CLAIMS

1. Chemical reaction encoding software for one-step, multi-step and equilibrium reactions, characterised in that it executes instructions corresponding to the following steps:

- a step (105) of receiving, upon a computer interface, a chemical reaction graph comprising at least one chemical reaction reagent (hereafter “reagent”) and at least one chemical reaction product (hereafter “product”),
- a first step (110) of encoding into a string of characters independent of atom numbering, by a computing device, said chemical reaction graph describing the structure of said at least one reagent and said at least one product,
- a step (115) of determination, by a computing device, of changing bonds within the encoding representative of the chemical structures of said at least one reagent and said at least one product,
- a second step (120) of encoding, by a computing device, in a single string of characters, for at least one changing bond determined, at least one character representative of an atom subject to the change of bond, at least one character representative of the type of changing bond determined and at least one character representative of an atom resulting from the change of bond, in which a changing bond is encoded by a set of two characters representative of the changing bond determined, the first character being representative of the reagent bond and the second character being representative of the product bond, each character being selected in a library of bijective characters wherein one character is representative of only one bond type and
- a step (125) of providing, upon a computer interface, the string of characters corresponding to the encoding of changing bonds of the chemical reaction and representative of the chemical structures of said at least one reagent and said at least one product.

2. Software according to claim 1, in which the second (120) step of encoding is configured to embed the two characters representative of the changing bonds determined in between two neutral tag characters representative of the presence of an encoding of said changing bonds.

3. Software according to any one of claims 1 or 2, in which multistep reactions, represented by a succession of change of bonds between two atoms, are encoded by a succession of single characters, each single character being representative of the successive state of a bond between said two atoms, the order of the characters being representative of the order of changes of bonds between said two atoms.

4. Chemical reaction encoding method (100) for one-step, multi-step and equilibrium reactions, characterised in that it comprises:

- a step (105) of receiving, upon a computer interface, a chemical reaction graph comprising at least one chemical reaction reagent (hereafter “reagent”) and at least one chemical reaction product (hereafter “product”),
- a first step (110) of encoding into a string of characters independent of atom numbering, by a computing device, said chemical reaction graph describing the structure of said at least one reagent and said at least one product,
- a step (115) of determination, by a computing device, of changing bonds within the encoding representative of the chemical structures of at least one said reagent and said at least one product,
- a second step (120) of encoding, by a computing device, in a single string of characters, for at least one changing bond determined, at least one character representative of an atom subject to the change of bond, at least one character representative of the type of changing bond determined and at least one character representative of an atom resulting from the change of bond, in which a changing bond is encoded by a set of two characters representative of the changing bond determined, the first character being representative of the reagent bond and the second character being representative of the product bond, each character being selected in a library of bijective characters wherein one character is representative of one changing bond type and
- a step (125) of providing, upon a computer interface, the string of characters corresponding to the encoding of changing bonds of the chemical reaction and representative of the chemical structures of said at least one reagent and said at least one product.

5. Method (100) according to claim 4, in which multistep reactions, represented by a succession of change of bonds between two atoms, are encoded by a succession of single characters, each single character being representative of the successive state of a bond between said two atoms, the order of the characters being representative of the order of changes of bonds between said two atoms.

6. Method (100) according to any one of claims 4 to 5, in which the first step (110) of encoding is configured to encode the chemical reaction graph into a line notation, the method further comprising, prior to the second step (120) of encoding, a step (130) of augmenting the line notation encoding.

7. Method (100) according to any one claims 4 or 6, in which the second step (120) of encoding comprises a step (121) of extracting, by a computing device, of a bond table for reagents and products from a computer memory, said encoding being performed as a function of said bond table.

8. Method (100) according to any one of claims 4 to 7, in which the second step (120) of encoding comprises a step of removing, from the first encoding resulting from the first step (110) of encoding, of at least one atom identifier from at least one reagent and/or product, each said atom being removed as a result of the step (115) of determination in the event said atom and the associated bonds are located in a product and/or reagent that remains unchanged from reagent the reaction stage to the product stage of the chemical reaction.

9. Method (100) according to any one of claims 4 to 8, which comprises a step (135) of obtaining the products of the encoded chemical reaction by performing said chemical reaction in a physical device.

10. Chemical reaction dataset augmentation method (300), characterised in that it comprises:

- a step (305) of receiving, upon a computer interface, a string of characters obtained according to the method of claims 4 to 9,
- a step (310) of reordering, by a computing system, the string of characters in order to shift at least one character representative of an atom and at least one

string of at least one character representative of a change of bond associated to the corresponding atom and

- a step (320) of outputting, upon a computer interface, an augmented string of characters corresponding to the reaction initially encoded by the received string of characters.

11. Augmentation method (300) according to claim 10, which further comprises a step (315) of associating, by a computing system, at least two strings of characters obtained according to the method of claims 4 to 9, each said string of characters being representative of the same chemical reaction graph.

12. Chemical reaction dataset preprocessing method (400), characterised in that it comprises:

- a step (405) of receiving, upon a computer interface, a dataset of at least two chemical reaction graphs comprising at least one chemical reaction reagent and at least one chemical reaction product,
- a step (100) of encoding of at least two chemical reaction graphs according to the method according to any one of claims 4 to 9,
- a step (410) of determining, by a computing system, a distribution of chemical reaction classes within the encoded dataset,
- a step (300) of augmenting the dataset, according to the method according to either claim 7 or 8, for at least one chemical reaction class as a function of the determined distribution and
- a step (415) of outputting, upon a computer interface, the preprocessed dataset.

13. Training method (500) for a classifier, transformer or regressor, characterised in that it comprises:

- a step (505) of inputting, upon a computer interface, a dataset of chemical reaction graphs encoded according to the method of claim 8,
- a step (510) of operating, by a computing system, a recursive neural network architecture configured to use, as input, the dataset of chemical reaction graphs to classify the chemical reaction bond evolution as a function of the input and
- a step (515) of outputting, upon a computer interface, a trained classifier, transformer or regressor.

14. Chemical reaction bond evolution prediction method, characterised in that it operates a classifier, transformer or regressor obtained by the method (500) according to claim 13.

15. Chemical reaction generation method, characterised in that it operates a classifier, transformer or regressor obtained by the method (500) according to claim 13.

16. Computer implemented classifier, characterised in that the classifier, transformer or regressor is obtained by the method (500) according to claim 13.

17. Computer program, characterised in that it comprises instructions to operate a method (500) according to either claim 13 to 15.