

Título: ESTIMATING DIFFICULTY OF LEARNING ACTIVITIES IN DESIGN STAGES: A NOVEL APPLICATION OF NEUROEVOLUTION

Nombre: Gallego Durán, Francisco José

Universidad: Universidad de Alicante

Departamento: Ciencia de la computación e inteligencia artificial

Fecha de lectura: 18/12/2015

Programa de doctorado: Doctorado en Ingeniería Informática y Computación

Dirección:

> **Director:** Faraón Llorens Largo

Tribunal:

> **presidente:** Ramón Rizo Aldeguer

> **secretario:** FRANCISCO JOSE GARCÍA PEÑALVO

> **vocal:** David Edward Kirkhouse Griffiths

Descriptor:

> INTELIGENCIA ARTIFICIAL

> INFORMATICA

> SOFTWARE

El fichero de tesis ya ha sido incorporado al sistema

> <http://hdl.handle.net/10045/53697>

Resumen: In every learning or training environment, learning activities are the basis for practical learning. Learners need to practice in order to acquire new abilities and perfect those previously gained. The key for an optimized learning process is correctly assigning learning activities to learners. Each learner has specific needs depending on previous knowledge and personal skills. A correct assignment for a given learner would be selecting a learning activity that closely matches learner's skills and knowledge. This brings up the concept of difficulty. Difficulty of a learning activity could be defined as the effort that a learner has to make to successfully complete the learning activity and obtain its associated learning outcomes. So, a difficult activity would simply require much effort to be successfully completed.

Learners presented with too difficult learning activities tend to abandon rather than performing required effort. This situation could be better understood as the learner perceiving the activity as an unbalanced investment-return ratio: too much effort for the expected learning outcomes. A similar case occurs when difficulty is too easy. In that case, effort perceived is low, but learning outcomes are perceived as even lower. If the activity does not pose a challenge for the learner is because the learner already masters the involved abilities, and that makes learning outcomes tend to zero. Both situations drive learners to losing interest.

To prevent this from happening, teachers and trainers estimate difficulties of learning activities based on their own experience. However, this procedure suffers an effect called the Curse of Knowledge: every person that masters an activity, becomes biased for estimating the effort required to master that same activity. Therefore, correctly estimating difficulties of learning activities is an error-prone task when expert-knowledge is used to estimate them. But estimating difficulty without carrying out the learning activity would probably yield even worse results.

In order to escape from this error-prone cycle, the first solution would be to measure the effort involved in successfully completing the learning activity. For that purpose, an objective effort measurement should be defined. This approach has been followed by many previous works and it is the general approach in the field of Learning Analytics. Although this approach yields many types of considerable results, it has an important drawback. It is impossible to have a measure without learners performing the learning activity. Therefore, at design stages of the learning activity, how does the designer know whether the activity is too hard/too easy? Is there a way to have an valid estimation of difficulty of a learning activity before handing it to learners?

This work proposes a new approach to tackle this problem. The approach consists in training a Machine Learning algorithm and measure the "effort" the algorithm requires to find successful solutions to learning activities. The "effort" will be the learning cost: the time the algorithm requires for training. After that, results obtained from training the Machine Learning algorithm will be compared to results measured from actual learners. Under the assumption that learning costs for Machine Learning algorithms and those for learners have some kind of correlation, results from comparing them should show that correlation. If that were the case, then the learning cost that Machine Learning algorithms invest in training could be used as an estimation of the difficulty of the learning activity for learners.

In order to implement this approach and to obtain experimental data, two Neuroevolution algorithms have been selected for the Machine Learning part: Neuroevolution of Augmenting Topologies (NEAT) and Hyper-cube-based Neuroevolution of Augmenting Topologies (HyperNEAT).

Implementing this proposed approach has yielded several contributions that are presented in this work:

1. A new definition of difficulty as a function, based on the progress made over time as an inverse measure of the effort/learning cost.
2. A similarity measure to compare Machine Learning results to those of learners and know the accuracy of the estimation.
3. A game called PLMan that is used as learning activity in the experiments. It is a Pacman-like game composed of up to 220 different mazes, that is used to teach Prolog programming, Logics and a light introduction to Artificial Intelligence.
4. An application of NEAT and HyperNEAT to learn to automatically solve PLMan mazes.
5. A novel application of Neuroevolution to estimate difficulty of learning activities at design stages.

Experimental results confirm that there exists a correlation between learning costs of Neuroevolution and those of students. Goodness of the presented results is limited by the scope of this study and its empirical nature. Nevertheless, they are greatly significant and may open up a new line of research on the relation between Machine Learning and humans with respect to the process of learning itself.

