Imitating Operations On Internal Cognitive Structures for Language Aquisition

Thomas Cederborg, Pierre-Yves Oudeyer

Imitating Operations On Internal Cognitive Structures for Language Aquisition

Thomas Cederborg and Pierre-Yves Oudeyer
INRIA, France
http://flowers.inria.fr

***Abstract*—The paper examines the problem of learning socio-linguistic skills through imitation when those skills involve both observable motor patterns and internal unobservable cognitive operations. This approach is framed in a research program trying to investigate novel links between context-dependent motor learning by imitation and language acquisition. More precisely, the paper presents an algorithm allowing a robot to learn how to respond to communicative/linguistic actions of one human, called an interactant, by observing how another human, called a demonstrator, responds. As a response to 2 continuous communicative hand signs of the interactant, the demonstrator focuses on one out of three objects, and then performs a movement in relation to the object focused on. The response of the demonstrator, which depends on the context, including the hand signs produced by the interactant, is assumed to be appropriate and the robotic imitator uses these observations to build a general policy of how to respond to interactant actions. In this paper the communicative actions of the interactant are based on hand signs. The robot has to learn several things at the same time: 1) Weather it is the first sign or the second sign that specifies the object to focus on (that is, requests an internal cognitive operation), and the same for the request of a movement type. 2) How many hand signs there are and how to recognize them. 3) How many movement types there are and how to reproduce them in different contexts. 4) How to assign specific interactant hand signs to specific internal operations and specific movements. An algorithm is proposed based on a similarity metric between demonstrations, and an experiment is presented where the unseen "focus on object" operation and the hand movements are successfully imitated, including in situations not observed during the demonstrations.**

I. Introduction

A robot that is to operate in human populated environments, such as a home or an office, must be able to take the social context into account and take simple directions. In addition to all other difficulties that must be overcome before a robot is able to function around humans it should be able to learn how it should act as a response to social and linguistic cues[1]. In the presented experiment the linguistic cues is that of a sign language where a hand signs requests either a type of hand movement or an object focus. The current object focus will be referred to as a state in an internal cognitive structure (where internal refers to the fact that it is not observable). Thus the operation to change this state is referred to as an operation on an internal cognitive structure (unlike the visible hand movements, this operation is not visible). Therefor the hand sign requests an operation on an internal cognitive structure, and this operation must be imitated even if it is not directly visible (and it is the imitation of this unseen operation that results in the strong generalization performance). One of the challenges is to simultaneously learn new signs and learn new types of actions. We will investigate the unlabeled case, where the robot encounters an unknown number of communicative signs, movement types and internal cognitive operations, neither of which has been seen before. In order to study this case we present an experiment where a robot learns how it should respond to communicative signs of one human by imitating the responses of another human. The experimental setup is shown in fig 1 where one human, called an interactant, performs a communicative action and then another human, called a demonstrator, reacts (performing two types of actions, one internal cognitive operation and one movement). The imitator then builds a model of how it should react to communicative gestures.

Experiments show that the model generalizes well to combinations of communicative signs that was not observed during demonstrations.

Both actions and gestures are continuous, never exactly the same, and the imitator is never presented with any form of symbolic representation of interactant hand signs or demonstrator actions. The imitator must therefore infer both the number of actions it has been demonstrated and the number of gestures it has observed from data (an unknown number of specific instances are observed for each of an unknown number of action types and each of an unknown number of gesture types). Both directly observable actions (hand movements) and inferred unseen actions (internal cognitive operations of focusing on an object) are imitated. The imitation of unobservable internal actions extends the type of communication that can be learned by imitation to include some words that are not direct action requests.

Related work

There are two related lines of work, imitation learning and linguistics. These fields are traditionally studied separately but the present paper argues that there are fruitful ways to combine them.

Imitation learning, sometimes referred to as programming by

[1]Being able to achieve world states and predict the consequences of its actions does not tell it which world states are preferable. A robot that knows how to make coffee, how to smash the coffee cup, and is able predict the response of humans to each action, still have no way of knowing which of these two actions is the appropriate response to a request for coffee unless more information is somehow provided (for example, as in the presented experiment, by observing the actions of a human, that is assumed to act appropriately. Receiving feedback from a human that is assumed to know what should be done is another approach).

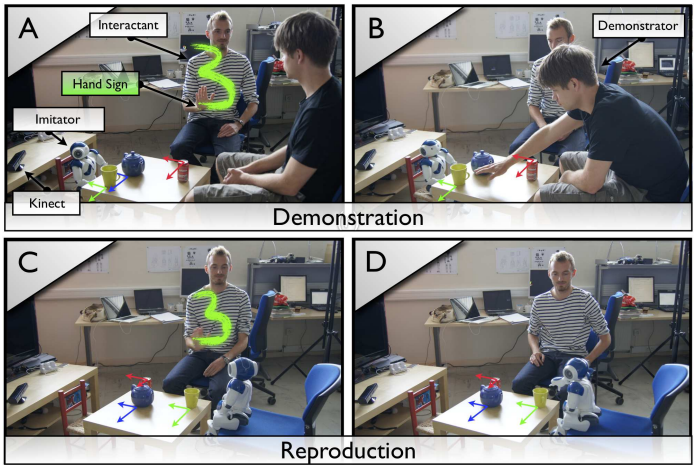

Fig. 1. A and B shows the demonstration phase while C and D shows the reproduction phase. In A the human interactant, in the middle, makes a communicative gesture. The human demonstrator, to the right, and the robotic imitator, to the left, observes the communicative gesture. In B the demonstrator, to the right, performs an action which is dependent on the observed gesture and the imitator observes. After several such demonstrations the imitator is able to build a model of how it should respond to the communicative gestures of the interactant based on how the demonstrator responded. Each demonstration and reproduction have its own object positions. In C the interactant performs a communicative gesture and the imitator observes it, and in D it reproduces the action. The hand trajectories are actually captured using a mouse instead of a kinect, and a simulated rather than physical robot is used (see section II-F for details).

demonstration or learning from demonstration examines the problem of learning sensorimotor tasks from demonstrations. For recent overviews of the field see [2], or [3] and for a recently proposed formalization of tele-operated imitators see [4]. The work presented in this paper would fit within this formalism and, as all such work, does not have to deal with the correspondence problem (see [5] for an early or [6] for a more recent explanation of this important problem in imitation learning). In [7], the question of how to find appropriate task spaces are considered, referred to as finding appropriate reference points. It examines a larger number and more diverse type of reference point based task spaces using demonstrations that are known to be of a single task. Most imitation learning research consider this single task setup, but see [8] for an exception where two different table tennis tasks are learnt from unlabeled demonstrations. Most research also concern tasks without a communicative component. For an exception, see [9], which also deals with multiple tasks and the problems of finding the number of gestures of an interactant and is perhaps the work that is most related to the present paper. It is however very unusual that these issues are dealt with in the field of imitation learning, and how to solve the related problems is largely an open research question.

A method that has been extensively used for learning a single sensorimotor task is Gaussian Mixture Regression (GMR). See for example [12], [13] or the recent book [14]. The main parameter is the number of Gaussians to use and this can be directly chosen but can also be estimated from the training data using the Bayesian Information Criteria (BIC), see [8] for such an experiment and [15] for an in depth explanation of BIC. The paper [17] examines how a task can be divided into subtasks, in a way that is related to the problem of finding the number of movements. Previous work [18] examined Incremental Local Online GMR (ILO-GMR) that can learn an open ended number of tasks from unlabeled demonstrations. In [18], the 2D position of an object was used to determine what task should be performed. This object position was used in a way which is similar to the communicative signs in the experiment presented here. The main difference is that the triggering regions could be kept well separated due to the fact that the object position was directly controllable (unlike the 3D point that a hand sign is transformed into in this experiment).

The task solved in the presented experiment is close to the task solved in [10]. There are also structural similarities in how the task is solved since the architecture presented in [10] is not based on separate systems for language and action. The difference between the proposed architecture and [10] is the use of imitation learning methods instead of rewards and the fact that the proposed architecture does not use a symbolic representation of the communicative acts. Furthermore the proposed architecture does not use neural networks. The generalization ability exhibited by the system in [10] is also exhibited by the proposed algorithm which is able to respond properly to novel combinations of linguistic commands. See [11] for another artificial neural network based approach to this type of task. However, in [11] the action and linguistic units are separate units and, like in [10], the linguistic inputs are symbolic.

Linguistics research have resulted in models of the evolution of language, for example using the setup of language games, see for example [19]. There has also been a move within the developmental robotics community (see for example [20]) in the direction towards viewing language in relation to the physical context of the speaker and hearer, as opposed to viewing language as independent from a physical reality (assumed to be connected later using some form of interface). Language is however still seen as a separate system and the research problem is framed as finding the link between this system and the sensorimotor system, or to find out how the two separate systems co-develop.

The proposed algorithm does not include a separate language system but is instead an imitation learning system whose context has been extended to include the communicative hand signs of an interactant. The number of different signs observed is not obtainable without looking at the effect that the actions of the interactant has on the demonstrator. If the field of imitation learning needs to include an interactant, perhaps the field of linguistics needs to include a demonstrator that acts in the world, an agent whose actions are modified by the interactants behavior. In linguistics, it is easier to see the need of learning to perform internal cognitive operations since so many sentences yield no external actions. There are difficult practical problems involved with imitating such operations but, as the present paper shows, this is possible if the demonstrator

acts according to a consistent policy, the internal state is changed in response to observable parts of the context (in this case the signs of the interactant) and the internal state modifies observable behavior.

Hidden Markov Models (HMM) is a mathematical structure that has been used in imitation learning and which would be suitable for the type of imitation learning presented here, modeling the current state of the demonstrator. Current research has, however, used HMM's as a way to represent how far along in the task the demonstrator is, a form of task time, as for example in [12]. In this paper we investigate a simple internal structure that does not change state during the demonstrations/reproductions, that has only 3 states and where the policy of the demonstrator is to perform a single operation with a deterministic outcome at a single point as a direct response to the environment. The structure of a more complicated internal structure and set of operations could be represented using an HMM (for example if a demonstrator could, with some probability, switch to focusing on a fast moving object). The problem in the current paper is finding out how one state effects behavior. Indeed, in a single demonstration there is no information about what the internal state is.

II. Algorithms

There are always three objects in the scene, one red, one blue and one green. Their position is set randomly at the start of each demonstration and at the start of each reproduction and is then kept constant during the demonstration/reproduction.

The demonstrator always focuses on exactly one of the objects in the scene, meaning that the only thing that will influence how the hand is moved is the position of the hand relative to the object that is focused on. Then it always performs one out of an unknown number of movements, defined in a coordinate system centered on the object focused on (for example: moving the hand in a circle around the object). We say that this trajectory is an instance of the movement that the demonstrator was trying to perform. For example a specific trajectory where the hand of the demonstrator moved roughly in a circle around the object is an instance of the "circle" movement. The imitator must estimate the number of movements that it has observed and, for each trajectory, determine what movement it was an instance of. This will be represented by hypothesizing a number of possible movements equal to the number of trajectories and assigning a probability $m_{m;t}$ (the probability that trajectory number t is an instance of movement number m) to each trajectory-movement pair. For example, if trajectories number A,B and C are the only instances of the circle movement, this would be represented by $m_{x;A}$, $m_{x;B}$ and $m_{x;C}$ being large and $m_{x;t}$ being small for all other values of t. It does not matter which hypothesized movement they are all instances of since a movement is completely defined by its members. For N demonstrations, this results in an NxN membership matrix denoted M. To infer which trajectories are instances of the same movement, a measure of similarity between trajectories is proposed, resulting in a similarity matrix (where hopefully trajectory number A will have higher similarity to trajectories number B and C than to other trajectories). Using the similarity matrix the values of the membership matrix is found by an iterative procedure referred to as the grouping algorithm[2]. The grouping algorithm proposed is novel but it is not a main contribution of the paper, and no claims of optimality are made. For completeness it is however presented in detail and thus takes up a large part of the text. After the grouping algorithm is done, and values for M have been found, the internal operation performed by the demonstrator in each demonstration is inferred. Given that the internal operations and the type of movement for each demonstration has been inferred this information is used to determine the word order of the sign language. In the experiment presented below the interactant always uses the first hand sign to request some form of internal operation and always uses the second hand sign to request some form of movement. The imitator infers this word order from the demonstrations.

The algorithm makes two basic assumptions. First it is assumed that for each movement there is a single low dimensional task space, valid during the entire movement, such that the policy of the demonstrator can be well specified in this space. Second, it is assumed that a gesture signals either the type of movement to be done or what internal operation to perform.

In figure 2 we see a graphical overview of how the demonstrations are acquired and in figure 3 how the reproductions are done.

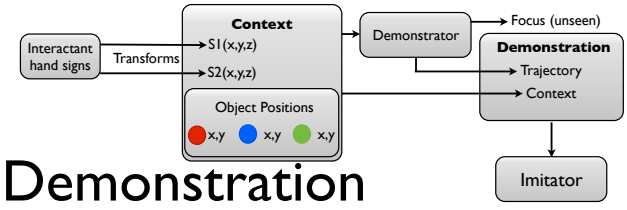

Fig. 2. The demonstrations are generated by a demonstrator reacting to hand signs of an interactant and the position of 3 objects (which together forms the context). Each of the hand signs are transformed into a 3D space and is observed by the imitator, along with the object positions and the trajectory. The internal operation must be reproduced by the imitator in order to achieve success, but it is not visible (the imitator must infer what this operation was from the data)

A. Demonstrations

In the experiment presented in this paper the syntax of the sign language that the imitator must find is as follows: the first sign requests the internal operation of a specific object focus; a "1" requests focus on the red object, a "2" the green and a "3" the blue object. The second sign requests a specific type of movement, defined in relation to the object specified[3]; a "4" requests performing the "triangle up" movement, a

[2]Since it results in groups of trajectories, where either a group is empty or the members of any specific group are all instances of the same movement.

[3]The policy maps hand positions in a coordinate system with (0,0) at the center of the object focused on

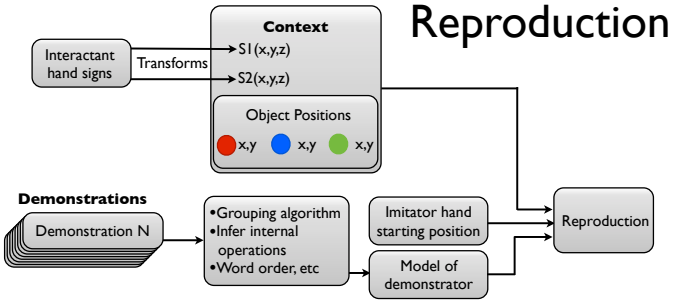

Fig. 3. The reproduction starts with the building of a model of the demonstrator using the algorithms detailed below. Then this model along with transformed hand signs of the interactant, object positions and the starting position of the imitators hand determines how the reproduction is done.

"5" the "triangle down" movement and a "6" requests the "circle" movement. For the imitator to find this word order it needs to infer the internal operations performed and it needs to know what trajectories are instances of the same movement. Each sign is transformed into a 3D point by comparing its similarity (after dynamic time warping) to three prototypes (trajectories of the signs "7","8" and "9"). There is nothing special or optimal about this particular transform, it is just a simple and fast way of transforming a continuous hand sign into a low dimensional space and any number of other methods could have been used (the only requirement on a transform from trajectories to three coordinates, is that there is some correlation between the type of trajectory and the resulting coordinates). During reproduction the imitator compares distances in this 3D space between the hand sign it is observing and the hand signs observed during demonstrations. If a larger number of hand signs is to be learnt it would be easy to create a transform to a higher dimensional space taking the distance to additional prototypes.

B. Trajectory distance $\Delta_{t;k;i;j}$

To determine which trajectories are instances of the same movement it is necessary to define some measure of distance between two trajectories. In each demonstration the three objects and the starting position is different. For each movement type, the policy of the demonstrator is determined by the hand position in the coordinate system centered on the object focused on. What object is focused on is not observable to the imitator, but two trajectories that are instances of the "circle" movement will only look similar if each is viewed in the coordinate system of the object that is focused on (the object encircled). For this reason the distance between two trajectories is defined relative to two coordinate systems; one used for trajectory 1 and the other one used for trajectory 2. Thus $\Delta_{A;B;1;3}$ is the distance between trajectory A seen in coordinate system 1 and trajectory B seen in coordinate system 3. If trajectory A is a circle around $0, 0$ in coordinate system 1 and trajectory B is a circle around $0, 0$ in coordinate system 3 they will probably look similar (trajectory A viewed in another coordinate system would still be a circle, but not around $0, 0$, and the two trajectories will not look similar).

For each of the N points in trajectory number t the closest points in trajectory k is selected (with distance measured using the respective coordinate systems). For each point p of trajectory number t, the closest point of trajectory number k is found, using the positions in the coordinate systems i and j respectively[4]. δ_p is defined as the angular difference in output of the two points. Then we have $D_{t;k;i;j} = \sum_{p=1}^{N} \delta_p^2/N$. Finally we have $\Delta_{t;k;i;j} = min(D_{t;k;i;j}, D_{t;k;i;j})$.

There are many possible ways of measuring similarity between two trajectories, given the coordinate systems to view them in and the paper makes no claim on the optimality of the specific similarity measure introduced. Like many other parts of the algorithm the important part is not how the specific part is implemented but instead how it is combined with the rest of the algorithm, with the details included only for completeness.

C. The grouping algorithm

The current estimate of the probability that trajectory number t is an instance of movement number m is denoted $m_{m;t}$. The suitable value of $m_{m;t}$ is completely determined by what movements the other trajectories are estimated to be instances of. The only thing that matters is that trajectories that are instances of the same movement are grouped together. Since the number of movements is unknown there are as many movements as trajectories (so that M is a NxN matrix for N demonstrations).

Given the similarity between trajectories there are many possible ways to divide them into subgroups and the iterative algorithm proposed is not claimed to be optimal (the reader that is not interested in exactly how similarities between trajectories is used to form groups whose members have high similarity can skip this section II-C). The basic principle of the grouping algorithm is that if two trajectories A and C are more similar to each other than other trajectories likely to be instances of movement x, then $m_{x;A}$ and $m_{x;C}$ will increase. If A and C are less similar than average, then $m_{x;A}$ and $m_{x;C}$ will decrease, and the magnitude of the change depends oh how much the similarity deviates from the other likely members.

The algorithm is described using pseudocode in 1. In order to save space, several variables (either used in the pseudocode or used to define other variables that are used in the in the pseudocode) are defined and explained below rather than in the pseudocode, such as: maximum trajectory similarity $\gamma_{t;k}$, joint memberships: $\omega_{t;k}$, weighted mean similarity ϖ_t and push strength $\xi_{t;k}$.

Maximum trajectory similarity $\gamma_{t;k}$. $\gamma_{t;k;i;j}$ is the inverse of the distance $\Delta_{t;k;i;j}$ and $\gamma_{t;k}$ is the maximum similarity between trajectories t and k, $\gamma_{t;k} = max_{i;j}(\gamma_{t;k;i;j})$ (for example, if trajectories A and C have the highest similarity when A is in coordinate system 1 and C is in coordinate system 2, $\gamma_{A;C}$ = $\gamma_{A;C;1;2}$, which is likely to be the case if

[4]So that if i=1, j=3, and point number p's position in coordinate system 1 is (0,0.4) then the point of trajectory k that is closest to (0,0.4) in coordinate system 3 is chosen (so that a point above the red object in trajectory t is compared to a point above the blue object in trajectory k).

Algorithm 1 Overview of the iterative grouping algorithm

Input: M_1, S, N

• M_1 *is the initial membership probabilities*

• S *is the number of steps (S=50 is used in the experiment presented below)*

• N *is the number of demonstrations*

for $s = 1$ **to** S **do**

  $M_{mod} \leftarrow M_s$ ($m_{m;t}$ refers to M_{mod})

  $M_{old} \leftarrow M_s$ ($m_{m;t;old}$ refers to M_{old})

  for $m = 1$ **to** N **do**

    for $t = 1$ **to** N **do**

      for $k = 1$ **to** N, $k \neq t$ **do**

        $m_{m;t} \leftarrow m_{m;k;old}\xi_{k;t} + (1 - m_{m;k;old})m_{m;t}$

      end for

    end for

  end for

  Rescale

  Preferring hypotheses with few movement types:

  $\forall : 1 < m < N, 1 < t < N$**:**

  $m_{m;t} \leftarrow m_{m;t} \times (\sum_{\tau=1}^{N} m_{m;\tau})^{1/4}$

  Rescale

  $m_{m;t} \leftarrow m_{m;t} + 0.0001$

  Rescale

  $M_{s+1} \leftarrow M_{mod}$

end for

note that if the push factor $\xi_{t;k}$ is positive $m_{m;t}$ will increase and if it is negative it will decrease in the central update step. Remember that a positive $\xi_{t;k}$ indicates that the policy similarity between t and k is higher than the weighted average. The rescaling makes the memberships of a single demonstration sum to 1

trajectory A is a circle around the red object and trajectory C is a circle around the green object).

Joint memberships $\omega_{t;k}$ is a measure of how probable it is that trajectories t and k are instances of the same movement according to the current state of the membership matrix M. It is calculated as: $\omega_{t;k} = (max_m(m_{m;t} * m_{m;k}))/(\sum_{\tau=1}^{N} max_{m;\tau}(m_{m;t} * m_{m;\tau}))$.

Weighted mean similarity ϖ_t is a measure of the weighted average similarity to trajectory t of trajectories that are likely to be instances of the same movement. $\varpi_t = \sum_{k=1}^{N} \omega_{t;k} * \gamma_{t;k}$.

Push strength $\xi_{t;k}$ is the strength with which trajectory t will affect the memberships of trajectory k in the movement groups that they are both probable members of. If it is positive the presence of trajectory k in a movement group will increase the membership of trajectory t and decrease it if it is negative. It is calculated as: $\xi_{t;k} = e^{((\gamma_{t;k}/\varpi_t)-1)}$, and we can for example see that $\xi_{t;k} = 1$ if the similarity between t and k is exactly the same as the average weighted similarity between t and the other trajectories that has high joint memberships with t. If the similarity $\gamma_{t;k}$ is bigger than the weighted average ϖ_t, the we will get a push strength $\xi_{t;k} > 1$ (and if the similarity $\gamma_{t;k}$ is smaller than the weighted average ϖ_t, we will get $\xi_{t;k} < 1$).

Inferring what object was focused on during each demonstration

When the grouping algorithm is successful we know what demonstrations include the same movements. The coordinate system in which a trajectory is the most similar to the other trajectories of the same movement is set as the coordinate system of that demonstration.

Finding the word order

The within group distances of the first signs and the second signs are compared and the one that has the biggest distance is assumed to designate the coordinate system. If this is successful the imitator knows which of the signs designates the coordinate system and which one designates the movement.

D. Finding the movement and the coordinate system during reproduction

The sign that has been found to designate movement is compared to the corresponding signs of all demonstrations and the group of the demonstrations whose sign is closest is assumed to be demonstrations of the correct movement.

The same is done to find the coordinate system: The sign that has been found to designate coordinate system is compared to the corresponding signs of all demonstrations and the coordinate system of the demonstration whose sign is closest is assumed to be the correct coordinate system.

E. Reproduction

At each timestep during the reproduction, the imitator finds the 50 points that are closest to the current state (measured in the coordinate system found) amongst those trajectories that are members of the movement found. The average of the output of these points is used. More sophisticated methods could easily be inserted here, for example ILO-GMR [18] or GMR [12] together with BIC [15]. Since low dimensional and accurate data is available after a successful grouping algorithm more sophisticated methods are not needed in this case. Again, the global power of the architecture lies in how simple algorithmic parts are combined together.

F. Simulating the setup

The imitator robot is simulated and is able to move its hand in any direction it wants which, if a physical robot is to be used, would require an inverse kinematics model that translates current joint configurations and desired hand directions to motor outputs. The simulated imitator was easier to perform experiments with and since the focus of the presented experiment is about learning what should be done rather than how to do it (in the language of [1] the "what to imitate" instead of the "how to imitate" question is the focus of the presented experiment) it was used in place of a physical robot. There are obviously limits to what types of behaviors a robot can learn to do in simulation before this starts to become a serious simplification, and if more advanced physical manipulations are to be investigated a physical robot will have to be used. The hand trajectories of the demonstrator as well as the communicative signs of the interactant are captured using

a mouse. Using for example a Kinect device would not reduce the quality of the trajectories and the presented approach was used due to its simplicity.

III. Experiment

In figure 4 the 12 demonstrations are shown relative to the three different objects. The appropriate response to six of the total nine possible combinations of communicative inputs are demonstrated, but the imitator successfully imitates in all nine combinations.

The algorithm finds the number of movements and correctly infers all the internal actions as well as the word order. Four separate reproductions are performed in each of the nine combinations, with no degradation in performance for the 3 tasks not demonstrated.

Similarity

In figure 5 we can see the 4 dimensional similarity matrix displayed graphically. Higher values mostly correspond to two trajectories that are instances of the same movement and under the correct hypotheses of object focus.

Maximum similarity

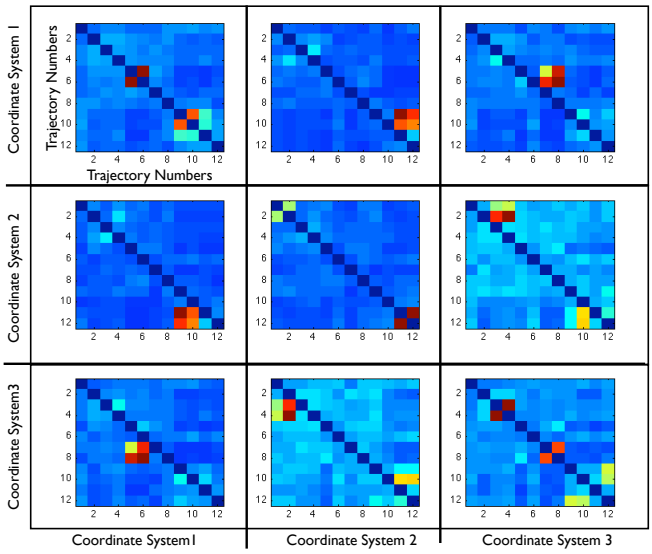

Fig. 5. Here we can see the similarity matrix with entries $\gamma_{t;k;i;j}$ presented graphically. The similarity goes from dark blue (lowest similarity) to dark red (highest similarity). The similarity of a trajectory with itself is undefined, but for the purpose of a graphical representation it must be given a value and is arbitrarily set to 0. t indexes the first trajectory (y-axis from 1 to 12 in each sub figure), k the second trajectory (x-axis from 1 to 12 in each sub figure), i the coordinate system used for trajectory k (indicating the sub figures row number) and j the coordinate system used for trajectory t (indicating the sub figures column number). For example (12,9,2,1) is the bright red of the sub figure in row 2 column 1 (indicating a high similarity between trajectories 12 and 9 in coordinate systems 2 and 1 respectively).

In figure 6 we can see the 2 dimensional maximum similarity matrix displayed graphically (trajectory number 1, number 2). In general the trajectories that are instances of the same movement have significantly higher similarity (1-4 for triangle up, 5-8 for triangle down, 9-12 for circle, as can be seen in fig 4).

Final groupings

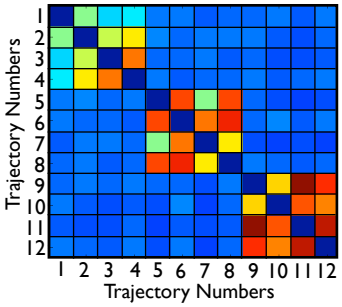

Fig. 6. Here we can see the maximum similarity matrix with entries $\gamma_{t;k}$ presented graphically. Ttrajectories 1 to 4 are of the triangle up movement, trajectories 5 to 8 are of the triangle down movement and finally trajectories 9 to 12 are of the circle movement. In general trajectories that are instances of the same movement have high similarity.

Using the maximum similarity matrix of figure 6 all trajectories was been grouped correctly together. The trajectories demonstrating movement 1 were all assigned to movement group 4, the trajectories of movement 2 to movement group 2 and the trajectories of movement 3 to movement group 11 (the number of the group is irrelevant as a movement group is completely defined by its members). These grouping results and the full 4D similarity matrix (fig 5) was then used to correctly infer the internal operation in each demonstration and the word order (as described in section II).

The grouping algorithm was tested an additional 50 times on the maximum similarity matrix from figure 6 and 49 times it was successful but one time it failed by grouping trajectories 1 to 8 in the same group (which would probably have lead to a failed reproduction of the two triangle movements but would not have compromised reproduction of the circle movement).

Reproductions

In figure 7 we can see 36 successful reproductions, where the top left, middle middle and bottom right each show 4 correct reproductions of an unseen task. The edges of the triangles are not as sharp as they should be and, when the starting position in the circle movement is far to the right of the object, the imitator initially makes a to big semi circle before falling into the correct small circle movement (more sophisticated methods for the reproduction could be used on the data obtained, but that is not the focus of the current paper and the reproduction ability was enough for our purposes). The three tasks not demonstrated is reproduces as well as the other tasks (top left, middle middle and bottom right), as should be expected from the structure of the algorithm.

IV. Conclusions and Future Work

We have shown that it is possible to simultaneously learn never before encountered communicative signs and never before encountered movements, without using labeled data, and at the same time learn new compositional associations between movements and signs. We have also shown that the actions learnt can include unseen internal operations (focus on object) of a demonstrator under a set of conditions. One condition was that the unseen operation is performed as a

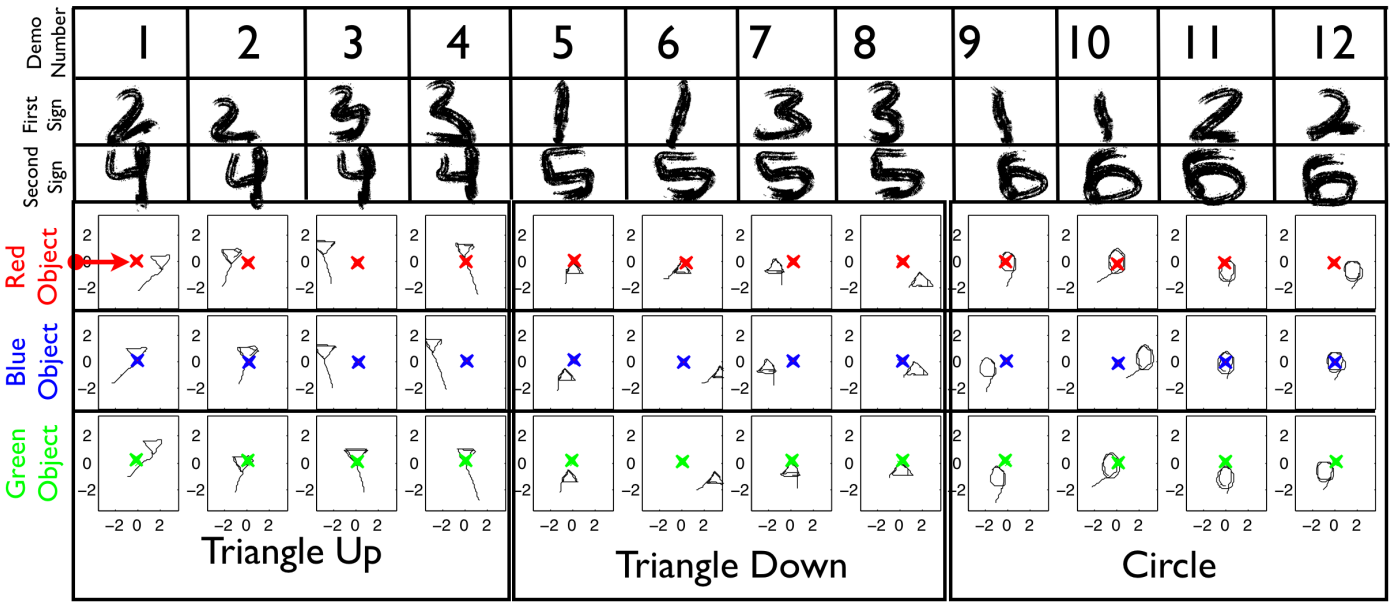

Fig. 4. Here we see the 12 demonstrations relative to the three objects. The demonstrator observes the first sign, the second sign and then performs the hand movement presented under. Each of the trajectories are shown relative to the three objects. One thing we can see is that demonstration number 11 almost seems to be making a circle around the red object in column 11 row 1. When we look at the similarity in 5 we can see that demonstration 11 indeed does look fairly similar to the other circles, even when viewing it in the incorrect first coordinate system and the other demonstration in their respective correct coordinate system (top left for comparison with demonstrations 9 and 10 and middle left for comparison with demonstration 12). Something similar happens with demonstration 4. This demonstrates one way in which this similarity measure could fail to inform the imitator; if two objects are close to each other it is difficult to know which one was focused on.

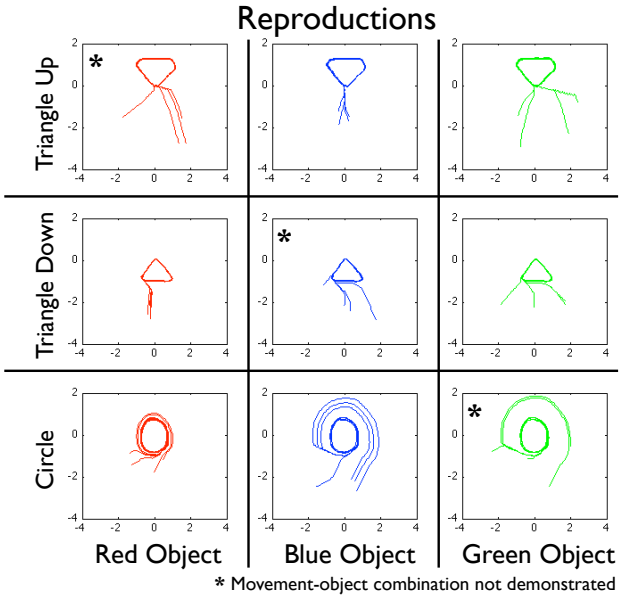

Fig. 7. Here we can see the 36 reproduction attempts. The rows indicate movement and the columns indicate coordinate system. In each case the signs given to the imitator led to correctly finding the correct data and the correct coordinate system and as we can see it went fairly well.

predictable response to a part of the context that is visible to the imitator. Another condition was that the operation resulted in a state that had a consistent influence on a policy of the demonstrator which determined actions that were observable by the imitator. We have further shown how imitating these internal operations resulted in a policy that is able to generalize correctly and results in successful reproductions in situations where there are no demonstrations.

From a linguistic point of view we have shown that, in some situations, observations of two speakers of a language can be used to find discrete communicative classes. That is; using the effect on an interlocutor of communicative acts to find how many classes the continuous behavior should be divided into (where behavior modification is only dependent on weather or not a communicative act is a member, not on its specifics). Traditionally, a communicative act is assumed to go through some recognition system and be represented as a symbol (so that finding the meaning of the symbol is the only remaining problem). The presented approach opens up for the investigation of borderline cases such as body language and facial expressions, where it might not be easy to find a recognition system transforming a continuous representation into symbols without looking at how they modify the behavior of an interlocutor/demonstrator.

One venue for future work is to use a real robot to perform the reproduction and to use a Kinect device to capture hand movements.

The grouping algorithm as presented is a batch computation but is suited for modification into an incremental version. When the demonstrations already seen have been grouped, a new demonstration can be checked for similarity with these established groups (for a new group of demonstrations those that are similar to an established group is added to it and then

the algorithm could be run on the remaining demonstrations to find the new groups).

The actual reproduction that is performed after the full model of the demonstrator is built has access to a small amount of relevant low dimensional data and several more sophisticated methods could be used, for example the well explored combination of BIC with GMM (allowing quick regression during reproduction). ILO-GMR [18] would allow the immediate incorporation of new data (if a new demonstration is close to a group of movements it can be immediately added to that group without the need to re build a model as the models are built on line).

One could add additional heuristics or information sources specific to the particular setting, such as the hand being on average closer to focused on objects, or add an estimate of what object the demonstrator is looking at to the demonstrations. In the current paper we do not use anything like this and indeed a single demonstration contains absolutely no information regarding the internal state. The imitator has no idea how a specific state of the imitator influences policy, and so a single demonstration gives no information about what internal state was. If it had such a model relating internal states to actions; it could simply infer the internal state directly and the problem would be trivial. The problem solved is to find a correlation model between internal states and observed behavior, which is very different from the standard problem of using a known correlation model to infer the exact state. The imitator builds a model of this correlation based on an assumption of consistency in the influence. Since the way in which an internal state influences actions is inferred from data, the algorithm should be usable even in situations where the structure of this influence is unknown. If, for example, the demonstrators internal state is modified in some way by the interactants tone of voice or body language and the programmers does not know how to encode the relationship between these different states and behavior (or even know how many states to use), the algorithm could in principle still be used (since the number of internal states and the correlation models are inferred).

Future work could also include learning situation-specific correlations from easy tasks, such as "the hand of the demonstrator is more likely to be close an objects that is focused on". The imitator could also learn that in some situations the internal state is correlated with the eye gaze of the demonstrator (looking at objects focused on). Being able to learn such correlations would increase autonomy and reduce reliance on the programmers predicting what correlations the imitator will find useful.

ACKNOWLEDGMENT

We would like to thank Adrien Baranes for help with the graphical presentation. This research was partially funded by ERC Grant EXPLORERS 240007

REFERENCES

[1] K. Dautenhahn and C. L. Nehaniv. The agent-based perspective on imitation, *Imitation in animals and artifacts*, pages 1-40. MIT Press, 2002.

[2] Aude Billard, Sylvain Calinon, Ruediger Dillmann, Stefan Schaal, Robot Programming by Demonstration, *Springer Handbook of Robotics.*, 2008, pp 1371-1394.

[3] B. D. Argall, S. Chernova, M. Veloso, and B. Browning, A survey of robot learning from demonstration, *Robotics and Autonomous Systems*, vol. 57, no. 5, pp. 469-483, May 2009.

[4] E. A. Billing and T. Hellstrm, A formalism for learning from demonstration, *Paladyn Journal of Behavioral Robotics*, vol. 1, no. 1, pp. 1-13, 2010.

[5] Nehaniv, C., and Dautenhahn, K., Of hummingbirds and helicopters: An algebraic framework for interdisciplinary studies of imitation and its applications, *Interdisciplinary approaches to robot learning.*,World Scientific Press, vol. 24, 2000, pp 136-161.

[6] A. Alissandrakis, C.L. Nehaniv, and K. Dautenhahn, Correspondence mapping induced state and action metrics for robotic imitation, *IEEE Transactions on Systems, Man and Cybernetics, Part B. Special issue on robot learning by observation, demonstration and imitation.*, vol. 4, 2007, pp 299-307.

[7] Komei Sugiura, Naoto Iwahashi, Hideki Kashioka, and Satoshi Nakamura, Learning, Generation, and Recognition of Motions by Reference-Point-Dependent Probabilistic Models, *Advanced Robotics*, Vol. 25, No. 5, 2011(to appear).

[8] Calinon, S. and D'halluin, F. and Sauser, E. L. and Caldwell, D. G. and Billard, A. G., Learning and reproduction of gestures by imitation: An approach based on Hidden Markov Model and Gaussian Mixture Regression, Robotics and Automation Magazine, vol. 17, 2010, pp 44-54.

[9] Yasser Mohammad and Toyoaki Nishida, Learning Interaction Protocols using Augmented Baysian Networks Applied to Guided Navigation, *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems* 2010.

[10] Tuci E., Ferrauto T., Zeschel A., Massera G., Nolfi S. (in press). An Experiment on behaviour generalisation and the emergence of linguistic compositionality in evolving robots, *IEEE Transactions on Autonomous Mental Development* 2011

[11] Ito, M., Noda, K., Hoshino, Y., and Tani, J, Dynamic and interactive generation of object handling behaviors by a small humanoid robot using a dynamic neural network model, *Neural Networks*, 19 (3), 2006, pp 323-337.

[12] Calinon, S. and Billard, A, Incremental Learning of Gestures by Imitation in a Humanoid Robot, *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI),.*, 2007, pp255-262.

[13] S. Calinon and F. Guenter and A. Billard, On Learning, Representing and Generalizing a Task in a Humanoid Robot, *IEEE Transactions on Systems, Man and Cybernetics, Part B* , vol. 37, 2007, pp 286-298.

[14] Calinon, S *Robot Programming by Demonstration: A Probabilistic Approach.*,EPFL/CRC Press , 2009.

[15] G. Schwarz, Estimating the dimension of a model, *Annals of Statistics*, vol. 6, no. 2, pp. 461-464, 1978.

[16] Duy Nguyen-Tuong and Matthias W. Seeger and Jan Peters, Real-Time Local GP Model Learning, *in From Motor Learning to Interaction Learning in Robots*, 2010, pp 193-207.

[17] Daniel H Grollman and Odest Chadwicke Jenkins, Incremental Learning of Subtasks from Unsegmented Demonstration, *International Conference on Intelligent Robots and Systems*, 2010, Taipei, Taiwan.

[18] Cederborg, T., Ming, L., Baranes, A., Oudeyer, P-Y: Incremental Local Online Gaussian Mixture Regression for Imitation Learning of Multiple Tasks, *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*.

[19] Steels L., Experiments on the emergence of human communication, *Trends in Cognitive Sciences*, 10(8), pp. 347-349, 2006.

[20] Cangelosi A., Metta G., Sagerer G., Nolfi S., Nehaniv C.L., Fischer K., Tani J., Belpaeme B., Sandini G., Fadiga L., Wrede B., Rohlfing K., Tuci E., Dautenhahn K., Saunders J., Zeschel A. Integration of action and language knowledge: A roadmap for developmental robotics. *IEEE Transactions on Autonomous Mental Development*, 2010, In press.