LEARNING MORE FROM LESS DATA: EXPERIMENTS WITH LIFELONG ROBOT LEARNING

Sebastian Thrun and Joseph O'Sullivan[1]

1 Introduction

Designing robots that learn by themselves to perform complex real-world tasks is a still-open challenge for the field of Robotics and Artificial Intelligence. While powerful and very general learning techniques readily exist that—at least in principle—enable robots to learn complex tasks, most of them suffer from their enormous sample complexity, a limitation that prohibits their application in most realistic robot applications. Learning more accurate functions from less data is thus a key issue in robot learning.

This paper investigates robot learning in a *lifelong learning framework*. In lifelong learning, the learner faces an entire collection of learning tasks, not just a single one. Thus, it provides the opportunity for synergy among multiple tasks. To obtain this synergy, the central question in lifelong learning is: How can a learner re-use knowledge gathered in its previous $n-1$ learning tasks to reduce the sample complexity in its n-th? In other words: How can the learner transfer knowledge across multiple tasks?

Recent research has produced a variety of methods that transfer knowledge across learning tasks (see literature review in [6]). Such methods learn domain-specific knowledge, and use it to guide generalization in subsequent learning tasks. However, these methods weigh previous learning tasks equally strongly—thus, they may fail when only a small subset of learning tasks is related appropriately. In this paper we describe a *selective* approach to lifelong learning, the *TC algorithm* (short for *task clustering*), which we just have begun to explore in our lab [7]. TC transfers knowledge across multiple tasks by adjusting the distance metric in nearest neighbor generalization. To increase robustness to unrelated tasks, TC arranges all learning tasks hierarchically. When a new learning task arrives, TC relates it to the task hierarchy, in order to transfers knowledge selectively from the most related tasks only. As a result, TC is more robust than its unselective counterpart (which would transfer knowledge from all available tasks). Thus far, TC has been successfully applied to perception tasks involving visual and ultrasonic input, using our mobile robot XAVIER[1].

[1]The authors are affiliated with Carnegie Mellon University, Computer Science Department, Pittsburgh, PA 15213-3891, USA

2 The TC Algorithm

The TC algorithm has been designed to support fast learning of large sets of similar classification tasks:

1. **Nearest neighbor generalization.** At the underlying function approximation level, the TC algorithm uses nearest neighbor for generalization (see e.g. [5]).

2. **Adjusting the distance metric.** TC uses a *globally weighted Euclidean distance metric*:

$$dist_d(x,y) \;=\; \sqrt{\sum_i d^{(i)}\left(x^{(i)}-y^{(i)}\right)^2}.$$

   Here d denotes an adjustable vector of weighting factors. TC transfers knowledge across tasks by adjusting d for some tasks, then re-using it in others. This is done by minimizing the distance between training examples that belong to the same class, while maximizing the distance between training examples with opposite class labels:

$$E(d) \;=\; \sum_{x,y} \delta_{xy}\; dist_d(x,y)$$

   where

$$\delta_{xy} \;=\; \begin{cases} 1 & \text{if } \mathit{class}(x)=\mathit{class}(y) \\ -1 & \text{if } \mathit{class}(x)\neq\mathit{class}(y) \end{cases}$$

   Notice that the idea of tuning a distance metric is well-established (see e.g. [1, 4, 2, 3]). TC applies these ideas to the transfer of knowledge (concerning the importance of input features) across multiple learning tasks.

3. **Estimating the task transfer matrix.** Obviously, using the E-optimal distance metric obtained for one task when learning another task will only improve the results when both tasks demand a similar feature weighting. To determine the degree to which tasks are related to each other, TC computes the *task transfer matrix*

$$C \;=\; (c_{n,m})$$

[1]World Wide Web: http://www.cs.cmu.edu/∼Xavier

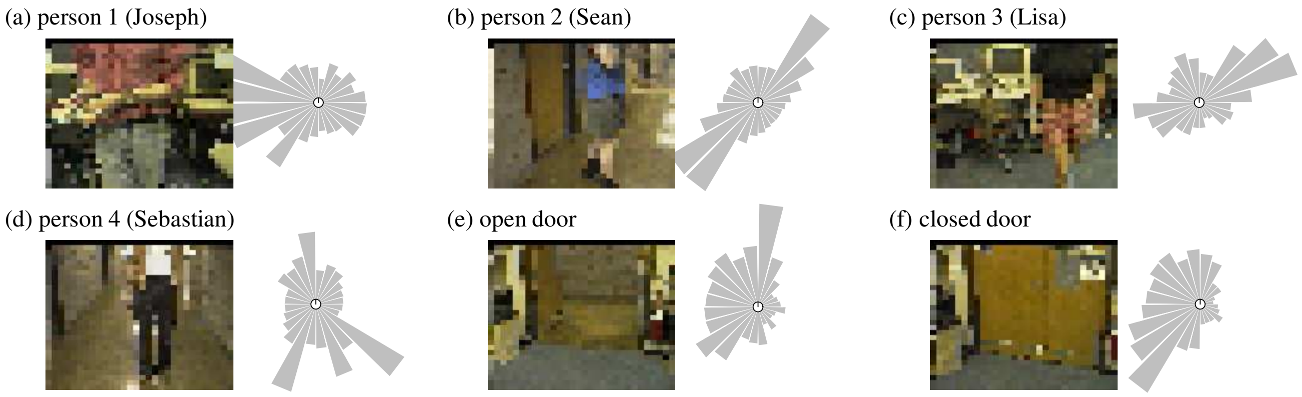

Figure 1: Examples from six databases (images and sonar scans).

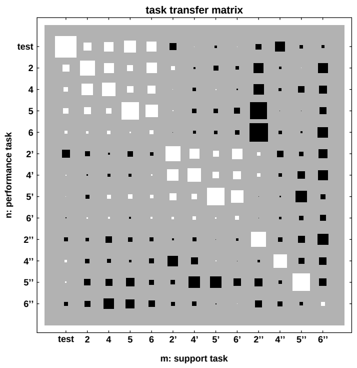

Figure 2: Task transfer matrix $(c_{n,m})$. White (black) values indicate that the error in task n decreases (increases) when the E-optimal distance metric of task m is used.

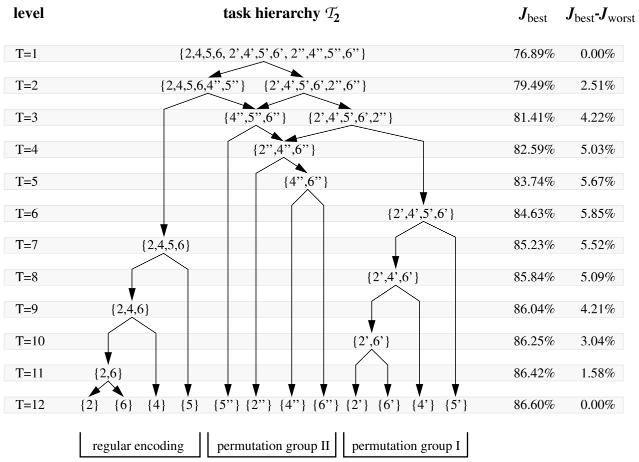

Figure 3: Task hierarchy. Notice that early on, the task hierarchy separates the three different task types. The related task cluster, $\{2, 4, 5, 6\}$, is identified when $T \geq 3$ clusters are available.

which contains a value $c_{n,m}$ for each pair of learning tasks n and m. The value $c_{n,m}$ is the *expected generalization accuracy* for task n when using m's E-optimal distance metric. Each element $c_{n,m}$ is estimated via k-fold cross-validation.

4. **Clustering learning tasks.** TC clusters all N learning tasks into $T \leq N$ bins, denoted by $A_1, \ldots, A_T$, by minimizing

$$J = \frac{1}{N} \sum_{t=1}^{T} \sum_{n \in A_t} \frac{1}{|A_t|} \sum_{m \in A_t} c_{n,m}.$$

J measures the averaged estimated generalization accuracy that is obtained when task $n \in A_t$ uses the E-optimal distance metrics of another task $m \in A_t$ in the same cluster. In other words, minimizing J groups those tasks together that are most related.

5. **Constructing the task hierarchy.** By repeating the clustering process for different values of T, TC constructs a hierarchy of task clusters (see example in the next section). Notice that each task cluster defines its own E-optimal distance metric (Step 2 above).

6. **Transfer to novel tasks.** When a new task arrives, TC first determines the most related task cluster, then uses the E-optimal distance metric of that task cluster for nearest neighbor generalization in the new task. In other words, TC selects the most related cluster of tasks and transfers knowledge from this cluster only.

3 Example and Results

Figure 1 shows a database collected using our mobile robot XAVIER. Here the "new" task involves distinguishing open from closed doors (Figure 1e,f). TC can exploit distance metrics learned in 12 previous learning tasks, all of which involve the recognition of people. Four of those tasks use the same encoding as the door status recognition tasks (hence are well-related), whereas the other eight tasks use different encodings (they are unrelated; in fact, transfer from those hurts the performance. See [7] for details). The task transfer matrix is shown in Figure 2. Each entry in this matrix indicates the effect of transferring knowledge from a task n to another task m. The first line corresponds to the "new" task. Here white boxes indicate an increase

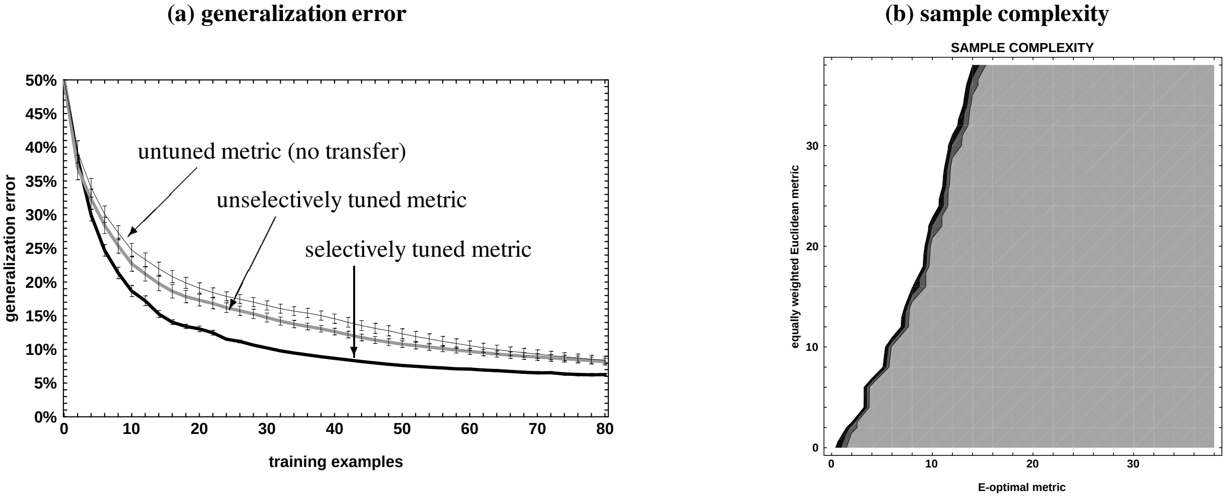

Figure 4: Performance results. (a) Error as a function of training examples. (b) Reduction in sample complexity. All results are averaged over 100 experiments. The bars indicate confidence intervals.

in generalization accuracy, whereas black boxes indicate a decrease. As can be seen in Figure 2, the testing tasks benefits most from the tasks called 2, 4, 5, and 6, which are the tasks with the same input encoding. Figure 3 depicts the entire task hierarchy, obtained by clustering the task space using different numbers of clusters T. Notice that TC discovers the different input encodings, which are found in different branches of the hierarchy.

Figure 4a shows performance results, comparing (1) regular nearest neighbor without transfer (thin curve in Figure 4a), (2) unselective transfer not using the task hierarchy (grey curve), and (3) selective transfer in TC (thick curve). Obviously, unselective transfer performs only slightly better than plain nearest neighbor, basically because most tasks are unrelated. If knowledge is transferred selectively, nearest neighbor generalizes much more accurately from less data. The reduction in sample complexity is captured in 4b, which shows the result of a series of statistical comparisons of this error obtained when using the regular vs. the TC distance metric, for varying numbers of training examples. In the grey area the regular weighted (E-optimal) distance metric is superior at the 95% confidence level. In the dividing dark region between both methods generalize about equally well. As the main result, TC requires only approximately a third of the training examples required without transferring knowledge.

4 Summary

In lifelong learning, the learner faces an entire collection of related learning tasks. This paper shows how hierarchical structure can be discovered in the space of learning tasks, and how it can be used to selectively transfer knowledge to other, new learning tasks, in order to boost generalization. The results indicate that (a) more can be learned from less data if knowledge is transferred from previous learning tasks, and (b) selectively transferring knowledge is superior to unselective transfer—unless, of course, *all* tasks are indeed appropriately related to the new learning task.

These results are well in tune with other results obtained in robot perception, robot control and game playing domains [6], which illustrate that a lifelong learner can generalize more accurately from less data if it transfers knowledge acquired in previous learning tasks.

References

[1] C. A. Atkeson. Using locally weighted regression for robot learning. In *Proceedings of the 1991 IEEE International Conference on Robotics and Automation*, pages 958–962, Sacramento, CA, April 1991.

[2] J. H. Friedman. Flexible metric nearest neighbor classification. November 1994.

[3] T. Hastie and R. Tibshirani. Discriminant adaptive nearest neighbor classification. Submitted for publication, December 1994.

[4] A. W. Moore, D. J. Hill, and M. P. Johnson. An Empirical Investigation of Brute Force to choose Features, Smoothers and Function Approximators. In S. Hanson, S. Judd, and T. Petsche, editors, *Computational Learning Theory and Natural Learning Systems, Volume 3*. MIT Press, 1992.

[5] C. Stanfill and D. Waltz. Towards memory-based reasoning. *Communications of the ACM*, 29(12):1213–1228, December 1986.

[6] S. Thrun. *Explanation-Based Neural Network Learning: A Lifelong Learning Approach*. Kluwer Academic Publishers, Boston, MA, 1996. to appear.

[7] S. Thrun and J. O'Sullivan. Clustering learning tasks and the selective cross-task transfer of knowledge. Technical Report CMU-CS-95-209, Carnegie Mellon University, School of Computer Science, Pittsburgh, PA 15213, November 1995.