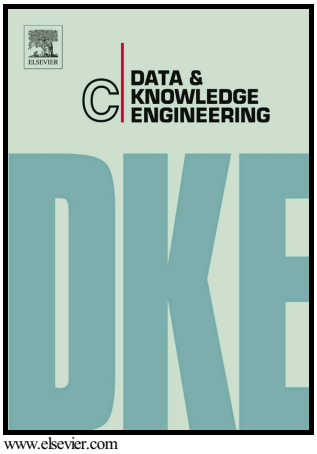
Author's Accepted Manuscript



A knowledge-based system for generating interaction networks from ecological data

Willem Coetzer, Deshendran Moodley, Aurona Gerber

PII: S0169-023X(17)30045-9
DOI: https://doi.org/10.1016/j.datak.2017.09.005
Reference: DATAK1616

To appear in: *Data & Knowledge Engineering*

Received date: 26 January 2017
Revised date: 15 September 2017
Accepted date: 26 September 2017

Cite this article as: Willem Coetzer, Deshendran Moodley and Aurona Gerber, A knowledge-based system for generating interaction networks from ecological data, *Data & Knowledge Engineering,* https://doi.org/10.1016/j.datak.2017.09.005

A knowledge-based system for generating interaction networks from ecological data

Willem Coetzer[1,2,5*], Deshendran Moodley[2,4], Aurona Gerber[2,3]

[1] SAIAB: South African Institute for Aquatic Biodiversity, Private Bag 1015, Grahamstown 6140, South Africa

[2] CAIR: Centre for Artificial Intelligence Research, CSIR Meraka, PO Box 395, Pretoria, 0001, South Africa

[3] Department of Informatics, University of Pretoria, Private Bag X20, Hatfield, 0028, South Africa

[4] Department of Computer Science, University of Cape Town, Private Bag X3, Rondebosch, 7701, South Africa

[5] School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Private Bag X54001, Durban 4000, South Africa

[*] Corresponding author. E-mail: w.coetzer@saiab.ac.za

Abstract

Semantic heterogeneity hampers efforts to find, integrate, analyse and interpret ecological data. An application case-study is described, in which the objective was to automate the integration and interpretation of heterogeneous, flower-visiting ecological data. A prototype knowledge-based system is described and evaluated. The system's semantic architecture uses a combination of ontologies and a Bayesian network to represent and reason with qualitative, uncertain ecological data and knowledge. This allows the high-level context and causal knowledge of behavioural interactions between individual plants and insects, and consequent ecological interactions between plant and insect populations, to be discovered. The system automatically assembles ecological interactions into a semantically consistent interaction network (a new design of a useful, traditional domain model). We discuss the contribution of probabilistic reasoning to knowledge discovery, the limitations of knowledge discovery in the application case-study, the impact of the work and the potential to apply the system design to the study of ecological interaction networks in general.

Keywords: Semantic heterogeneity, ontologies, Bayesian network, knowledge discovery, semantic architecture, interaction network, ecological interactions

Funding: This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.