"Identify the Champion"

An Organisational Pattern for Programme Committees

Version 2.4 — July 29, 1998[1]

Oscar Nierstrasz

Software Composition Group, University of Berne[2]

Contents

1. Name

Identify the Champion.

2. Intent

Make the paper review and selection process for a scientific conference more efficient by focusing programme committee members' attention on whether or not they will "champion" a submitted paper during the programme committee meeting.

3. Context

Review and selection process for submissions to a scientific conference by a programme committee (PC) under the guidance of a PC Chair (PCC). The PCC collects submissions and distributes them to individual PC Members for review. Review forms are collected and sorted, and submissions are ranked prior to a PC meeting where the papers are discussed and either accepted or rejected for presentation at the conference. A PC meeting typically lasts 1-2 days, during which, say 20-30 papers are selected for presentation at a conference from a much larger number, say 50-200, of submissions. Each submission is typically reviewed by 3 or 4 PC Members.

4. Problem

Although, in principle, each submission is discussed at the PC meeting, in practice papers are only discussed if there is a "champion" who is in favour of having the paper accepted. If a paper has a champion, but no detractor (who is strongly against acceptance), then it is likely to be ac-

---

2. *Author's address:* Institut für Informatik (IAM), Universität Bern, Neubrückstrasse 10, CH-3012 Berne, Switzerland. *E-mail:* `oscar@iam.unibe.ch`. *WWW:* `http://www.iam.unibe.ch/~oscar/`

cepted. On the other hand, paper with a detractor but no champion is certain to be rejected. Submissions with both champions and detractors are likely to be strongly debated, and a submission with neither a clear champion nor a detractor is a "borderline" paper that is likely to be rejected unless (for example) there is a shortage of good submissions.

A common problem that arises during PC meetings is that the scores assigned to papers are typically *not* a good indicator of whether the paper will be championed or not. PC Members and external reviewers typically focus on trying to reach an objective evaluation of a submission rather than on how they will behave during the PC meeting. As a consequence, much time can be lost during the PC meeting trying to determine which papers have a chance of being accepted and should be discussed.

5. Forces

While the identification of champions (and detractors) clearly helps to structure and organize a PC meeting, it must be made explicit in the review and selection process from the very beginning, or a number of problems can arise to make the meeting less efficient:

- *Meandering discussion:* If no conscious effort is made to quickly identify champions, much time is wasted discussing papers that have no chance of being accepted. Very often, when a paper comes up for discussion, a PC Member will start by saying, "Well, I didn't like this paper because ..." This is not very useful, first of all because it does not tell the rest of the PC what the paper is about. Second, it does not lead to effective decision making, since the purpose of the meeting is to *accept* papers, not to reject them (i.e., it is more productive to concentrate on discussing papers that have a chance of being accepted than those that don't). Long unfocused discussions with delayed decisions may exhaust all reviewers. In the end, the decision taken may depend on who has the most stamina.

- *Inconsistent ratings:* Typically review forms ask PC Members to rate a paper according to a subjective scale, such as from 1 to 10, or from "strong reject" to "strong accept." Unfortunately these ratings are highly subjective, and may depend on the context of the other papers reviewed. One reviewer's "strong accept" may mean the same as another reviewer's "weak accept," just as one person's 9 may be worth another's 6. As a consequence, it can be hard to determine whether certain papers have a champion or not on the basis of the review forms alone, and much time can be lost during the meeting trying to identify which papers have a real chance of being accepted.

- *Nonsensical ranking:* Papers are typically ranked by some magical number calculated from a weighted sum of the various subjective scores on the review forms. This rarely leads to a useful grouping of papers, except for those at the very top and the very bottom of the list. Worse, it may lead to nonsensical groupings in which a paper with a "strong accept" from a reviewer with "low confidence" may be ranked equally to a paper with a "weak accept" from a reviewer with "high confidence."

- *Unprepared champions:* Very often a PC Member is surprised to discover at the PC meeting that he or she is the only champion for a submission. An unprepared PC Member may buckle under negative criticism of the paper and withdraw support. PC Members often pass on papers for evaluation to "subreviewers." This can be an efficient way to review large numbers of papers, as long as the PC Member carefully checks the papers and

reviews before the meeting. If the review is positive, the PC Member may end up being an unprepared champion.

- *Inexpert champions:* Sometimes during the PC meeting it turns out that the only champions for a paper have low confidence because they are not experts in the problem domain, while the experts either were not assigned the paper, or are only lukewarm about acceptance. In these situations it is often necessary to get an additional last-minute review from another expert.

- *Missing champions:* Worse, it may turn out that the only champion for a paper is not present at the meeting. It is inevitable that some PC Members — despite best intentions — will be unable to attend the meeting (or will need to leave early). Absent champions are almost useless, unless there is some alternative way for them to provide input to the meeting. This may require additional last-minute reviews, or long-distance phone conferences.

- *Late or missing reviews:* For one reason or another, some PC members will be unable to provide their reviews on time. A submission with an inadequate number of reviews can be neither accepted nor rejected and must be reviewed "on the fly" at the PC meeting.

6. Solution

Organize the review forms, the ranking and sorting of reviews, detection of conflicts, and the review meeting itself around the identification of champions. Use rating schemes with explicit operational meaning, such as "I will champion this paper", rather than implicit, subjective meaning, such as "strong accept", or "better than average", or "5". Group papers around presence or absence of champions and detractors rather than ranking them by weighted scores. Drive all discussions and decisions by identifying the champion.

6.1   Roles and Responsibilities

Author	
Responsibilities	Collaborators
Write and submit papers.	PCC.

PCC	
Responsibilities	Collaborators
Select the "best" papers for presentation at a conference.	Authors, PC Members, PC Meeting.
Handle all communication with Authors.	Authors.
Distribute papers for evaluation to PC Members.	PC Members.
Convene and run the PC Meeting.	PC Members, PC Meeting.

PC Member	
Responsibilities	Collaborators
Review submitted papers. Decide whether to be a Champion or Detractor for each paper.	PCC.

Champion	
Responsibilities	Collaborators
Clearly document all points in favour of (and against) acceptance of a paper on the review form.	
Champion a paper at the PC Meeting. Be prepared to defend the paper for acceptance.	PCC, PC Meeting.

Detractor	
Responsibilities	Collaborators
Clearly document all points against (and in favour of) acceptance of a paper on the review form.	
Argue against acceptance of a given paper at the PC Meeting.	PCC, PC Meeting.

PC Meeting	
Responsibilities	Collaborators
Collectively decide which papers to accept.	PCC, PC Members.

Papers and Review Forms are passive objects not represented explicitly here. The responsibilities are purely to record the communications of the Authors and the PC Members respectively.

6.2 Dynamics
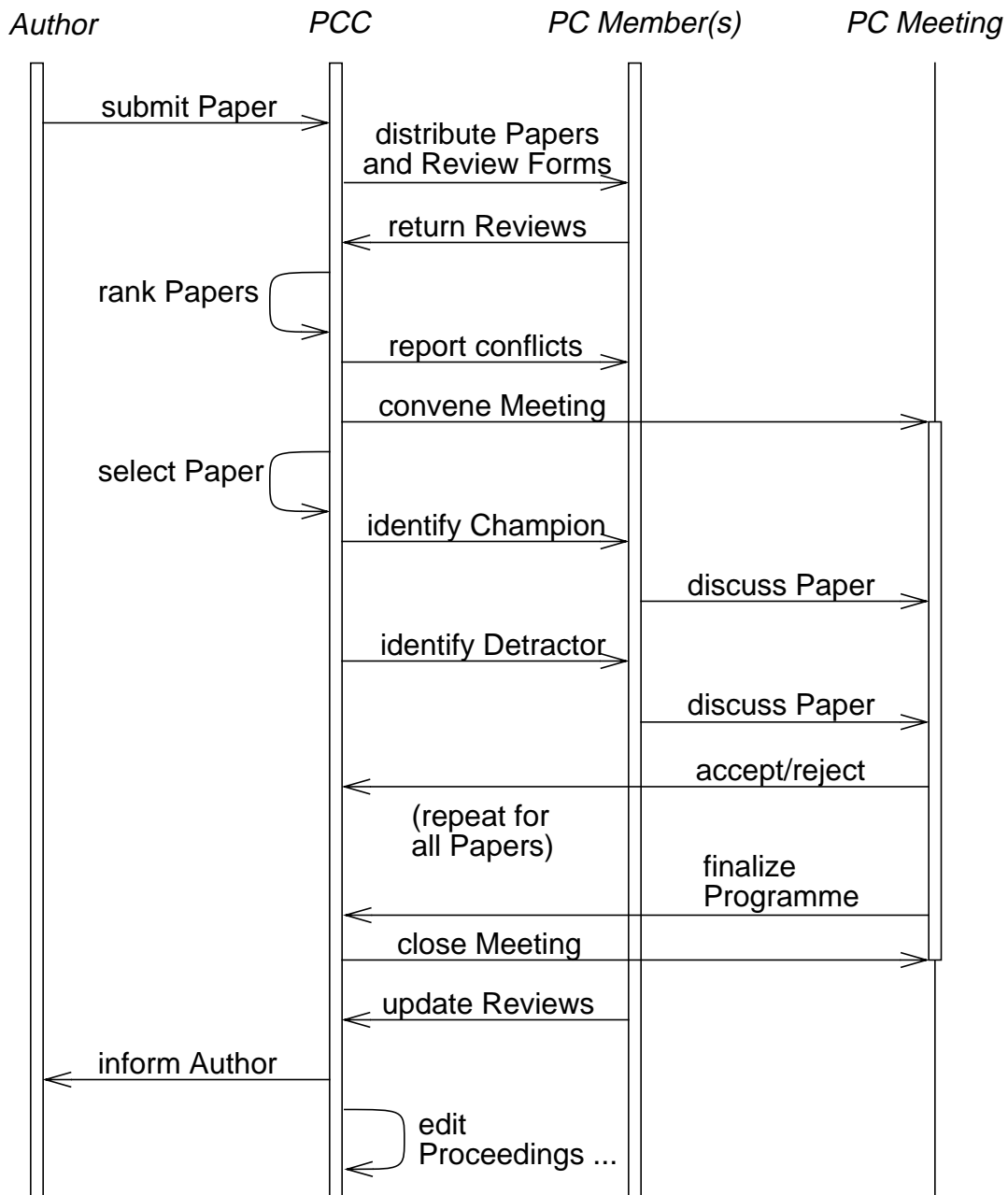
All communications are asynchronous. For simplicity, only a single scenario is presented in which Champions and Detractor roles are subsumed by the PC Member role. PC Members spring into being when the PCC formally invites them, and end their duties when the conference is over.Only a single PC Member is shown above. (Collectively, the PC Members are the same thing as the PC.) Papers and review forms, being passive objects, are not represented here. Distribution of papers may be preceded by electronic distribution of abstracts, allowing PC Members to "bid" for papers to review. PC Members may return papers to the PCC in case of conflict of interest, or lack of expertise.

6.3   Implementation

Most PCs end up applying the pattern "identify the champion" in some form or another during the PC meeting itself. In order to apply it effectively, however, it must be made explicit in the other phases of the selection process, namely during reviewing, ranking of papers, conflict detection, and the discussions of individual papers. A good way to achieve this is to have PC Members select the papers they want to review and to make the notion of championing explicit in the review form. The rest then follows naturally.

Distribution of Submissions

Papers can be distributed by the PCC alone on the basis of domains of expertise of the PC Members and the subject matter of the papers (keywords can help to match them up). Another common approach is to require Authors to "pre-register" their intent to submit a paper at least a week in advance, providing essential information that includes the title, keywords and an abstract. This information can be distributed to the PC, and PC Members can "bid" for submissions they are interested in reviewing. Although the second scheme is not necessarily more reliable than the first, it encourages PC Members to bid for papers that they feel they might like to champion.

Review Form Design

It is very tempting to introduce very fine-grained scales of appreciation on review forms, such as scales from 1-10 for various criteria, including originality, soundness, presentation, etc. These kinds of rating typically have the opposite of the desired effect, namely they waste the reviewers' time and they make it more difficult to tell who is willing to champion a paper.

In practice the only ratings that are really critical for the PC meeting are (i) a score indicating whether the paper should be accepted, and (ii) a score indicating the reviewer's expertise. The other issues, though important, normally appear in the written commentary (if they are relevant) and are not essential to running the PC meeting.

The most important thing about the paper's score is to make the *operational* semantics of the score clear. It frequently happens that a PC Member gives a paper a strong accept "because it was the best of the papers I had to review," but not because it was particularly good. The essential semantic categories are:

A: Good paper. I will champion it at the PC meeting.
B: OK paper, but I will not champion it.
C: Weak paper, though I will not fight strongly against it.
D: Serious problems. I will argue to reject this paper.

These four positions cover the interesting ones taken by PC Members during discussion. Finer gradations of appreciation are typically uninteresting. Note that it is not important how the scores are labelled — for example, they may still be numeric (i.e., from 1 to 5 or from 0 to 10), or textual (i.e., strong/weak accept/reject), but their meaning must be clear.

The most important positions are A and D, as these are, respectively, the champions and detractors. B and C are fence-sitters, but will also supply arguments for or against. The difference

between a B and a C is that B is basically in favour of a paper, but is not willing to champion it, whereas C is not impressed by a paper, but could be convinced if someone else champions it.

Separate ratings of high or low confidence are not especially useful, since low confidence tends to show up anyway as a B or C position.

A separate rating for the reviewer's expertise, on the other hand, is essential to detect the *inexpert champion* situation. The following ratings are used *only* for conflict detection and *not* to rank papers:

X: I am an expert in the subject area of this paper.

Y: I am knowledgeable in the area, though not an expert.

Z: I am not an expert. My evaluation is that of an informed outsider.

Note that it is not necessarily the intention that all reviewers be experts. In conferences with broad scope, it can be useful to have some non-expert reviews to evaluate a paper's accessibility to a general audience. Only in rare situations, however, should a non-expert consider championing a paper.

The scores and expertise ratings would normally *not* be revealed to authors, as this information is purely procedural, and has no function after the selection process is done.

The details of the remaining parts of the review form will vary, but the most important ones for applying this pattern are the following:

Summary: A champion should always summarize a paper before presenting arguments in favour of acceptance. It is not easy, however, to remember details of 20 papers. It is therefore a useful exercise to briefly summarize each paper reviewed in a few short sentences. This information should be returned to the authors.

Points in favour/against acceptance: Whether a reviewer decides to champion a paper or not, the arguments for or against acceptance should be explicitly presented during the meeting. Listing them explicitly on the form makes it easier to get to the point during the meeting. Remarks concerning originality, presentation, soundness, etc. should be listed here, but need not be assigned a numerical score. This information should also be passed on to the authors.

Additional comments for the authors: It helps to separate the arguments that are pertinent to the selection process from additional comments only of interest to the authors. (In general, it is good policy to pass on all information that will help authors to improve their papers, whether they are accepted or not.)

Additional comments for the PC: If the above three parts are returned to authors, a separate part should be available for additional remarks that should be raised during the meeting, but not returned to authors.

The review form may also contain other parts useful for running the meeting or the conference itself: Paper number, Authors, Title, Referee. Should the author be invited to present a demo? Is the paper suitable for receiving an award? etc.

Ranking and Sorting

The purpose of ranking and classifying papers is to give some structure to the PC meeting by grouping together papers that are likely to require the same kind of debate. Whereas numerical rankings typically fail to achieve this, grouping by presence or absence of champions quickly gets to the point.

A particularly simple and effective way to group papers is to assign a two-letter code to each paper, consisting of the highest and the lowest scores, and to sort the papers by this code. This yields 10 groups of papers, of which 7 are interesting:

AA, AB: All reviews are positive, and there is at least one champion. These papers will almost certainly be accepted.

AC: This means that all reviews are A, B or C. This is a likely accept, since there is at least one champion, and no strong detractor. The only question is whether the reservations of the C review are serious or not.

AD: This is a serious conflict, and will certainly lead to debate. Note that this does not distinguish between cases where, for example, we have three As and one D, or one A, one B, one C and one D. In practice, the important positions are the extremes.

BB: All reviewers are fence-sitters. Everyone likes the paper, but no one is willing to be a champion. The discussion should determine whether the B's are really A's or not.

BC: These tend to be borderline papers, since no one is willing to be either a strong advocate nor a detractor. Such papers are often put on a "slush pile" and resurrected or discarded after the rest of the programme has been defined.

BD: These papers are likely to be rejected. There is no strong champion, but there is a strong position against acceptance. Such a paper might still be accepted if the B decides after all to champion it.

CC, CD, DD: These papers are almost certain rejects. Papers may be resurrected from this group only under exceptional circumstances, for instance, if it turns out that none of the reviewers were experts, but another PC Member who is an expert in the domain reads the paper during the meeting and decides to champion it.

Note that this classification scheme works independently of the number of reviews each paper receives. What is significant are the high and low scores.

Even though papers are grouped rather than ranked numerically, it is still useful to assign a numerical rank to each paper, based on its classification and, say, alphabetical order of authors or of paper title, for easy identification during the meeting. ("What paper are we discussing now?" "We just finished with paper number 12 ranked 46 on the list, and we're going back to paper 6 ranked 22.")

Conflict Detection

Since PC meetings are expensive to organize (think not only of the travel and hotel costs, but of the salaries paid for those attending!) and cannot be repeated, and the selection process depends so heavily on the identification of champions, it is important to detect potential problems *before*

the meeting takes place. This means that reviews should be returned to the PCC well in advance — typically at least a week before the PC meeting.

To reduce delays, to facilitate analysis of the results, and to permit automatic preparation of review packages, it is a good idea to distribute review forms electronically.

The following situations should be detected early to help PC Members better prepare for the meeting:

- *Missing reviews:* This is the most basic problem to check for. PC meetings often start with only one or two reviews received for some of the papers. Each paper should receive a minimum of three reviews for a fair review process.
- *AD classification:* A PC Member who is a champion or detractor for such a paper should be prepared for a debate. It is useful to know this in advance. If the review was written by a subreviewer, it is essential to be warned in advance so the PC Member can validate or overturn the review.
- *Inexpert champions:* It may be that all the A and B ratings come from non-experts. Typically non-experts will back down from an "accept" position if there is dissent from an expert. In such cases it can be useful to solicit an extra review from an expert in advance of the meeting.
- *Missing champion (or detractor):* The PCC should check whether anyone who cannot be present at the meeting happens to be the only champion or detractor for a paper. An email exchange in advance between the reviewers may help them come to a consensus in advance of the meeting, or at least to clarify the source of disagreement.
- *Absent reviewers:* This is a variation of the above, in which none of the reviewers are present to present their views of the paper. In a large conference with 200 submitted papers and 20 or 30 PC Members, it is almost inevitable that a couple of papers will fall into this category. These papers should be discussed by email. If necessary, another review should be solicited by an attending PC Member.
- *Low overall expertise:* If none of the reviewers is an expert, then the selection process can break down regardless of the scores given by the reviewers. In such cases the PCC should solicit an expert review in advance of the meeting.

Finally, papers submitted by PC Members, or for which PC Members have a conflict of interest (i.e., papers submitted by close colleagues) require special treatment during the meeting, but do not constitute problems as such.

Running the Meeting

It is good to set some ground rules to keep discussions focused. For each paper, the champion, if one exists, or the closest there is to a champion, should introduce the paper by briefly summarizing it and presenting the points in its favour. Then, the detractor (or whoever has the strongest negative points) should speak next. Finally the remaining reviewers can back up these arguments, or fill in missing points. If there is a "detractor," then the champions and detractors typically play the roles of defence and prosecution in a trial, and the rest of those present play the role of the jury. Frequently either the champions or the detractors become convinced by the arguments of the other, and a consensus is quickly reached.

PC Members should also be reminded what are the criteria for acceptance. These may be more stringent, or more lax, depending on the nature of the conference, or may be quite specialized. Typically, an accepted paper should have a clear, original contribution, and fulfil the usual criteria of readability, completeness, etc. Originality is a strong criterion, and many papers fail to be accepted if they do not clearly demonstrate new results. A champion/detractor should address these specific criteria.

If there is no clear champion for a paper, the discussion should focus on checking why no one wants to champion it (i.e., to try to smoke out a reluctant champion). If no champion can be identified, the paper can be quickly rejected.

Delaying a decision on a paper is almost always a bad idea. A decision should only be delayed if something will happen in between that may change the outcome, i.e., if an expert will check the paper for originality. Borderline papers (BC grouping) may be delayed until the other papers have been considered.

Certain papers, such as those submitted by PC Members, may be subject to more stringent criteria, such as "must be better than the average paper accepted." This should be translated to something more precise, such as "must be championed by all reviewers", or, "there must be no detractor."

It is highly recommended to supply each PC Member attending the meeting with copies of all the reviews for which they do not have conflicts. This makes it easier for everyone present at the meeting to actively participate in the decision-making process, even for papers they have not personally reviewed.

7. Examples

This pattern is invariably applied in some form during PC meetings for computer science conferences, though it is typically not well integrated into the rest of the review process. In particular, most review forms do not make the notion of championing explicit.

The ICSE 98, ECOOP 98, OOPSLA 98 and FSE 6 PCs have implemented review procedures essentially as is described here.

Dick Kemmerer, the PCC of ICSE 98 reports [personal email communication]:

> As far as using the approach at ICSE98 goes, ... I had six categories instead of four. ... The categories were: ... Will argue for acceptance (A); Inclined to accept (B); Not opposed to acceptance (C); Not opposed to rejection (D); Inclined to reject (E); Will argue for rejection (F)
>
> Having the two middle categories caused some problems, and I would use only four if I were to do this again.
>
> Before the meeting I separated the papers into three groups:
>
> > Group 1 (Likely Accepts): AA, AB, AC
> >
> > Group 2 (Mixed): AD, AE, AF, BB, BC, BD, BE, BF, CC, CD, CE, CF
> >
> > Group 3 (Likely Rejects): DD, DE, DF, EE, EF, FF
>
> In addition, all PC member papers were postponed until all other papers were decided.
>
> We first discussed the Group 1 papers and they were all accepted with little time devoted to them. Next we rejected all of the Group 3 papers without discussing them, although I stated that any PC member that wanted to bring one of them up was welcome to do so.

This did not happen. The middle group as you predicted took the most time.

I also distributed reviews for papers with conflicting reviews to the reviewers beforehand for online discussion. When reviews were changed I entered the updated review before the PC meeting. I also informed each PC member what papers they would champion before the PC meeting, so that they were prepared.

8. Resulting Context

Forces Resolved

Identifying the champion forces PC Members to focus on their behaviour during the PC meeting rather than on their subjective impressions while reading the paper.

Asking PC Members to "bid" for papers to review reinforces championing.

Grouping papers according to whether or not there are champions or detractors gives better focus to discussions than rankings by weighted scores.

PC Members who delegate papers to subreviewers are aware that they must be prepared to play the role of champion/detractor on the basis of the review. They are warned in advance if they are the only champion for a paper.

Identifying champions in advance helps everyone be better prepared for the PC meeting.

Champions who are unable to attend the meeting have a better chance to influence the meeting if they are identified explicitly well in advance.

PC papers should be accepted only if there is at least one champion and there are no serious (expert) detractors.

Explicitly encouraging PC Members to champion papers provides opportunities to draw reluctant champions out in the discussion. (Each reviewer can be explicitly asked, "Are you willing to champion this paper?")

Discussions tend to be shorter and more focused if discussions can only take place when a champion is identified. Delays only take place if there is a chance that a new champion can be identified.

Consequences

"Identify the champion" can stifle debate if applied too rigorously. One must be careful not to discourage reluctant champions.

Related Patterns

Dick Kemmerer also points out a variant of "Identify the Champion" that has some distinct drawbacks:

Several weeks before the PC meeting I told the PC that we would be following a procedure close to what you described in the paper, and gave them all a pointer to the paper, asking them to read it before the meeting. I immediately got an objection from one pc member who said he did not want to use the "champion" approach. ... I mention this because there are folks out there that have a preconceived notion of the "champion" ap-

proach. At these PC meetings a paper would not be accepted *unless someone was willing to put their name on it* (i.e., their name appeared on the paper as "recommended by"). ... The problem seems to be that people are less likely to want their name to appear in print as being the endorser than to be a champion as per your approach.

"Shepherding" is a pattern of guiding a paper (or rather, its authors) through rough terrain so that it can reach ground where it can be truly championed without reservations. Some conferences (like PLoP) make heavy use of shepherding, whereas others avoid it. A variant is "Conditional Acceptance" in which the PCC exercises discretion over publication of the final version of the paper. Yet another variant is "Mentoring," in which authors who seek advice in preparing a paper for submission may be assigned a "mentor," who is usually a PC Member.

"Write to the program committee," is a pattern that authors can apply to increase the chances that their papers will be accepted [Kent Beck, OOPSLA 93 Panel]. The idea is to write in such a way as to win over a potential champion by catching his or her attention and providing good ammunition to argue for acceptance during the PC meeting. A good understanding of the dynamics of PC meetings helps in applying this pattern. A related tactic is to try and identify specific PC Members who are likely to review the paper, and write in such a way as to win them over as champions. (It almost goes without saying that this tactic can easily backfire!)

Acknowledgements