Noldus
Information Technology

The Observer

Using Theme for data analysis

Theme™ is a software package for pattern detection in sequential data. It was originally developed for the analysis of social interactions. Theme uses raw data as input. Theme's algorithms are based on relationships in the order and timing of behaviors. These relationships are critical in any goal directed action. Thus the time structures detected by Theme often reflect the purpose of behavior.

Theme's pattern statistics offer sensitive measures of organization. Theme calculates numbers of detected patterns, length, diversity, amount of interaction between subjects and many more. They make it possible to measure 'qualitative' differences in a quantitative and objective way. Besides that, Theme can be used for in-depth studies of the structure of behavior. Theme can detect which events and behaviors drive the behavior, which follow, and which are just random.

Data from The Observer® are often very suitable for analysis with Theme. Many behavioral scientists focus on highly structured phenomena such as social interactions or the execution of tasks. Others focus on sequential behaviors that reflect stress or the effect of pharmacological substances, such as stereotypies. In any case there are time series of discrete and meaningful behaviors, which is exactly what Theme requires as input.

Step 1: Import and inspect your data

The first step in a Theme analysis is importing the data into Theme. From The Observer you import the configuration and the observations you want to analyze. Then you inspect the data. Depending on your research question you may want to exclude observations if they are much longer or shorter than the rest, if they contain many more or fewer data than others, if they contain long empty periods, etc.

Step 2: Rough-cut analysis

The first real analysis step is intended to get a quick overview of the time structures in the data. In this step you set the search parameters to relatively insensitive values for fast searching. Theme will leave out a lot of detail, and give only the more abundant and more significant patterns. Make sure you set all the observations to the same length for statistical validity.

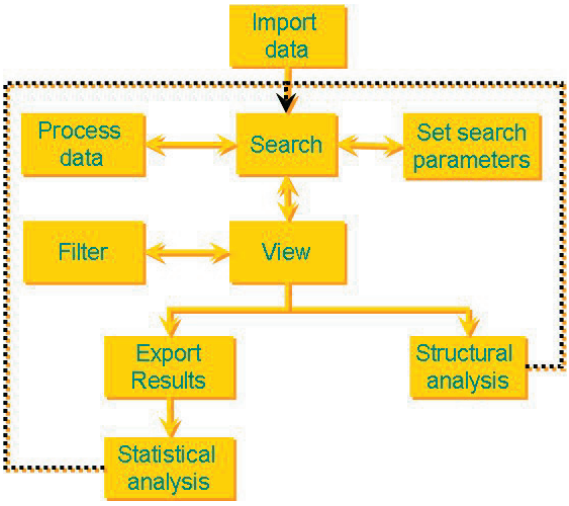

Figure 1. *Using Theme on your Observer data.*

After the search you have a table with many statistics for each observation, such as the number of patterns and their diversity. Before you do anything else, you check if these results could be a coincidence. For that you use a randomization function that was built into Theme. In most behavioral data you will find that the results are very significant.

The next step is exporting the statistics to a statistical package. With that you test for differences between your groups or treatments. Make sure you use non-parametric tests, because normal distributions are rare in behavioral data and even more so in Theme results. If you know what you are looking for, this statistical testing may be the endpoint of your analysis.

Step 3: Refine your search

After step 2 you probably have an idea of the important differences between your groups and treatments. You may have found that one group shows more or longer patterns than the other, etc. Or you did not find significant differences, but you think you will if you analyzed

the data in a different way. In either case you may want to refine your search, to get a clearer picture.

There are various options for refinement. You may want to change the input data by filtering out certain items, or by removing detail. Theme offers a number of features for that. You may also want to change the search settings, so that your results get a higher resolution. Or you may want to filter out part of the resulting patterns. You could focus on patterns of a minimum length, or patterns that contain a particular behavior. E.g. in a soccer match, you may be interested only in long attack patterns.

Step 4: zoom in on interesting patterns

Once you are satisfied with your search criteria you can zoom in on particular patterns that you find interesting. You may want to know if one soccer team shows more long attack patterns than the other, which players tend to work together, if there are patterns that lead to a goal, etc. You can filter your results and test statistically, or search for similar patterns in other datasets. For each pattern you get all important characteristics, such as the exact times of all the events, time intervals, and significance levels.

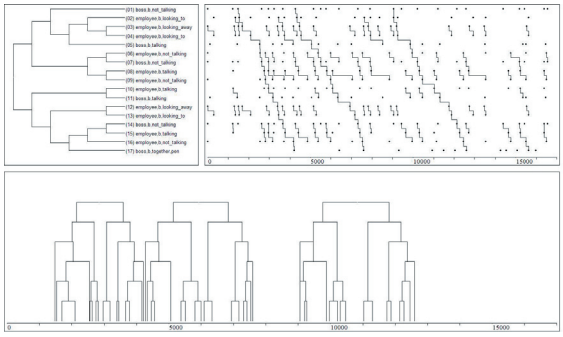

Figure 2. *Pattern diagram.*

Step 5: higher order analysis

If you want to make the most of your data set, you may want to go on with higher order analysis. Higher order analysis reveals structures and features that go beyond individual patterns. You may use these analyses to answer questions such as 'how do tasks alternate' or 'can I predict what is going to happen next time'.

An interesting feature is the Markers analysis: find behaviors that are linked one-to-one with patterns. Every time the behavior occurs, the entire pattern unfolds. The marker is a central element in the behavior, like the exchange of rings in a wedding mass. Event types in patterns analysis reveal the importance of individual behaviors in the behavior as a whole: some tend to start patterns, some tend to end them. Some occur in many different patterns, some in few. And some occur in none at all, even though they are quite abundant.

Network analysis reveals the relationships between individual behaviors, and which link to which. It is a summary matrix of all the time relationships between behaviors in the observation. Finally there is composition analysis: Theme finds the smallest set of alternating patterns that together cover the highest proportion of observation time. This analysis uncovers the higher level organization of behavior: how tasks or other activities alternate, and what proportion of the time is involved.

Step 6: other searches

After finishing your analysis there are still other possibities to explore. There may be highly significant patterns that occur once per observation or less. A Christmas dinner occurs only once a year but it is very complex and regular. You will never find such patterns, unless you join all observations into one. Theme can do so, based on independent variables or for the entire project.

By default Theme looks for events that are relatively close to each other in time. That is not the whole story: some events may be far apart, but at regular intervals. Lunch follows breakfast within about five hours. You may miss that with default settings. If you also want such patterns you can include 'free' patterns in your search.

International Headquarters
Noldus Information Technology bv
Wageningen, The Netherlands
Phone: +31-317-473300
Fax: +31-317-424496
E-mail: info@noldus.nl

North American Headquarters
Noldus Information Technology Inc.
Leesburg, VA, USA
Phone: +1-703-711-0440
Toll-free: 1-800-355-9541
Fax: +1-703-771-0441
E-mail: info@noldus.nl

We are also represented by a worldwide network of distributors and regional offices.

For the nearest regional office or distributor, please visit www.noldus.com/contact

Due to our policy of continuous product improvement, information in this document is subject to change without notice. The Observer is a registered trademark of Noldus Information Technology bv. Theme is a trademark of PatternVision Ltd.

© 2006 Noldus Information Technology bv. All rights reserved.

www.noldus.com