**CPSC 311 – Term Project Proposal**

Expression to Diagram (ETD language)

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Student #** | **CS ID** | **E-Mail** |
| Zhe Li | 88792486 | d3q1b | lizhe918@alumni.ubc.ca |
| Maxwell Yang | 96245428 | r0g2b | shukany@alumni.ubc.ca |
| Eric Yan | 22548424 | g5h2b | ericy676@students.cs.ubc.ca |
| Yifei Wang | 26530162 | w9a2b | yifei.wang.27@gmail.com |
| Hao Wang | 97951503 | x9e2b | h.wang.81@alumni.ubc.ca |

**Introduction**

The Expression to Diagram (or simply ETD pronounced “Leco”) is a domain-specific language designed to autonomously generate different types of diagrams such as circuits and binary trees from user-provided logical expressions. To elaborate, the user can enter a logical expression in a concrete syntax similar to the Racket language into a Racket file. The ETD then parses the given expression into an abstract syntax and proceed to interpret it into a corresponding diagram. We opted to use Racket because we think making a diagram is suitable for functional programming, and it is better to choose a language that we all know. The primary motivation behind the development of this language is to provide a clear visual representation of unfamiliar academic concepts for people. The language may also be used by course staff to help the development of new exam/homework questions, as well as solutions to some problems like logical circuit design and binary tree problems.

**Project Goals**

Our core goal is for the language to be able to convert simple expressions that have only tree structures (without loops) into their corresponding diagram. It is easy to think about them as a recursive structure and process them recursively. We think the EBNF of this language should be as simple as the picture below. The id represents the type of a node such as and-gate and tree-node. It also supports the expression that allows us to avoid very long duplicated expressions.

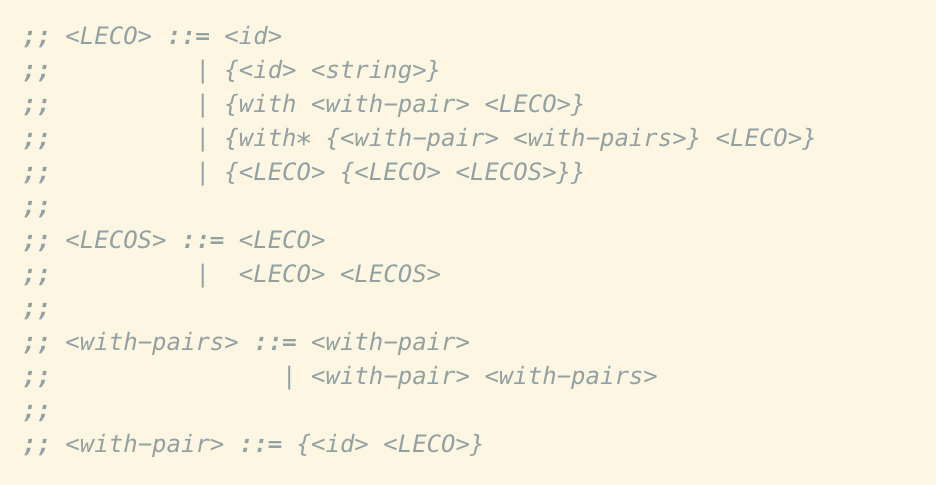


Figure 1. The EBNF of LECO

For our full goal, we want to generate complex diagrams with loops such as graphs and circuits that use some outputs as inputs. It seems that it is still easy to connect the nodes, but we think it is very complex to leave spaces for these structures and show the user a clear diagram without overlapping or too many crosses. Also, it is not easy to design and deal with the expression for the diagrams with loops, even the regular logical expression does not support the circuit with loops, and we might need to do something like “parse by reference” to achieve this goal. Another idea that we want to achieve is to have different coloring for different lines to show the relationship between nodes, which may also show the status such as the signal on the line. We think this is not very hard, but it can be time-consuming.

**Milestone Plans**

To realize the language, we split the project into multiple milestones. The first step in our strategy is to perform sufficient background research on relevant topics. The central research topic would be ideas of drawing diagrams in functional programming. This milestone would involve researching the data structure of the diagram, how to draw things recursively, diagram generation, image creation for functional programming (or recursive algorithms). All group members will contribute to the research of these core topics above.

Furthermore, as the Racket 2htdp/image library is the only way we use to draw on the screen, we will dedicate 1-2 group members to focus on the documentation, tutorials, and demos involving the use of the library. Some other topics that we might also research include syntax choices, impact in educational settings, and possible visual designs. Once we have completed most of the research (1 week before background research report submission), we will compile our results into a single background research document highlighting our research on the topic.

The next milestone of our project is proof of concept and plan. At this stage, we will clarify our “minimal” core goals and provided some of our more “ambitious” full goals. To demonstrate our capability for accomplishing the core goals, we will elaborate on how to achieve some critical components of the project by providing written explanations, diagrams and code snippets. Some of the exciting parts that currently come to us include how to construct the parse, how to define and store each node, how we might recursively draw the diagram. We will also provide possible solutions to achieve the “ambitious” full goals. They will be in a similar format as the minimal goals. Considering the importance of achieving the core goals first, all members of the group will collaborate on providing methods to achieve them. We will also attempt to build a language that supports most of the features that we specified in the main goals. After completing this phase, we will move on to the full goals.

For the poster session, we will make a poster that illustrates how our project works. We will show the several types of diagrams generated by our project and describe step by step on how we approach the final version. The first part of the poster is parsing and desugaring. We won’t spend much time on this part since the parsing tool we use is essentially the same as the one used in the lecture. The next part is interpretation. We will give examples of how we convert the expression to a specific data structure State. This data structure is used to record all the print stuff in the lecture, but here we will store the location of nodes and relationships between them. The last part will be the process of rendering. If the State stores a string, it can be printed with a straightforward expression, but what we have is the information of a diagram, so what we need is to draw every node and line depend on their types on the screen. Our poster will mostly be composed of diagrams such as logic circuits and a flow chart of how we approach the final project. We will try to make our poster presentation clear and exciting.

**Literature Review**

In the article about graphical data flow programming language (Hunt, 1990), the author created a language that can generate flow charts, including a full set of data types and control flow operators. Similar to the logic circuits, the language of the data flow chart involves a distinct definition of each operator and the way to connect them. We can obtain some useful design concepts from this language.

Another article (Ren, 2002) is about the language of generating abstract syntax trees. The author introduced the design and implementation of the parser and implementation of the tree generator. The design of the language was beginning with the general architecture of the language, and then the author introduced the implementation of the parser detailed with examples. He also mentioned some critical issues to consider about the parser, such as supporting grammar, which gives us some inspiration for designing our language. The author also provided source codes of this language, which is considered to be a useful reference.

**References**

Hunt, N. (1990, November). IDF: A graphical data flow programming language for image processing and computer vision. In 1990 IEEE International Conference on Systems, Man, and Cybernetics Conference Proceedings (pp. 351-360). IEEE.

Ren, C. L. (2002). Parsing and abstract syntax tree generation in the GIPSY system (Doctoral dissertation, Concordia University).

(n.d.). Retrieved from <https://docs.racket-lang.org/teachpack/2htdpimage.html?q=2htdp>.

(n.d.). Retrieved from https://futtetennismo.me/posts/algorithms-and-data-structures/2017-12-08-functional-graphs.html.