Journal of Environmental Science and Management Special Issue 2-2016: 47-57 ISSN 0119-1144

Exploring Utility of Formal Concept Analysis Approach to Coral Reef Assessment

ABSTRACT

This study explored the use of formal concept analysis (FCA), a data mining technique, to analyze coral reef transect data (in terms of life forms) and comparing its results to the standard assessment analysis. Utilizing the quadrat-life form as the object-attribute pair, the results derived from the context was analyzed to assess the coral reefs in the study site which consisted of three stations. Data from Station 1 and Station 2 showed the dominance of Acropora digitate and Acropora branching life forms, respectively. Some life forms were absent from both Stations 1 and 2 but all life forms were present in Station 3 with eight life forms having the highest occurrence. Station 3 had the highest diversity of life forms while Station 2 had the highest live coral cover. This study showed how FCA can be used to generate new knowledge from transect data that can be verified by traditional coral reef assessment results, a possible complement to standard coral reef assessment analytical tools. FCA approach shines when it deals with large data sets from many different sources, which may pave the way for data-driven ecological assessment analysis studies such as those already being done for agriculture.

Key words:*data mining; attribute implications; concept lattices; formal concept analysis; coral reef assessment; photo transect data*

Val Randolf M. Madrid[1*]
Lei Kristoffer Lactuan[1]
Jonel M. Corral[2]

[1] Institute of Computer Science College of Arts and Sciences, University of the Philippines Los Baños (CAS-UPLB), College, Laguna 4031
[1] Institute of Biological Sciences, CAS-UPLB

*corresponding author:
vmmadrid@up.edu.ph

INTRODUCTION

The Philippines has the most extensive coral reef in the world, second only to Indonesia (*Burke et al. 2001*). Touted to be the tropical rainforest of the sea, coral reefs attract a diverse array of marine organisms providing a source of food and shelter for a large variety of species including fish, shellfish, fungi, sponges, sea anemones, and turtles. As fishing grounds, they are thought to be 10 to 100 times as productive per unit area as the open sea. In the Philippines, an estimated 10-15 percent of the total fisheries come from coral reefs.

Thanks to technological improvements in the area of digital image processing, cheaper computation, mass storages devices, and faster computer processing speed, these tools have been harnessed by marine scientists to archive and analyze coral reef communities (*Konotchick et al. 2006*).

Coral Reef Assessment Approach

A common technique used by marine scientists is to determine the sampling stations used for coral reef assessment based on the reconnaissance survey (manta tow) prior to assessment. Then, three 30-m transect lines are laid in each station. Each transect is composed of 30 quadrats (1×1 m in size) having 25 equal grids. Photo transect technique is then employed wherein photos of the substrate per quadrat are taken using an underwater camera. Coral identification will be done by the marine scientist with the aid of a photo viewer software. All taxa within each quadrat are then identified up to the lowest possible taxonomic level using standard field guides for coral identification (*White 2001*).

The data garnered from the photos are plotted in a spreadsheet for further study. Given the volume of photo and video transects that are created each year for monitoring our coral reefs, however, it would be difficult for marine scientists to analyze all of them. Therefore, it would be cost-effective for marine scientists to train and employ machines to assist in mining new knowledge and discovering new insights from past accumulated data.

Data-Driven Approach

Data mining has begun making waves as a subfield of computer science. It is the process of sifting through huge volumes of data ranging from satellite imagery to surveillance videos to twitter feeds. It has become relevant ever since we used computers to store data and it also makes access to data up to 30 years ago possible. It has plenty of applications from combating terrorism to fighting hunger and poverty (*Kim et al. 2014*).

Formal concept analysis (FCA), a mathematical lattice-based framework, has gained momentum as a data mining technique. The major feature of FCA is that it organizes the information in the form of a mathematical lattice structure and visualizes the sub-super concept hierarchy. FCA aims to analyze the data represented as a formal context. The context table consists of rows corresponding to objects and columns corresponding to attributes (places, features) (*Aswani and Srinivas 2010*).

Data-driven agriculture, which uses ICT and data mining, has already been utilized with the aim of helping farmers have better crop yield and direct policy makers toward best practices in climate change adaptation (*Jimenez et al. 2016*). In fact a team of CGIAR scientists won the UN Big Data Climate Challenge 2014 for their work in rice production in Latin America. By mining historical and current data from weather patterns, agricultural records, rice yield, rice variety, and rice farming practices, the team was able to come up with recommendations for rice farmers, which included choosing more productive rice varieties, suitable planting times and cropping duration for specific sites, and seasonal forecasts. Their method enabled farmers to increase their yield by 1–3 MT (*Cornish 2014*).

The main objective of this study is to explore the application of a data mining technique, such as FCA, to analyze coral reef data (in terms of life forms) from underwater images. By using FCA along with traditional methods, we can either verify existing or discover new relationships between various coral life forms in a single quadrat as well as between groups of quadrats. By incorporating an information technology platform into coral reef assessment, results may be used as basis or starting point to discover emergent properties or attributes in the data that would not surface in traditional approaches. It could also lead to different applications for data-driven ecological assessment and analysis by incorporating present and past accumulated data.

Formal Concept Analysis

FCA is a familiar method for object-attribute data analysis introduced by Rudolf Wille in 1980 (*Ganter et al. 2005*), which was inspired by how humans think and generate knowledge Suppose a set of fruit objects *F* ={"apple", "orange", "banana", "mango", "avocado", "peach"} and set of color/shapes/taste attributes C= {"round", "yellow", "sour", "green", "sweet"} to act as sample dataset in which to apply FCA. Then the context can be defined, it is a triplet of the form (*F, C, A*) where *F* represents a set of fruit objects, *C* the set of color/shape/taste attributes, and A as incidence binary relation between *F* and *C*. For example, the statement "an apple is round" can be translated in FCA as given an object a with an attribute *r* then it can be denoted as *aAr* where object "apple" has the attribute "*round*".

Consider set *G* = {"banana", "mango"} and set *M* = {"sweet", "yellow"}, set G′ is defined as the set of attributes common in the objects of set *G*. Set *M′* is defined as the set of objects sharing the attributes in set *M*. A concept of the context (*F, C, A*) is a pair (*G, M*) where $G \subseteq F$, $M \subseteq C$, $G'=M$ and $M'=G$. *G* is the extent and *M* the intent of the concept (*G, M*). A concept is also identified by its extent and its intent. The extent consists of all objects belonging to the concept while the intent contains all attributes shared by the objects. The relation between these objects and attributes is defined as formal context (*Ganter et al. 2005*).

A change in the formal context will certainly affect the concepts and the structure, which, in turn, will affect the relationship of the object with the associated attributes and also the basis to derive a conclusion. The nature of concept hierarchies can be observed along with concept ordering. For example, set *W* = {"peach"} and set *X* = {"round", "sweet", "yellow"}, then the ordering of concept (*W, X*) and concept (*G, M*) is $(G, M) \leq (W, X)$ because attribute set *X* contains the attribute set *M*. The idea of concept hierarchies and ordering means that fewer objects share more and more attributes. FCA can also generate knowledge in the form of attribute implications. An attribute implication is valid in a formal context whenever all objects of that context satisfy it. For example, for set *A1*= {"yellow"} and set *A2* ={"sweet"}, FCA can say $A1 \rightarrow A2$ or "yellow" implies "sweet" because the objects "banana", "mango", and "peach" support it. There are usually two bases of implications used for finding implication sets—Duquenne-Guigues (DG) base and Luxenberger base (*Wormuth and Becker 2004*). The DG base discusses about the implications (association rules) with 100% confidence while the Luxenburger base includes association rules with less than 100 % confidence (*Yevtushenko 2000*). In this paper only DG base is applied because of the 100% object support, which means there are actual examples of objects and attributes that support the implications.

FCA in Data Mining

Since FCA deals with data, there have been studies using FCA for knowledge discovery and data mining. The idea is to discover knowledge or ideas from a set of data stored in databases. This process is called knowledge discovery in databases (KDD). *Valtchev et al.* (*2004*) found that FCA was an appropriate framework for KDD because knowledge extracted may be used for decision making.

Stumme (*2002*) used FCA in KDD to observe FCA's conceptual clustering as applied to analyzing very large databases. Introducing the notion of iceberg lattices demonstrated by their algorithm TITANIC, *Stumme* (*2002*) found that iceberg concept lattices are starting points for computing condensed sets of association rules without loss of information.

Kaytoue et al. (*2011*) experimented on gene expressions using two FCA-based methods comparing them in terms of efficiency and readable results in mining numerical data from gene expressions. One method used standard FCA techniques and relied on a particular scaling, while the other relied on pattern structures. The second method proved to be more computationally efficient and had more readable results.

AA survey (*Poelmans et al. 2013*) of FCA research showed a wide range of applications, which include discovering best practices for software engineering by analyzing static source codes of programmers (*Eisenbarth et al. 2003*); criminal profiling from the unstructured text of police reports (*Poelmans et al. 2010*); identifying ecological traits to assess water quality by mining macroscopic plant species living in water bodies data (*Bertaux et al. 2009*); improving healthcare practices by mining time-series data, questionnaires, and journals in medicine (*Jay et al. 2006*); identifying biomarkers for breast cancer (*Gebert et al. 2008*); and analyzing structural-activity relationships of chemical compounds in chemistry (*Stumpfe et al. 2011*).

MATERIALS AND METHODS

Photo Transect Locale

The data used in the study were taken from three sampling stations, which were established to assess the coast of Barangay Ilijan, Batangas City as part of an initiative to develop a coastal resources management plan. The sampling protocol and the setup of the phototransect approach were designed by the marine scientists in the study (*Anit et al. 2014*) commissioned by the Batangas City LGU. This paper, in turn, used the coral assessment data gathered by that study for creating the FCA concept lattices and implications.

Three sampling stations were established in the study area according to the intensity of human disturbance (**Figure 1**). Station 1 is uninhabited and with greyish, coralline substrate. Station 2 is where the coastal community is concentrated. This station also covers the area where the bivalves seeding project in 2003 was done through the initiatives of Kepco Ilijan Corp. (KIELCO) and Apercu Consultants Inc. This seeding project was designed to make the waters of Ilijan more conducive to marine life (*Anit et al. 2014*). Station 3, aside from being populated by a fishing community, is located close to a KIELCO power plant.

The photo transect data from the three sampling stations were processed and inputted into a spreadsheet file with the aid of a coral field guide. Since identification of coral life forms is more useful in determining coral reef health than taxonomic data, the following abbreviations of 19 life forms categories were used:

1. *Acropora* digitate (ACD)

Figure 1. Map of Ilijan Bay with the 3 sampling stations.

2. *Acropora* coral branching (ACB)
3. *Acropora* coral submassive (ACS)
4. *Acropora* coral tabulate (ACT)
5. Coral branching (CB)
6. Coral encrusting (CE)
7. Coral foliose (CF)
8. Massive coral (CM)
9. Mushroom coral (CMR)
10. Coral submassive (CS)
11. Dead coral (DC)
12. Dead coral with algae (DCA)
13. Gorgonians (GOC)
14. *Halimeda* (HA)
15. Macroalgae (MA)
16. Rock (R)
17. Sand (S)
18. Soft coral (SC)
19. Sponge (SP)

Transect Photo Quadrat

Data was collected from the three sampling stations with Station 1 having 30 quadrats in Transect 1, 29 quadrats in Transect 2, and 29 quadrats in Transect 3. Station 2 had 3 quadrats in Transect 1, 30 quadrats in Transect 2, and 19 quadrats in Transect 3. For Station 3, there were 30 quadrats in Transect 1, 30 quadrats in Transect 2, and 30 quadrats in Transect 3. **Figures 2, 3,** and **4** show sample quadrats from the three stations. The data was plotted using Concept Explorer (version 1.3) with each row corresponding to each quadrat in a transect image and each column to the coral life form found in each quadrat. A 30-quadrat transect would have 30 object rows and 19 attribute columns. Each station has three transects each and FCA was used on the data from the three transects to describe the three stations.

Concept Explorer 1.3

Concept Explorer (version 1.3), which is a Java-based tool that provides basic functionalities in FCA, was used. This is an open-source software developed by Serhiy Yevtushenko and some contributors (*Yevtushenko 2000*) and provides basic functionalities such as context editing, building concept lattices from context, finding bases of implications, finding bases of association rules, and performing attribute exploration. The spreadsheet data was imported into the software (**Figure 5**).

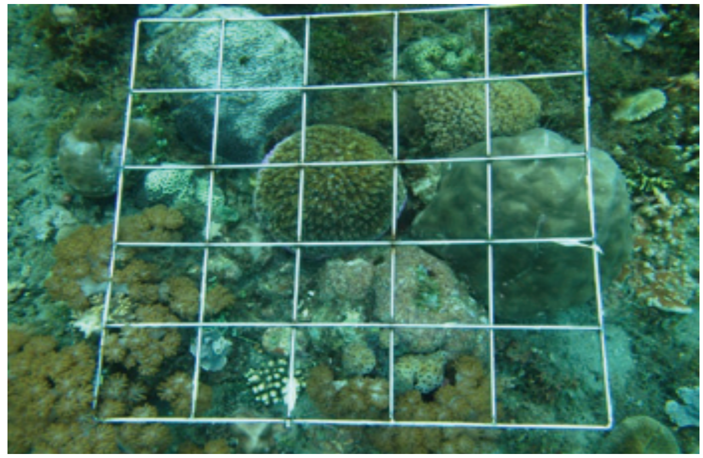

Figure 2. Sample transect photo with quadrat taken from Station 1.

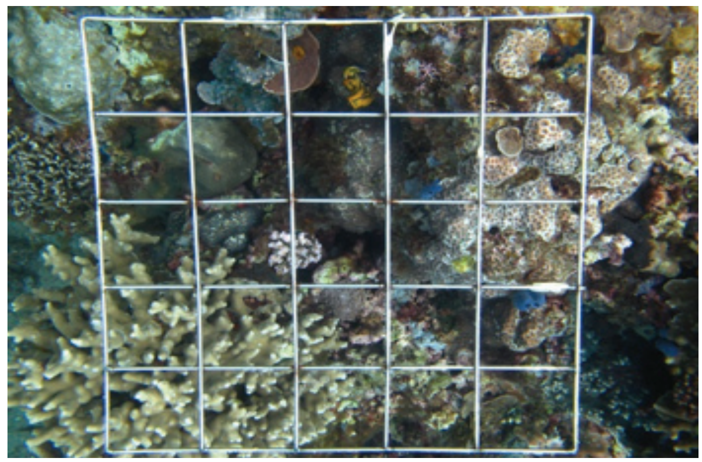

Figure 3. Sample transect photo with quadrat taken from Station 2.

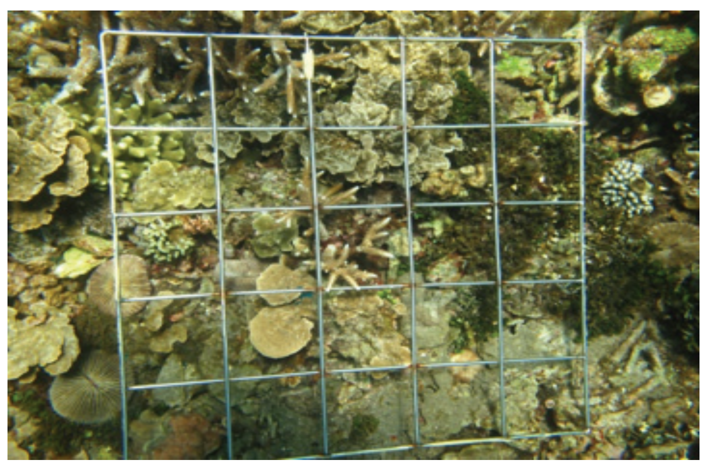

Figure 4. Sample transect photo with quadrat taken from Station 3.

RESULTS AND DISCUSSION

The sample context tables are generated for Station 1-Transect 1, Station 2-Transect 2, and Station 3-Transect 3 (**Figure 6**). It can be observed that in Station 1 there were just six attributes in the highest concept hierarchies while Station 2 had seven. Station 3 had the highest count with 11 attribute concepts at the top hierarchy. This can be further seen in the following concept lattice diagrams based on the context tables (**Figures 7, 8 and 9**).

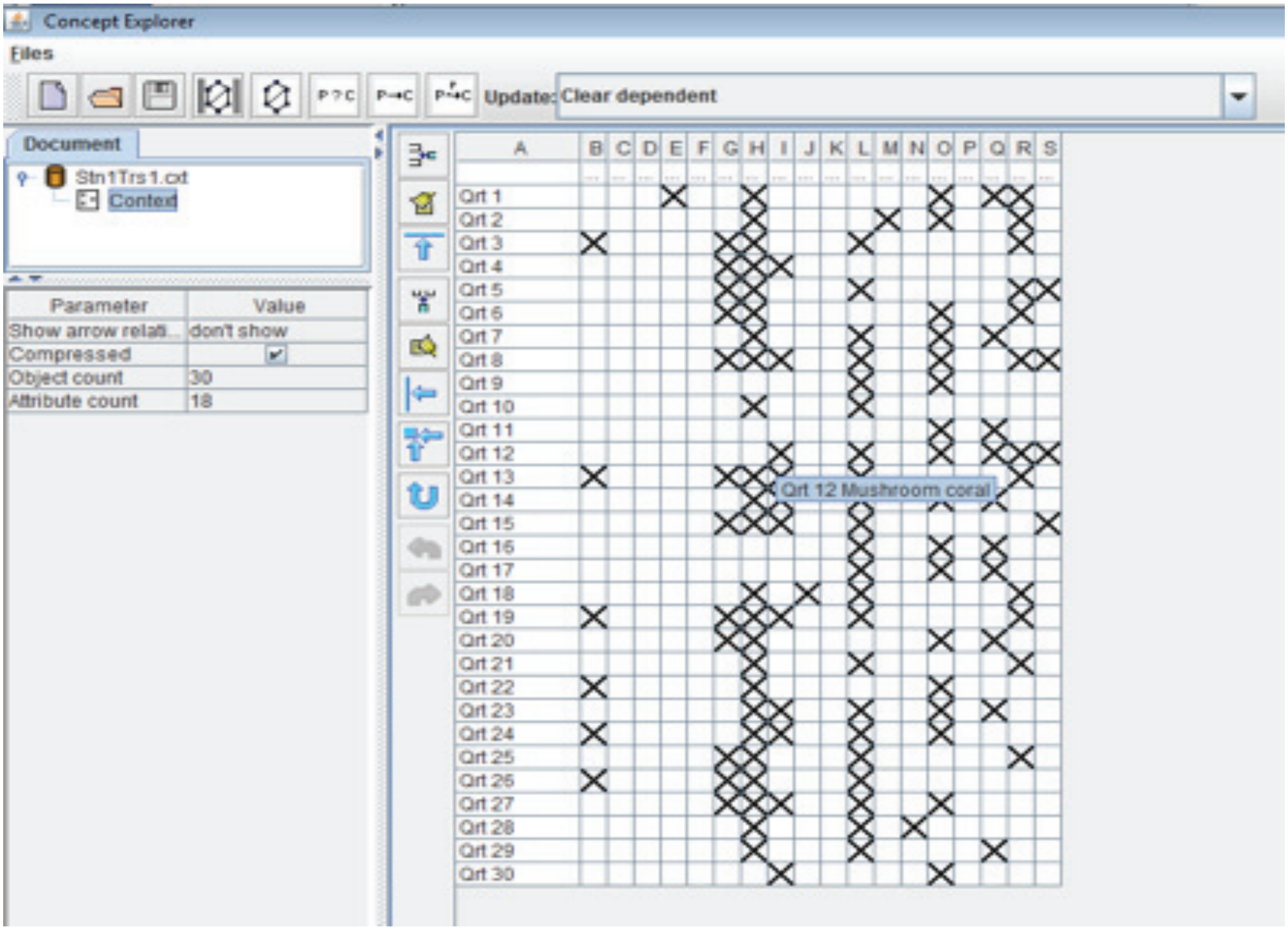

Figure 5. Concept Explorer 1.3 screenshot of binary relationship matrix between each quadrat having coral life form attributes from Transect 1, Station 1.

Quadrats	ACD	ACS	ACT	CB	CE	CF	CM	CMR	CS	DC	DCA	GOC	HA	MA	R	S	SC	SP
	Attributes																	
Qrt 1	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	1	1	0
Qrt 2	0	0	0	0	0	0	1	0	0	0	0	1	0	1	0	0	1	0
Qrt 3	1	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	1	0
Qrt 4	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0
Qrt 5	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	1	1
Qrt 6	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0	1	0
Qrt 7	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0	1	0	0
Qrt 8	0	0	0	0	0	1	1	1	0	0	1	0	0	1	0	0	1	1
Qrt 9	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0
Qrt 10	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0
Qrt 11	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0
Qrt 12	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	1	1	1
Qrt 13	1	0	0	0	0	1	1	1	0	0	1	0	0	0	0	0	1	0
Qrt 14	0	0	0	0	0	0	1	1	0	0	1	0	0	1	0	1	0	0
Qrt 15	0	0	0	0	0	1	1	1	0	0	1	0	0	0	0	0	0	1
Qrt 16	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	0	0
Qrt 17	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	0	0
Qrt 18	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	1	0
Qrt 19	1	0	0	0	0	1	1	1	0	0	1	0	0	0	0	0	1	0
Qrt 20	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	1	0	0
Qrt 21	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0
Qrt 22	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0
Qrt 23	0	0	0	0	0	0	1	1	0	0	1	0	0	1	0	1	0	0
Qrt 24	1	0	0	0	0	0	1	1	0	0	1	0	0	1	0	0	0	0
Qrt 25	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	1	0
Qrt 26	1	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	0
Qrt 27	0	0	0	0	0	1	1	1	0	0	1	0	0	1	0	0	0	0
Qrt 28	0	0	0	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0
Qrt 29	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0
Qrt 30	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0

Quadrats	ACD	ACS	ACT	CB	CE	CF	CM	CMR	CS	DC	DCA	GOC	HA	MA	R	S	SC	SP
	Attributes																	
Qrt 1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Qrt 2	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Qrt 3	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Qrt 4	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0
Qrt 5	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0
Qrt 6	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0
Qrt 7	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0
Qrt 8	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0
Qrt 9	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
Qrt 10	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Qrt 11	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Qrt 12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
Qrt 13	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0
Qrt 14	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Qrt 15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Qrt 16	0	0	1	1	0	1	0	1	0	0	1	0	0	1	0	0	0	0
Qrt 17	1	0	0	1	0	1	0	1	0	0	1	0	0	1	0	0	0	0
Qrt 18	0	0	0	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0
Qrt 19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Qrt 20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Qrt 21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Qrt 22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Qrt 23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Qrt 24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Qrt 25	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Qrt 26	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Qrt 27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Qrt 28	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
Qrt 29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Qrt 30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Quadrats	ACD	ACS	ACT	CB	CE	CF	CM	CMR	CS	DC	DCA	GOC	HA	MA	R	S	SC	SP
	Attributes																	
Qrt 1	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0
Qrt 2	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	1	0
Qrt 3	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	1	0	0
Qrt 4	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	1	0
Qrt 5	0	0	0	0	0	0	1	0	0	0	0	1	0	1	0	0	0	1
Qrt 6	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0
Qrt 7	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Qrt 8	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0
Qrt 9	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Qrt 10	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0	0
Qrt 11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Qrt 12	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1	1	0
Qrt 13	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Qrt 14	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Qrt 15	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0
Qrt 16	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	1	0
Qrt 17	0	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Qrt 18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Qrt 19	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	1
Qrt 20	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Qrt 21	1	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0
Qrt 22	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Qrt 23	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0
Qrt 24	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Qrt 25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Qrt 26	1	0	0	1	0	1	0	1	0	0	1	0	0	1	0	0	0	1
Qrt 27	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Qrt 28	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Qrt 29	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0
Qrt 30	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0

Figure 6. Context tables of Station 1-Transect 1, Station 2-Transect 2, and Station 3-Transect 3.

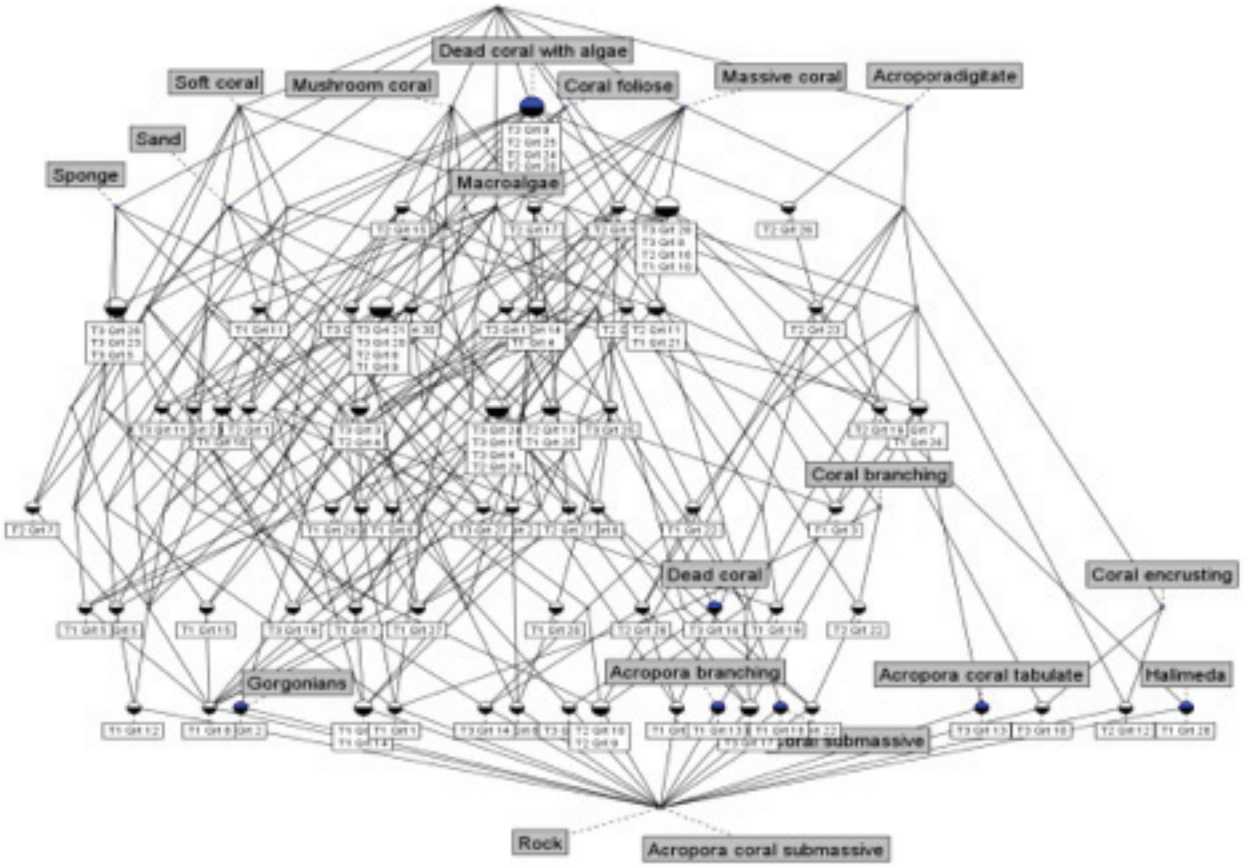

Figure 7. Concept lattice diagram of Station 1.

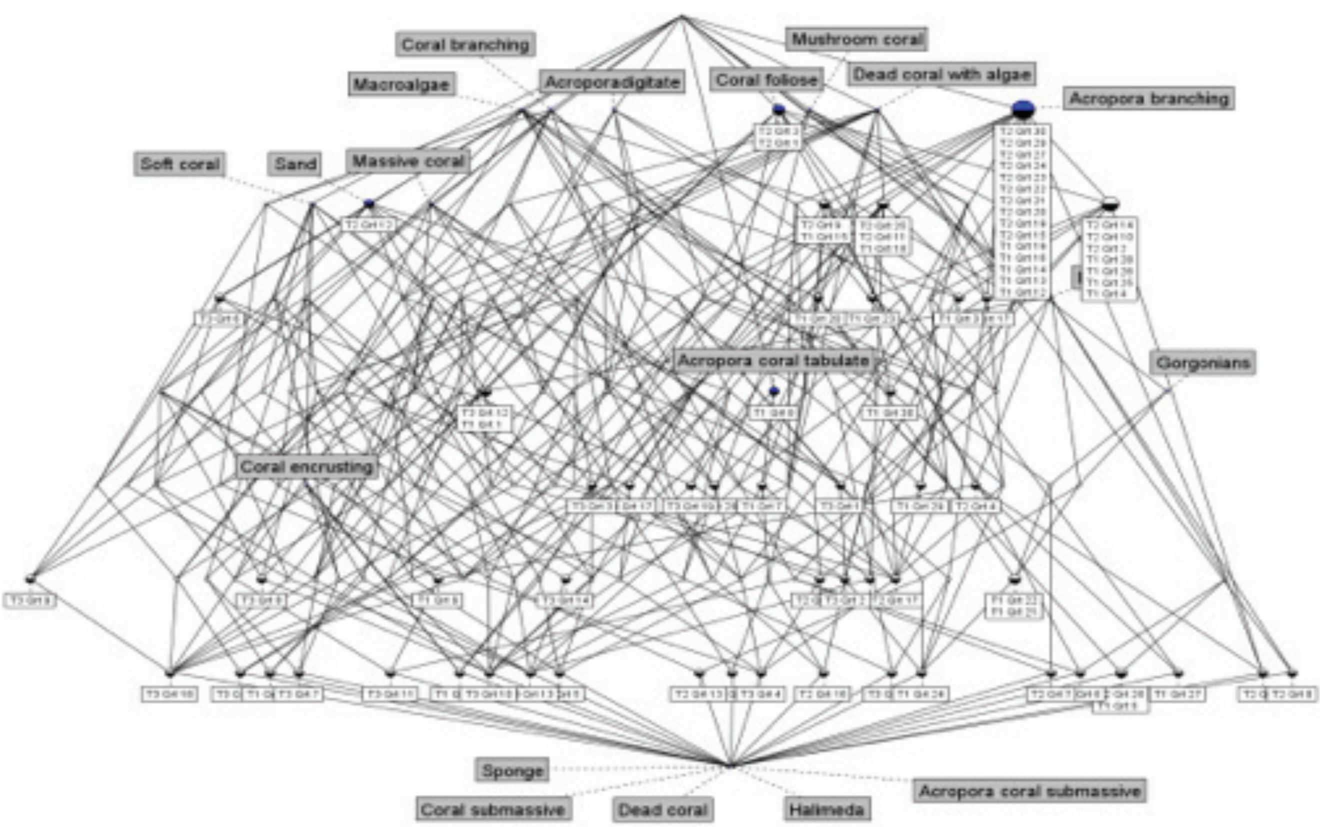

Figure 8. Concept lattice diagram of Station 2.

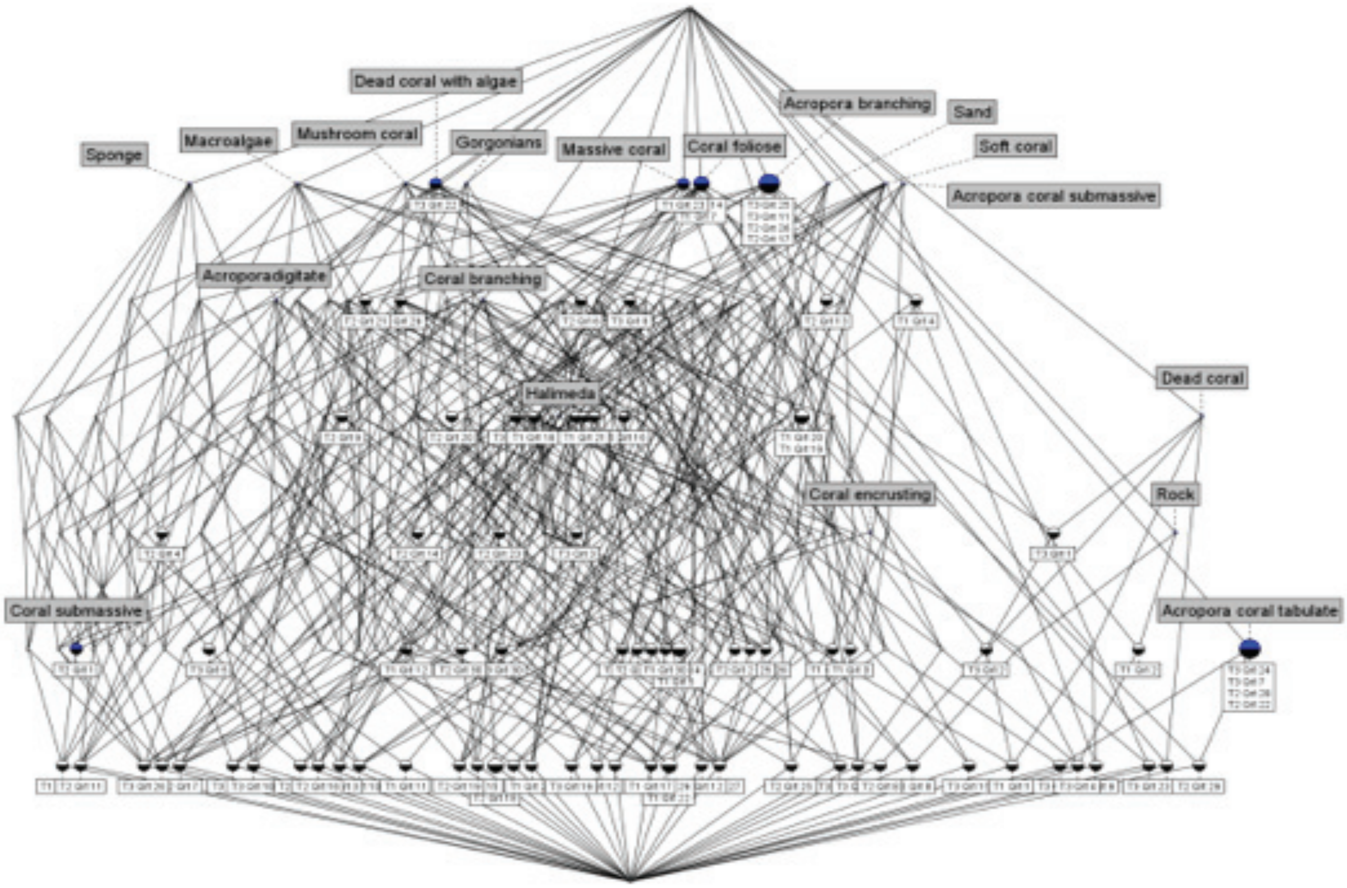

Figure 9. Concept lattice diagram of Station 3.

In the concept hierarchy diagram, the concept nodes, along with the attached objects and attributes, go from most general at the top to more specific at the bottom. The concept hierarchy diagram also generates all possible concepts that can be mined. The concept nodes could either be blue-black circle, white-black circle, or empty small circle. The attributes (coral lifeform) are represented by grey rectangles and objects (quadrat) are represented by white rectangles. A blue-black concept node means there are actual objects attached to the attribute(s) it defines exclusively. The white-black node means there are objects attached to the attribute(s) it defines, but there are also other attributes present in the object that are attached to other concepts.

In the concept hierarchy diagram, the concept nodes, along with the attached objects and attributes, go from most general at the top to more specific at the bottom. The concept hierarchy diagram also generates all possible concepts thatcan be mined. The concept nodes could either be blue-black circle, white-black circle, or empty small circle. The attributes (coral lifeform) are represented by grey rectangles and objects (quadrat) are represented by white rectangles. A blue-black concept node means there are actual objects attached to the attribute(s) it defines exclusively. The white-black node means there are objects attached to the attribute(s) it defines, but there are also other attributes present in the object that are attached to other concepts.

For Station 1, the total number of concepts calculated was 125, given 88 total quadrats. It can be seen from the top hierarchy of the concept lattice that the most commonly found coral life forms were soft coral, dead coral with algae, coral foliose, mushroom coral, and *Acropora* digitate. Dead coral with algae was the most common with 64 quadrats having this life form. It can be also seen that there were four quadrats that contained only dead corals with algae. Massive coral, coral foliose, and *Acropora* digitate were also found in 59 quadrats. At the bottom of the diagram were rock and *Acropora* submassive. This means that those coral life forms were not found in that transect area.

For Station 2, there were 169 concepts calculated given 79 total quadrats. The life forms at the top hierarchy indicated the presence of macroalgae, *Acropora* digitate, coral branching, coral foliose, dead coral with algae, *Acropora* coral branching, and mushroom coral. *Acropora* coral branching was the most common, being present in

54 quadrats; 15 of the 54 quadrats had *Acropora* coral branching exclusively. Coral foliose and dead coral with algae were found in 27 and 29 quadrats, respectively. Of the three stations, Station 2 also had the highest live coral cover as evidenced by the dominance of Acropora coral branching seen in the concept lattice's object extent. This coincided with the assessment report that Station 2 had a 58% live coral cover. There was also high occurrence of sand and rocks. But the sponge, dead coral, coral submassive, *Halimeda*, and *Acropora* submassive life forms were notably absent.

For Station 3, only 247 concepts were calculated given 90 total quadrats. The most common coral life forms found were *Acropora* coral branching, *Acropora* submassive, soft coral, massive coral, coral foliose, dead coral with algae, macroalgae, mushroom coral, sponge, gorgonian, and sand. *Acropora* coral branching had an extent of 40 quadrats with 4 exclusive quadrats while dead coral with algae, massive coral, and coral foliose had the extent of 33 quadrats, 44 quadrats, and 24 quadrats respectively. *Acropora* coral tabulate was also seen to have four exclusive quadrats but only had an extent of six quadrats. All life forms were present in Station 3.

In addition to the concept lattice, FCA also produced implications, which present the relationships between attributes. Using the DG basis as tool, the attribute implications were derived from the context with object support greater or equal to two (**Tables 1, 2** and **3**).

Table 1. Duquenne-Guigues basis table for Station 1, implications with 100% support.

Implication	**Support ≥ 2**
Coral foliose, Soft coral ==> Massive coral	12
Acropora digitate, Coral foliose ==> Massive coral	11
Acropora digitate, Macroalgae ==> Massive coral	8
Acropora digitate, Soft coral ==> Massive coral	7
Massive coral, Mushroom coral, Soft coral ==> Coral foliose	5
Coral foliose Sponge ==> Massive coral	5
Mushroom coral, Dead coral with algae Sand ==> Macroalgae	4
Massive coral, Soft coral Sponge ==> Coral foliose	4
Acropora digitate, Mushroom coral ==> Massive coral Dead coral with algae	3
Dead coral with algae Macroalgae Soft coral ==> Mushroom coral	3
Acropora digitate, Sponge ==> Massive coral Macroalgae	3
Mushroom coral, Sponge ==> Dead coral with algae	3
Coral foliose, Massive coral, Macroalgae Sponge ==> Soft coral	3
Coral encrusting ==> *Acropora* digitate Massive coral	2
Coral branching, Massive coral ==> Macroalgae	2

Table 2. Duquenne-Guigues basis table for Station 2, implications with 100% support.

Implication	**Support ≥ 2**
Massive coral Sand ==> Soft coral	8
Coral encrusting ==> Soft coral	4
Acropora digitate, Mushroom coral ==> *Acropora* branching Dead coral with algae	4
Coral branching, Mushroom coral ==> Coral foliose	4
Coral branching, Coral foliose, Dead coral with algae ==> Macroalgae	4
Coral branching, Coral foliose, Macroalgae ==> Dead coral with algae	4
Coral foliose, Mushroom coral, Macroalgae ==> Dead coral with algae	4
Mushroom coral, Dead coral with algae Macroalgae==> Coral foliose	4

Table 3. Duquenne-Guigues basis table for Station 3, implications with 100% support.

Implication	**Support ≥ 2**
Acropora branching, Gorgonians ==> Massive coral	5
Acropora branching, Dead coral with algae Soft coral==> Massive coral	5
Acropora branching, Sponge ==> Massive coral	5
Acropora coral submassive, Mushroom coral ==> *Acropora* branching Massive coral	4
Acropora coral submassive, *Acropora* branching, Dead coral with algae ==> Massive coral	4
Massive coral, Dead coral with algae Gorgonians ==> *Acropora* branching	4
Coral foliose, Massive coral, Macroalgae ==> Sand	4

Table 4. Frequency count and cover of coral genera in the three stations.

Family	Genus	Lifeform	Frequency Count			GRAND	Coral Cover		
			Station 1	Station 2	Station 3	TOTAL	Station 1	Station 2	Station 3
Acroporidae	Acropora	ACB	1	935	464	1400	0.05	47.34	20.62
		ACT	4	118	74	196	0.18	5.97	3.29
		ACD	0	0	6	6	0.00	0.00	0.27
		ACS	0	0	146	146	0.00	0.00	6.49
	Montipora	ACD	5	23	94	122	0.23	1.16	4.18
		CF	7	86	0	93	0.32	4.35	0.00
Agariciidae	Coeloseris	CM	2	20	51	73	0.09	1.01	2.27
Dendrophyllidae	Turbinaria	CF	54	89	65	208	2.45	4.51	2.89
	Euphyllia	CM	5	0	0	5	0.23	0.00	0.00
Faviidae	Diploastrea	CM	9	6	59	74	0.41	0.30	2.62
	Caulastrea	CM	7	0	0	7	0.32	0.00	0.00
	Colpophyllia sp	CM	1	0	1	2	0.05	0.00	0.04
	Echinopora	CF	21	49	93	163	0.95	2.48	4.13
	Favia	CM	46	9	20	75	2.09	0.46	0.89
	Favites	CM	138	15	88	241	6.27	0.76	3.91
	Leptoria	CE	2	0	5	7	0.09	0.00	0.22
	Leptastrea	CM	0	0	8	8	0.00	0.00	0.36
	Montastrea	CM	87	0	0	87	3.95	0.00	0.00
		CE	2	0	0	2	0.09	0.00	0.00
	Goniastrea	CM	24	0	0	24	1.09	0.00	0.00
	Solenastrea	CM	0	3	0	3	0.00	0.15	0.00
	Oulophyllia	CM	0	0	12	12	0.00	0.00	0.53
	Platygyra	CM	0	0	4	4	0.00	0.00	0.18
Fungiidae	Fungia	CMR	2	1	11	14	0.09	0.05	0.49
	Australogyra	CMR	0	2	16	18	0.00	0.10	0.71
	Ctenactis	CMR	17	0	1	18	0.77	0.00	0.04
	Cycloseris	CMR	64	15	80	159	2.91	0.76	3.56
	Lithophyllon	CMR	1	0	0	1	0.05	0.00	0.00
	Halomitra	CMR	0	0	12	12	0.00	0.00	0.53
Meandrinidae	Meandrina	CM	3	0	3	6	0.14	0.00	0.13
Merulinidae	Merulina	CF	1	0	0	1	0.05	0.00	0.00
Milleporidae	Millepora	CF	1	0	0	1	0.05	0.00	0.00
Mussidae	Lobophyllia	CM	6	9	0	15	0.27	0.46	0.00
	Symphyllia	CF	4	2	23	29	0.18	0.10	1.02
		CM	10	0	0	10	0.45	0.00	0.00
	Isophyllia	CM	9	0	0	9	0.41	0.00	0.00
	Mussa	CE	4	0	0	4	0.18	0.00	0.00
Oculinidae	Galaxea	CM	27	0	0	27	1.23	0.00	0.00
Pectiniidae	Pectinia	CF	6	0	0	6	0.27	0.00	0.00
Pocilloporidae	Pocillopora	ACD	72	29	12	113	3.27	1.47	0.53
	Stylophora	CB	0	0	21	21	0.00	0.00	0.93
Poritidae	Porites	CB	3	47	64	114	0.14	2.38	2.84
		CM	98	0	59	157	4.45	0.00	2.62
		CS	3	0	0	3	0.14	0.00	0.00
Siderastridae	Siderastrea	CM	4	2	8	14	0.18	0.10	0.36
Trachphyllidae	Trachyphyllia	CF	1	0	0	1	0.05	0.00	0.00
	Unknown	CM	0	7	0	7	0.00	0.35	0.00
		DC	23	0	35	58	1.05	0.00	1.56
		DCA	1045	263	400	1708	47.50	13.32	17.78
Non-hard coral	Soft Coral	SC	101	117	271	489	4.59	5.92	12.04
	Gorgonians	GOC	6	86	35	127	0.27	4.35	1.56
	Halimeda	HA	5	0	44	49	0.23	0.00	1.96
	Macroalgae	MA	666	242	358	1266	30.27	12.25	15.91
	Sand	S	198	269	160	627	9.00	13.62	7.11
	Sponge	SP	19	0	72	91	0.86	0.00	3.20
	Rock	R	0	80	10	90	0.00	4.05	0.44

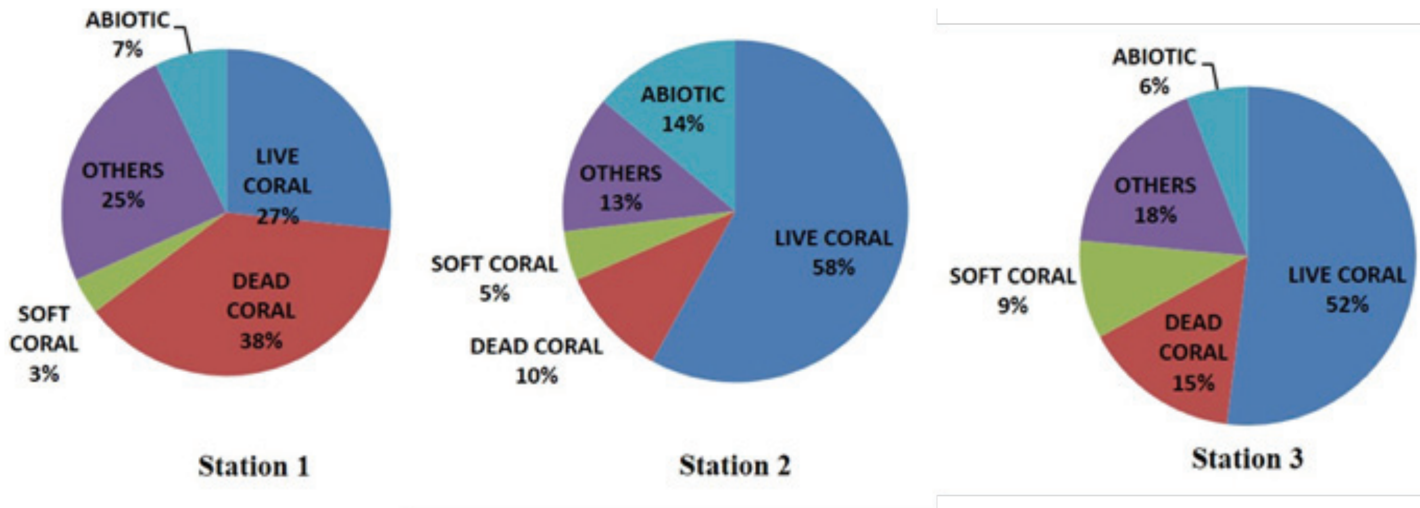

Figure 10. Proportion of coral and other benthic life in the three stations.

For Station 1, it was interesting to note that the implication with the highest object data support implied that in quadrats where coral foliose, *Acropora* digitate, and soft corals were found, massive corals were also present. It was also found out that coral foliose were contiguous to massive coral life form types. Massive corals with macroalgae were also found to exhibit attributes of dead corals with algae. Moreover, it was found that most of the massive corals were covered by algae or already dead.

For Station 2, massive coral and sand appeared together with patches of soft coral. Dead coral with algae and macroalgae appeared together as well. Numerous branching corals showed dead coral with algae life forms. This could be the result of increased human activity that has been seen in Station 2.

Station 3, which had the highest diversity of life forms including the non-hard corals, can be considered as having the healthiest coral reef among the three stations as implied from its concept lattice and implication table. Station 2's concept lattice and implication table would suggest that it had the highest live coral cover compared to Stations 1 and 3 but had the highest cover of rock and sand (abiotic cover) and higher dead coral with algae compared to Station 3. In the original assessment report, Station 1 was found to be more diverse in terms of number of species compared to Stations 2 and 3. On the other hand, their respective concept lattices and implications would imply that Station 3 is more diverse. This is because the concept lattices were generated using only lifeform information and different coral species may share the same lifeform category (**Table 4**). Moreover, the distribution of the object (quadrat) extent and attribute (lifeform) intent in Station 3's concept lattice and implication table would suggest that Station 3 had lower cover of sand and rocks (abiotic), low cover for dead coral with algae, with the rest of the lifeforms being evenly proportioned as compared to Stations 1 and 2. This implies that Station 3 was the healthiest of the three stations based on the FCA concept lattice and implications. This finding is similar to those of the original coral assessment report (**Figure 10**).

When applied to the same data from the coral reef assessment, the FCA approach generated similar analysis by analyzing the structure of the concept lattice and the spread of the object extent and the object count of each formal concept. The added bonus for FCA was it also generated implications regarding how different coral lifeforms group together. For example, in the Station 1 implication table, the top six implications with the highest object support mostly had coral foliose and *Acropora* digitate lifeforms, implying the presence of the massive coral lifeform (**Table 1**). Perhaps these lifeform relationships could be explored further to determine if lifeforms just grow in random together or whether certain conditions have to be met before they do.

CONCLUSION AND RECOMMENDATIONS

In this paper, we presented how FCA was applied to coral reef transect data gathered from coral reef assessment of three sampling stations. The FCA approach used in this paper was to treat each quadrat as "object" and the lifeforms identified as the "attributes". The implications obtained using FCA, which showed relationships between these attributes were also analyzed.

Despite limited data, the FCA approach, to some extent, can draw a similar conclusion to that of the traditional assessment approach. It would also be interesting to see how FCA can be applied not just in coral reef assessment but in other areas in ecological studies where voluminous data and variables are generated and analyzed. Moreover, FCA like other data mining algorithms, can easily integrate additional data and variables in the analysis as long as the design of the context tables are well defined.

For FCA to be truly useful and relevant in natural resources management, further refinement of the data mining approach through improved design of the context tables and more replications and time series coral reef assessment data should be done. If the object and attribution structure were expanded to include fish, seagrass, plankton, and water quality information, more nuanced implications may be discovered.

REFERENCES

Anit, K.P.A, J.M. Coral, C.B. Gigantone, V.R.M. Madrid, L. Mercado, K.C.E. Pardo, and R.R. Sazon. 2014. "Assessment of the Coastal Ecosystem of Barangay Ilijan, Batangas City, Batangas, Philippines. A Case Study for ENS 204 (Issues on Aquatic Resources)."

Aswani Kumar, Ch, and S. Srinivas. 2010. "Mining Associations in Health Care Data Using Formal Concept Analysis andand Singular Value Decomposition." Journal of biological Systems 18 (4): 787–807.

Bertaux, A., F. Le Ber, A. Braud, and M. Tremolieres. 2009. "Identifying Ecological Traits: A Concrete FCA-Based Approach." International Conference on Formal Concept Analysis edited by S. Ferre, 224-236. Berlin: Springer.

Burke L., Y. Kura, K. Kassem, C. Revenga, M. Spalding, and D. Mcallister. 2001. Pilot Analysis of Global Ecosystems: Coastal Ecosystems. World Resources Institute, Washington DC.

Cornish, Sarah. 2014. "Big Data Climate Challenge Winners Announced." Accessed June 25, 2016. http://www.unglobalpulse.org/big-data-climate-challenge-winners-announced

Ganter, B., and R. Wille. 1999. Formal Concept Analysis: Mathematical Foundation. Springer Verlag, Berlin Heidelberg.

Gebert, J., S. Motameny, U. Faigle, C.V. Forst, and R. Schrader. 2008. Identifying Genes of Gene Regulatory Networks Using Formal Concept Analysis. *Journal of Computational Biology* 15 (2): 185-194.

Jay, Nicolas, François Kohler, and Amedeo Napoli. 2008. "Using Formal Concept Analysis for Mining and Interpreting Patient Flows within a Healthcare Network." In *Concept Lattices and Their Applications*, 263-268. Springer: Berlin Heidelberg.

Jiménez, Daniel, Hugo Dorado, James Cock, Steven D. Prager, Sylvain Delerce, Alexandre Grillon, Mercedes Andrade Bejarano, Hector Benavides, and Andy Jarvis. 2016. "From Observation to Information: Data-Driven Understanding of On-Farm Yield Variation." PloS One 11 (3): e0150015.

Kaytoue, Mehdi, Sergei O. Kuznetsov, Amedeo Napoli, and Sébastien Duplessis. 2011. "Mining Gene Expression Data with Pattern Structures in Formal Concept Analysis." *Information Sciences* 181 (10): 1989–2001.

Kim, Gang-Hoon, Silvana Trimi, and Ji-Hyong Chung. 2014. "Big-data Applications in the Government Sector." *Communications of the ACM* 57 (34): 78–85.

Konotchick, T.H., D.I. Kline, D.B. Allison, N. Knowlton, S.J. Belongie, D.J. Kriegman, and B.G. Mitchell. 2006. "Computer Vision Technology Applications for Assessing Coral Reef Health." *EOS, Transactions, American Geophysical Union* 87 (36).

Poelmans, Jonas, Dmitry I. Ignatov, Sergei O. Kuznetsov, and Guido Dedene. 2013. "Formal Concept Analysis in Knowledge Processing: A Survey on Applications." *Expert Systems with Applications* 40 (16): 6538–6560.

Poelmans, J., P. Elzinga, S. Viaene, G. Dedene. 2010. "A Method Based on Temporal Concept Analysis for Detecting and Profiling Human Trafficking Suspects." Proceedings of the IASTED International Conference on Artificial Intelligence, Insbruck, Austria, 15-17 February.

Stumme, Gerd. 2002. Efficient Data Mining Based on Formal Concept Analysis." In *Database and Expert Systems Applications*, 534–546. Berlin: Springer.

Stumpfe, D., E. Lounkine, and J. Bajorath. 2011. Molecular Test Systems for Computational Selectivity Studies and Systematic Analysis of Compound Selectivity Profiles. In Chemoinformatics and Computational Chemical Biology: 503-515.

Valtchev, Petko, Rokia Missaoui, and Robert Godin. 2004. "Formal Concept Analysis For Knowledge Discovery and Data Mining: The New Challenges." In Proceedings of the International Conference on Formal Concept Analysis, 352–371. Berlin: Springer.

White, A.T. 2001. Philippine Coral Reefs: A Natural History Guide Manila: The Bookmark Inc.

Wormuth, Bastian, and Peter Becker. 2004. "Introduction to Formal Concept Analysis." In Paper presented at the 2nd International Conference of Formal Concept Analysis, February 2004.

Yevtushenko, Serhiy A. 2000. "System of Data Analysis 'Concept Explorer' (In Russian)." In Proceedings of the 7th National Conference on Artificial Intelligence KII-2000, 127–134.