Architecture: assembly and compiler optimisation

In this practical you will undertake various tasks related to understanding the assembly language of x86-64, in the AT&T syntax (also known as GNU Assembler Syntax, or GAS). You will be analysing both unoptimised and optimised assembly.

For the highest marks, you will also be required to analyse a piece of assembly of the ARMv8 architecture (or to be precise AArch64, which is the 64-bit execution state of ARMv8). This will involve independent study.

This practical is worth 25% of the continuous assessment portion of this module and is due 21:00 on Wednesday Feb 24th.

Background

You are supplied with C source code to do binary search in a sorted array, in `array.c`, `main.c`, and `array.h`. The provided `Makefile` compiles these to an executable `main`, as well as to several assembly files of x86-64, with varying degrees of optimisation, with `array0.s` being unoptimised and `array3.s` being most optimised.

(For the assembly files to be of the expected kind, you need to run the `Makefile` on a machine with a x86-64 architecture, preferably one of the machines in the school. Because not all versions of `clang` produce the same assembly, I have provided two files `array0-commented.s` and `array2-commented.s`, which are copies of the `array0.s` and `array2.s` that were obtained by running the `Makefile` on `klovia`. Please take these two files in what follows, to ensure everyone does the same work and marking is consistent.

This assignment is broken up into three parts, which are best done in the indicated order.

1 Analysing unoptimised assembly code

You are to add comments to the assembly in `array0-commented.s`, to help a human reader understand it.

- The comment character in assembly is `#`. From this character onward, the rest of the line is ignored by the assembler.

- Comments should come after instructions on the same line. If necessary, you can add additional lines with no instructions to extend a comment.
- Your commented file should not change the assembly code at all, so that it could still be used as input to the assembler.
- Each assembly language instruction for the function `contains` should have some kind of comment. This comment needs to indicate succinctly what the purpose of the instruction is *in its context*, for example by referring to variables or other expressions or statements in the C program, or by relating the instructions to aspects of the calling convention. Some comments may be as short as for example `# %eax = x+y`, as pseudo-formal notation to mean that the purpose of the instruction is to put `x+y` in register `%eax`, where `x` and `y` would be variables from a source C program.
- You do *not* need to comment on assembler directives, whose names start with a period.
- You may see instructions you haven't seen in the lectures. A web search may help to find more information.
- In unoptimised code, beware that some instructions may not do anything useful in their context. Such instructions are generally removed by compiler optimisations (as you will see in Parts 2 and 3 below).

After studying the assembly, you should be able to determine the structure of the stack frames, and describe your findings in the report. By means of a table or list of bullet points, list all elements of the stack frames of function `contains`, with their index relative to the base pointer. Which values are stored where? Where is the return address stored, and where the saved base pointer? Where are 64-bit values used and where 32-bit or 16-bit or 8-bit values? Are any bytes in the stack frames unused? If so, which, and why are these bytes unused?

2 Recursion to iteration

Now study `array2-commented.s`, which was obtained by optimisation level 2. You will see that recursion has turned into iteration.

Add code commenting to `array2-commented.s`. You can also further explain some of the observed compiler optimisations in the report. For example, how is division by 2 realised, and why is this a correct way of doing division by 2 here?

The code commenting and the accompanying text in the report should convince the marker that you understand why the code does the same computation as before.

3 ARMv8

Now study the provided file `array1-arm-commented.s`, which is assembly of a very different architecture. It was also compiled from `array.c`, but now with optimisation level 1. In order to understand this, some independent study may be required. A good starting point is `https://modexp.wordpress.com/2018/10/30/arm64-assembly/`. A quick reference is `https://courses.cs.washington.edu/courses/cse469/18wi/Materials/arm64.pdf`.

Add code-commenting as before, and describe in more detail in the report what you have learned about the ARMv8 architecture as it relates to `array1-arm-commented.s`. This may include, but is not limited to:

- What do the various instructions in `array1-arm-commented.s` do?
- What addressing modes are used and what do they mean?
- How do you see the 64-bit ARM calling convention reflected in the assembly?
- Where are 64-bit values used and where 32-bit or 16-bit or 8-bit values?
- What is the layout of the stack frames of the `contains` function? Do you notice any interesting differences with the stack frames you found in Part 1?
- How is division by 2 realised, and why is this a correct way of doing division by 2? (Ask yourself how integer division is defined in C (from the C99 standard onwards).)

Submission

Submit a zip file containing:

- versions of `array0-commented.s` and `array2-commented.s` extended with your code commenting,
- if you completed Part 3, a version of `array1-arm-commented.s` extended with your code commenting, and
- your report as a PDF file.

Marking

0-9 Work that fails to demonstrate an understanding of the meaning and purpose of assembly.

10-14 Completion of Parts 1 and 2, explaining the purpose of each instruction in its context, but failing to express good understanding of overarching principles such as stack frames, calling conventions, and compiler optimisations.

15-17 Completion of Parts 1 and 2, demonstrating thorough understanding of the instructions, as well as the overarching principles.

18-20 Completion of all three parts to a high standard.

Rubric

There is no fixed weighting between the different parts of the practical. Your submission will be marked as a whole according to the standard mark descriptors published in the Student handbook at:

> `https://info.cs.st-andrews.ac.uk/student-handbook/learning-teaching/feedback.html#General_Mark_Descriptors`

You should submit your work on MMS by the deadline. The standard lateness penalties apply to coursework submitted late, as indicated at:

> `https://info.cs.st-andrews.ac.uk/student-handbook/learning-teaching/assessment.html#lateness-penalties`

I would remind you to ensure you are following the relevant guidelines on good academic practice as outlined at:

> `http://www.st-andrews.ac.uk/students/rules/academicpractice`