partial\delta P(s|\lambda)} = \sum_{i=1}^{n_{new}}\frac{\delta(\lambda,\lambda^{new})P(\omega_i^{new}|s)}{P(\omega_i^{new}|\lambda^{new})} \quad (22)$$

Note that $\delta(\omega,\omega_i^{new})$ and $\delta(\lambda,\lambda^{new})$ are Kroencker deltas. Although setting the derivatives equal to zero provides two solutions of $\boldsymbol{\delta P}$ which have different sign from each other, an appropriate solution is selected by Eqn.19 and Eqn.20 such that $\delta\phi(\boldsymbol{\delta P})$ can become positive.

Therefore the optimization of the objective function under the constraint condition through Lagrange multipliers yields the changes of the model parameters. Incremental learning of the motion language module can be implemented.

$$P(\omega|s)^{(i+1)} = P(\omega|s)^{(i)} + \delta P(\omega|s) \quad (23)$$

$$P(s|\lambda)^{(i+1)} = P(s|\lambda)^{(i)} + \delta P(s|\lambda) \quad (24)$$

The superscript i means that the model parameter is derived by i-th iterative updating.

IV. EXPERIMENTS

The incremental learning was tested on human motion data and sentences. The motion data is obtained by using a optical motion capture system. The motion capture system measures positions of 34 markers attached to a performer. The joint angles can be estimated from the measured Cartesian positions by inverse kinematics computation based on a humanoid robot with 20 degrees of freedom. The human behaviors are represented by the sequences of 20 joint angles. The motion data can be recognized as motion symbol λ^{new}. The motion data is also expressed by a sentence $\boldsymbol{\omega}^{new}$. The unknown pair of the motion symbol and the sentence is incrementally learned by motion language module. In this experiment we use a motion language module with 25 motion symbols, 50 hidden variables and 40 words ($N_\lambda = 25, N_s = 50, N_\omega = 40$) and a natural language module with 5 nodes. We set both ϵ and η to 0.0001. Note that we use 50 training pairs, each of which is a combination of a motion symbol and a sentence such as motion symbol λ_1 and a sentence "a player runs".

The motion language module incrementally learns 5 new pairs of a motion symbol and a sentence such as

{ λ_{21} : a hitter starts running },
{ λ_{22} : a runner jumps },

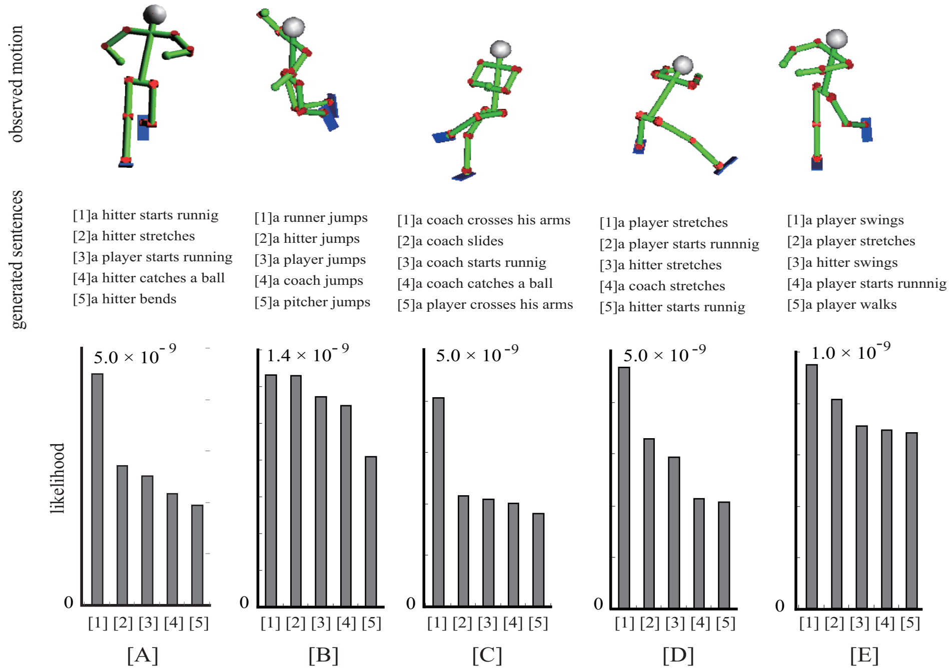

Fig. 4. An obeserved motion is recognized as a motion symbol. Sentences are associated with the motion symol. A sentence of "a hitter starts runnnig " is the most likely associated with running motion in [A]. A sentence of "a runner jumps" is the most likely generated by jumping motion in [B]. A sentence of "a coach crosses his arms" is the most likely generated by crossing-arm motion [C]. [D] shows that a sentence of "a player stretches" is associated with stretching motion . [E] shows that a sentence of "a player swings a bat" is generated by swinging motion with the largest likelihood.

TABLE II

COMPARISON BETWEEN WITHOUT AND WITH INCREMENTAL LEARNING FOR TRAINING DATA. SEVERAL SENTENCES ARE GENERATED FROM EACH MOTION SYMBOL THROUGH MOTION LANGUAGE MODULES AFTER AND BEFORE INCREMENTAL LEARNING.

♯ Motion Symbol	Training Sentences	Generated Sentences After Increment Learning	Generated Sentences Before Increment Sentences
1	A runner runs	A runner runs	A runner runs
	A player runs	A player runs	A player runs
	A hitter runs	A hitter runs	A hitter runs
2	A player shakes hands	A player shakes hands	A player shakes hands
	A runner shakes hands	A runner shakes hands	A runner shakes hands
	A hitter shakes hands	A hitter shakes hands	A hitter shakes hands
3	A hitter swings	A hitter swings	A hitter swings
	A player swings	A player swings	A player swings
	A hitter swings a bat	Swing a hitter	Swing a hitter
4	A pitcher applauds	A pitcher applauds	A pitcher applauds
	A player applauds	A player applauds	A player applauds
	A coach applauds	A coach applauds	A coach applauds
5	A pitcher crouches	A pitcher crouches	A pitcher crouches
	A player crouches	A player crouches	A player crouches
	A pitcher crouches on a mound	A mound crouches	A mound crouches

{ λ_{23} : a coach crosses his arms},
{ λ_{24} : a player stretches},
{ λ_{25} : a hitter swing a bat}

Note that motion symbols, $\lambda_{21}, \lambda_{22}, \lambda_{23}, \lambda_{24}, \lambda_{25}$, represent motion pattern of running, jumping, crossing one's arms, stretching and swinging a bat respectively. Table.I shows probabilities that the motion symbol generates the sentence through the motion language modules. One motion language does not incrementally learn the 5 new pairs of the motion symbol and the sentence and another model does. The probabilities are expressed by Eqn.9 and logarithms of the probabilities are shown. Table.I verifies that the proposed incremental learning algorithm improves the new association between a motion symbol and a sentence.

We also tested the validity of the incremental learning for integration of the motion language module and the natural language module. Fig.4 shows 5 sentences with the largest likelihood that each motion pattern generates a sentence as computed in Eqn.7. Fig.4(A) shows that a sentence of “a hitter starts running” is the most likely associated with running motion as the motion language model learns a pair of motion symbol of running and a sentence of “a hitter starts running”. The 2nd associated sentence is “a hitter stretches”. It implies that the incremental learning makes the connection of the motion symbol and the word “hitter” strong. In Fig.4(B) a sentence of “a runner jumps” is appropriately generated from jumping motion. Other generated sentences includes the word “jump” since the incremental learning improves the association between the motion symbol corresponding to “jump” and the word “jump”. In Fig.4(C), a sentence of “a coach crosses his arm” is correctly generated from crossing-arms motion as incrementally learned. 5th sentence of “a player crosses his arms” also includes the words “cross his arms”. 2nd, 3rd and 4th sentences do not have the words “cross his arms” but have the word “coach”. These generated sentences proves that unknown association between the motion symbol λ_{23} and every word in the sentence “a coach crosses his arms” is incrementally learned. Fig.4(D) also shows that a sentence of “a player stretches” is correctly associated with stretching motion. In Fig.4(E) the our framework does not the most likely generate a sentence of “a hitter swings a bat” but a sentence of “a player swings”, where the subject is different from the learned sentence and a word of “bat” is not associated. The word of “bat” is rarely generated in the natural language module since there are a few sentences with “bat” learned by the natural language module. The natural language module is required to represent sentences more appropriately even if the sentences are not learned so much. However sentences including word of “swing” are generated from swinging motion.

Table.II shows that several sentences are composed from a motion symbol by using a motion language modules with and without incremental learning. The motion language module with incremental learning generate the same sentences as one without. The incremental learning can be achieved without destroying association between previously trained motion symbols and sentences.

V. Conclusion

This paper develops a novel approach towards incremental learning of motion language module, which stochastically represents association between motion symbols and words through two kinds of parameters : a probability that a motion symbol generates a hidden variable and a probability that a hidden variable generates a motion symbol. There parameters in the motion language module are optimized for a training pair of a motion symbol and a sentence such that the derivative of the variation of the probability that the motion symbol generates the sentence is maximized. The optimization under the stochastic constraints can be solved using Lagrange multiplier method.

The incremental learning was tested on unknown pairs of a motion symbol and a sentence. The proposed method improves the probability of association between the motion symbols and sentences. The experiments of generating sentences from motion patterns demonstrate the validity of the incremental learning of unknown relation between motion symbols and sentences.

However, the proposed framework represents the meaning of a sentence by words which consists of the sentence. For example, a sentence “a pitcher throws a ball” generates the same motion symbol as a sentence “a ball throw a pitcher”. We need to use dependency parsing results or word classes in the motion language module in order to deal more accurate meanings.

VI. Acknowledgments

This research was supported by Category S of Grant-in-Aid for Scientific Research (20220001), Japan Society for the Promotion of Science.

References

[1] F. D. Saussure, “Course in general linguistics.” McGraw-Hill Book Company, 1966.

[2] C. S. Peirce, “The collected papers of Charles Sanders Peirce.” Harvard University Press, Cambridge, MA, 1931-1958.

[3] T. Deacon, “Symbolic species : The co-evolution of language and the brain.” W.W.Norton and Company Inc, 1997.

[4] M. Haruno, D. Wolpert, and M. Kawato, “Mosaic model for sensorimotor learning and control,” *Neural Computation*, vol. 13, pp. 2201–2220, 2001.

[5] J. Tani and M. Ito, “Self-organization of behavioral primitives as multiple attractor dynamics: A robot experiment,” *IEEE Transactions on Systems, Man and Cybernetics Part A: Systems and Humans*, vol. 33, no. 4, pp. 481–488, 2003.

[6] H. Kadone and Y. Nakamura, “Symbolic memory for humanoid robots using hierarchical bifurcations of attractors in nonmonotonic neural networks,” in *Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2005, pp. 2900–2905.

[7] T. Ogata, S. Matsumoto, J. Tani, K. Komatani, and H. G. Okuno, “Human-robot cooperation using quasi-symbols generated by RNNPB model,” in *Proceedings of the IEEE International Conference on Robotics and Automation*, 2007, pp. 2156–2161.

[8] T. Inamura, I. Toshima, H. Tanie, and Y. Nakamura, “Embodied symbol emergence based on mimesis theory,” *International Journal of Robotics Research*, vol. 23, no. 4, pp. 363–377, 2004.

[9] A. Billard, S. Calinon, and F. Guenter, “Discriminative and adaptive imitation in uni-manual and bi-manual tasks,” *Robotics and Autonomous Systems*, vol. 54, pp. 370–384, 2006.

[10] G. W. Taylor, G. E. Hinton, and S. Roweis, “Modeling human motion using binary latent variables,” in *Proceedings of the Conference on Neural Information Processing Systems*, 2006, pp. 1345–1352.

[11] Y. Sugita and J. Tani, “Learning semantic combinatoriality from the interaction between linguistic and behavioral processes,” *Adaptive Behavior*, pp. 33–52, 2005.

[12] T. Ogata, M. Murase, J. Tani, K. Komatani, and H. G. Okuno, “Two-way translation of compound sentences and arm motions by recurrent neural networks,” in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2007, pp. 1858–1863.

[13] W. Takano and Y. Nakamura, “Integrating whole body motion primitives and natural language for humanoid robots,” in *Proceedings of the IEEE-RAS International Conference on Humanoid Robots*, 2008, pp. 708–713.

[14] ——, “Statistically integrated semiotics that enables mutual inference between linguistic and behavioral symbols for humanoid robots,” in *Proceedings of the IEEE International Conference on Robotics and Automation*, 2009, pp. 646–652.

[15] K. Takeuchi and Y. Matsumoto, “Hmm parameter learning for japanese morphological analyzer,” in *Proceedings of the 10th Pacific Asia Conference on Language, Information and Computation*, 1995, pp. 163–172.