How to Write Research Reports and Papers with High-Quality
— *Some Personal Perspectives*

Qing Li (李青)
Dept of Computer Science,
Multimedia software
Engineering Research Centre (MERC),
City University of Hong Kong

Outline

* **Motivations**
* **Writing research reports and papers**
  - **Some FAQs**
  - **Some rules & pitfalls**
* **How to prepare…**
  - **Warm-ups**
  - **Some "tricks"**
* **How do we evaluate others' papers**
  - **Case studies**
* **Summary (with hints & reminders)**

Motivations

* **Writing research reports and papers with high-quality papers**
  => increasingly important!
  ○ **A "must" for academic staff in many institutes**
    * to advance research study with evidence
    * to enrich education via research
    * to achieve career goals, etc.

So, like it or not, we all have to!

Some FAQs：

Q1: How could we write high quality reports and papers from our research?
  * Are there "rules" to obey?
  * Are there "pitfalls" to avoid?
  * Are there "tricks" to learn?

Q2: How can we be better prepared to get into the starting line?

Q3: How do we evaluate (others') research reports and papers?

What are "poor quality" ones?

1. Loosely connected and/or flat-structured

2. Results inadequately demonstrated

3. Poor presentations (not just errors)

4. Bibliography incomplete or outdated

5. For research papers, the topics too old and/or not promising

…

Writing Good Research Reports and Papers -- Some "Rules"

R1: Topic should be something new!
- quite like cooking: you need good valuable/nutritious ingredients (fresh ideas with good output)

- "delta" papers not welcomed by top-tier journals & conferences!

Writing Good Research Reports and Papers -- Some "Rules"

R2: Need to be well abstracted!
_ not a summary of the report/paper
_ not yet another introduction
_ not the ToC

As an example, for a paper titled "OS_RANK: Structure Analysis for Ontology Ranking", which of the following two versions of "abstract"is more proper?

Abstract -- V1

The Semantic Web has greatly changed the current Web by using ontologies to describe information resources. In order to exploit the rich semantics of the Semantic Web and facilitate the ontology constructions, methods are needed to rank the ontologies on the Semantic Web. In this paper, we propose and test OS_RANK for ranking the ontologies on the Semantic Web. By analyzing the structure and semantics of an ontology, OS_RANK ranks the ontologies according to whether or not an ontology has a cover of the query term. The advantages of OS_RANK include (1) the implicit and complex semantics on the Semantic Web can be explicitly represented and referred; (2) We can set an adjustable value of the ontology cover of the query concepts, so that users can gear the weights of the ranking measures according to the needs of their applications.

Abstract -- V2
Ontologies have been shown as beneficial to representing domain knowledge, and are quickly becoming the backbone of the Semantic Web. The need for ontology ranking techniques is crucial as the ontology reuse becomes increasingly important. In this paper, a new approach called OS_RANK has been proposed for ranking ontologies on the Semantic Web. In particular, query terms provided by users are regarded as containing special information about domain knowledge of interest. Each ontology candidates are analyzed separately and ranked with respect to the structure and semantics. Experiments are performed and the results show that OS_RANK is cost-effective.

Writing Good Research Reports and Papers -- Some "Rules"

R3: Presentation should be proper!
- Again like cooking: you need good skills/techniques to put things
  in right order, with proper tone,
  in high clarity, with right heating

- Exhibit "hierarchical" structure
  -> Keypoints must be clear, and
  -> logic flow should be hill-climbing like yet very smooth

Writing Good Research Reports and Papers -- Some "Rules"

R4: Result(s) must be convincing!
We need to hold firm standpoints, with strong evidence and proofs
_ theoretically
_ conceptually
_ empirically

R5: References must be fully cited!
(What's wrong with this list below?)

"Pitfalls" to Avoid?

* Little/no idea on who your readers are
   > ask yourself about whom your report or paper is intended to, why and how would they like to read it!

E.g., *if the report is to be read by your boss, what aspects would he/she be concerned (interested in) most?*
If you publish a paper in an XYZ journal, will the majority of the readers be PG students and academic peers only, or from industrial sector also?

7. REFERENCES

[1]   Li, Q and ZHUANG, Y, "A Semantic Data Modeling Mechanism for a Multimedia Database", Multimedia Information Retrieval and Management , D Feng, W C Siu and H Zhang et al. (eds) , Springer-Verlag, Germany, p305, (Chapter 9, LNCS), 2003.

[2]   Bai, J., J. Y. Nie, G. Cao, and H. Bouchard. 'Using Query Contexts in Information Retrieval', Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval, ACM, 2007.

[3]   Rahurkar, M, and S Cucerzan. 'Predicting When Browsing Context Is Relevant to Search', Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval, ACM, 2008.
…
[13] Kritikopoulos, A., and M. Sideri. 'The Compass Filter: Search Engine Result Personalization Using Web Communities', Intelligent Techniques for Web Personalization, 229-240.

[14] Huang, X., F. Peng, A. An, and D. Schuurmans.'Dynamic Web Log Session Identification with Statistical Language Models', Journal of the American Society for Information Science and Technology Vol. 55, No. 14, 1290-1303, 2004.
…

"Pitfalls" to Avoid?

* **Plain and unattractive introduction…**

> **A clear introduction with good motivation**
=> **(almost) half of your success!**

> **It should not be too lengthy…**
=> **let reader/reviewer see the main body in less than 1/3 of the paper!**

We show next a "good" example…

---

1. INTRODUCTION

How to collect and effectively manage various kinds of recipes on the Web, ranging from purely text-based to multimedia-flavoured (with image/photo, audio and/or video clips), becomes a meaningful and important issue, especially for people who are interested in cooking and dining, and also for people who want to study cooking from academic/research perspectives.

In comparison with traditional data types, recipes have their distinct characteristics which make most (if not all) of the conventional data models unsuitable/inapplicable for such data, including:

* **Loosely structured**: recipes are usually semi-/loosely- structured data, with various levels of details and granularities. Some are more detailed/ wordy than others.
* **Behavior oriented**: recipes are not only data-intensive, but also behavior oriented, in that the main part of a recipe is about the procedure to follow in cooking a dish.
* **Constraints bound**: recipes are usually also bound by various constraints which are applicable to either individual actions or a sequence of actions. Such constraints often contribute to the "skillful" way in cooking a dish successfully.

Aiming at providing practically useful services for users, we are engaged in developing an adaptive, personalized multimedia recipe database system. Figure 1 illustrates a layered view of the overall system architecture, which consists of three layers – Data Collection and Management Layer [1], Recipe Management Layer and User Interaction Layer.
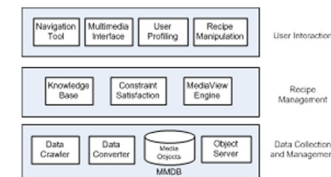

Figure 1. Layered System Architecture: an Overview

A user of our system may learn how to make a dish through studying a recipe retrieved from the database; he/she may also want to know if and how an existing recipe can be adapted so as to accommodate real-world situations (e.g., due to lack of some ingredients), or to better fit his/her own taste preference by replacing some cooking step or method. It would be more than desirable if some advices (on-line or via human experts) are available for people to consult in advance, so that many undesirable/erroneous actions can be avoided. Even better, if the system can work like a professional advisor to teach users how to identify a particular cooking style, a relevant cooking skill or how to evaluate one's cooking craft, users will get more benefits.

The **main contributions** of this paper include the following:
- We describe our approach on recipe data modeling (which is the first such work to the best of our knowledge). This model encompasses static attributes as well as dynamic behaviours including cooking procedures and constraints.
- We present a novel perspective of evaluating the "quality" of a recipe through constructing and analysing its *cooking graph*, in which both action flows and data (ingredient) flows are captured.
- An intelligent mechanism is advocated to handle real-time dynamic recipe adaptation cases so as to provide users with flexible and feasible solutions whenever an exception occurs.

The rest of the paper is organised as follows. Section 2 reviews some related literature work. In section 3, the details of our *Dish_Master* recipe model are presented. Section 4 further teaches users to understand the cooking procedure by using our recipe model, and how the system helps when users meet difficulties during recipe execution and/or adaptations. We describe an experimental prototype in section 5, and conclude the paper with summary and future research directions in section 6.

"Pitfalls" to Avoid?

* **Over ambitious in scope…**
> No one can solve and report all the problems in one paper!!
> So try to avoid covering too many, just focus on one or two, and do it well and in-depth!

* **Making a wrong assumption…**
> Or is it a refutation of a hypothesis on the problem?
> An explicit claim, please!
(and strong evidence to support your claim!)

"Pitfalls" to Avoid?

> **Contributions** vs. **Research Steps**
  _the former needs to have "novelty", with
    * **strong basis**
    * **convincing or promising results**

  _the latter is about what you (plan to) do
    * only highlight your research approach
    * more like what you <u>report</u> on your work

  _IMPORANT: Don't mix them up!

"Pitfalls" to Avoid?

> **Over-claim or under-claim…**
  Does your technique perform better than its rivals in any/all of the following dimensions?
  _behavior: X has a higher success rate than Y, or produces better quality output
  _coverage: X is applicable to a wider range of applications than Y
  _efficiency: X runs faster or uses less space than Y
  _ useability: users find X to be easier to use than its rivals

"Pitfalls" to Avoid?

> **Duplicate submission and publication?**
  Conference vs. Journal versions…
    _Journalizing a published conf. paper is fine, if
    Ǝ enough (eg, 30%) new stuff
    _ the other way around is not OK!

> **Outcomes from the same (large) project?**
  Again, some limited overlap is fine, but must be with enough (>70%) new materials

> **In all cases, NO plagirism !!**

HOW To Prepare?

* **Warm-ups**
  > **Find and read your peers' papers!**
    >> **Where to find**
      (main stream technical journals, conference proc., with high citations if possible;
      <u>Google scholars</u> provide a very convenient source on citations…)
    >> **How to read**
      (the "reverse process" of writing; from crude reading to intensive reading!)

HOW To Prepare?

* Warm-ups
> **Digest your peers' papers…**
>> **Try to ask questions!!**
(if no question can be raised, then you don't understand their work)
>> **Discover their works' limitations**
(could be their assumptions, their scope, and/or their results)
>> **Identify possible "entry points"**
(which you're comfortable to work on, and confident to surpass)

HOW To Prepare?

* Getting into the basic requirements…
-- per typical *System* (also *Application*) *Oriented* Paper Structure

> **Title:** be indicative and concise!
(Normally, no more than 2 lines…)

> **Abstract:** not a summary of your paper, but of your research problem, your claim, and evidence!

HOW To Prepare?

* Diving into the basic requirements…
-- per typical *System* (also *Application*) *Oriented* Paper Structure

> **Introduction:** *// Very Important!*
_Motivation (possibly w/ an example)
_Significance and Expected Contributions
_Paper Structure

Key points: Clarity, Straight forward Focus, Proper Claims

HOW To Prepare?

> **Related Work:** // can be in 2nd, or 2nd last section
_review on related, representative works (of others) supporting the motivation

_A good place to show your width and depth of knowledge in the related field!

_A proper place to acknowledge and show your respect of your peers' work if wanted

_need to differentiate from your own approach and/or contributions
Key points: Critical Review, Adequate Citation

HOW To Prepare?

> **Problem Statement & Algorithm Design:**
-- **The Main Body** for highlighting your contributions…

  * need to explain and define your ideas in detail, *preferrably with examples*

  * for computer science papers, eg., algorithm(s) in pseudo code OK, but with some explanations, analysis and/or discussions

  * Good to add *relevant* lemma, observations, properties, theorems; *not to mix with definitions!*

HOW To Prepare?

> **Evaluation:**
_evidence to support your claimed contribution(s) **objectively**

_experimental results over real and/or synthetic data are **required** unless it's a theory/conceptual paper (for the latter, either theoretical proofs or case studies need to be provided)

Key points: Extensive, Intensive, Objective

HOW To Prepare?

> **Conclusion:**
_summary of what you have done
_discussion of the significance
_mention of **some** future work issues

> **References:**
_Enumeration of cited works related to your reported research
_Format needs to be consistent and follows the conventions
_Make sure no "dangling" references!

Writing Good Research Reports and Papers -- Some "Tricks"

T1: Design well before writing
- Assess the purpose: what you want the readers to obtain and take away

T2: Avoid using long sentences
- short sentences more preferred! No need to overly show off your English proficiency!
- all good ideas can be simply explained and presented, or else…

Some "Tricks"

T3: **Use positively tensed sentences as many as possible**
- **Only in very special cases** you'll need to use passively tensed ones

T4: **Provide adequate figures (图文并茂)**
- **Often, a picture is worth more than a thousand of words!**
- **On crude reading, a reviewer may get delighted when some figures/photos and diagrams are found at his/her glance, or get depressed when none is found**

Some "Tricks"

T5: **Put enough time and emphasis on the design of your report/paper, in terms of the topic, approach, the structure, and, the Introduction!!**

T6: **Be aware of who your "enemy" are**
> **check who're the PC members or editorial board members, and cite their relevant work adequately!**

How do we evaluate a paper…

* *Writer's own view*
    _ more intimate
    _ more tolerant
    _ more understanding, and
    _ less presumptious
    _ less intruding…

=> **More like treating our own baby!**

How do we evaluate a paper…

* *Reviewer's View*
    _ **more critical and picky**
    _ **less patient**
    _ **much more concerned with**
        > **readability**
        > **overstatements**
        > **completeness of**
            * *solutions*
            * *citations*
            * *acknowledgement*

How do we evaluate a paper…

* *Editor's View*
  _ Quite often, his/her decision is entirely dependent on the review reports, with some exceptions…

  _ more concerned with
  > relevance (to the journal)
  > interests (of the journal audience)
  > impact to the journal
  * *citations of past papers from the same journal !!*

A typical review form…

IEEE Transactions on Knowledge and Date Engineering Review Form

Section I. Overview

 A. Reader Interest

  1. Which category describes this manuscript?
     E.g. Research/Technology

  2. How relevant is this manuscript to the readers of this periodical? Please explain your rating under Public Comments below.
     E.g. Relevant

A typical review form…

IEEE Transactions on Knowledge and Date Engineering Review Form

B. Content

 1. Please explain how this manuscript advances this field of research and/or contributes something new to the literature.
  E.g., Four feature selection metrics are presented and tested on two data sets against three metrics previously known from the literature

 2. Is the manuscript technically sound? Please explain your answer under Public Comments below.
  E.g., Partially

A typical review form…

C. Presentation

 1. Are the title, abstract, and keywords appropriate? Please explain under Public Comments below.

 2. Does the manuscript contain sufficient and appropriate references? Please explain under Public Comments below.

 3. Does the introduction state the objectives of the manuscript in terms that encourage the reader to read on? Please explain your answer under Public Comments below.

 4. How would you rate the organization of the manuscript? Is it focused? Is the length appropriate for the topic? Please explain under Public Comments below.

 5. Please rate the readability of the manuscript. Explain your rating under Public Comments below.

A typical review form…

Section II. Summary and Recommendation

 In an effort to adhere to our strict page budget and maintain a healthy yet short publication queue, we are trying to better enforce our long standing page limitations and formatting guidelines with our authors. In order to help them adhere to these guidelines, we need your support. One way you can help us meet these objectives is by suggesting ways to maintain the lengths of their manuscripts should you decide to ask the author to add new content

 Evaluation   Please rate the manuscript. Please explain your answer.
 E.g. Fair

 Recommendation
 Accept With No Changes
 Author Should Prepare A Minor Revision
 Author Should Prepare A Major Revision For A Second Review
 Revise and resubmit as "new"
 Reject

A typical review form…

Section II. Summary and Recommendation

Comments
Confidential Comments (authors will not see these comments)
E.g. I don't think their work/presentation is serious…
(or I've seen similar work published at xyz conference before…

Public Comments (these will be made available to the author)
Please give detailed justifications and explanations for your assessments, including positive and negative aspects of the manuscript. Without clear, detailed support, it is hard for authors and editors to make good use of the review.

Miscellaneous

M1: If not getting accepted?
 * **For a research paper to a conference**
   > **Some conferences allow authors to rebut**
   > Most don't, so you need to revise according to review comments, and try some other again!
 * **For a research paper to a journal**
   > **Usually a paper does not get accepted in its first attempt; needs to be revised**
     >> **major vs. minor**
   > **How to write responses to reviews?**
     >> **must be complete (point to point)**
     >> **try to be polite and thankful**

Response to Reviewer Comments

Comments from Reviewer 1:
In this paper, a coalitional game is employed to model the community-based cooperation among Web services, and a distributed coalition formation algorithm is proposed. The topic looks interesting to the readers. The paper is well organized, though some parts are hard to read. References are sufficient and appropriate too. However, this paper suffers from some drawbacks.

First, in Section 2.1, it defines availability as the probability that the response is received within a specified time frame, and associates correctness with reliability. In fact, availability specifies uptime percentage, and reliability is relative to malfunction, not correctness.

Response: A new definition of service availability proposed in [1] is used in the revised version.
[1] N. Milanovic and B. Milic, "Automatic Generation of Service Availability Models," IEEE Transactions on Services Computing, 4(1): 56-69, 2011.

Comments from Reviewer 1 (cont'd):
Second, the demand curve in Fig. 2 is problematic too. It is used to deduce formula 1, which is fine in itself. But, it cannot conclude that the demand of Web service j is no less than that of Web service i. As the availability of Web service j is higher than that of Web service i, it makes sense that its price is higher than that of Web service i. However, it cannot lead to the conclusion above, as they differ in quality of service, i.e., availability.

Response: Thank you for pointing out this. We have fixed the problem and reorganized the corresponding paragraphs.

Third, it claims that the proposed algorithm leads to 76% improvement in the average availability per Web service, when compared with non-cooperation. But, this argument is problematic, as the paper defines a coalition as a Web service in Section 5.1. In fact, as a coalition contains several functionally equivalent Web services, it is no surprise that the average availability per coalition could be increased via redundancy. However, it cannot draw a conclusion that the average availability per Web service is increased.

Response: We agree with your comment that redundancy can always improve the availability of coalitions. Actually, this is one advantage of our coalition formation algorithm. More importantly, our algorithm does not sacrifice WS autonomy and can improve the average payoff per WS during the process of coalition formation.

Miscellaneous

M2: Proper usage of footnotes and appendices
* **It is sometimes convenient to use footnotes**
    > **on something additionally explanatory, or maybe orthogonal but good to add on…**
    > **Must not be used excessively!**
* **For appendices, normally used to include extra details in case of further interests/inquiries**
    > **Examples like proofs of some theorems or lemmas, procedural source code (partial), etc.**
        >> **still count towards the paper length!**
    > **Research reports seldom use appendices (why?)**

Some useful hints (for practice):

* **When read other authors' paper:**
    _ take notes (on good & bad things)
    _ raise questions
    _ learn from it
* **When read your own paper:**
    _ pull yourself out and pretend you are not the author
    _ raise questions
    _ take notes (on good & bad things)

Some useful hints (for Attention):

* **English obstacles:**
    _ "a" and "the"
    _ plural and singular
    _ correct tense, please!
* **References:**
    _ **Develop the habit to cite, and cite completely!**
        > complete details
        > acknowledge adequately
    _ **Be consistent in format!!**

Summary (and reminders)

* Choose an interesting/promising topic, and NO "delta" paper (contribution/value)

* Start with a good introduction+motivation

* Don't be too ambitious (scope)

* Produce and present convincing results

Summary (and reminders)

* Your research report / paper should be unique in the field (novelty, technicality, significance, clarity, well-structured, etc.)

* Reading and citing relevant and updated papers from premium forums (adequate references)

* Last (but not least):
    Practice, Persistence, Progress!!!

Thank You!

itqli@cityu.edu.hk