Title
Can Symbolic Algorithms Model Cognitive Development?

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 18(0)

Author
Ling, Charles X.

Publication Date
1996

Peer reviewed

Can Symbolic Algorithms Model Cognitive Development?

Charles X. Ling
Department of Computer Science
The University of Hong Kong, Hong Kong
ling@cs.hku.hk
(On leave from The University of Western Ontario)

Abstract

Symbolic decision-tree learning algorithms can provide a powerful and accurate transition mechanism for modeling cognitive development. They are valid alternatives to connectionist models.

Symbolic Decision-Tree Learning

In general, a learning algorithm is "symbolic" if it directly operates on a symbolic representation during the learning period, and generates a data processing structure based directly on the symbolic representation. Additionally, one might wish to require that attributes describing training examples should be specified at a symbolic level that corresponds directly to symbolic descriptions in the learning domain. According to this definition, decision-tree learning algorithms, such as ID3 (Quinlan, 1986) and C4.5 (Quinlan, 1993), operating on letters and phonemes in modeling past tense acquisition of English verbs is symbolic (Ling & Marinov, 1993; Ling, 1994). Similarly, decision-tree learning operating on descriptive information based on balance scale problems is symbolic (Schmidt & Ling, 1996a, 1996b). On the other hand, a decision tree which utilizes "sub-symbolic" information (i.e., distributed representation) may not be called symbolic. Connectionist models of these same tasks are not either.

Because of continuous weight modification and/or gradual structure change (i.e., generative connectionist algorithms) during learning, connectionist methods are often regarded as a natural approach to modeling developmental process of cognition, and often perceived as superior to their symbolic counterparts. Our work demonstrates that symbolic decision-tree learning algorithms can provide a powerful and accurate transition mechanism for modeling cognitive development. Therefore, symbolic learning methods and models are valid alternatives to connectionist approaches.

Incremental and Developmental

There are two independent dimensions on which we can classify decision-tree learning for modeling development. One is called *incremental*, and the other is *developmental*. These combine to produce four possible classes of model: {incremental, non-incremental} × {developmental, non-developmental}. Note that this classification of model is applicable to other learning algorithms as well.

Incremental learning (a term often used in the machine learning community) requires that the algorithms take training examples one at a time, and only update the decision trees when new examples are received (rather than constructing a new one based on a single expanded set of examples). Although the well-known decision-tree learning algorithms ID3 and C4.5 are non-incremental, there are incremental versions, such as ID5 (Utgoff, 1989), that are guaranteed to produce the same results as the non-incremental versions, if applied to the identical set of examples. Because of this, and because of high computational efficiency (taking just minutes on datasets of thousands of examples), non-incremental decision-tree learning algorithms are often preferred in developmental modeling *for the ease of implementation* (Ling & Marinov, 1993; Schmidt & Ling, 1996a).

Incremental learning is desirable for constructing detailed, internal process of learning and development. However, incrementality plays little role in demonstrating end results and model predictions.

Developmental learning is to construct a series of decision trees in which successors expand upon predecessors, based on a fixed set of training examples (which may have accumulated incrementally). Early decision trees in the series are small, and can only accommodate a relatively small number of examples. The decision trees late in the series are larger, and the overall error rates are generally reduced. There are several possible methods for decision-tree developmental modeling. One obvious approach is to limit the maximum depth of the decision trees constructed, and to increase this limit gradually This method produces trees of uniform depth, but the error rates associated with different leaves are uneven. Another approach is to limit the number (or percent) of errors that leaves can tolerate, and to decrease that limit gradually. In this case, errors at different leaves would be uniform, and parts of trees with higher errors would be expanded more deeply first. A simple way to accomplish this with C4.5 is to manipulate a built-in parameter m, which has the effect of limiting the number of errors allowed in leaves. Note that one can easily combine these strategies. For example, one could gradually decrease the percent of errors allowed on leaves, and if the growth of the next decision tree is abrupt, control the maximum depth of the trees and produce several trees with an increasing depth.

We propose the following hypothesis for modeling cognitive development with developmental decision-tree learning. Early in development, children have limited mental abilities, which include learning capacity and memory capacity: they cannot explain a large number of examples observed, and at the same time, their memory for storing learned hypothesis is limited. Young children's poor performance can be modeled with the small decision trees early in the series by developmental decision-tree learning algorithms. Note that since most decision-tree algorithms (such as C4.5) choose the most discriminant attribute as the root of (sub)trees, such small decision trees are likely the *best* small trees that get as much regularity out of the training examples as possible. This reflects that children maximize learning, even with their limited mental capacity. Late in development, more examples can be accommodated (learning capacity is improved) as more complex regularities are learned (and memory capacity is increased). This can be modeled by large decision trees late in the series generated by the developmental strategies discussed above.

We have successfully used the developmental strategies discussed to model cognitive development on the balance-scale task (Schmidt & Ling, 1996a, 1996b). The set of C4.5 decision-trees demonstrated the major psychological phenomena (orderly stage progression, U-shaped development, and torque difference effect) observed in children, thereby providing evidence that C4.5 can act as a transition mechanism for modeling developmental phenomena.

Competence and performance models

Incremental and developmental decision-tree learning algorithms still lack certain features as development models. For example, decision-tree learning algorithms are deterministic, and they fail to produce individual variations. In the balance scale model, for example, stage skipping, regression, and individual differences, which mark the human developmental data, did not exhibited in C4.5 modeling. It appears too "precise", producing perfect stage progressing each run.

To answer this criticism, we need to first review the difference between *competence* and *performance models* (Chomsky, 1968). Competence is the ability of an idealized subject to execute the task at hand. This ideal is not affected by situational variables, memory span, or perceptual limitations. In reality, competence is revealed only indirectly through a subject's performance, which is always influenced by situational factors. The development model on the balance scale task using C4.5 is intended as a competence model. It is more concerned with characterizing the knowledge structures and the learning process underlying human performance.

Simple measures could be taken to augment the competence model to make it accountable for the intricacies of the human performance data. For example, random sampling can be taken from the entire training set. In addition, one can add probability in building decision trees (by not always choosing the best nodes). This added variability would still clearly be based on the data, and can model human performance (with non-deterministic behavior, ignorance, and individual differences).

Summary

Symbolic decision-tree learning algorithms are incremental and developmental, and are adequate for competence modeling of cognitive development. They can also be augmented to reflect individual human performance.

Acknowledgments

I would like to gratefully thank Billy Schmidt for contributions in the development of the ideas presented here, and for his careful reading and feedbacks on this paper.

Reference

Chomsky, N. (1968). *Language and Mind.* New York: Harcourt, Bruce and World.

Ling, C. X. (1994). Learning the past tense of English verbs: the Symbolic Pattern Associator vs. connectionist models. *Journal of Artificial Intelligence Research, 1*, 209 – 229.

Ling, C. X., & Marinov, M. (1993). Answering the connectionist challenge: a symbolic model of learning the past tense of English verbs. *Cognition, 49*(3), 235–290.

Quinlan, J. (1986). Induction of decision trees. *Machine Learning, 1*(1), 81 – 106.

Quinlan, J. (1993). *C4.5: Programs for Machine Learning.* Morgan Kaufmann: San Mateo, CA.

Schmidt, W. C., & Ling, C. X. (1996a). A decision-tree model of balance scale development. To appear in *Machine Learning.*

Schmidt, W. C., & Ling, C. X. (1996b). A symbolic model of cognitive transition: The balance scale task. In *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society.* (This volume).

Utgoff, P. E. (1989). Incremental induction of decision trees. *Machine Learning, 4*, 161 – 186.