Parallel Design Patterns

Assessed Coursework, 2025
PART ONE

The deadline for PART ONE of the assessed coursework is Friday 14th of February at 4:00pm. You will submit this via the PDP learn pages.

About the coursework

The assessment for Parallel Design Patterns is split into two parts. This first submission ("part one") comprises of a short report. I expect that a complete answer could be expressed succinctly in **three or four pages** however there is no page limit, so feel free to go beyond this if required. The second submission ("part two") builds on this same problem.

Overview of the problem

You will solve a problem similar to one often faced by developers working at EPCC. The human brain is an important area of research, allowing us to learn more about how signals in our brains are processed. However, due to the large number of neurons in the brain, and the even larger number of connections (86 billion neurons and 100 trillion connections are estimated to be in the human brain by the Harvard Medical School), modelling its behaviour is a computationally intensive task. To this end, a medical research company have asked you for help parallelising their in-house brain simulation code, which models how signals propagate through the brain.

A more detailed description of the model is provided later in this document.

Your tasks

For this, coursework part one, you should analyse the problem and write a report answering the following key questions:

1. **[weighting: 60%]**
   a. Select **two** patterns from the parallel algorithm structure/strategy design space that you believe could be used for this problem. **For both patterns** discuss how the pattern's context and forces relate to the problem, the advantages and disadvantages of your choice of pattern, and why you consider it suitable.
   b. From the two patterns you have described in part a, which one would you select for this problem and why?
   c. Given the pattern you have selected in part b, what aspects of the pattern and its solutions might be relevant to the choice of implementation language or hardware platform?
   d. Given the pattern selected in part b, what aspects of your design might ultimately limit the overall parallel performance and scaling achieved when it is implemented, and what trade-offs might need to be considered to address these?
2. **[weighting: 25%]**
   a. Select **two** patterns from the parallel algorithm structure/strategy design space that you think are **inappropriate** for this problem, **for both patterns** describe the disadvantages of the pattern in relation to this problem and why it would not be a suitable choice here.
3. **[weighting: 15%]**
   a. Based on the problem described in this handout, what approach(es) might you use to help evaluate the medical research company's existing serial code?
   b. What techniques would be useful in measuring the performance and scaling of your parallel design and code as it develops?

并行设计模式

2025 年课程评估
第一部分

评估课程作业第一部分的截止日期为 2 月 14 日星期五下午 4:00。您将通过 PDP 学习页面提交此作业。

关于课程

并行设计模式的评估分为两部分。第一次提交（“第一部分”）包括一份简短的报告。我希望完整的答案可以在三到四页内简洁地表达出来，但没有页数限制，所以如果需要，可以随意超过这个限制。

第二份提交的文件（“第二部分”）也是针对同一问题而提出的。

问题概述

您将解决的问题与 EPCC 开发人员经常遇到的问题类似。人脑是一个重要的研究领域，它使我们能够更多地了解大脑中信号的处理方式。然而，由于大脑中神经元的数量庞大，连接数量甚至更多（哈佛医学院估计人脑中有 860 亿个神经元和 100 万亿个连接），对其行为进行建模是一项计算密集型任务。为此，一家医学研究公司请您帮助并行化他们的内部大脑模拟代码，该代码模拟信号如何在大脑中传播。

本文档后面将提供该模型更详细的描述。

你的任务

对于这门课程的第一部分，你应该分析问题并撰写一份报告来回答以下关键问题：

1. [权重：60%]
    a. 从并行算法结构 / 策略设计空间中选择两种你认为可以用于解决该问题的模式。对于这两种模式，讨论模式的背景和作用与问题的关系、你选择该模式的优缺点以及你认为它合适的原因。b. 从你在部分 a 中描述的两种模式中，你会选择哪一种来解决该问题，为什么？c. 鉴于你在部分 b 中选择的模式，该模式及其解决方案的哪些方面可能与实现语言或硬件平台的选择有关？d. 鉴于在部分 b 中选择的模式，你的设计的哪些方面可能最终限制实现时实现的整体并行性能和扩展性，以及可能需要考虑哪些权衡来解决这些问题？

2. [权重：25%]
    a. 从并行算法结构/策略设计空间中选择两种您认为不适合此问题的模式，对于这两种模式，描述该模式与此问题相关的缺点以及为什么它在这里不是一个合适的选择。

3. [权重：15%]
    a. 根据本讲义中描述的问题，您可以使用哪些方法来帮助评估医疗研究公司现有的串行代码？ b. 在开发过程中，哪些技术可用于测量并行设计和代码的性能和扩展性？

Details of the brain simulation model

The model that the medical research company have described exhibits the following behaviour:

- The brain is represented as a graph
  - Neurons are nodes in the graph, and connections are edges between these nodes
  - A small number of additional nodes represent nerves. These serve as end-points where signals are either inputs to (i.e. an external stimulus) or outputs from (i.e. the brain's response) the model
  - Individual connections (graph edges) can either be unidirectional (one way), or bidirectional (both ways)
  - Whilst there will be connections linking a specific neuron to many others in the brain, there is no guarantee that there is a connection between every pair of neurons
  - The graph is provided to the model via an input file, which is read on initialisation and remains fixed throughout code execution.
- Nerves generate random signals at random intervals
  - This signal gets sent to any connected neurons
  - All signals have a type associated with them
- Neurons receive signals from connected nodes (neurons and nerves), manipulate these signals and transmit this to connected nodes (neurons/nerves)
  - The neurons are of different types which is driven by their location in the brain
    - This type determines how frequently signals are generated by the neuron and the neuron modifies signals that are received
  - The neurons also have xyz-coordinates representing their geographic locations in the brain
  - Neurons (the number of them, and their type) are fixed and do not change throughout the simulation
  - Neurons maintain an internal state, which changes depending on the number of signals they have received in a specific timeframe
    - This state impacts how the neuron will modify signals that it receives before sending them on
  - The value of a signal determines how many nodes (neurons or nerves) the signal is sent to. After a signal has been manipulated by the neuron then the higher this value the more nodes it will be sent to.
- Edges connect two neurons or a neuron and a nerve
  - The connections have several weightings, one for each possible message type
    - This weighting will modify the value of the signal that is transmitted just before it is sent
  - These connections have a maximum value of a signal that can be transmitted through them
    - The value of a signal therefore determines how many nodes (neurons or nerves) that it will be sent to. Anything larger than the capacity of an edge will need to be transmitted in a round robin fashion to multiple neurons/nerves
    - For signals with large values, these will wrap around nodes if the number of nodes connected to the neuron is exhausted
- The input configuration file also defines nerve inputs (firing of signals at specific nerves) at specific points in time during the simulation. This simulates some sort of input to the brain
- All outputs from nerves are logged as these represent outputs from the brain, as well as the number of neuron firings that occur in each part of the brain

大脑模拟模型的细节

该医疗研究公司描述的模型表现出以下行为：

- 大脑以图形表示 o 神经元是图形中的节点，连接是这些节点之间的边 o 少量其他节点代表神经。这些节点作为端点，信号要么输入（即外部刺激），要么输出（即

  大脑的反应）模型

  o 单个连接（图边）可以是单向的（一种方式），也可以是双向的（两种方式） o 虽然会有连接将特定的神经元与大脑中的许多其他神经元连接起来，但不能保证每对神经元之间都有连接 o 图通过输入文件提供给模型，该文件在初始化时读取并在整个代码执行过程中保持不变。

- 神经以随机间隔产生随机信号 o 该信号被发送到任何连接的神经元 o 所有信号都有与之相关的类型

- 神经元从连接的节点（神经元和神经）接收信号，操纵这些信号并将其传输到连接的节点（神经元/神经）o 神经元有不同的类型，这取决于它们在大脑中的位置

  这种类型决定了神经元产生信号的频率，并且神经元会修改接收到的信号 o 神经元还具有 xyz 坐标，代表它们在大脑中的地理位置 o 神经元（它们的数量和类型）是固定的，在整个模拟过程中不会改变 o 神经元保持内部状态，该状态会根据它们在特定时间范围内接收到的信号数量而变化

  此状态影响神经元在发送之前如何修改所接收的信号 o 信号的值决定了信号发送到多少个节点（神经元或神经）。信号被神经元处理后，该值越高，发送到的节点就越多。

- 边连接两个神经元或一个神经元和一个神经元 o 连接有多个权重，每个权重对应一种可能的消息类型

  这种加权将在信号发送之前修改信号的值 o 这些连接具有可以通过它们传输的信号的最大值

  - 因此，信号的值决定了它将被发送到多少个节点（神经元或神经）。任何大于边缘容量的信号都需要以循环方式传输到多个神经元/神经

  - 对于值较大的信号，如果连接到神经元的节点数量耗尽，这些信号将环绕节点

- 输入配置文件还定义了模拟过程中特定时间点的神经输入（向特定神经发射信号）。这模拟了对大脑的某种输入

- 所有神经输出都被记录下来，因为它们代表大脑的输出，以及大脑每个部分发生的神经元放电次数

Details of the problem size

The medical research company have input configurations on several different scales. Currently they run their serial simulation code with the following parameters, and this is classed as a small configuration:

- 124 neurons and 3086 connections

Their medium configuration size is as follows, currently the serial simulation is much slower when simulating this:

- 170 neurons and 5335 connections

Ideally, the medical research company would like to run with the following parameters, but these are currently not possible due to the excessive runtime required:

- A large problem with 272 neurons and 10249 connections
- An even larger problem with 1058 neurons and 46958 connections
- They are also keen on simulating some of the smaller problem sizes (e.g. small and medium) for more simulated minutes too

问题规模的详细信息

该医学研究公司有几种不同规模的输入配置。目前，他们使用以下参数运行串行模拟代码，这被归类为小配置：

- 124 个神经元和 3086 个连接

它们的中等配置大小如下，目前串行模拟在模拟这个时要慢得多：

- 170 个神经元和 5335 个连接

理想情况下，医学研究公司希望使用以下参数运行，但由于所需运行时间过长，目前无法实现：

- 具有 272 个神经元和 10249 个连接的大问题
- 一个更大的问题，有 1058 个神经元和 46958 个连接
- 他们也热衷于模拟一些较小的问题规模（例如小型和中型），以获得更多模拟时间