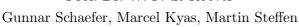CHRISTIAN-ALBRECHTS-UNIVERSITÄT ZU KIEL
Institut für Informatik und Praktische Mathematik

Prof. Dr. W.-P. de Roever

Gunnar Schaefer, Marcel Kyas, Martin Steffen

Coma

Informal requirement spec. (1)

Termin: 26. 10. 2004

Abstract

The document describes the first informal specification for the "Fortgeschrittenenpraktikum" in the winter term 2004/05. It is also available via the website. The requirement specification is being updated and refined during the semester according the the project's progress and the decisions taken. To reflect the development, the specification is qualified with a versioning number.
version: 1
status: transient

Contents

1 Introduction

The document describes *informally,* i.e., in plain English, the functionality (or rather the purpose) of *Coma*, a tool to assist in the distributed preparation, organization and processing of a conference, workshop, or similar event.

At the current stage of affairs, the rest of the document sketches the intention of the project, the informal functionality etc. We are here not explicit about specific technologies or how to realize the task

2 General purpose

Scientific conferences nowadays are in general prepared, advertised, and managed via the web. In particular, part of the organization is done in a "distributed" and loosely couple manner, i.e., the actors in the organization work at various locations, time zones, etc.

At an rather abstract level, the goal is to allow a group of experts, to collaborate to find an agreement in selecting from a set of contributions to the conference.

3 Phases and events

The organization of the conference takes a limited amount of time, which is fixed at the beginning. We distinguish the following main phases:

1. assembly

2. paper submission

3. reviewing, consensus finding about the selection of the submission

4. registration

The phases themselves are ordered linearly, the exact duration or distance of the single events should be adaptable. In the following, we discuss the purpose of the phases the outcome of each. Most (but not all) phases are separated by events.

3.1 Event: Configuration

Configuration serves a rather simple goal, namely the preparation and setup for the rest of the phases. After installation, its includes to fix basic data of the conference, which means it's name, it acronym, the date (begin and end) and the location of the conference. Furthermore, the chairs of the program committee are (probably) known.

3.2   Phase: Assembly

The assembly is done by the *program chairs.* The result of the assembly phase is the *program committee,* i.e., the collection of experts or individuals that will, in later phases, decide about the program.

The program committee is not *self-assembled.* It is the task of the program committee to *invite* members via email.

Invitees are free to decide, whether they wish to participate, which means they should actively *acknowledge* participation or they should reject. Before they acknowledge, they are called program committee candidates, or *candidates* for short. By default, candidates are not members unless they positively acknowledge participation. They can register with the committee by using a web-interface. By acknowledging, they also fill in further relevant personal data, such as affiliation (=professional address), full name, preferred email address, home page.

3.3   Event/document: Call for papers

Once the program committee is fixed the conference is "advertised", i.e., authors are invited to contribute to the event. The document, basically an url i.e., the "web-page" which contains this advertisement is called the *call for papers.* It must announce all information about the conference relevant for authors, which

1. basic conference data

   - date
   - name, acronym
   - location

2. important dates for authors

   - optional: deadline for abstracts
   - deadline for submission (cf. Section 3.5)
   - notification date (cf. Section 3.7)
   - date for the final, i.e., a possibly corrected version of the paper.

3. reference to the submission page

4. short description as free form text

5. names, affiliations and perhaps other information about the organizing people.

3.4   Phase: submission

The effect of the announcement is that authors decide to contribute to the conference. The contributions are called *papers.* The paper is *uploaded* by an author onto the appropriate web page, the act of doing so is the *submission.*

A paper has at least one *author,* but may have more than one. One author may have more than one paper. One author of a paper is considered as the main author, known as the *corresponding author.* This is the author the organization of the conference deals with.

An author can decide to *retract* a paper. In this case, no older versions (if any) of the paper are restored, but the paper is removed completely.

3.5   Event: submission deadline

The submission phase is terminated by the *submission deadline,* a pre-announced time after which no submission is possible. Until that time an author can submit many versions of the same paper. Only the last one before the deadline counts, i.e., each later version overwrites the earlier one.[1]

It is possible for an author to remove a paper even after the deadline (see Section 3.4).

3.6   Phase: Reviewing

Reviewing is the process of selecting from the submission a set of papers. The selection is done by the program committee in a joint effort using the web server. The (informal and conflicting) goals are:

quality of paper selection: Since in general there are more papers than time on the conference, the best papers are to be selected.

load balance for reviewers: In order to form an opinion about a paper, a reviewer must read and understand it. This workload must be distributed in a fair manner for the members.

load balance for papers: Each paper gets (in general) more than one reviewer, to make the decision less random. Each paper should get an equal share of the reviewing task.

quality of reviewer selection: Each reviewer gets the papers that he wants to and/or he is the best expert for.

The above specification obviously is informal and unprecise as it allows a number of interpretations. We do not fix an exact specification here, because there are probably many plausible solutions. Instead we discuss aspects of the mentioned goals. In the specification task, we would like more explicit proposals to solve this problem.

3.6.1   Assignment of papers

One task is the *assignment* of papers to reviewers. It is to be expected that there are more papers than reviewers, and furthermore one should cater for the case that each paper gets more than one reviewer. Preferably, and unlike the selection of the papers, the assignment is done *automatically,* i.e., without general discussion.

Furthermore, the assignment should be "fair" wrt. the reviewers and wrt. the papers, in that the *load* is equally shared. An easy and not very useful solution would be to make a random assignment, under the side condition of approximate load balance. The disadvantage is that, in general, the members of the committee have slightly different fields of expertise, and preferably a member evaluates papers in a field he is a strong expert of.

Assignment by topic Papers are classified according to a finite list of *topics.* The topics are predefined for the conference, and the *author* must pick those he feels his paper fits in. He might choose more than just one topic. Also each reviewer, beforehand, chooses a number of topics which he prefers to read papers from. Once the papers are

---

[1]We cannot prevent that an author acts stupid and submits a later version of the same paper erroneously as a new paper, because we cannot check the semantics of the paper.

in, the software tries to take the preferences of the reviewers into account, but of course still maintaining load balance concerning the numbers of papers per reviews and the numbers of reviewers per paper

Assignment by paper This approach does not rely on predefined topics.[2] Each referee shortly looks at the list of papers and declares preferences (or dislikes) according to some schema. It this might be very simple like "I want 2, 17, and 42". Also, it should be possible to state: "I cannot review this paper.".[3] Again in this scheme, the selection mechanism should take the choices into account, but adhering to the side condition of balance. In other words: if someone only picks one paper, it does not mean he will get only one. If 15 people find paper 76 very interesting, it does not mean that paper 76 gets 15 reviewers.

One can imagine to combine those approaches, or to make it a chooseable alternative.

3.6.2 Selection of papers

In general there are more papers than there's time, so the intention is, of course, to pick the best of them. To talk about finding the "best papers" is misleading, though, because this uses the idealistic assumption that there are best papers and one just does not know yet which ones they are. On the other hand: even if it is more than questionable whether there is a globally and universal quality scale to be applicable to the papers, it does not mean that some papers are better than others, in the sense that most everyone would agree on that. The task is to come fast and efficiently to an agreement about this issue.

Let's assume two fundamentalistic approaches, which sheds light about the range of possibilities. Both sketched approaches are in practice not very useful and should be avoided. In order to talk about the best papers, one obviously assumes an (imaginary) *linear order* which needs to be determined by consensus and now the question is, how to reach this order.

Discuss everything One standpoint is: all participants discuss all papers in a free-form manner until all agree on some order, and this fixes the best papers. This solution is impractical: A *rational* agreement, i.e., an agreement based on common understanding, would require that all reviewers read all papers (which one wants to avoid ...). And even if all papers are read and discussed by all committee members, to reach at a common order lead to endless dispute.

Discuss nothing The opposite standpoint is: There is *no discussion* at all. Each reviewer gives the paper(s) he reviews a numerical value, say a mark. At the end the marks for each paper are averaged,[4] the results are ordered linearly, and then the best are chosen.[5] That is the most efficient solution but it might easily lead to bad decisions. In general, there is more than one reviewer per paper but there are in most cases not more than 4, and this makes the mean value of ratings rather random.

---

[2]This does not mean that there could not be predefined topics.

[3]Typically this is the case, if one sees that it's a paper by some friend/colleague ... so that one fears that one does not have an unbiased opinion.

[4]In general a paper gets reviewed by more than one expert.

[5]For the discussion we ignore the fact that there might be equal outcomes.

As perhaps from the discussion, there will not be a clean, mathematically optimal solution for the problem; basically the decision finding requires human intelligence and social interaction. The trick will be to *assist* in this social process, to make it more efficient than free discussion but more rational than random selection.

Possible states   Next we discuss which general states a paper can have during the reviewing phase. *Ultimately,* the judgment for each paper will be: "*accepted*" or "*rejected*", which is when all the papers are decided. A proposal for a schema of states could be:

$$
\begin{array}{rcl}
\text{status} & ::= & \text{decided} \mid \text{undecided} \\
\text{decided} & ::= & \text{accepted} \mid \text{rejected} \\
\text{undecided} & ::= & \text{unreviewed} \mid \text{unclear} \\
\text{unclear} & ::= & \text{conflict} \mid \text{inconclusive}
\end{array}
$$

There, a paper is yet undecided if it's not yet reviewed or the situation is unclear. Two causes are that there is a serious disagreement about the quality. For instance if one reviewer thinks the is very good in one category, and another one says it's very bad in the same category, this is an indication that one better looks at this point again. One could distinguish from that a situation where a paper is in the "so-la-la"-range. In general, most submissions are in the middle-field. In this case there might not be enough statistical evidence to distinguish between two contributions with slightly different ratings.

Decision herding   The core of a solution is to *focus* the process. Certain discussion is unavoidable/wanted, but the participants should focus (or rather helped to be able to focus) on the right, i.e., discussion-worthy things. Since the goal of the discussion is to find an agreement, discussion-worthy things, in first approximation, are those which are not yet decided.

Basically, the ones taking part in the discussion, must be assisted to get a good overview of the status of the debate and what things profitably to do next. This includes some form of *visualization* (which might be as simple as a table) of relevant information. Relevant information could include:

per reviewer: A reviewer will (if not "forced" otherwise) concentrate on "his" papers perhaps his reviews. So he should be presented "his" part of the task first. If not restricted, he might of course look also at other parts/aspects of the information.

"executive info": short, high-level overview over the status and progress (how many papers are decided, how many still to be discussed.)

delays: Indication about missed deadlines (someone has not yet send his reports or similar)

chosen focus: Some papers are chosen (for instance by the chair) to be discussed next.

Of course, the access to the information must obey the restrictions concerning the various user groups mentioned in Section 4.1.

Criteria As said before, the reviewers study the paper to come to an opinion about the quality of the contribution. In general one does not wish (only) a uniform single numerical value, but a (reasoned) rating in various categories. Those categories could include

overal rating: A single numerical value which expresses the overall quality of the paper, taking all aspects into account

originality: how new is the result/content?

soundness: Is the technical content sound or are there serious errors in the argumentation/proofs/results ...

relevance: how good does the paper fit into the theme of the conference

style: How well is the paper written? How sloppy is it? Is the English (or German ...) ok?

confidence: How confident is the reviewer about his own opinions? This depends on whether he understood most of it, whether he considers himself an expert in the topic etc.

The list should be adaptable per conference, but the above could be taken as *default.*

3.7 Event: Notification

Notification is the event which informs the authors about the final decision of the reviewing process. There are only two possible outcomes, namely *yes* or *no* for each paper. Besides the binary decision, the author is informed about the "opinion" concerning his submission. Concerning what information the authors are allowed to see, cf. also Section 4.

3.8 Event/document: Call for participation

As the call for papers (cf. Section 3.3), the call for participation is basically an advertisement. This time the addressees are not the potential authors, but potential participants of the conference itself. The call for participation contains similar information about the conference, but as additional information of course the program, i.e., the list of accepted papers with authors etc.

3.9 Phase: registration

This phase is characterized by the interaction of *participants* of the conference with the tool. Users can register with the conference, i.e., announce their participation. Again this will be done via some interface. The registration should be acknowledged.

The participant provides the usual personal information (name, title, affiliation). Furthermore, he is offered a number of options he must choose from:

- in which role he registers: as student (reduced fee) or as full participant

- preferred method of payment (+ credit card info, if applicable)

- dietary restrictions (vegetarian, etc)

Furthermore, if a user registers before a predefined deadline (*"early registration"*), the fee is reduced.

4   Users

By users, we mean people interacting with the tool/the service. People interact in various roles. They might include:

administrator: The administrator is not part of the scientific organization of the conference, i.e., he does not take part *semantical* decisions concerning the content. In a certain sense, he's outside the game.

chair: program committee chair (*chair* for short). There might be more than one chair. He or she is chairing the program committee and is therefore part of the committee. A chair is in some sense the "dictator" concerning semantical issues which means he is granted more privileges than the others.[6] Basically he can change all data all the time, for instance, he can remove or add people from the committee, throw out papers, etc. Again, whether it's smart to do so excessively is another question, but sometimes his intervention is necessary. A simple example: one of the reviewers does not do his work (for instance does not read the assigned papers), the chairman must be able to throw him out, which means for the tool: remove him from the data-base etc.[7]

reviewer: = program committee member

author: each paper must have at least one author, but there might be more. For the software, one author plays are more important role than the others, namely the one who interacts with the organizers. This is the **corresponding author**.

participant of the conference, i.e., a person who attend the conference and listens to the presentations etc.

outsider: That's the general public. An outsider does not interact at all, except having a look at the public web pages.

4.1   Restrictions and side conditions

The groups of people are not disjoint, in other words, an individual can interact with the service in different roles (at different times). Crucial is to obey certain *rules* concerning who (respectively who in which role) is allowed to do what or to see what. This also changes during time.

  Here is a number of informal restrictions.

1. a not-yet-accepted paper may only be seen by the author who submitted it.

2. the identities of submitting authors must be unknown to the outside and also among the authors themselves

---

[6]Whether it's wise to act as a dictator, overruling all the others in the committee, taking lonely decisions, or undoing decisions taken in consensus is a different question and outside the possibility of our modeling.

[7]Note that those semantical changes of the chair are different from what the administrator could do. The chair should be able to make the changes *within* the software, while the admin in some sense is outside. The admin can probably also change data, for instance by directly querying the data base (or wiping out the data base from the file system, for that matter ...) but that's something else.

3. a reviewer must not review his own paper, i.e., he must of course not influence the decision

4. an author never sees the discussion on his paper

5. an author never sees the identity of his reviewers

6. an author sees the verdict once all decisions are taken

7. a reviewer does not see the reviews of his colleagues, until he has sent his own review

8. a chair can see everything all the time.

9. it is expected that at least one author of an accepted paper participates at the conference, i.e., each accepted paper must be presented by an author

10. *optional:* a reviewer must not know the identity of the authors of the papers he reviews. This is also known as *double-blind* review.

5   Further requirements

In this section, we list some further requirements, which are not so much functional specification, but rather general requirements or also

5.1   Adaptability, parameterization, and options and defaults

Since the intention is that the tool is not usable for just one conference, but for many, certain things must be adaptable per conference. We call one conference an *instance* of a conference. Each instance of a conference is characterized by number of basic data, which must be provided in the configuration phase. At later phases, additional information (papers, reviewers etc.) are provided.

All in all, the whole thing is required to offer quite an amount of things the users can *adapt*. Part of these adaptability is described and discussed here and is part of the standard scenario of reviewing. There are much more things imaginable, for instance: besides the program chair, someone wishes to have a financial chair and a local organizer listed on the web-page, someone things the background color of the generated web-pages should be green, and weirder things ...).

To make the tool useful in practice, a good balance between hard-coded decisions and adaptable variations will be crucial. A couple of things should be kept in mind here:

Modesty: don't get carried away! The fact that one can easily include another battery of check-buttons to adapt this and than and other does not mean the user will like this.

Empty options/switching off features take care that options or features can also be switched off. If the user don't want a requistration phase in the, he may completely switch it off, for instance. This is related also to the point "modesty".

A useful technique in this context is also the choice of *good defaults.* There are always users, who want to fiddle with everything and they should be given freedom to a certain extent. There are also others, who don't bother, *provided* the default setting are usable. Therefore, making good defaults enhance and keep the number of necessary variation points beyond that small, enhances the usability quite a lot.

5.2   Webserver & download version

The software should be usable also as some web server managing more than one conference. The alternative is, that one user (in the role of the admin) downloads the source code, install it, adapts it for his or her needs and uses it for organization. The consequence of the "webserver" requirement is simple: more than one conference must be held in the data base at the same time, without any information flow (other than the information concerning the general public and other than by the administrator) between the two conferences.[8]

5.3   Changeability, non-standard scenarios

A general remark which applies to almost everything described above. namely: we concentrated on the smooth, standard evolution of the event.

In general, the tool must be prepared that this is not the case, which basically means, that choices once made must be *changeable*. Basically, all data in the data base must be "manually" adapted if need be, the problem is that is must not unnecessarily interfere with the process. Possible non-default scenarios are:

- a reviewer does not review, and is removed.

- an additional review about a paper is delivered, or a paper is re-assigned.

- the chair (or the committee) decides that one particular paper is rejected/accepted despite the fact that perhaps the numerical valuation or the status of the discussion would indicate otherwise. The valuation of the rest of the papers must not be "confused" by that.

---

[8]Actually, the download version is more challenging in some sense, it has higher requirements on platform independence and easy installation. The webserver can be installed and "massaged" by the experienced admin until it runs stable and then rot.