

New algorithm found for learning languages

September 6 2005

Cornell University and Tel Aviv University researchers say they've developed a method for enabling a computer program to learn languages.

The researchers said the algorithm they've formulated allows a computer to scan text in many languages, including English and Chinese, and autonomously and without previous information infer the underlying rules of grammar.

The rules can then be used to generate new and meaningful sentences, they said. The method also works for such data as sheet music or protein sequences.

"The algorithm -- the computational method -- for language learning and processing that we have developed can take a body of text, abstract from it a collection of recurring patterns or rules and then generate new material," explained Shimon Edelman, a computer scientist and professor of psychology at Cornell.

"This is the first time an unsupervised algorithm is shown capable of learning complex syntax, generating grammatical new sentences and proving useful in other fields that call for structure discovery from raw data, such as bioinformatics," he said.

The study appears in the Proceedings of the National Academy of Sciences.

Copyright 2005 by United Press International

Citation: New algorithm found for learning languages (2005, September 6) retrieved 22 May 2024 from <https://phys.org/news/2005-09-algorithm-languages.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.