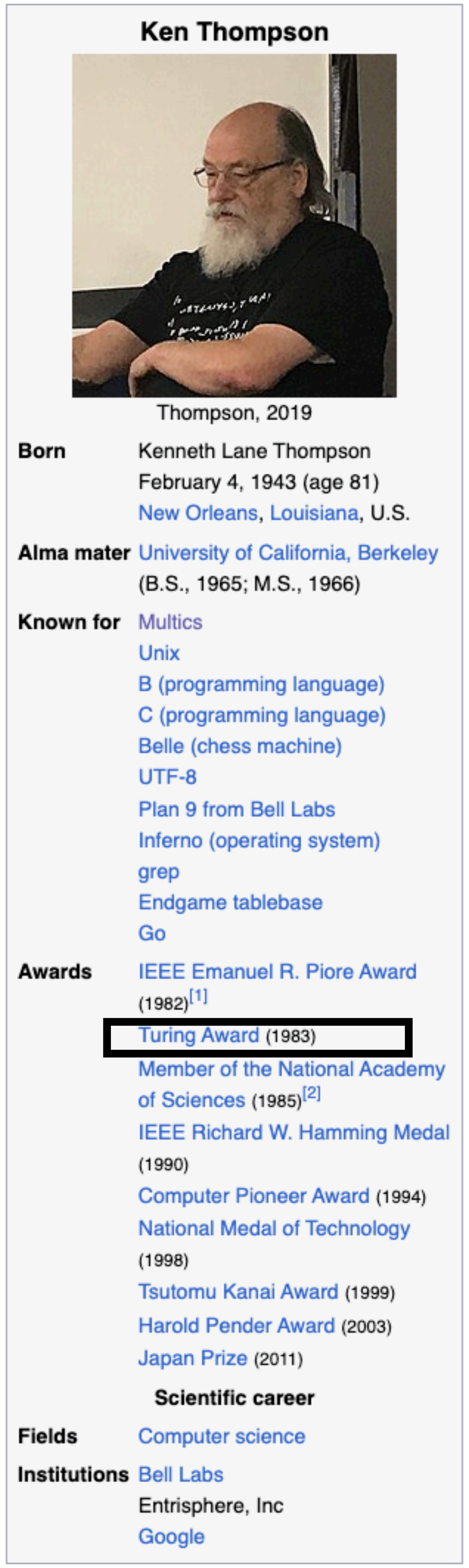
CS202 (003): Operating Systems
Trusting Trust

Last time



Ken Thompson

Thompson, 2019

Born	Kenneth Lane Thompson February 4, 1943 (age 81) New Orleans, Louisiana, U.S.
Alma mater	University of California, Berkeley (B.S., 1965; M.S., 1966)
Known for	Multics Unix B (programming language) C (programming language) Belle (chess machine) UTF-8 Plan 9 from Bell Labs Inferno (operating system) grep Endgame tablebase Go
Awards	IEEE Emanuel R. Piore Award (1982)[1] Turing Award (1983) Member of the National Academy of Sciences (1985)[2] IEEE Richard W. Hamming Medal (1990) Computer Pioneer Award (1994) National Medal of Technology (1998) Tsutomu Kanai Award (1999) Harold Pender Award (2003) Japan Prize (2011)

Scientific career

Fields	Computer science
Institutions	Bell Labs Entrisphere, Inc Google

Did you do the reading?

To what extent should one trust a statement that a program is free of Trojan horses? Perhaps it is more important to trust the people who wrote the software.

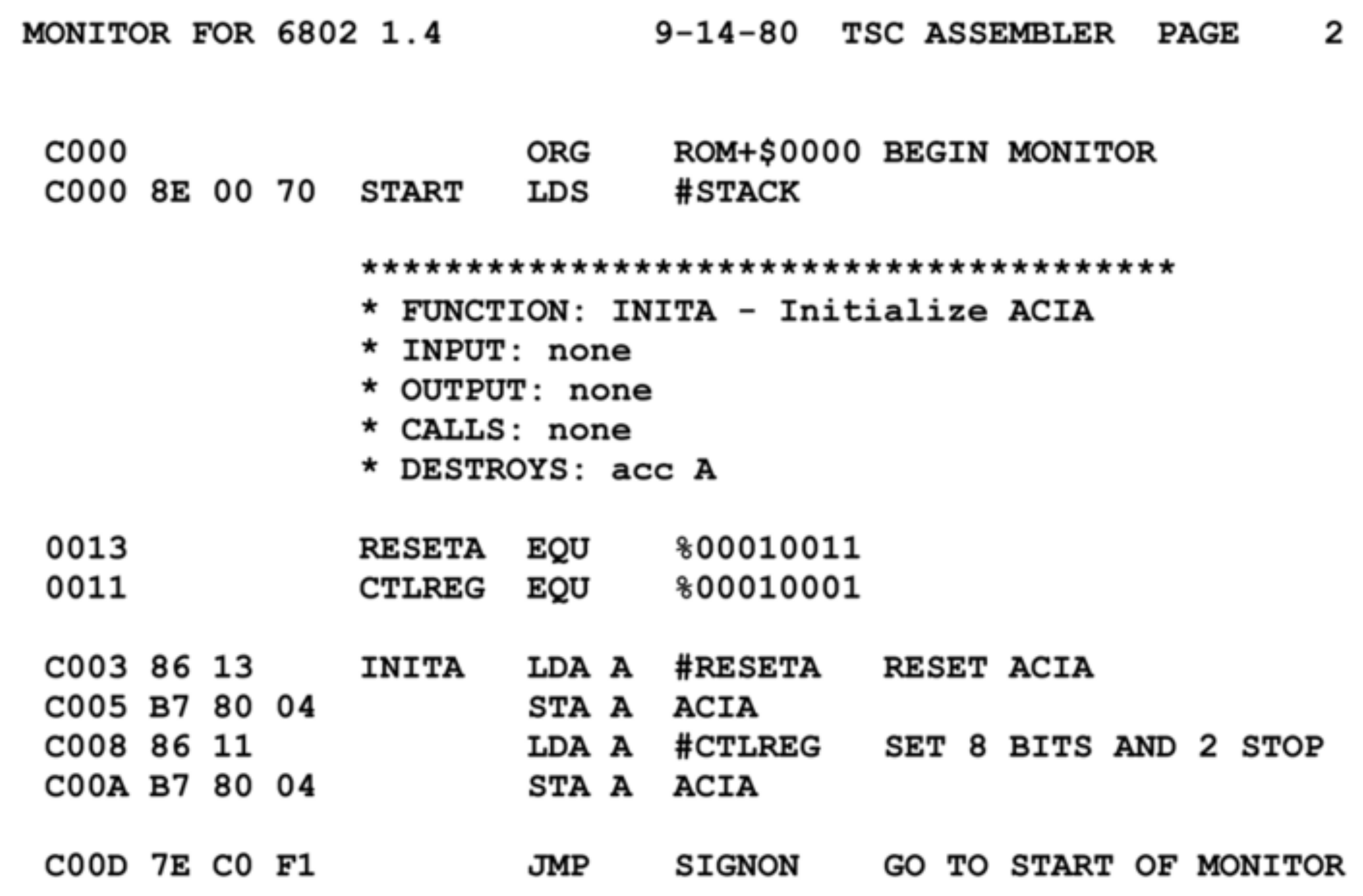
https://en.wikipedia.org/wiki/Ken_Thompson

Forget about what you read for a sec…

Java

C

Compiler $\xrightarrow{\hspace{3cm}}$



```
MONITOR FOR 6802 1.4          9-14-80  TSC ASSEMBLER  PAGE    2

 C000                    ORG    ROM+$0000 BEGIN MONITOR
 C000 8E 00 70  START    LDS    #STACK

                ****************************************
                * FUNCTION: INITA - Initialize ACIA
                * INPUT: none
                * OUTPUT: none
                * CALLS: none
                * DESTROYS: acc A

 0013           RESETA   EQU    %00010011
 0011           CTLREG   EQU    %00010001

 C003 86 13     INITA    LDA A  #RESETA   RESET ACIA
 C005 B7 80 04           STA A  ACIA
 C008 86 11              LDA A  #CTLREG   SET 8 BITS AND 2 STOP
 C00A B7 80 04           STA A  ACIA

 C00D 7E C0 F1           JMP    SIGNON    GO TO START OF MONITOR
```

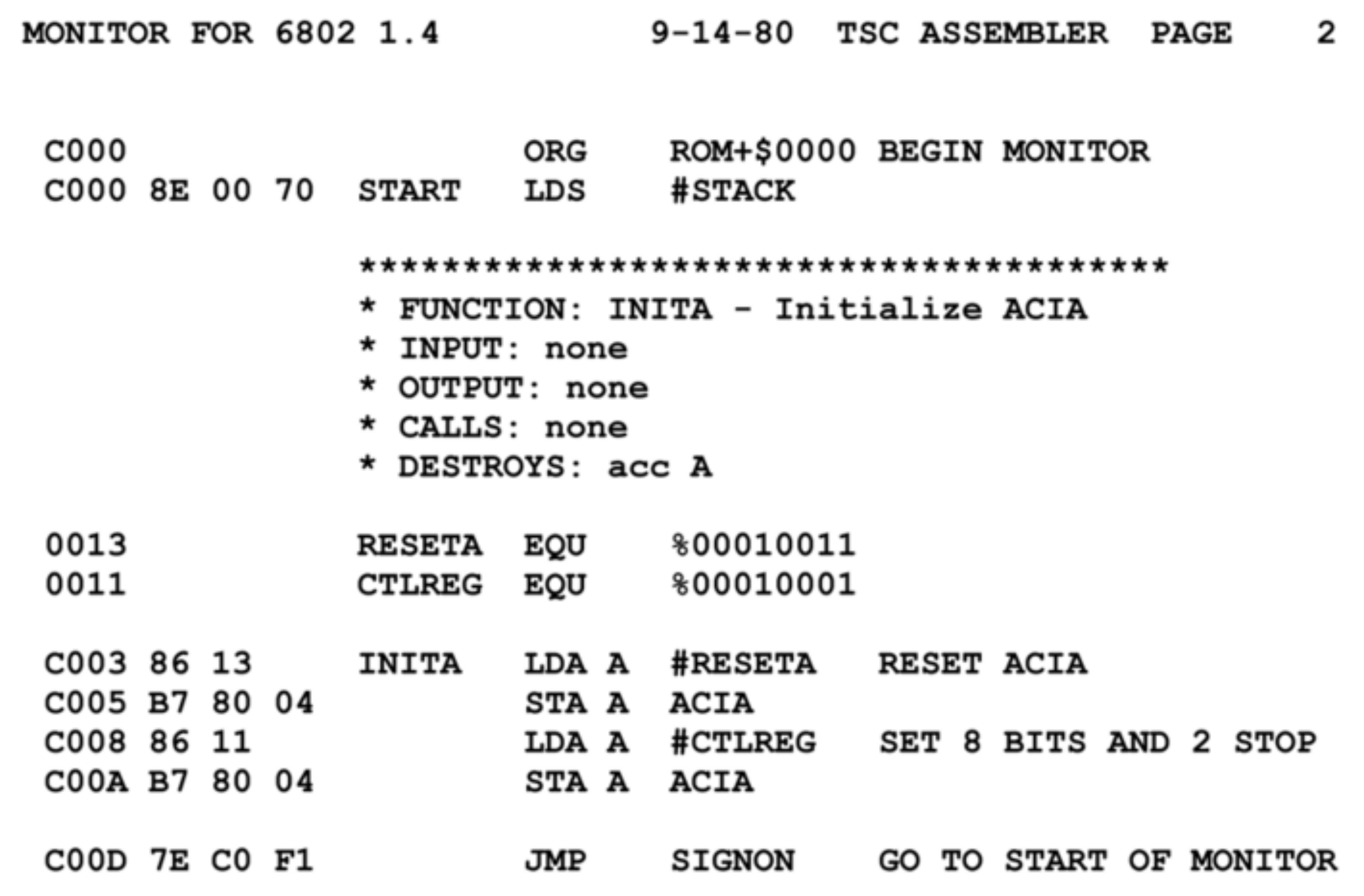
Compiler is a program.
So what does this program written in?

Forget about what you read for a sec…

Java

Compiler written in

C



```
MONITOR FOR 6802 1.4          9-14-80   TSC ASSEMBLER   PAGE    2

 C000                   ORG     ROM+$0000 BEGIN MONITOR
 C000 8E 00 70   START  LDS     #STACK

                 ***************************************
                 * FUNCTION: INITA - Initialize ACIA
                 * INPUT: none
                 * OUTPUT: none
                 * CALLS: none
                 * DESTROYS: acc A

 0013            RESETA EQU     %00010011
 0011            CTLREG EQU     %00010001

 C003 86 13      INITA  LDA A   #RESETA    RESET ACIA
 C005 B7 80 04          STA A   ACIA
 C008 86 11             LDA A   #CTLREG    SET 8 BITS AND 2 STOP
 C00A B7 80 04          STA A   ACIA

 C00D 7E C0 F1          JMP     SIGNON     GO TO START OF MONITOR
```

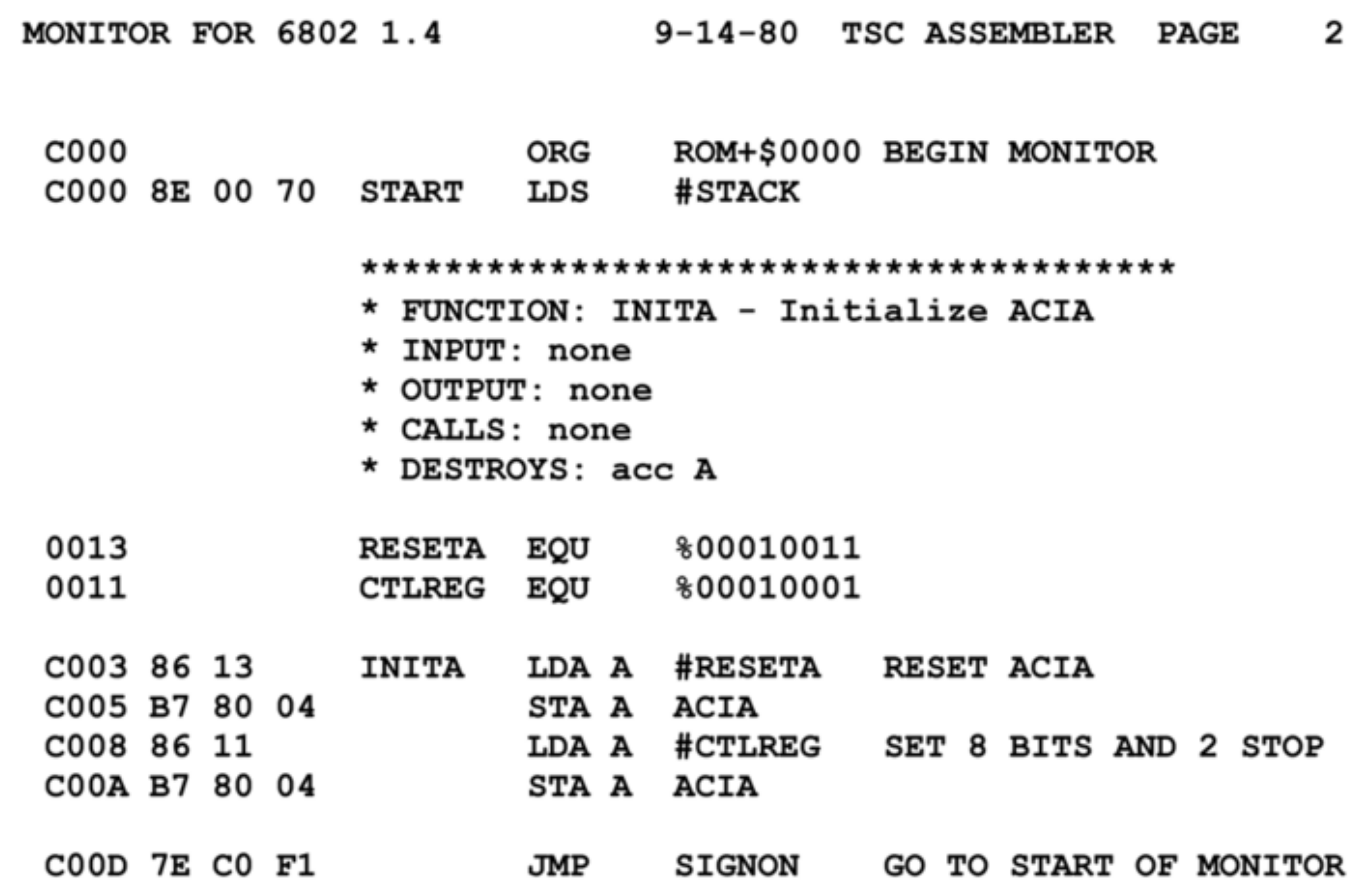
How does compiler know how to translate different types of language features (conditionals, loops, classes) into another language?

Forget about what you read for a sec…

Java

Compiler written in

C



```
MONITOR FOR 6802 1.4          9-14-80  TSC ASSEMBLER  PAGE    2

 C000                   ORG    ROM+$0000 BEGIN MONITOR
 C000 8E 00 70  START   LDS    #STACK

               ***************************************
               * FUNCTION: INITA - Initialize ACIA
               * INPUT: none
               * OUTPUT: none
               * CALLS: none
               * DESTROYS: acc A

 0013           RESETA  EQU    %00010011
 0011           CTLREG  EQU    %00010001

 C003 86 13     INITA   LDA A  #RESETA   RESET ACIA
 C005 B7 80 04          STA A  ACIA
 C008 86 11             LDA A  #CTLREG   SET 8 BITS AND 2 STOP
 C00A B7 80 04          STA A  ACIA

 C00D 7E C0 F1          JMP    SIGNON    GO TO START OF MONITOR
```

How can we add new language features to Java?

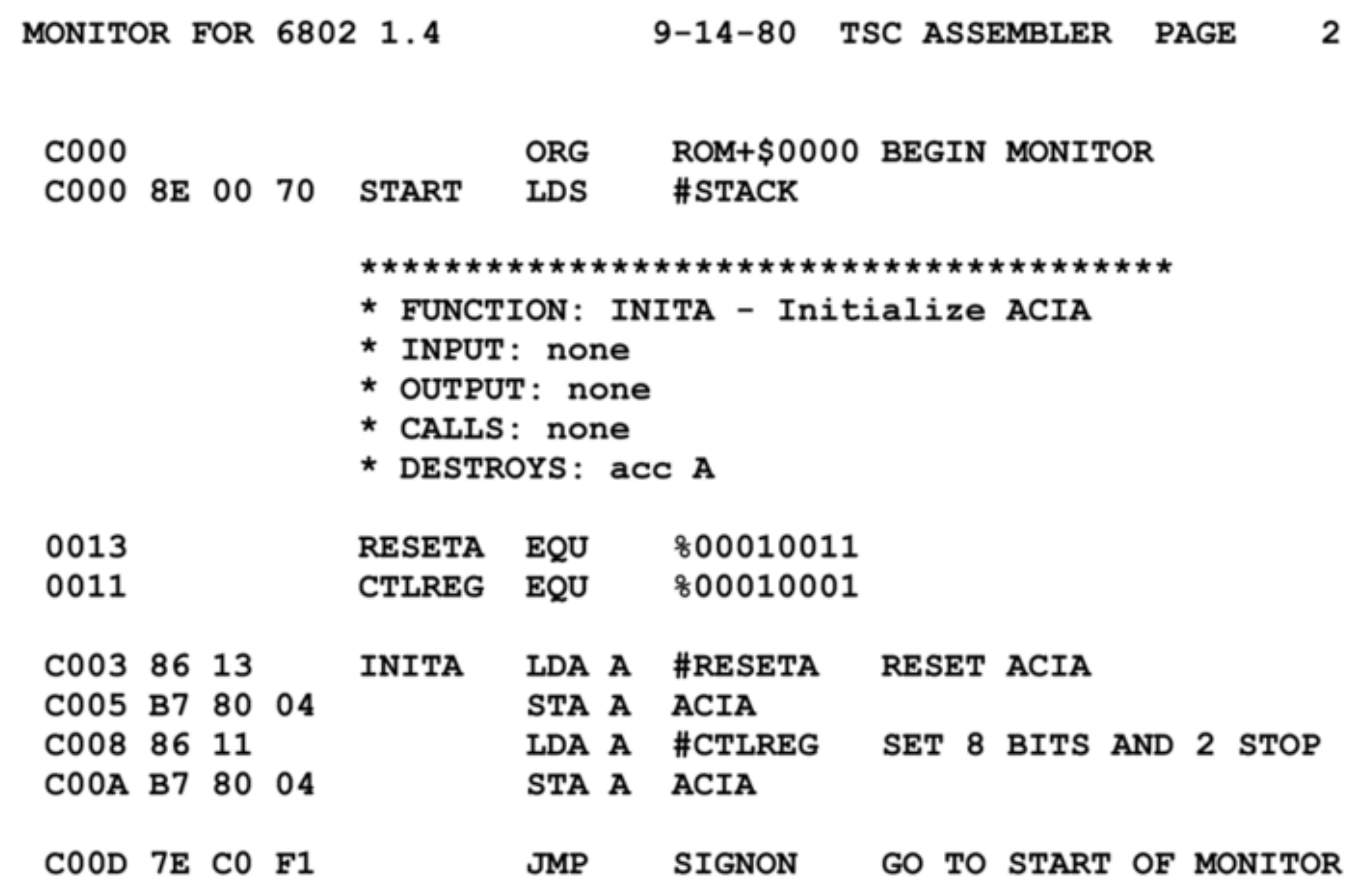
Forget about what you read for a sec…

Java

A **new** compiler written in

C

$\longrightarrow$

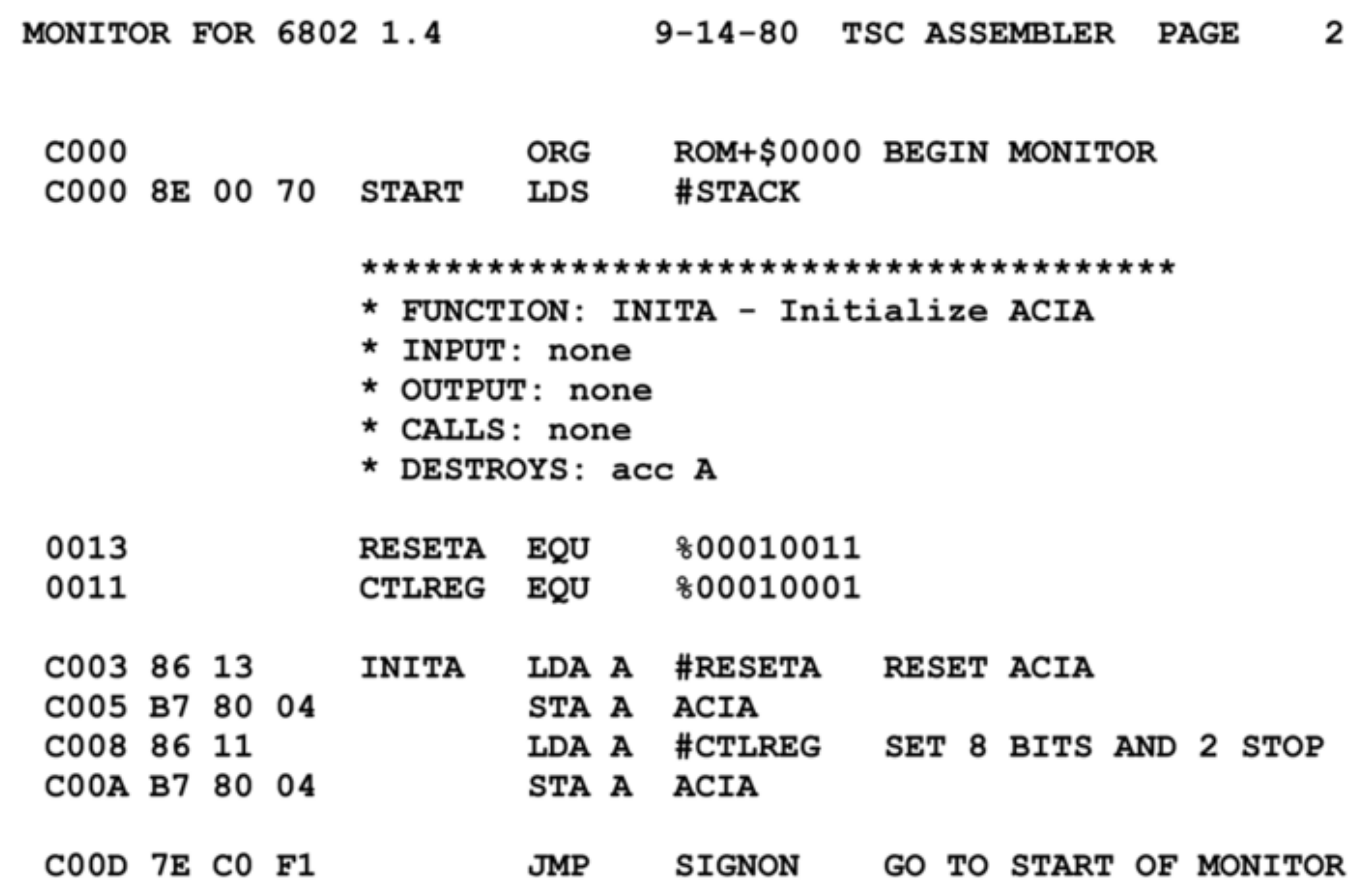


```
MONITOR FOR 6802 1.4          9-14-80  TSC ASSEMBLER  PAGE    2

 C000                ORG    ROM+$0000 BEGIN MONITOR
 C000 8E 00 70  START   LDS    #STACK

                ****************************************
                * FUNCTION: INITA - Initialize ACIA
                * INPUT: none
                * OUTPUT: none
                * CALLS: none
                * DESTROYS: acc A

 0013           RESETA  EQU    %00010011
 0011           CTLREG  EQU    %00010001

 C003 86 13     INITA   LDA A  #RESETA   RESET ACIA
 C005 B7 80 04          STA A  ACIA
 C008 86 11             LDA A  #CTLREG   SET 8 BITS AND 2 STOP
 C00A B7 80 04          STA A  ACIA

 C00D 7E C0 F1          JMP    SIGNON    GO TO START OF MONITOR
```

How can we add new language features to Java?

Forget about what you read for a sec…

Compiler written in



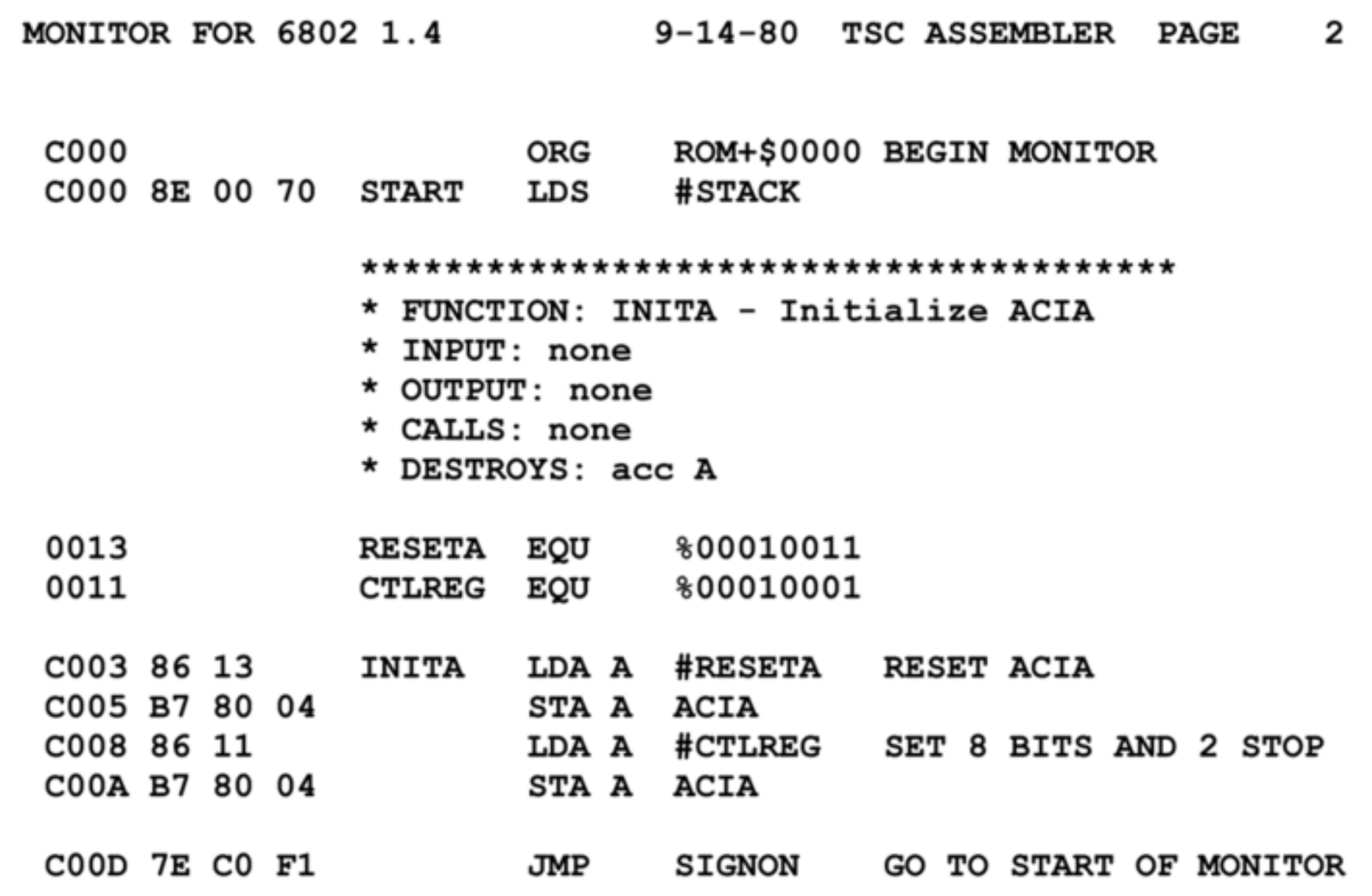
```
MONITOR FOR 6802 1.4          9-14-80   TSC ASSEMBLER   PAGE    2

 C000                    ORG    ROM+$0000 BEGIN MONITOR
 C000 8E 00 70   START   LDS    #STACK

                 ***************************************
                 * FUNCTION: INITA - Initialize ACIA
                 * INPUT: none
                 * OUTPUT: none
                 * CALLS: none
                 * DESTROYS: acc A

 0013            RESETA  EQU    %00010011
 0011            CTLREG  EQU    %00010001

 C003 86 13      INITA   LDA A  #RESETA    RESET ACIA
 C005 B7 80 04           STA A  ACIA
 C008 86 11              LDA A  #CTLREG    SET 8 BITS AND 2 STOP
 C00A B7 80 04           STA A  ACIA

 C00D 7E C0 F1           JMP    SIGNON     GO TO START OF MONITOR
```

How can we add new language features to C?

Forget about what you read for a sec…

C

Old compiler written in C



```
MONITOR FOR 6802 1.4          9-14-80   TSC ASSEMBLER   PAGE    2

 C000                    ORG    ROM+$0000 BEGIN MONITOR
 C000 8E 00 70   START   LDS    #STACK

                 ***************************************
                 * FUNCTION: INITA - Initialize ACIA
                 * INPUT: none
                 * OUTPUT: none
                 * CALLS: none
                 * DESTROYS: acc A

 0013            RESETA  EQU    %00010011
 0011            CTLREG  EQU    %00010001

 C003 86 13      INITA   LDA A  #RESETA    RESET ACIA
 C005 B7 80 04           STA A  ACIA
 C008 86 11              LDA A  #CTLREG    SET 8 BITS AND 2 STOP
 C00A B7 80 04           STA A  ACIA

 C00D 7E C0 F1           JMP    SIGNON     GO TO START OF MONITOR
```

How can we add new language features to C?

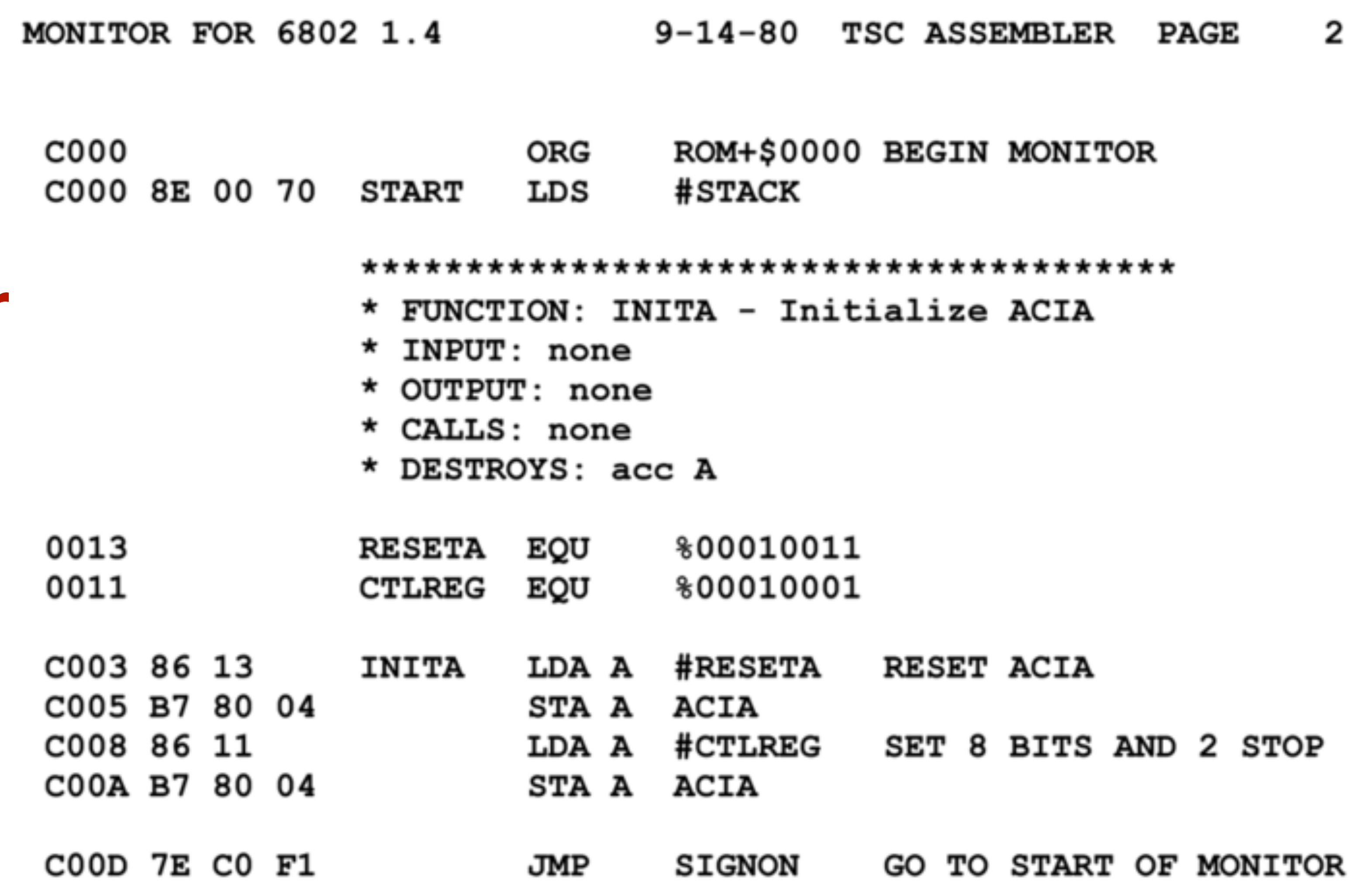
Forget about what you read for a sec…

C

New compiler written in

C

compiled using the old compiler



```
MONITOR FOR 6802 1.4           9-14-80  TSC ASSEMBLER  PAGE    2

 C000                ORG    ROM+$0000 BEGIN MONITOR
 C000 8E 00 70  START  LDS    #STACK

                ***************************************
                * FUNCTION: INITA - Initialize ACIA
                * INPUT: none
                * OUTPUT: none
                * CALLS: none
                * DESTROYS: acc A

 0013           RESETA EQU    %00010011
 0011           CTLREG EQU    %00010001

 C003 86 13     INITA  LDA A  #RESETA   RESET ACIA
 C005 B7 80 04         STA A  ACIA
 C008 86 11            LDA A  #CTLREG   SET 8 BITS AND 2 STOP
 C00A B7 80 04         STA A  ACIA

 C00D 7E C0 F1         JMP    SIGNON    GO TO START OF MONITOR
```

How can we add new language features to C?

“Bootstrapping”: the technique for producing a self-compiling compiler

Some more context

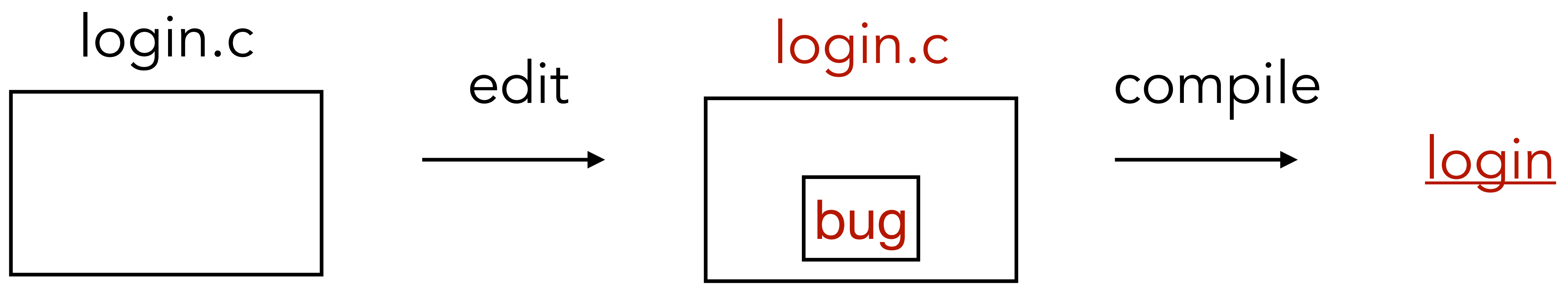
Earlier version of Unix were distributed with a full set of binaries and source for those binaries.

It is common for people to make change in one source file and recompile all their programs

How did Thompson add a bug to the login program without leaving a trace?

Goal

Have no source files hint at the bug, and meanwhile, the bug will persist across all recompilations



Anyone looking at login.c will realize something is wrong!

Goal

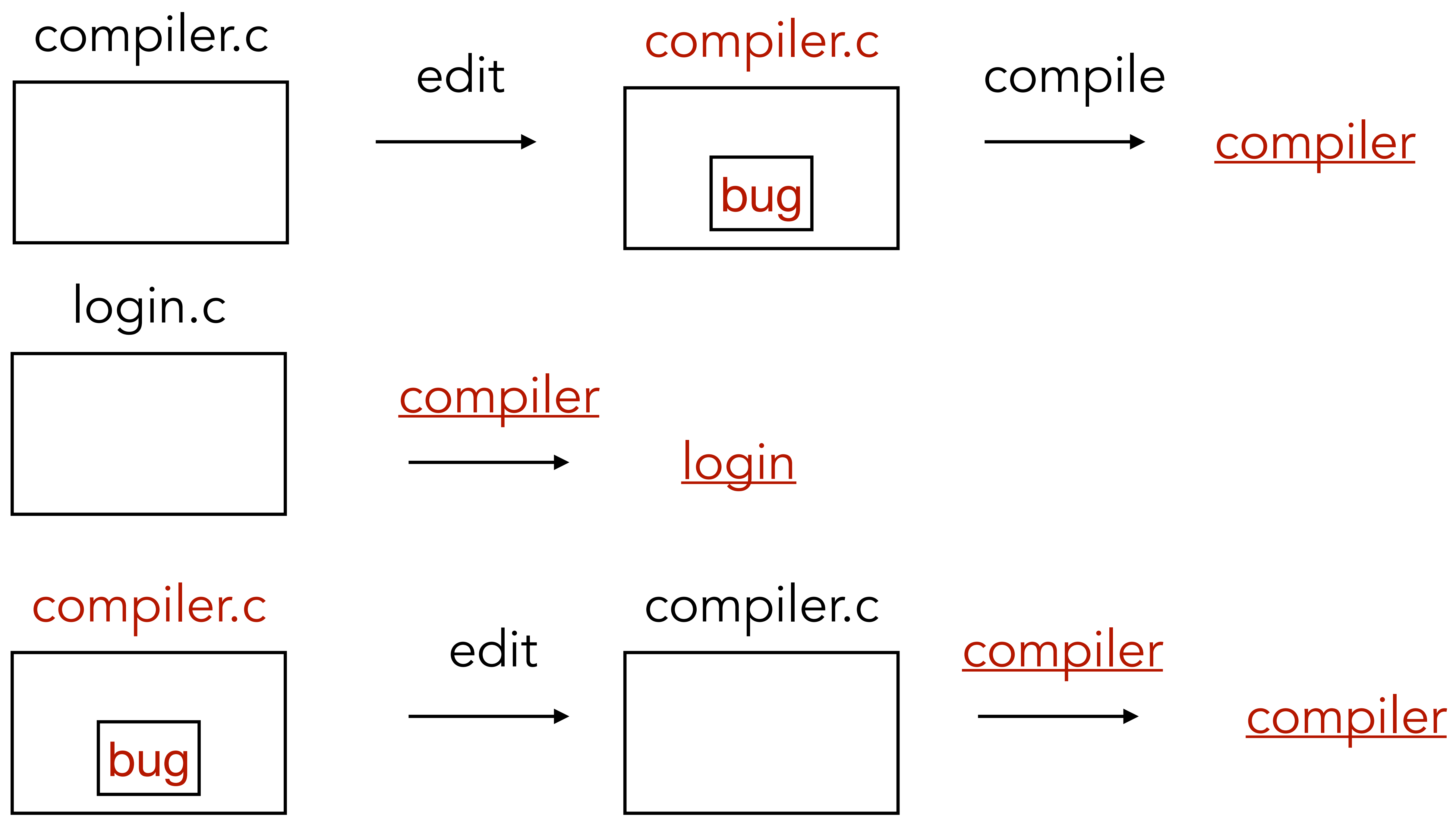
Have no source files hint at the bug, and meanwhile, the bug will persist across all recompilations

login.c
login

If you recompile locally, login will be bug-free again

Goal

Have no source files hint at the bug, and meanwhile, the bug will persist across all recompilations



Done!

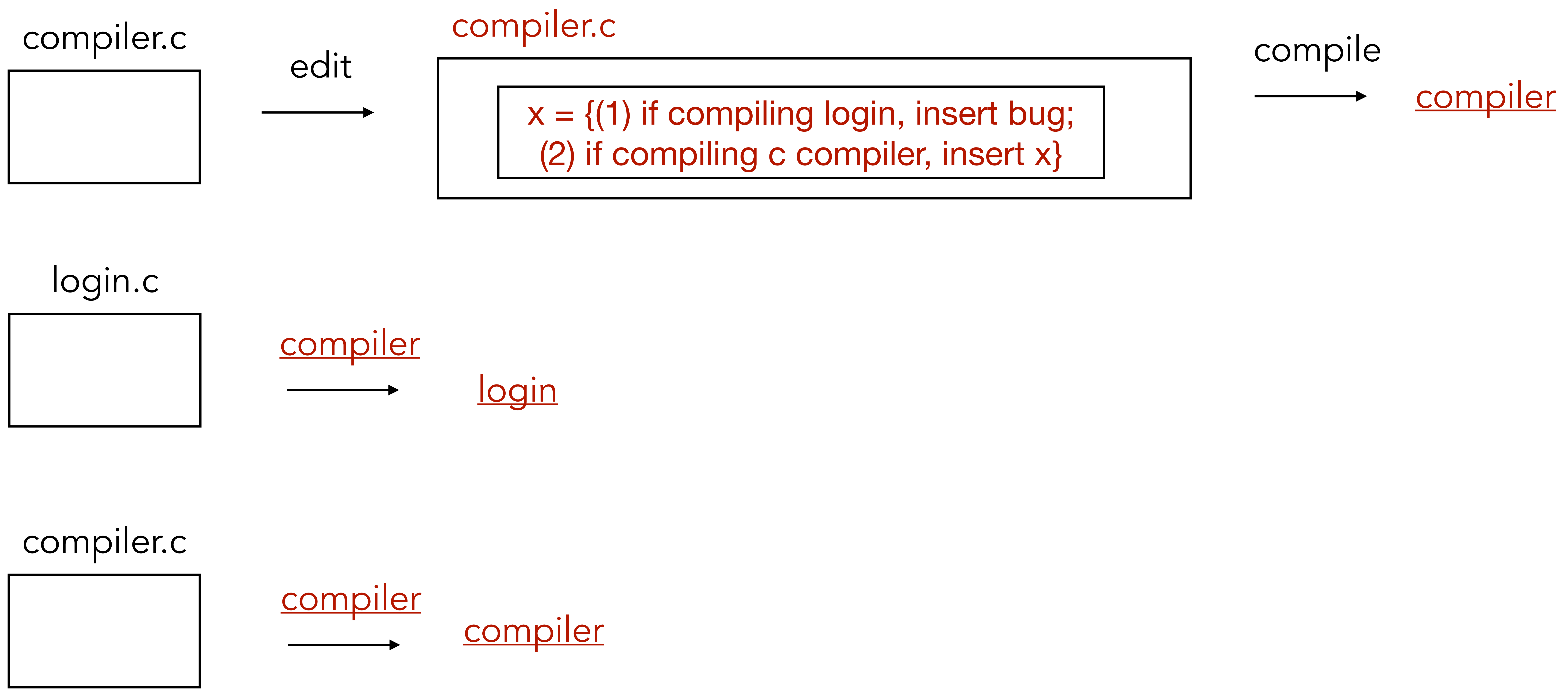
How can Ken figure out this attack?

Self-reproducing program: a computer program that takes no input and produces a copy of its own source code as its only output. (Quine)

"yields falsehood when preceded by its quotation" yields falsehood when preceded by its quotation.

https://en.wikipedia.org/wiki/Quine%27s_paradox
https://rosettacode.org/wiki/Quine

Actual attack



Done!

Implications

You can’t trust code that you did not totally create yourself!

Protections and security in Unix

U(ser)ID and G(roup)ID

Files and directories are access-controlled:
system stores with each file who owns it (in inode)

Root (UID 0)

Has all the permissions: read any file, do anything, …

Some legitimate actions require more privileges than UID

How should users change their passwords (root-owned)?

Each process has a real and effective UID/GID

Real is user who launched the program, effective is owner/group executables, used in access checks

Setuid

a program that is run in with **raised privilege level**

effective uid = real uid