Writing Your Research Question

Raimund Kirner
Department of Computer Science, University of Hertfordshire

version: 20.08.2021
(constructive feedback welcome: r.kirner@herts.ac.uk)

In this document we discuss how to write the research questions, for a Master's or PhD thesis. The document can also be a guide for Bachelor's thesis, even though there the requirements on the research novelty are usually a bit more relaxed. The main research question is sometimes also called research theses.

When to Write the Research Questions:
The research questions serve as a guide for the planned research. Thus, the research questions should be written in the early phase of the thesis. For example, within the English PhD program it should be written before submitting a report for the registration assessment, as this helps the examiners to judge the soundness of the research direction.
The research questions might very well change during the course of the research project, as newly obtained knowledge might require amending of the research direction to produce novel research based on the state of the art.

Where to Place the Research Questions:
The research questions should be placed in a separate subsection in the introduction of the report/thesis, so that it can be easily identified in the table of contents. In the final thesis, the conclusion should also make a direct reference to the main research question and all its individual sub-questions separately to discuss how the finished thesis has answered them.

Formulating the Research Questions
The research questions consist of two major parts:
1. The main research question (also called research thesis)
2. About 3 or more sub-questions

The main research question should describe the overall scope of the research in a short and concise form. It shall enumerate all the important factors that together makes the research novel. It is important to be clear about what exactly the novelty of the research is. For example, it might be the case that method M1 and M2 have been already developed and analysed before, but in the proposed research the novelty is to use methods M1 and M2 in combined form. Or, for example, method M1 has been used before in application domain A1, but in the proposed research method M1 will be applied to application domain A2. It should be also stated, what aspect of the research work is used to measure the success of the project, and how success is defined. For example, one might be interested in the runtime complexity of an analysis algorithm, and as such hope that methods M1 and M2 in combined form provide a lower runtime complexity than methods M1 or M2 themselves for the chosen application domain A2.

The main research question might be formulated in question form or as a postulation to be investigated, with the latter typically called research thesis.

It might be tempting to use words like "novel" in the formulation of the main research question. However, this would be just a filler word with no information, as it is a requirement that the research question addresses a novel problem/approach. It is rather important to describe in concise form the overall technical approach that makes it novel.

The *sub-questions* should focus on particular aspects of the envisioned research that are seen as particularly challenging to the success of the project. As such, the sub-questions are not meant to cover aspects that just require hard work but can be done fairly well within the given time frame. The sub-questions should rather focus on the critical intellectual parts, where it is fairly unclear how to implement them, i.e., the high-risk parts of the research.

Discussion of an Example Research Question:
In the following we discuss an example research question[1]. The original formulation of the research thesis and sub-questions is in italics and blue colour.

Although Reactive Stream Programs (RSPs) are in general well suited for many-core architectures, they work in tandem with their environment, in which the load imposed by the environment may vary over time. Keeping this in mind, this thesis is motivated by the following research question:
(The above text just provides a short connection with what has been written before in the introduction)

> *Is it possible to improve the adaptive resource utilisation and improve the energy efficiency of RSPs on many-core platforms by exploiting knowledge about the states of the system?*

(The *application domains* here are RSP (A1) on many-core platforms (A2). The applied *method* is the exploitation of knowledge about the states of the system (M1). The *success criteria* are the improvement of the resource utilisation by adaptation (C1) and the improvement of the energy efficiency (C2). What is left open in this formulation of a main research question is on what types of RSPs the proposed method is applied to. This could be further improved by specifying on what RSP system the proposed method is applied. Limiting the focus is actually important, as even though the method might lead to improvements in a generic sense, it is useful to state in what scope these improvements have been analysed. To summarise the structure, this example of a main research question uses application domains A1, A2, method M1, and success criteria C1, C2.)

---

[1] the example has been taken from the following PhD thesis: "RA-LPEL: A Resource-Aware Light-Weight Parallel Execution Layer for Reactive Stream Processing Networks on The SCC Many-core Tiled Architecture", Nilesh Karavadara, University of Hertfordshire, 2016.

This question is split into the following sub-questions:
(In this research project five sub-questions have been identified as critical and important for the success of the project)

1. *What is an efficient way to port an existing RSP execution layer to the SCC?*
   (The research takes the implementation of an existing RSP and ports it to the SCC processor as a concrete application domain. As this involves a relatively complex code, it is important to study how the porting can be done within the given time frame of the research)

2. *What are meaningful performance-indicators to identify the workload situation of an RSP on a many-core processor?*
   (As part of the research, the "knowledge about the states of the system" from the research thesis has to be concretised in order to find a concrete method.)

3. *Is it possible to have these performance-indicators to be independent of any specific hardware feature of that many-core processor?*
   (While the application domain A2 of the research question is many-core platforms, the concrete implementation is done on the SCC processor. To answer the research thesis in its more generic form, it has to be investigated whether the method can be achieved by utilising HW properties that are likely to have their equivalent in other many-core processors.)

4. *What are the adequate strategies to optimise the performance per watt of RSPs on many-core platforms?*
   (The "energy efficiency" mentioned in the research thesis is interpreted as performance per watt, for which optimisation strategies are sought.)

5. *Is it possible to design Dynamic Voltage/Frequency Scaling (DVFS) strategies that are light-weight and simple, but still adequate to provide substantial reduction in energy consumption of RSPs?*
   (A method to improve "resource utilisation" and "energy efficiency" is only of practical value, if the caused overhead is more than compensated by the resulting gain. Thus, the research is geared towards methods that are implementable with light-weight overhead.)

Research Questions are Central for the Time-effective Supervision of the Research:
Writing the main research question and sub-questions in the early phase of the research is very helpful to develop a clear plan for the research project to be suitable for the given resource and time limitations. It allows to discuss the research thesis and its sub-questions early on with the supervisor.

The main research question allows to communicate quickly what the scope of this research project is, while the individual sub-questions describe the concrete research challenges that have been identified for this research project. Once the main research question and the sub-questions are formulated, it is possible to discuss with the supervisor the overall feasibility of the research project for the given time frame, even without much else of the research project already been described.

For the individual sub-questions it has to be checked if their answer is really to be investigated within this research project. If not, then this concrete sub-question should be removed, as it is not within the scope of the research project. At the same time, it can be checked early on whether the individual sub-questions are really within the scope of the overall main research question. If any sub-question is without the scope of the main research question, then one must decide whether to extend the scope of the main research to cover the sub-question, or whether to reduce the sub-question to something that is within the scope of the main research question.

Adaptation of the Research Questions:
The formulation of the research questions does not have to be set in stone, as it is quite likely that their amendment is necessary during the course of the research project to reflect the obtained knowledge.
In the final report/thesis it should be also discussed in the conclusion how the stated research thesis and sub-questions have been answered by the performed research. To improve readability, each sub-question should be addressed individually in a separate paragraph.