

CS328, Winter 2002, Test 1

Time 115 min. Open notes/books. Use extra paper as needed, but make sure to identify each answer.

YOU MUST RETURN THIS PAGE. NAME _____

- 1 You have the source for a P compiler for machine M (just the source), and you have P interpreter running on M, and you have the machine M available.
 - a) Show what you have
You write a program PrgP in language P.
 - b) Show how to execute by interpretation
 - c) Show how to generate machine code in M for the program PrgP, using the compiler P->M that you have
 - d) Is there any inefficiency in step c) and how would you improve that? Show all step.

- 2 You have an assembler where each instruction takes 1 argument. The instruction takes 8 bits, and the argument takes 16 bits (total of 3 bytes). Data is in 2 bytes. Assembly arguments can be either symbols or immediate values, and there is no literal pool (machine code arguments will have immediate values as well as addresses). For example, we may write either
Add X
Add 20
 - a) How many different instructions can we have? (different operations)
 - b) Assuming no base register, how big a program can be?

- 3 Show machine code, in dec digits for individual bytes, for this program
LOAD X
SUB 10
STORE Y
BOGUS
JUMP EXIT
X CONST 200
EXIT: STORE X
STOP
Y CONST 100
STOP and BOGUS take no arguments and use 1 byte, other opcodes take 1 byte and each argument takes 2 bytes. Assume opcodes are Add=00, Load=10, Sub=20, Stop=30, Jump=40, Store=50, Bogus=60. Data storage is allocated in the same segment exactly where it appears in the program except all data must be even aligned on even byte boundary and takes 2 bytes. Immediate values are placed in a literal pool, 1 byte each, no alignment, and the pool is allocated at the end of the target.

- 4 Suppose we deal with the binary alphabet {0,1}. Tokens come in separated by WS. We want to recognize tokens of the form 0^21^2 , where $0^2/1^2$ indicates 0/1 repeated 2 times. Design a DFA.

- 5 Redo the above for 0^n1^n , for any $n > 2$. If you cannot, why? Would it help if a nondeterministic FA was allowed? Why?
- 6 We have a language where the only tokens are unsigned integers, with the usual WS separators. There are two data types, `char` and `long`. `char` means number is representable in 1 byte, all other numbers are of type `long`. Design a tokenizer DFA that will recognize whether a token is `char` or `long`.
 - a) what is the alphabet
 - b) what are the tokens
 - c) show the DFA as a graph
 - d) give an example when an error would take place