6 Goal, Question, Metric

Heiko Koziolek

University of Oldenburg, Germany

This chapter gives an overview over the Goal-Question-Metric (GQM) approach, a way to derive and select metrics for a particular task in a top-down and goal-oriented fashion.

6.1 Motivation

Defining metrics for a measurement or prediction approach in a *bottom-up* manner is generally considered problematic. Without inclusion of the context and the goal of measuring or prediction, it is usually unclear which metrics shall be selected and how the selected metrics can be interpreted. If lots of data has been measured without a predefined plan for the evaluation, it will usually be very difficult to select the relevant data and to draw proper conclusions from the measurements.

Therefore, embedding metrics into a goal-oriented framework is widely regarded as a good practice. The Goal-Question-Metric (GQM) approach presents such a framework. It was developed by Basili and Weiss during the 1980s [42] and later extended by Rombach [41]. Opposed to the bottom-up approach, metrics are defined *top-down* in GQM. First, specific goals are stated, then questions are asked, whose answers will help attaining the goals. The metrics are defined in a third step to provide a scheme for measuring.

By applying GQM, goals and metrics are tailored for a specific measurement setting. Stating goals in advance leads to a selection of only those metrics that are relevant for achieving these goals. This reduces the effort for data collection, because only necessary data needs to be recorded, nothing more, nothing less. The interpretation of the metrics after measurement is rather effortless, because GQM creates an explicit link between the measured data and the goals of measuring before data collection. This way, misinterpretations of data can be avoided.

Although the GQM approach was originally used to improve software products and development processes, the underlying concepts are generic and applicable in any measurement setting.

The next section describes the process of applying the GQM-paradigm in detail. Another section explains how goals, questions, and metrics should be defined.

6.2 Goal-Oriented Measurement

The method of applying the GQM paradigm consists of four phases [481]: planning, definition, data collection, interpretation (see Fig. 1).

For the initial *planning* phase, first a GQM-team is established and the desired improvement area (e.g., reliability, performance, security, etc.) is identified. Afterwards, the team selects and characterises the product or process to be studied. The result of this

I. Eusgeld, F.C. Freiling, and R. Reussner (Eds.): Dependability Metrics, LNCS 4909, pp. 39–42, 2008.

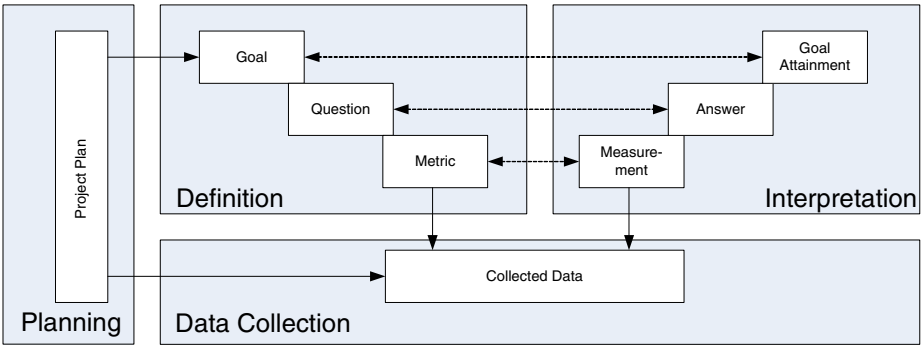

Fig. 1. The 4 phases of the GQM-method [481]

phase is a project plan that outlines the characterisation of the product or process, the schedule of measuring, the organisational structure, and necessary training and promotion activities for people involved in measurements.

During the *definition* phase, measurement goals are defined. A template for goal definition will follow in the next section. For the purpose of defining goals, GQM interviews may be conducted with people involved in the process or product under study. Based on the goals, questions are stated, asking for specific quality attributes and making certain aspects of the goals more concrete. For each question a hypothesis with an expected answer should be defined. Afterwards, metrics are defined for each question and checked on consistency and completeness. Results of this phase are a GQM plan, a measurement plan, and an analysis plan.

The actual measurement takes place in the *data collection* phase. Data collection may be performed manually or electronically and may involve automated data collection tools. A measurement support system consisting of spreadsheets, statistical tools, database applications and presentation tools should be established for this phase.

In the *interpretation* phase, the collected data is processed according to the metrics defined before to gain measurement results. The measurements can then be used to answer the questions, and with the answers it can be evaluated if the initial goals have been attained.

6.3 GQM Paradigm

A GQM plan consists of goals, questions, and metrics in a hierarchical structure (see figure 2). Before measuring, the elements are defined in a top-down manner. After measuring, the plan can be used bottom-up to interpret the results.

Goals are defined on a conceptual level and later made operational by questions. For clearness, each GQM plan should contain one goal. Goals can be derived by studying the policy and the strategy of the organisation that applies GQM. Interviewing relevant people and checking available process or product descriptions may also help in defining goals. If goals are still unclear, first modelling the organisation might be necessary to derive them.

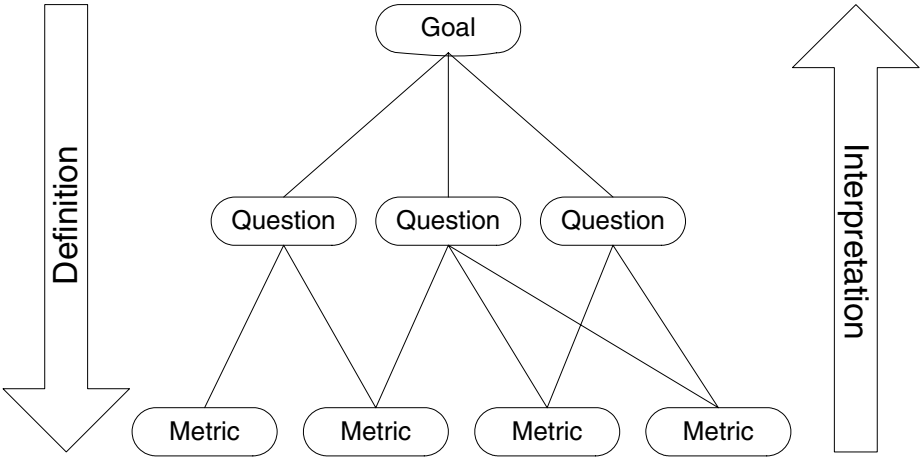

Fig. 2. The GQM-paradigm [481]

Goals must be documented in a structured way, preferably using the following template. Goals are defined for a purpose (e.g. understanding, improving, controlling), for the object under study (e.g. process, product), for a specific issue or quality attribute (e.g. reliability, security, performance), from a perspective (e.g. user, developer), and within certain context characteristics (e.g. involved persons, environment, resources, organisations). For example, a goal could be:

> "Improve (=purpose) the reliability (=issue) of product X (=object) from the viewpoint of the user (=perspective) within organisation Y (=context)."

Several *questions* are usually asked to achieve each goal. While goals are on a more abstract and conceptional level, questions shall refine goals and make them operational. By answering the stated questions, it should be possible to conclude if a goal has been reached. Questions should be defined on an intermediate level between goals and metrics. They should neither be too abstract, in which case it would be difficult to reveal the relationship to the collected data, nor should they be too detailed, so that an interpretation of the answers towards the goal would be difficult. If questions are stated in a quantitative way, data can be collected by measurements. If questions are stated in a qualitative way, questionnaires are necessary to answer them. Expected answers to the questions are formulated as hypotheses. Hypotheses increase the learning effect from measurements, because they allow comparing knowledge before and after measurements. For example, a question could be:

> "What is the probability of failure on demand for function Z in product X?"

A matching hypothesis would be:

> "The probability of failure on demand for function Z in product X is below 0.0001 percent."

Table 1. Example GQM plan

Goal	Purpose	Improve
	Issue	the reliability
	Object	of product X
	Viewpoint	for the user
	Context	within organisation Y
Question	Q1	What is the probability of failure on demand for function Z?
Metric	M1	POFOD
Question	Q2	How erroneous does function Z behave?
Metric	M2	MTTF
Metric	M3	MTTR
...	...	...

After these steps, several *metrics* can be defined for each question. One metric may be used to answer different questions under the same goal. Metrics are on a quantitative level making it possible to assign numbers to a quality attribute. Metrics are means to map reality into comparable values. They must be defined to answer the asked questions and be able to approve or disprove the stated hypotheses. Already existing data can be used to define metrics. A metric may be objective, so that different people measuring the metric would gain the same results (e.g. the number of jobs completed during a time span). Instead, subjective metrics may produce different results depending on the persons measuring (e.g. the readability of a text). Objective metrics are suitable for more mature objects, while subjective metrics can be used for more unstable or unclear objects.

A simple GQM plan can be found in Tab. 1. Any GQM plan needs permanent evaluation and improvement. Imprecise goals need to be corrected and no more matching metrics need to be updated.