# LAGRANGE - an equation discovery system: User's manual

## 1 Introduction to LAGRANGE

LAGRANGE [1, 2] is an equation discovery system capable of discovering differential as well as ordinary algebraic equations involving more than two variables from observational data only. This manual contains only brief instructions for installing and using the equation discovery system LAGRANGE. For additional information consult [1, 2].

## 2 Installation

LAGRANGE is implemented in the C programming language using the Gnu `gcc` compiler. It was developed on a Sun workstation working under the SunOS operation system. It was successfully compiled and run on Hewlett Packard (HP-UX 10.01), Sun (SunOS 4.1.3) and PC (under the Windows 95/NT and Linux operating systems, compiled using the `gcc` compiler) workstations.

Once you have retrieved the distribution file `lg-www.tar.Z`, you have to extract the files with the following command:

```
% uncompress lg-www.tar.Z ; tar xf lg-www.tar
```

To build the executable file `lg`, type `make all` in the directory `Lagrange/src`. Later, after the software starts working, you can remove the compiled code (`.o` files) and other intermediate files by typing `make clean` in the same directory.

The directory (`Lagrange/doc`) contains postscript versions of the papers referenced in this manual and the manual itself. Several example domain files are included within this distribution in the third directory `Lagrange/Inputs`.

## 3 Using LAGRANGE

LAGRANGE can be run using the following command:

% `lg [-R` $t_R$`] [-S` $t_S$`] [-o` $o$`] [-d` $d$`] [-r` $r$`] domain_file_name`

Parameters $t_R$, $t_S$, $o$, $d$ and $r$ are optional. If an option is omitted, the default value for the corresponding parameter is used. For a description of the parameters consult [1, 2]. We only comment on the use of the options `-R` and `-S`. The following rules are applied.

- If both parameters are given, LAGRANGE uses both significance criteria with corresponding threshold values for evaluating equations. If none of them are given, LAGRANGE uses both criteria with the default threshold values.

- If only one parameter is given, LAGRANGE uses only the corresponding significance criterion for evaluating equations.

We use the obligatory parameter `domain_file_name` to determine the file with the description of the input domain. A domain input file consists of three parts:

1. The time step $h$ between two successive equidistant time points in the table of measurements, given in part 3.
2. List of the names of system variables. Optionally, units of measure for each variable can be provided. The first character of the variable name has to be a letter, the remaining characters can be letters, digits or the underline character (_). If the measurements for the time derivative of the variable $X$ is provided within the table, it should be names $X'$. In that case LAGRANGE will not calculate the numerical derivative of the system variable $X$.
3. Table of measured values of the dynamic system variables used for generating equations.

The formal definition of the input files format in the form of a Bacchus-Naur grammar is given in Appendix A.

The output of LAGRANGE first lists the parameters settings. Than, the list of discovered equations along with their significance criteria values, and finally the number of all equations considered during the equation discovery process.

## References

[1] Džeroski, S. and Todorovski, L. (1993) Discovering dynamics In *Proc. Tenth International Conference on Machine Learning*, pages 97–103. Morgan Kaufmann, San Mateo, CA.

[2] Džeroski, S. and Todorovski, L. (1995) Discovering dynamics: from inductive logic programming to machine discovery. *Journal of Intelligent Information Systems*, 4: 89–108.

## A Lagrange input files format

The Bacchus-Naur grammar specifying the format of the Lagrnage input files is as follows:

$$
\begin{array}{lcl}
\mathit{InFile} & \longrightarrow & \mathit{TimeStep}\ \ \mathit{VarDecls}\ \ \mathit{Measures} \\[1em]
\mathit{TimeStep} & \longrightarrow & \mathbf{number} \\[1em]
\mathit{VarDecls} & \longrightarrow & \mathit{VarDecls}\ \ \mathit{VarDecl} \mid \mathit{VarDecl} \\[1em]
\mathit{VarDecl} & \longrightarrow & \mathit{VarId}\ \ \mathit{VarType}\ ; \\[1em]
\mathit{VarId} & \longrightarrow & \mathbf{identifier}\ \mid \mathbf{identifier}\ ' \\[1em]
\mathit{VarType} & \longrightarrow & :\ \mathit{VarDim} \mid \epsilon \\[1em]
\mathit{VarDim} & \longrightarrow & [\ \mathit{Product}\ /\ \mathit{Product}\ ] \mid [\ \mathit{Product}\ ] \mid [\ ] \\[1em]
\mathit{Product} & \longrightarrow & \mathit{Product} * \mathbf{identifier} \mid \mathbf{identifier} \\[1em]
\mathit{Measures} & \longrightarrow & \mathit{Measure}\ ;\ \mathit{Measures} \mid \mathit{Measure}\ ; \\[1em]
\mathit{Measure} & \longrightarrow & \mathit{Measure}\ \mathbf{number} \mid \mathbf{number} \\[2em]
\mathbf{number} & \longrightarrow & \mathbf{sign\ digits} \mid \mathbf{sign\ digits}\ .\ \mathbf{digits} \mid \mathbf{sign\ digits}\ .\ \mathbf{digits\ E\ sign\ digits} \\[1em]
\mathbf{digits} & \longrightarrow & \mathbf{digits\ digit} \mid \mathbf{digit} \\[1em]
\mathbf{identifier} & \longrightarrow & \mathbf{letter\ string} \\[1em]
\mathbf{string} & \longrightarrow & \mathbf{string\ alphanum} \mid \epsilon \\[1em]
\mathbf{alphanum} & \longrightarrow & \mathbf{letter} \mid \mathbf{digit} \mid \_ \\[1em]
\mathbf{letter} & \longrightarrow & \mathbf{a} \mid \mathbf{A} \mid \mathbf{b} \mid \mathbf{B} \mid \ \ldots\ \mid \mathbf{z} \mid \mathbf{Z} \\[1em]
\mathbf{digit} & \longrightarrow & \mathbf{0} \mid \mathbf{1} \mid \ \ldots\ \mid \mathbf{9} \\[1em]
\mathbf{sign} & \longrightarrow & + \mid - \mid \epsilon
\end{array}
$$