Social learning of skills and language

Paul Vogt[1,2] and Evert Haasdijk[1,2]

[1] Communication and Information Sciences, Tilburg University
[2] Artificial Intelligence, Vrije Universiteit Amsterdam
The Netherlands
{p.a.vogt,e.w.haasdijk}@uvt.nl

Abstract. In this paper, we explore how human-like social learning can be implemented in artificial life models. We focus on the social learning of both skills and language and we illustrate our considerations and design issues based on our developments in the NEW TIES project. We conclude that our assumptions regarding autonomy, embodiment and situatedness impose many limitations and, consequently require difficult design choices.

1 Introduction

In the past decade, many computational studies have investigated how languages can evolve in populations of autonomous agents (see, e.g., [1, 2] for overviews). Simultaneously, a lot of effort has been put in computational studies regarding the social learning of behaviours, e.g., through imitation learning [3]. In this paper, we investigate the possibility of merging these two endeavours such that a population of agents evolve a language through which they can learn skills from each other.

Social learning can be categorised in similar ways as individual learning, but where conspecific demonstrators provide the exemplars [4]. Roughly three types of learning can be distinguished: (i) learning from a single stimulus, (ii) learning from the relationship between two stimuli and (iii) learning from action-response. In social learning, the demonstrator may increase the probability that an observer attends to or interacts with a stimulus. The demonstrator may also act as an unconditioned or discriminative stimulus to provoke a matching response from the observer, or act as a model for imitation or copying [4].

Non-human animals typically learn from stimuli representing some overt behaviour. Although humans, too, learn this way, with language they have evolved the additional capability to explain to others why they act the way they do. Hence, humans – through language – have the ability to learn from each other's decision processes. This form of social learning may be much more powerful than learning from observing the behaviour of others.

In artificial life and robotics, many studies have focused on social learning at a lower level (i.e. learning directly from stimuli) [3], but no studies are known in which individuals learn from other individuals who tell them what they are doing and why. In this paper, we explore how the human explicit form of social learning can be studied in artificial life models. Although the model for social learning we explore assumes that a language exists, we also explore requirements necessary in case such a language has to emerge in the population. (Note that language development, too, is a form of social learning.) The purpose of the paper is to present and discuss problems and possible solutions arising when developing such models.

We take it as given that agents communicate. Although it is interesting to investigate how agents can learn to communicate as such (i.e., how they can learn that communication is valuable), this is beyond the scope of this paper.

This paper aims to present our approach to design and implement social learning of skills and language based on the NEW TIES project. While doing so, we also touch upon more general issues of social learning. In the next section, we introduce the NEW TIES project. Issues regarding social learning of skills are discussed in Section 3, while Section 4 discusses issues regarding social learning of language. Section 5 concludes the paper.

2 The NEW TIES project

The NEW TIES project[3] aims to set up a simulation platform in which a cultural society evolves through evolution, individual learning and social learning [5]. The artificial society lives in a grid world containing various food sources, tokens, places, building bricks and agents.

The project aims to model agents anthropomorphically, thus imposing strict autonomy, (virtual) embodiment and situatedness. This aim limits our options in designing agent interactions (e.g., agents cannot communicate unless they are within each other's vicinity), perception (e.g., they cannot see inside each other's heads) and learning mechanisms (e.g., no supervised learning).

Agents receive visual stimuli regarding the objects in their visual field and output actions that move them around, pick up or put down objects, eat, communicate or interact otherwise with others (e.g., mating, or giving or taking objects to/from other agents). Actions of agents are decided upon by their individual controllers, which are implemented as *decision Q-trees* (DQTs) that can adapt through evolutionary, individual or social learning.

A DQT is a decision tree that contains test nodes, bias nodes and action nodes (Fig. 1). The test nodes test whether the agent has categorised a concept from its current context, which includes visual stimuli, objects it may carry, internal states, and interpretations of communication acts. The concepts are formed from one or more categories that relate to some feature of an object, such as *shape*, *colour*, *direction*, *sex* or *action*. If the test on a test node succeeds, the agent traverses to the next node in the left branch; otherwise it traverses to the next node in the right branch.

The bias node allows the agent to select from a multiple of branches based on a specific bias towards some attitude, such as socialness, agressiveness or

[3] NEW TIES stands for New Emerging World models Through Individual, Evolutionary and Social learning. http://www.new-ties.org.

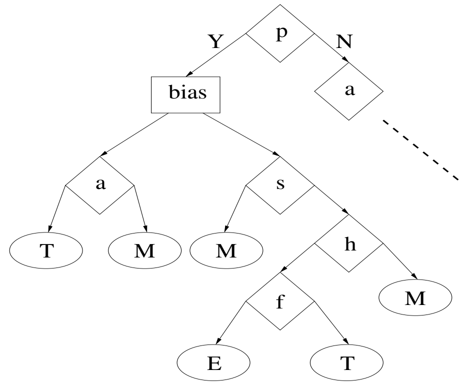

Fig. 1. A simplified example of a decision Q-tree (DQT). The diamonds represent test nodes, the rectangles bias nodes and the ovals action nodes.

food-mindedness. These biases are determined genetically through evolution and ontogenetically through individual learning. The attitudes are related to certain actions. For instance, socialness is related to talking, agressiveness to hitting and food-mindedness to eating. The biases, which are values between 0 and 1, determine the probability with which a certain branch is visited in the tree.

The leaves of the DQT are action nodes, which include simple actions, such as move, turn-left or turn-right, first order predicates such as eat(x) or hit(y), and second order predicates such as give(a,o). The arguments can be any object, but if, e.g., an agent attempts to eat a non-food item, this action will fail in the world.

The agent traverses the DQT until an action node is reached. The agent then performs the action, which will cost a certain amount of energy. When the agent's energy drops below zero, the agent dies, which may also happen if the agent reaches a certain age.

Agents can mate with agents of the opposite sex, thus producing offspring. The offspring is produced with a genome that is the result of cross-over between the genomes of its parents and some mutation like in standard genetic algorithms. The genome, which does not change during lifetime (so evolution is non-Lamarckian), carries information regarding both the initial structure of the DQT and the genetic biases. The selection criterion for mating is not following any standard genetic algorithm and does not use an explicit fitness function. The agent can mate at any time (provided they are near to each other), so evolution is asynchronous. The longer an agent lives, the more offspring it can produce, thus the agent's fitness is implicit.

Individual learning is achieved by reinforcement learning. Agents can insert or delete nodes in the DQT and change the learnt biases. Each node is assigned a Q-value, which alters based on rewards (primarily energy gain or loss). Inserting new nodes can be random, but may also be based on information received from other agents, thus amounting to social learning, which we will explain later.

Simulations in NEW TIES are challenge-driven, which means that world scenarios are designed that pose a problem for the population's survival. For instance, the world may contain two types of food: edible and non-edible plants and in order for the agents to survive they have to learn that eating the non-edible poisonous plants is not good.

In the remainder of this paper, we will ignore evolutionary and individual learning. Instead, we will focus our attention on social learning of skills and language.

3 Social learning of skills

We consider social learning (SL) as an adaptation method that operates at the same level as evolutionary learning (EL) and individual learning (IL), i.e. it operates on the same search space [5]. Agents use SL to exchange knowledge pieces (behaviour rules), thus learning from each other. Other aspects of SL can be considered, such as observing the behaviour of other individuals and imitating that (e.g., only eating what other conspecifics eat) or exchanging information about the state of the environment and using that to make your own decisions (e.g., telling another individual where to find food). However, for the purpose of this paper, we will restrict our discussion to the SL of behavioural rules that are transmitted explicitly through communication.

As mentioned, the agents' controllers are decision trees (DQT), inherited (for the initial population: generated) at birth. Through EL, only the inherited controller is passed on (i.e. non-Lamarckian); agents do not inherit knowledge that their parents may have gained through experience (e.g., modifications to the controller due to IL); they can only recombine the controllers that their parents had at birth (with some mutation added). This means that, without some additional method of spreading the knowledge through the population of agents, everything an agent learns through experience (IL) will be lost when it dies.

This is where social learning comes into play: anything an agent learns during its lifetime can be taught to other agents, so this knowledge does not necessarily die with the agent that originally discovered it. With agents exchanging knowledge pieces through SL, the population as a whole becomes a knowledge reservoir – not only for EL, but also for IL. Social learning does not, however, distinguish between learnt and inherited knowledge, so it can function even without individual learning.

In combination with EL we expect faster convergence with EL maintaining diversity (by clinging to only genetic modifications). Our hope and expectation is that, IL and SL combined will provide mutual leverage, enabling the spread of individually acquired knowledge and turning the population into a repository for acquired as well as innate knowledge.

With SL in place, the population as a whole will have characteristics similar to that of an evolutionary system, even without EL - think, e.g., of 'memes' [6] or cultural evolution [7]. Moreover, although SL does not use an explicit fitness function as in standard genetic algorithms, population selection works similar to Darwinian selection: ill-adapted individuals tend to die, hence cannot further distribute their knowledge, while well-adapted individuals tend to survive and have more opportunities to distribute their knowledge. In addition, mechanisms are imaginable that can improve adaptation, e.g., by letting individuals choose to learn from fit conspecifics rather than from unfit ones.

When designing algorithms on social learning, some considerations need to be taken into account. The most crucial considerations concern

1. the type of information that agents pass on,
2. the decision mechanism of agents to decide who to learn from, and
3. the mechanism with which new knowledge is incorporated.

In our model, the type of information exchanged should be either the whole controller (DQT) or parts thereof. Exchanging the whole DQT would allow the receiver to incorporate a lot of knowledge at once, which may be good or bad, but from a natural point of view this is highly unrealistic; we cannot exchange all the contents of our brains, even if we wish to, simply because the bandwidth of our communication channel (the evolved language) is too narrow.

So agents will exchange parts of the DQT and therefore an agent has to decide which part to exchange. If the agent has some notion of the quality of parts of its DQT (e.g., the Q-values of nodes), it can decide to use a particularly useful (or often used) sub-tree. Since communicating sub-trees successfully requires a complex syntax and since we intend to use SL together with an evolved language whose syntax is not likely to be sufficient, we opt to have agents exchange (parts of) a single path. In particular, agents will exchange knowledge concerning the path they traversed leading to the action node at that time step. In addition, since the language will only be able to express concepts that imply a test node, the transfers would be of the form [hungry, food, eat] or [no food, move].

When an agent receives a fragment of knowledge, it decides if it will discard or accept it, based on

1. its natural proclivity to learn socially (implemented as the socialness bias),
2. the relative fitness of the sender (e.g., as a function of energy, age or perceived DQT utility), and
3. (at a later stage) the level of trust or social ties of the other agent.

The idea is that agents who are evolutionary more biased to being social are more likely to learn from other agents. This would allow us – at a later stage – to investigate some interesting interactions between EL and SL. If SL is good for the population, it is expected that the socialness gene yields a high bias.

Agents can determine the fitness (e.g., through age and energy level – both observable as discrete values such as old/young or high/low) of the source of a particular piece of knowledge. The agent can use that to estimate the usefulness of that piece. This would be even better if a path were passed to an agent along with some measure of its utility (e.g., Q-value).

Agents can keep track of their social network, allowing them to determine which agents they interact with on a frequent basis. It would make sense if agents are more inclined to learn from other agents in their social group, thus allowing forms of group selection to emerge.

The considerations regarding the incorporation of knowledge focus on integrating a received DQT path into an agent's own DQT. We considered three options:

1. De- and recomposition of decision trees using Fourier spectrum cf. [8].
2. Analysing/synthesising the partitions of the state-action space, or
3. Listing all possible paths for the DQTs and then comparing/combining these.

The latter option seems most promising as it is most amenable to DQTs – it is unclear how the first two options should deal with bias nodes. When incorporating a received DQT path, agents select the most similar path in their own DQT and insert the received path as an alternative. This is achieved by inserting a bias node that branches into the original and the newly inserted paths.

The current state of the project is that the above mechanisms have been implemented and preliminary experiments have been carried out. In these experiments, the knowledge is exchanged through *explicit meaning transfer*, i.e. by sending the contents of test nodes directly. In the near future we intend to have the knowledge exchanged through an intermediate language that the agents evolve by engaging in language games, as will be discussed in the next section.

The preliminary results seem to indicate that SL combined with EL outperforms EL only. In addition, we carried out simulations where a small proportion of the population was pre-designed with controllers that could solve the problem at hand (teachers) and the rest was initialised with a random controller (learners). The results revealed that the pre-designed knowledge spread through the entire population rapidly.

4 Social learning of language

Social learning of language is based on Steels' *language game* model [9] that allows a population of to develop (or evolve) a language from scratch. The language game model was adapted to fit the quasi-realistic environment of our project [10]. In our model, whose details are beyond the scope of this paper, agents can at each moment initiate a language game by producing an expression, which is received by one or more agents (a.k.a. hearers) within a certain range.

The speaker (i.e. the initiator) chooses a set of concepts to communicate and tries to encode an expression conveying this set using its acquired private lexicon that contains word-meaning associations. In turn, the hearers try to decode this broadcasted message into a set of concepts, which may or may not be the same.[4] When an agent's private lexicon is insufficient for encoding or decoding an expression, the agent can, respectively, invent or adopt a new word-meaning association; thus expanding its lexicon. (Note that at the agent's birth, its lexicon is empty.) Using a number of heuristics described in [10], agents adapt the weights of word-meaning associations, thus aligning their lexicons such that a shared lexicon can emerge.

This alignment is primarily based on *cross-situational learning* [12–14] in which associations of words with meanings are reinforced when both occur in the present context. The success of learning, thus depends on the size of the context [15] and on the consistency with which agents name a particular meaning [16]. In our model, the context size may be reduced using pointing gestures that identify the object – not the exact meaning – about which the speaker communicates. In addition, associations may be reinforced or weakened using feedback signals based on agents' self-evaluations of success. However, in [10] it was shown that this did not contribute much to the success of emergence.

Since we assume full autonomy, (virtual) embodiment and situatedness of the agents, the problems they face are similar to those faced by real robots (see, e.g., [2, 14] for discussions). In our project, concepts relate to properties of objects, such as shape, colour, direction, distance, weight, sex, energy level or actions (the latter three only apply to agents). Developing a shared lexicon regarding observer invariant properties, such as shape and colour, is relatively straightforward. However, words about observer dependent properties, such as direction and distance, are much harder to align, because what is left of me, need not be left of you. This depends on the relative locations and headings of the agents participating in the language game, who cannot be at the same location at the same time.

This fact, which is typically ignored in many simulations on language evolution (e.g., [17, 13, 16]), also has the consequence that agents do not tend to see the same objects. Not taking into account what the other can or cannot see, may cause the speaker to communicate something that the hearer does not see. This in itself may not be harmful if the hearer knows the meaning of the word (it may even be beneficial when the hearer learns, e.g., the location of a food source), but when it still needs to learn the meaning of the word (be it by adopting the association or aligning its weight), this poses problems. This is especially the case, because cross-situational learning assumes that words are about meanings that are present in the current context. Combined, these problems regarding spatial relations and different visual stimuli, had the effect that in the simulations reported in [10] yielded a *communicative accuracy* hardly exceeding 60%. (Communicative accuracy is the rate with which hearers can successfully decode the words expressed by speakers – i.e., they retrieve the same concepts intended by speakers – averaged over 50 time steps.)

Recent tests with only 2 agents in which they only communicated about concepts concerning observer independent properties yielded communicative ac-

[4] For now we assume these concepts are predefined, though these could be constructed ontogenetically using the *discrimination game* framework [11].

curacy above 90%. Simulations that also included observer dependent properties yielded a success of around 50%. Hence, it seems that mismatches regarding spatial relations have a larger effect on the success of language development than mismatches in objects agents see (i.e. mismatches in context). To overcome this problem, we are currently implementing methods in which agents can take perspective, which have proven to be successful in robotic experiments [18].

It is necessary to take perspective because the required autonomy and embodied does not allow agents to have access to each other's stimuli, nor do they have a common absolute frame of reference. Taking perspective can be implemented using straightforward triangulation and from two sides: the speaker taking perspective of the hearer or the hearer taking perspective of the speaker. (Note that this requires that agents can see each other, which is a rare event with their current controllers not dedicated to communication, but to random movement and foraging.) In addition, an agent can take perspective of the other for two reasons: 1) to transform spatial relations and 2) to determine which objects the other agent can see. Since the objective is to communicate about parts of an agent's own decision tree (see previous section), it is undesirable to have the speaker change the concepts regarding spatial relations as these make up its controller. It is better to have – as a convention – the hearer change perspective, so interpreted concepts can directly be applied to change its own DQT. It may, however, be desirable to have the speaker take perspective in order to know what the hearer can see, especially when it is aware that the hearer may not know the word. This awareness can be the result of a feedback signal sent by the hearer, but may also be inferred, e.g., in case the hearer is a young agent or was never seen before.

In the simulations of [10], agents only communicated about properties of visible objects, whereas for SL the agents require to communicate about paths through their DQTs. Such paths do include such properties, but additionally contain internal representations and actions. The latter are important to communicate as these form the consequence of the rules to be learnt. The speaker expresses its own action, which is visible to other agents.

Since paths in the DQTs may become relatively large (e.g., larger than 10 nodes), expressing a complete path would only be feasible once both agents have acquired (nearly) all necessary word-meaning pairs. Otherwise, learning the proper meaning becomes too hard as the uncertainty becomes too high. Suppose you hear the speaker communicating [see-agent, has-food, am-hungry, take, food, agent] (meaning something like *if see an agent who has food and I am hungry, then I take this food from the agent*), but you only know three words. Then you have to learn three novel words at once and all could potentially mean many different things. Using a pointing gesture to reduce the learning context may not be a solution, as the novel words may relate to different objects.

Two solutions are possible here: First, it is possible to allow agents to engage in rich multimodal interactions such that hearers can, for instance, request clarifications of words, which are then (if necessary one by one) provided by the speaker accompanied by pointing gestures. This allows agents to fill in gaps in their knowledge, but requires rather sophisticated interaction protocols, which may include aligning their visual fields. The latter may also be achieved by letting agents construct a cognitive map of their surroundings. Second, and more straightforward, is letting agents communicate short (e.g., one word) utterances to young or unfamiliar agents and longer utterances to older or more familiar agents. We decided to start testing with the straightforward method, and if this does not yield satisfactory results, we will move on to incorporate more sophisticated dialogues.

Social learning of skills, as described in the previous section, and currently implemented with explicit meaning transfer poses an upper limit to what can be achieved using an acquired language as described in this section. A learnt language leaves room for misunderstanding and ambiguity, especially during its acquisition, thus limiting the accuracy of knowledge transfer. Social learning of skills will have to deal with this inherent inaccuracy. For instance, agents could assign some level of confidence to the interpretation of received messages when they consider incorporating the knowledge they contain.

5 Conclusions

In this paper we explore how we can design an artificial life model of social learning of both skills and language. Regarding skills, we focus – unlike many previous studies [3] – on social learning where the stimuli come from verbal transmission of rules that make up agents' controllers. This, thus, implements a human-like form of social learning. Most methods presented have been implemented and are currently being tested in preliminary experiments.

The strict application of autonomy of the agents, their virtual embodiment and situatedness impose a number of restrictions to the model. We restrict agents to use language-like communication for social learning of skills, which reduces the bandwidth for knowledge transfer. In addition, inaccuracies in language learning reduce the efficiency of knowledge transfer. These inaccuracies are exacerbated by the fact that agents have different frames of reference. To overcome these differences, agents need to align perspectives.

We are currently working on integrating the social learning of skills and language, such that the learnt language will be used as the medium for knowledge transfer. Once this is achieved, it is possible to start investigating how the population can learn to communicate and ultimately learn to learn. The NEW TIES platform does allow for such experiments and is publicly available.[5]

Acknowledgements

The NEW TIES project is supported by an EC FET grant under contract 003752. We thank all members of the project for their invaluable contributions. Opinions and errors in this manuscript are the authors' responsibility; they do not necessarily reflect those of the EC or other NEW TIES members.

[5] http://www.new-ties.org.

References

1. Cangelosi, A., Parisi, D., eds.: Simulating the Evolution of Language. Springer, London (2002)
2. Vogt, P.: Language evolution and robotics: Issues in symbol grounding and language acquisition. In Loula, A., Gudwin, R., Queiroz, J., eds.: Artificial Cognition Systems, Idea Group (2006)
3. Nehaniv, C.L., Dautenhahn, K., eds.: Imitation and Social Learning in Robots, Humans and Animals. Cambridge University Press, Cambridge (2007)
4. Heyes, C.M.: Social learning in animals: Categories and mechanisms. Biological Reviews **69** (1994) 207–231
5. Gilbert, N., den Besten, M., Bontovics, A., Craenen, B., Divina, F., Eiben, A., Griffioen, R., Hévézi, G., Lörincz, A., Paechter, B., Schuster, S., Schut, M., Tzolov, C., Vogt, P., Yang, L.: Emerging artificial societies through learning. Journal of Artificial Societies and Social Simulation **9(2)** (2006)
6. Dawkins, R.: The Selfish Gene. Oxford University Press, Oxford (1976)
7. Boyd, R., Richerson, P.J.: Culture and the Evolutionary Process. University of Chicago Press, Chicago (1985)
8. Park, B., Ayyagari, R., Kargupta, H.: A Fourier analysis based approach to learning decision trees in a distributed environment. In: Proceedings of First SIAM International Conference on Data Mining. (2001)
9. Steels, L.: Emergent adaptive lexicons. In Maes, P., ed.: From Animals to Animats 4: Proceedings of the Fourth International Conference On Simulating Adaptive Behavior, Cambridge Ma., The MIT Press (1996)
10. Vogt, P., Divina, F.: Social symbol grounding and language evolution. Interaction Studies **8(1)** (2007) 31–52
11. Steels, L.: Perceptually grounded meaning creation. In Tokoro, M., ed.: Proceedings of the International Conference on Multi-Agent Systems, Menlo Park Ca., AAAI Press (1996)
12. Siskind, J.M.: A computational study of cross-situational techniques for learning word-to-meaning mappings. Cognition **61** (1996) 39–91
13. Smith, A.D.M.: Establishing communication systems without explicit meaning transmission. In Kelemen, J., Sosík, P., eds.: Proceedings of the 6th European Conference on Artificial Life, ECAL 2001. LNAI 2159, Berlin Heidelberg, Springer-Verlag (2001) 381–390
14. Vogt, P.: Bootstrapping grounded symbols by minimal autonomous robots. Evolution of Communication **4(1)** (2000) 89–118
15. Smith, K., Smith, A., Blythe, R., Vogt, P.: Cross-situational learning: a mathematical approach. In Vogt, P., Sugita, Y., Tuci, E., Nehaniv, C., eds.: Symbol grounding and beyond, Springer (2006)
16. Vogt, P., Coumans, H.: Investigating social interaction strategies for bootstrapping lexicon development. Journal for Artificial Societies and Social Simulation **6(1)** (2003) http://jasss.soc.surrey.ac.uk.
17. Kirby, S., Smith, K., Brighton, H.: From UG to universals: linguistic adaptation through iterated learning. Studies in Language **28**(3) (2004) 587–607
18. Steels, L., Loetzsch, M.: Perspective alignment in spatial language. In Coventry, K.R., Tenbrink, T., Bateman, J.A., eds.: Spatial Language and Dialogue. Oxford University Press (2007)