LangSec: Fourth Workshop on Language Theoretic Security

The Language-theoretic approach (LangSec) regards the Internet insecurity epidemic as a consequence of ad hoc input handling. LangSec posits that the only path to trustworthy software that takes untrusted inputs is treating all valid or expected inputs as a formal language, and treating the respective input-handling routines as a recognizer for that language. Only then can any correctness guarantees be assured for the input-handling code. Ambiguity of message/protocol specification is insecurity; ad hoc parsing is an engine of exploitation; overly complex syntax can make judging security properties of input impractical or even undecidable.

Treating input-handling code as an automaton allows the defender to reason about its behavior. The more limited computational power of the automaton, the easier the reasoning. The root cause of many bugs, memory corruptions, and exploitation is trying to validate inputs with inappropriate automata (e.g., much of XSS is due to “validating” context-free HTML with regexps). The recognizer automaton should be just as powerful as warranted by the message format, and no more; unnecessary complexity is computational power given to the attacker.

LangSec is mission assurance for connected software and hardware exposed to attacks via malicious inputs--through a practical data and code co-design methodology and filtering of legacy formats down to safe subsets. LangSec explains why ad hoc “input sanitization”, “sanity checking”, and other advice to be more careful with inputs is not enough, and why numerous secure programming initiatives have not ended input-driven exploitation. LangSec is also a code and protocol auditing methodology.

The goal of the workshop is to bring more clarity and focus to two complementary areas: (1) practical assurance of input-handling code and (2) analysis of input-related weaknesses, vulnerabilities, and resulting exploitation mechanisms. The LangSec Workshop solicits input on these and related topics from the software verification, programming languages, and offensive research communities.

Organizing Committee

Sergey Bratus, Dartmouth College
Daniel 'TQ' Hirsch, P3KI GmbH
Felix ‘FX’ Lindner, Recurity Labs / Phenoelit
Michael E. Locasto, SRI
Meredith L. Patterson, Mautinoa Technologies / Upstanding Hackers, Inc.
Anna Shubina, Dartmouth College
Julien Vanegue, Bloomberg