Assignment of bachelor's thesis

Title:	Compilation of a typed R-like language to WebAssembly
Student:	Bilguudei Baljinnyam
Supervisor:	Pierre Donat-Bouillud, Ph.D.
Study program:	Informatics
Branch / specialization:	Computer Science 2021
Department:	Department of Theoretical Computer Science
Validity:	until the end of summer semester 2026/2027

Instructions

R is an open-source programming language and software environment widely used for statistical computing and data analysis. Originally developed for statisticians, it has become the de facto standard in many scientific disciplines including bioinformatics, data analysis, finance, or data mining. In R, vectors are first-class citizens, as even 1 is represented as a vector of size 1. R is a dynamic language and is notoriously difficult to compile; it is currently only JIT-compiled in experimental projects [1]. WebAssembly [2] is an open, safe and portable compilation target for programming languages, enabling deployment in the web or locally. Currently, the R Interpreter can be compiled to WebAssembly [3] but R programs themselves cannot be.

The goal of the thesis is to define and compile a statically-typed subset of R, or R-like. It should be simple enough to be able to compile it, but expressive enough to express operations on vectors and perform statistical operations. The target architecture will be WebAssembly.

1) Research the R language and WebAssembly.
2) Define and formalize a typed subset of R. It should at least support vector operations, simple control-flow instructions, and function calls.
3) Write a compiler from this subset to WebAssembly
4) Write a runtime for the subset.
5) Evaluate the compiler and the runtime.

---

[1] Flückiger, O., Chari, G., Yee, M. H., Ječmen, J., Hain, J., & Vitek, J. (2020). Contextual dispatch for function specialization. Proceedings of the ACM on Programming Languages, 4(OOPSLA), 1-24.
[2] https://webassembly.org/
[3] https://github.com/r-wasm/webr