On Generation of New Ideas for PhD Research in Computer Science and Engineering: An Analysis

Authors:

Vladimir Blagojevic

Dragan Bojic

Miroslav Bojovic

Slavko Gajin

Veljko Milutinovic

Bosko Nikolic

Jelica Protic

Igor Tartalja

Milo Tomasevic

Table of Contents

- 1. Introduction
- 2. Other Relevant Studies (10)
- 3. Classification, Explanation, and Some Popular Examples (10)
- 4. Examples from Nobel Prize Research (10)
- 5. Examples from Turing Award (10)
- 6. Examples from the UCLA List of Top500 Computer Scientists of the World
- 7. Numerical Values from Google Scholar
  (distribution of 10 methods in Top10 papers of 10 different fields)
- 8. Examples from Authors of this Research (10)
- 9. Experiences with PhD Students of the Authors of this Research (10)
- 10. Conclusions

Introduction: General Strategy

Stage #1

- 1. Topic selection and collection of 20-40 research papers
- 2. Analysis and description (7S+F1infra+F2algo)
- 3. Classification criteria and classification tree/cube
- 4. Ideas for future research, along the lines of 10 methodological paths
- 5. Survey paper, in Serbian
- 6. Survey paper, in English, for an IEEE or an ACM conference
- 7. Survey paper for a journal (ACM, IEEE, SCI)
- 8. Springer book 50-125
- 9. .ppt equivalent (preconference tutorials + teaching for Bologna credit)
- 10. Living thru the incubation period

Stage #2

- 11. Doing a research paper (1 of 10) along the lines of 10 method paths
- 12. Publishing (SRB.con + ENG.con + ACM/IEEE/SCI.journal)

Stage #3

- 13. Writing an EU project proposal, with appendices from the above
- 14. Approaching the US companies, with knowledge from the above

Stage #4

- 15. If the project was successful, prepare a business plan for a startup
- 16. Work on turning the EU project prototype into a market product

Ten Idea Generation Methods*

Mendeleyevization: Inductor versus Catalyst (M1 vs M2)

- **Definition:**

  If one of the classification class includes no examples,
  it first has to be checked why is that so.
  If it is so because it makes no sense, an appropriate explanation is in place.
  If it is so because the technology or the applications are not yet ready
  for such an approach, one can act in the same way as the famous chemists Mendeleyev:
  Empty positions in any classification are potential avenues leading to new inventions.
  We refer to such an approach as: Mendeleyevization.

A: Mendeleyevization (Inductor versus Catalyst) – M1 vs M2

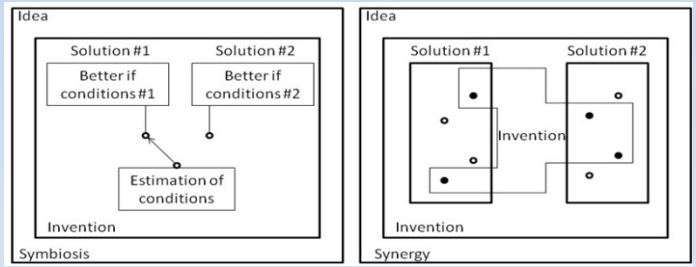
- **Examples:**

  As far as M1/M2, the famous classification of computer systems by Mike Flynn (SISD, SIMD, MISD, MIMD)
  initially included no examples of the MISD type.
  Later on, a DFT machine (generated using the M1 method)
  was categorized as an MISD machine [Milutinovic86A],
  as well as one pipelined machine (generated using the M2 method), namely [Milutinovic87C];
  the DFT served as an inductor, and pipeline as a catalyst.
  Other popular examples are related to various signal processors and process accelerators.

Hybridization:
Symbiosis versus Synergy (H1 vs H2)

- **Definition:**

  Sometimes two classification classes can be combined, in order to obtain a hybrid solution (hybridization).
  Hybrid solutions can be symbiotic (measuring the conditions in the environment
  and switching from one approach to the other, so that each approach is active all the time
  while the conditions are such that it provides better performance compared to the other approach)
  or synergistic (creating a new approach, which, for each particular solution element
  takes the better solution element of two different approaches).



B: Hybridization (Symbiosis versus Synergy) – H1 vs H2

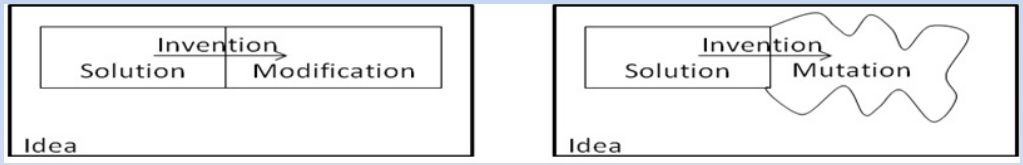
- ***Examples:***

  As far as H1/H2, the essence of [Milutinovic85] is that two algorithms are combined into one
  on the either-one-or-the-other basis (using the H1 method),
  and on a combine-inherent-details basis (using the H2 method) in [Milutinovic87B].
  Other popular examples include hybrid computers or computers that use special purpose accelerators,
  when appropriate data/process patterns are located.

Transdisciplinarization: Modifications versus Mutations (T1 vs T2)

- ***Definition:***

  Often times, good new ideas get generated if algorithms, procedures,
  ways of thinking, re ported from one field to another field,
  along the lines of transdisciplinary research methodologies (transdisciplinarization).



C: Transdisciplinarization (Modification versus Mutation) – T1 vs T2

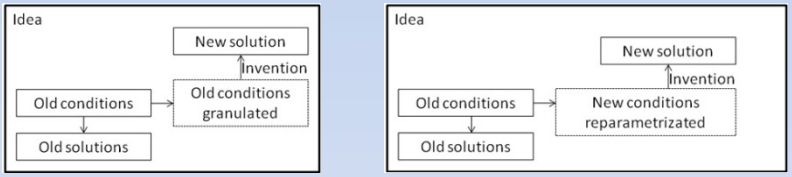
- ***Examples:***

  As far as T1/T2, [Milutinovic86B] ports algorithms from Silicon to GaAs (using the T1 method),
  and introduces appropriate modifications along the process,
  while [Milutinovic87A] creates a proposal for a novel computer architecture (using the T2 method),
  along the analogies with a biological honeycomb.
  Other popular examples include porting of the FFT from seismic signal processing
  to speech signal processing,
  or introduction of mathematical neural networks inspired by biological neural networks.

Remodelling:
Granularization versus Reparametrization (R1 vs R2)

- ***Definition:***

  Sometimes it is simply the best to take a research direction different (even opposite)
  compared to what others take (retrajectorization using remodeling).
  The different (opposite) research direction makes sense either if a more detailed set of parameters is in place
  (granularization, due to model changes because of application changes),
  or because parameters of the environment have changed permanently
  (reparametrization, due to model changes because of technology changes).
  The two alternatives are referred to as granularization and reparametrization.



D: Remodeling (Granularization versus Reparametrization) – R1 vs R2

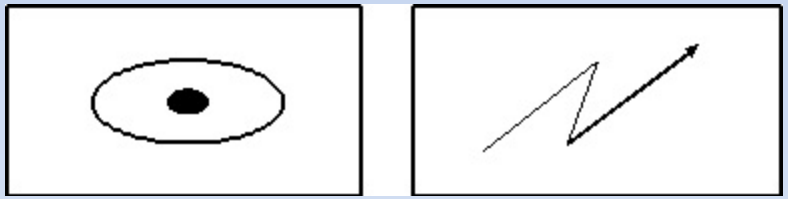
- ***Examples:***

  As far as R1/R2, [Milutinovic88] offers a new algorithm (using the R1 method)
  that makes sense if an environment is represented with a more detailed model,
  while [Milutinovic89] offers a new solution in a changed environment (using the R2 method),
  when a design has to be ported from Silicon to GaAs
  (from the performance/price point of view, the best adder design for Silicon, the carry-lookahead adder,
  is among the worst ones for GaAs, and the opposite: the worst adder for Silicon, the ripple-carry adder,
  is among the best ones for GaAs; this is because GaAs gate delays depend on fan-out
  and ripple-carry adders are characterized with only the minimal fan-out,
  while the fan-out of the carry-lookahead adders depends on the word size, and can grow infinitely).
  Other popular examples are related to concept modeling in AI based on graphical representation with icons
  (in a model with fewer icons, one can make a conclusion which is different, and often times even opposite, compared
  to a conclusion made from a model with a much larger number of icons);
  also, when the environment changes (for example, the ratio of processing speed to communication speed changes),
  a different type of supercomputing network becomes optimal.

Unorthodoxization: ViewFromAbove versus ViewFromInside (U1 vs U2)

- ***Definition:***

  This category encompasses the approaches that are difficult to classify:
  Sometimes one sees something that others did not see for decades or centuries (ViewFromAbove)
  or one gets stroked by an idea of a genius with no ground in existing research (ViewFromInside).

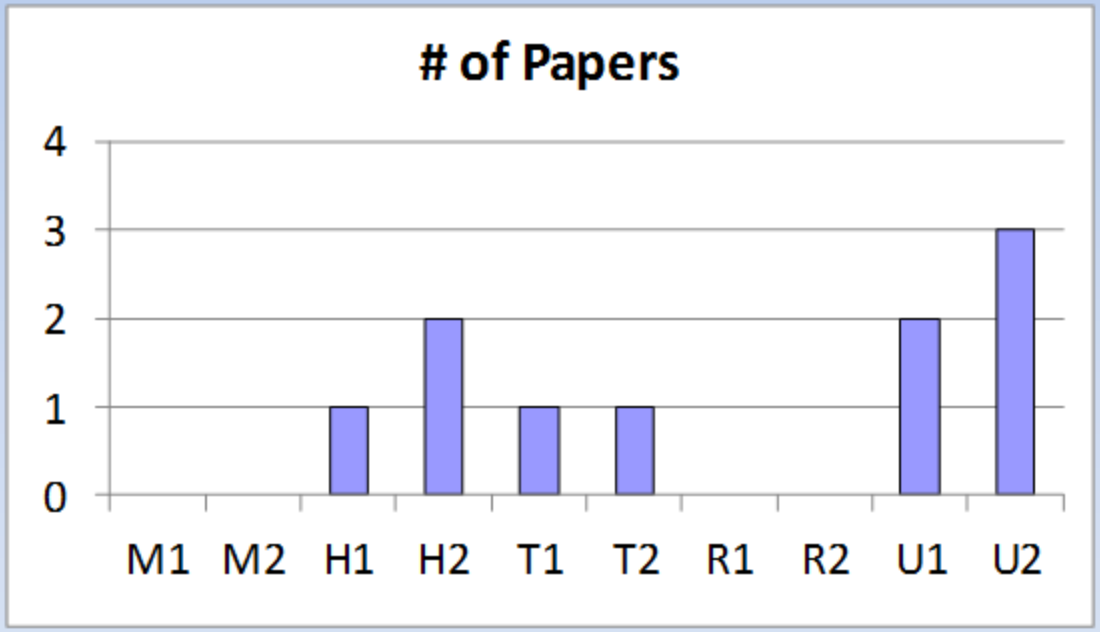


E: Unorthodoxization (ViewFromAbove versus ViewFromInside) – U1 vs U2

- ***Examples:***

  As far as U1/U2, [Milutinovic2000] generated an innovation after trying to make a holistic view (U1),
  and [Milutinovic2001] introduces an idea after an effort is made
  to understand the intrinsic details of the problem (U2).
  Other popular examples include the contributions of Nobel Laureates Martin Perl and Jerome Friedman.

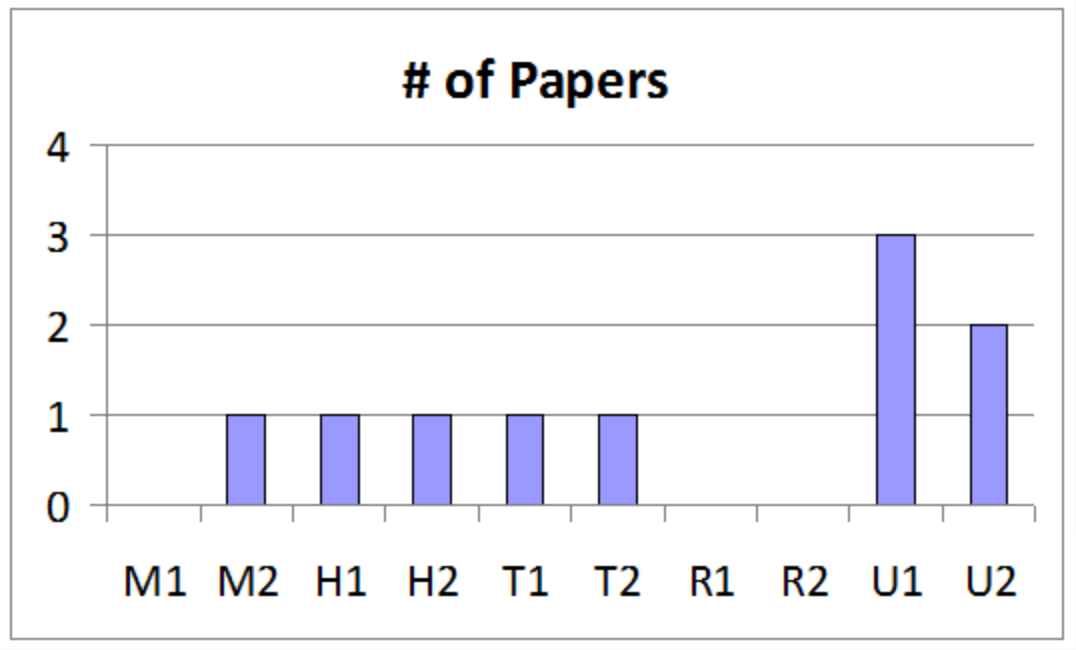
Examples from the Turing Award



Number of Turing Awards based on the given innovation method

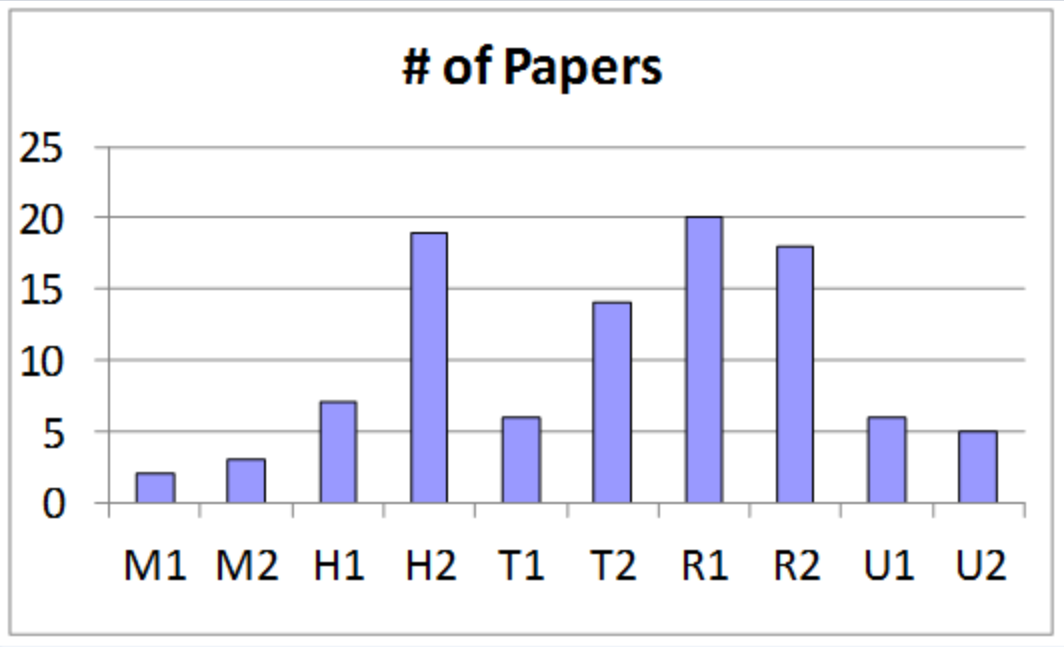
Alan J. Perlis (1966), Maurice V. Wilkes (1967), Richard Hamming (1968), Marvin Minsky (1969), James H. Wilkinson (1970), John McCarthy (1971), Edsger W. Dijkstra (1972), Charles W. Bachman (1973), Donald E. Knuth (1974), John Backus (1977)

Examples from the Nobel Laureate Research



K. Arrow,
L. Cooper,
P. DeGennes,
J. Friedman,
S. Glashow,
H. Kroto,
E. Maskin,
M. Perl,
B. Richardson,
K.Wilson

Examples from the List of Top 500 Computer Scientists of the World



Experiences with PhD Students of the Authors of this Research

Researcher	Research domain	Method
Drazen Draskovic	mutation algorithms for genetic search [Draskovic2012]	H1
Bojan Furlan	opinion mining for social networks [Furlan2011]	H1
Nemanja Kojic	data mining for wireless sensor networks [Kojic2012]	U1
Marko Misic	interconnection networks for multiprocessor systems [Misic2011]	R2
Milos Cvetanovic	system software for wireless sensor networks [Cvetanovic2008]	H1
Zaharije Radivojevic	application software for wireless sensor networks [Radivojevic2008]	H1
Zarko Stanisavljevic	computing infrastructure for distant education [Stanisavljevic2011]	H1
Zivojin Sustran	cache management for multiprocessor systems [Sustran2012]	T2
Djordje Djurdjevic	of computer graphics for mission applications [Djurdjevic2011]	R1
Sasa Stojanovic	hybrid computing for supercomputer architecture [Stojanovic2012]	H1

Classified References Used in the Educational Process

- **M1: Mendeleyevization/Inductor**
  [Milutinovic86a] *Milutinovic, V., Fortes, J., Jamieson, L.,A Multiprocessor Architecture for Real-Time Computation of a Class of DFT Algorithm,IEEE Transactions on Acoustics, Speech, and signal Processing, Aol. ASSP-34, No. 5, October 1986, pp. 1301-1309.(impact factor 1.463/1992).*

- **M2: Mendeleyevization/Catalyst**
  [Milutinovic87c] *Milutinovic, V.,A Simulation Study of the Vertical-Migration Microprocessor Architecture,IEEE Transactions on Software Engineering, Vol. SE-13, No. 12, December 1987, pp. 1265-1277.*

- **H1: Hybridization/Symbiosis**
  [Milutinovic85] *Milutinovic, V.,A Microprocessor-Oriented Algorithm for Adaptive Equalization,IEEE Transactions on Communications, Vol COM-33, No 6, June 1985, pp. 522-526.(impact factor 1.512/2010).*

- **H2: Hybridization/Synergy**
  [Milutinovic87b] *Milutinovic, V., Lopez-Benitez, N., Hwang, K.,A GaAs-Based Microprocessor Architecture for Real-Time Applications,IEEE Transactions on Computer, VolC-36, No 6, June 1987, pp. 714-727.(impact factor 1.822/2010).*

- **T1: Transdisciplinarization/Modification**
  [Milutinovic86b] *Milutinovic, V.,GaAs Microprocessor Technology,IEEE Computer, Vol 19, No. 10, October 1986 (Invited, Guest Editor's Introduction), pp. 10-15.(impact factor 2.205/2010).*

- **T2: Transdisciplinarization/Mutation**
  [Milutinovic87a] *Milutinovic, D., Milutinovic, V., Soucek, B.,The Honeycomb Architecture,IEEE Computer, Vol. 20, No. 4, April 1987 (Open Channel), pp. 81-83.(impact factor 2.205/2010).*

- **R1: Remodeling/Granularization**
  [Milutinovic88] *Milutinovic, V., A Comparison of Suboptimal Detection Algorithms Applied to the Additive Mix of Orthogonal Sinusoidal Signals, IEEE Transactions on Communicatiions, Vol. COM-36, No. 5, May 1988, pp. 538-543.*

- **R2: Remodeling /Reparametrization**
  [Milutinovic89] *Milutinovic, V., Bettinger, M., Helbig, W.,Multiplier/Shifter Design Trade-offs in a 32-bit Microprocessor,IEEE Transactions on Computers, Vol. 38, No. 6, June 1989, pp. 847-880.(impact factor 1.822/2010).*

- **U1: Unorthodoxization/ViewFromAbove [Milutinovic2000]** *Milutinovic, V., Cvetkovic, D., Mirkovic, J., "Genetic Search Based on Multiple Mutation Approaches," IEEE Computer, 2000. (impact factor 1.822/2010).*

- **U2: Unorthodoxization/ViewFromInside [Milutinovic2001]** *Milutinovic, V., Ngom, P., Stojmenovic, I., "STRIP --- A Strip Based Neural Network Growth Algorithm for Learning Multiple-Valued Functions," IEEE Transactions on Computers, 2001. (impact factor 1.822/2010).*

A Short Course for PhD Students in Science and Engineering: ”How to Write Papers for JCR Journals”

(A) survey papers

(B) research papers

Major Contributions of the Two Paper types

Major contributions of the two paper types are as follows:

a) for a survey paper:
   1) A novel classification of existing approaches to the problem, using a well thought set of classification criteria.
   2) Presentation of each approach using the same template and the same type of figures, so an easy comparison is possible.
   3) Some wisdom related to future research trends.

b) for a research paper:
   1) Introduction of a new idea.
   2) Comparison of that idea with the best one from the open literature, using the previously built tools, with appropriate modifications.
   3) In addition to a performance oriented comparison, any research paper also has to include a complexity oriented comparison.

1. Survey Papers

- Selection of the topic for a survey must satisfy the following requirements:

1) The field is newly emerging.
2) Popularity of the field will grow over time.
3) A critical number of papers with new algorithms/approaches does exist (at least twenty to forty).
4) A survey paper does not exist.
5) The PhD student worked before in a related scientific field.
6) The PhD student is enthusiastic about the particular field of his/her tutorial paper.

1. Survey Papers

- With the binary (or n-ary) criteria, one can create either a tree-like classification or a cube-like classification, as indicated in Figures 1 and 2 [Vukasinovic2012].

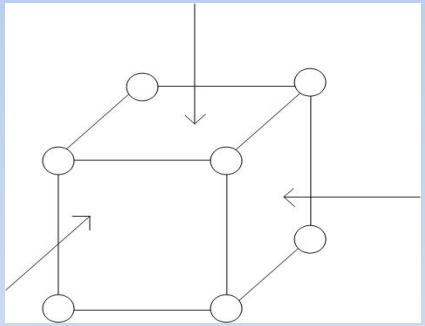


FIGURE 1. A tree-like classification:
Classes are only at the leaves of the tree.

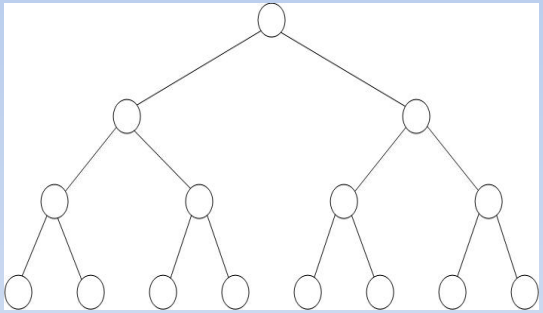


FIGURE 2. A cube-like classification:
Classes can exist also at points inside the cube, as pointed to by the three arrows.

- With a tree-like classification, one can classify only the approaches that entirely belong to a specific class.
  With a cube-like classification, one defines a space in which inner points include, to some extent, characteristics of all existing classes

- What is useful, is to prepare a figure which includes the following:
  1) The classification criteria.
  2) The classification.
  3) The technical mnemonics.
  4) The symbolic mnemonics.
  5) The number of selected examples per class.
  6) The full list of references of selected examples.
  7) The vector of relevant characteristics.

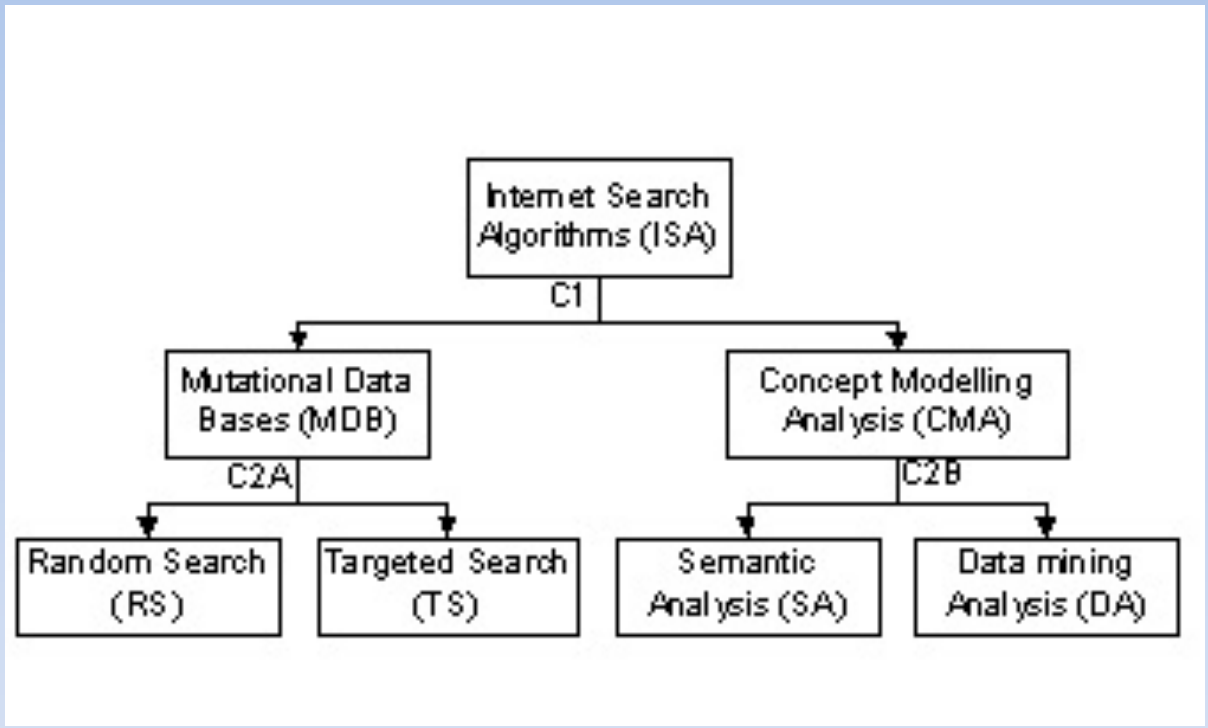


Figure 3. Classification of Internet Search Algorithms

Legend: C1 (criterion #1) = Retrieval-oriented vs Analysis-oriented
C2A (criterion #2, in the MDB path) = Random Search vs Targeted Search
C2B (criterion #2, in the CMA path) = Semantics-oriented vs Datamining-oriented

Technical Names	Random Search (RS or RS/MDB)	Targeted Search (TS or TS/MDB)	Semantic Analysis (SA or SA/CMA)	Data-mining Analysis (DA or DA/CMA)
Symbolic Names	Lion	Jaguar	Tiger	Panthera
Number of Surveyed Contributions	4	4	4	4
References	[Nikolic2011a] Nikolic, B., "Expert Systems," WUS Austria Educational Publishing and University of Belgrade, Classroom Textbook, June 2011.	[Milutinovic2000a] Milutinovic, V., Cvetkovic, D., Mirkovic, J., "Genetic Search Based on Multiple Mutation Approaches," IEEE Computer, November 2000, vol. 33, issue: 9, pp. 118-119.	[Nikolic2011b] Furlan, B., Sivacki, V., Jovanovic, D., Nikolic, B., "Comparable Evaluation of Contemporary Corpus-Based and Knowledge-Bases Semantic Similarity Measures of Short Text," JITA, vol. 1, no. 1, pp. 65-72, ISSN: 2232-962, June 2011.	[Milutinovic2000b] Milutinovic, V., Knezevic, P., Radunovic, B., Casselman, S., Schewel, J.,"Obelix Searches Internet Using Customer Data,"IEEE Computer, July 2000, vol. 33, issue: 7, pp. 104-107.
	[Nick2001] [ ] [ ]	[Simon2009] [Mirkovic1999] [Chen1997]	[Gordon2006] [Leroy2003] [Wang2006]	[Al-Dallal2009] [Hu2007] [Freitas2001]
-ability#1 ... -ability#N				

TABLE A: SUMMARY OF THE APPROACHES THAT LED TO THE CLASSIFICATION PROPOSED IN THIS PAPER

Survey Papers

- When presenting each particular example, one can use the template presented next
  1) Seven Ws about the survey example (Who, What, When, Where, Why, for Whom, hoW).
  2) Essence (it is extremely difficult to give entire essence in only one sentence).
  3) Structure (at this place, one can insert a call to a figure,
     like in Figure 4 from [Draskovic2012]).
  4) Some relevant details.
  5) Example (here one can call a figure that explains an example using a pseudo-code,
     like in Figure 5 [Draskovic2012];
     ideally, the same application case should be used for all surveyed examples).
  6) Pros and cons.
  7) Author's opinion of this example and its potentials.

- For short surveys, each template element is a sentence.
  For long surveys, each template element is a paragraph.
  For books, each template element can be a page, or more.

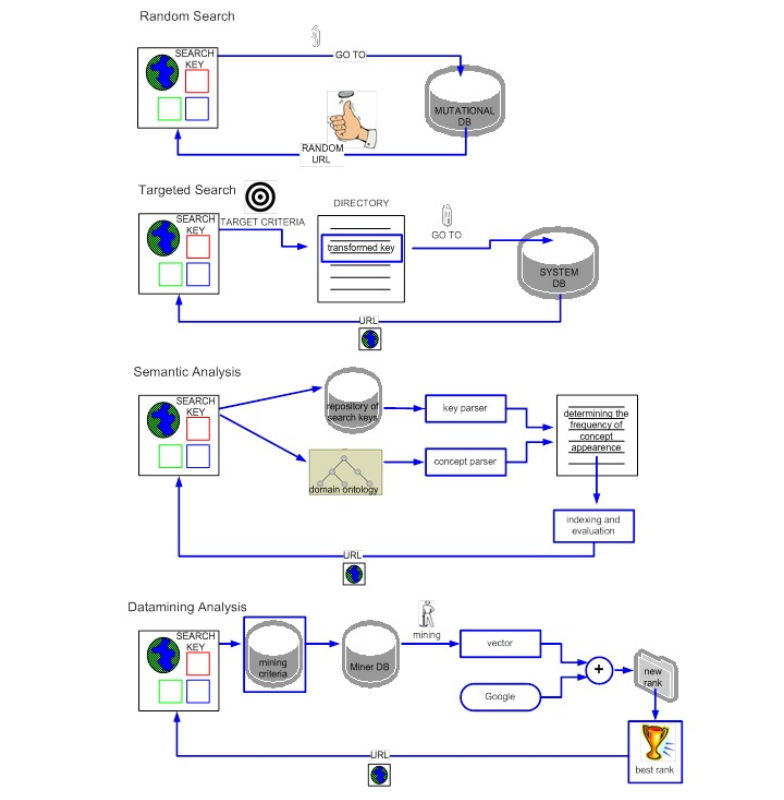


FIGURE 4: Generalized Structure of the four Search Classes
Legend:
DB = Database;
URL = type of URI that is used to describe
the location of a specific document;

PseudoCode
Procedure#X;
Procedure#Y;
Procedure#Z;
...
EndPseudoCode

FIGURE 5: PseudoCode
that demonstrates behavior of an example,
in the case of a specific application;
it is advised that the same application is used
with all examples.

2. Research Papers

- The major purpose of the research paper is to describe an innovation
  and to demonstrate that, under certain conditions, it has a better performance and/or complexity, compared to the best one from the open literature.
  The major steps in the process are:

  1) To create an invention.

  2) To perform a rigorous analysis, to demonstrate that the invented solution is better than the best one from the open literature under a specific set of conditions, and to show what these conditions are and for how much is it better.

  3) To write the paper using a methodologically correct template.

Research Papers

- As far as the presentation of the research results, the students are told that each research paper should contain the following twelve sections:

1) Introduction
2) Problem statement
3) Existing solutions
4) The proposed solution
5) Details
6) Axioms, conditions, and assumptions of the analysis to follow
7) Mathematical analysis
8) Simulation analysis to show performance
9) Implementation analysis to show complexity
10) Conclusion
11) Acknowledgments
12) Annotated references

1. Introduction

- The minimum introductory text should contain the following three paragraphs:

  a) About the general field of this research.
  b) About the specific field of this research.
  c) About the viewpoint of this research, as well as the goals of this research.

2. Problem statement

- The following elements are obligatory:

  a) Problem definition.
  b) Why is the problem important.
  c) Why will the importance of the problem grow over time.

3. Existing solutions

- Existing solutions and their drawbacks,
  looking from the viewpoint defined in the introduction,
  and having in mind the goals defined in the introduction.
  Elements of this section are:

  a) A brief classification of the best solution from the open literature.
  b) Short description of each relevant solution.
  c) A detailed criticism of each presented solution,
     especially in domains in which the proposed solution
     is expected to be better.

4. The proposed solution

- The proposed solution and its essence, and why is it supposed to be better compared to the best solution from the open literature;
  elements of this section are:

  a) Philosophical essence of the proposed solution.
  b) Why the proposed solution is without drawbacks of existing solution(s).
  c) What is the best methodology to prove the superiority of the proposed solution, and under what conditions that holds.

5. Details

- This section should contain details of the best one among the existing approaches and of the proposed solution. The relevant details should be grouped into categories. For example:

  a) Hardware details.
  b) System software details
  c) Application software details.

6. Axioms, conditions, and assumptions of the analysis to follow

a) Axioms refer to axiomatic standpoints.

b) Conditions refer to realistic specifiers of the environment.

c) Assumptions refer to simplifications that make the analysis easier, without jeopardizing on the quality of the final result.

7. Mathematical analysis

a) Axioms, conditions, and assumptions are described mathematically.

b) Closed or open form formulae are derived for the major performance measures.

c) Closed or open form formulae are derived for the major complexity measures.

8. Simulation analysis to show performance

a) Simulator, logical structure and user interface are described.

b) Simulation experiments are described.

c) Simulation results are discussed.

9. Implementation analysis to show complexity

a) Implementation strategy is discussed for the chosen technology.

b) Implementation details and complexity are presented.

c) If a prototype is implemented, show some characteristic measurement.
If a prototype is not implemented, give some implementation guidelines.

10. Conclusion

a) Summary of what was done
   and to what extent are the initial goals achieved.

b) To whom is that of benefit.

c) Newly open problem for further research.

11. Acknowledgments

a) To all those who patiently listened to your ideas
and especially to those who volunteered to share with you
some of their own ideas for further benefit of your research.
Also, it is obligatory to cite the relevant work
of all those who volunteered the improvement ideas.

b) To all those who helped provide the infrastructure for your research.
If this is related to one or more research project, list them.

c) To all those who suffered by taking everyday life responsibilities from you,
so you could dedicate more of your time to research.

12. Annotated references

- The references are more useful if listed in groups.
  Each topic requires different grouping.
  The grouping that seems most appropriate for this paper includes:

  a) References related to methodology.

  b) References related to examples.

  c) References related to success of past students.

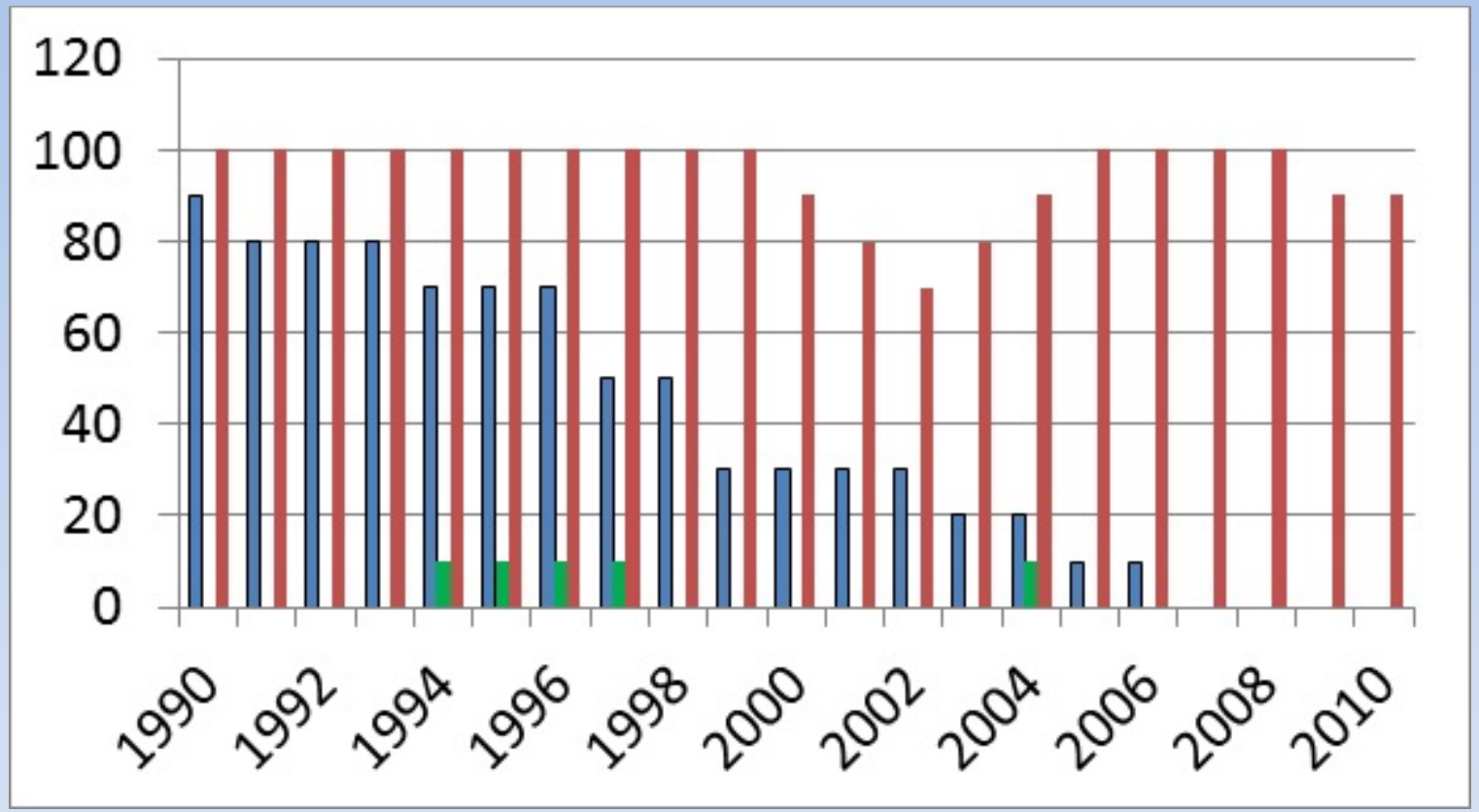


Figure 6: Success rates of paper submissions: (A) Normalized percentage of class with permission to submit.
(B) Normalized paper acceptance rate of those permitted to submit.
(C) Among the student with accepted papers, how many enrolled a PhD program later.
Explanation: Curve A is almost monotonically decreasing due to the fact that the class was not formal in the beginning and only the best students opted to take informal lectures. After the class became formal, the student body became larger, and consequently, not all of them were extraordinary. Curve B had ups and downs, with peeks separated about twelve years apart, which could be a consequence of the motivation ups and downs of the teacher. Curve C is at hundred percent, except in years characterized with an industry boom.

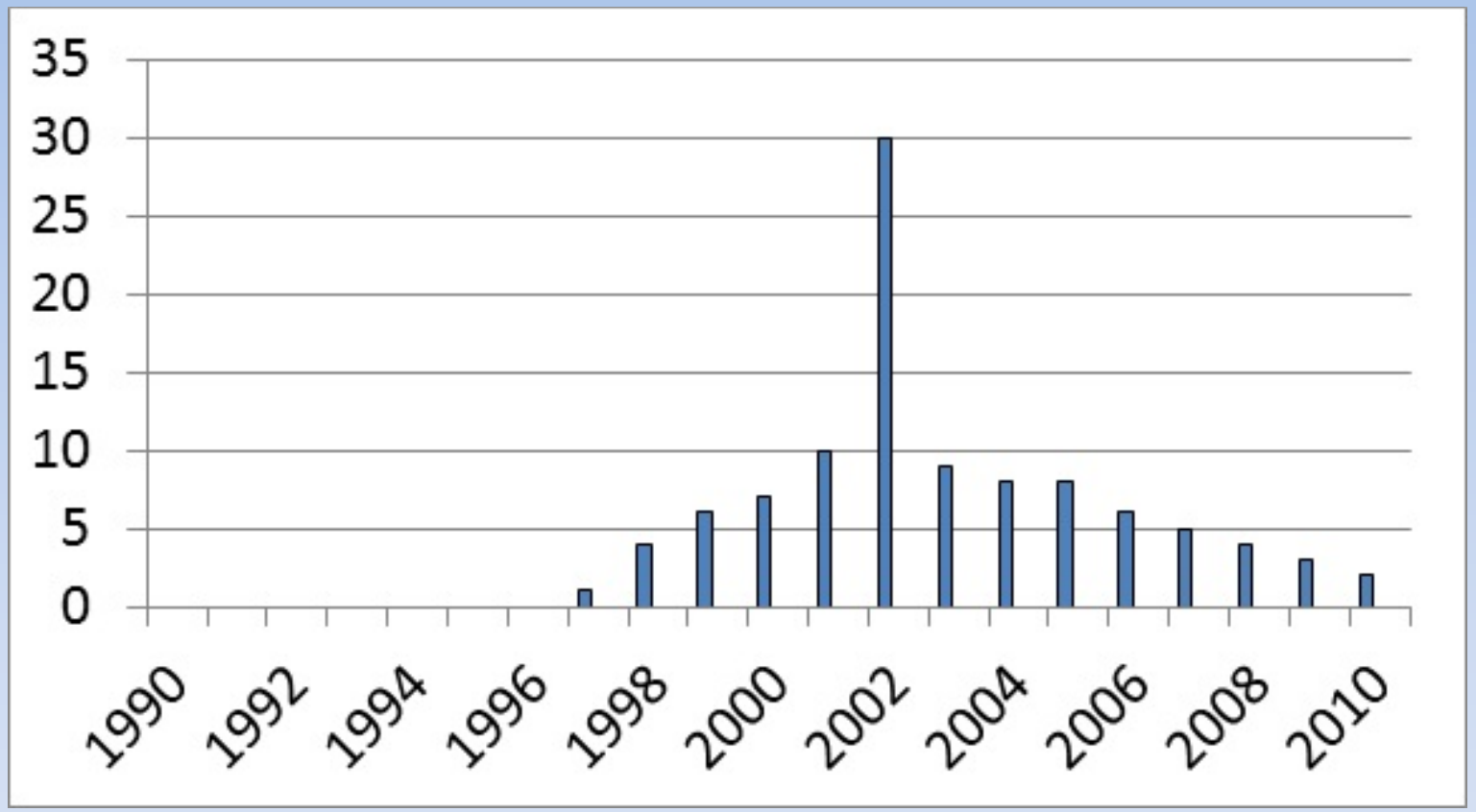


Figure 7: Citation analysis for the ten most referenced papers.
Explanation: It is clear that in some years the production was of a higher quality; that seems to be coinciding with years in which the world's top industry was giving donations to the department.
The Y axis refers to the total number of citations for the top 10 papers

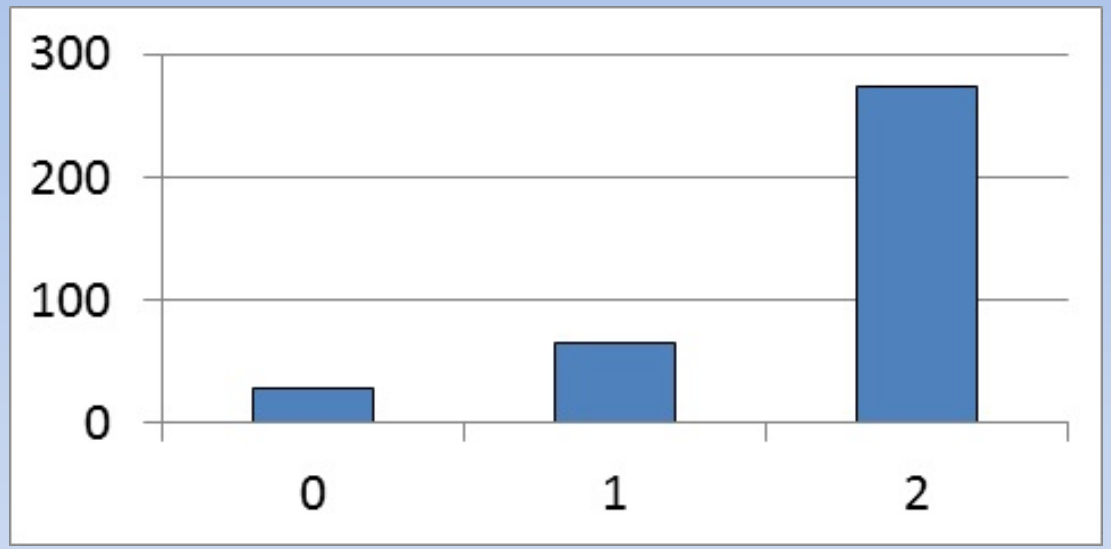


Figure 8: Impact of the existence of another survey paper.
Explanation: This figure gives a result which was absolutely unexpected.
The expectation was that existence of a survey would decrease citations of our survey, but it happened absolutely the opposite. This means that the quality is more important than the pre-existence of another survey paper on the same subject.
The paper with 2 preceded survey papers was the paper by Protic at al [Protic 1996].
The paper with one preceded survey was the paper by Tomasevic at al [Tomasevic1993].
The paper with no preceded survey was the paper by Jovanovic at al [Jovanovic1999].

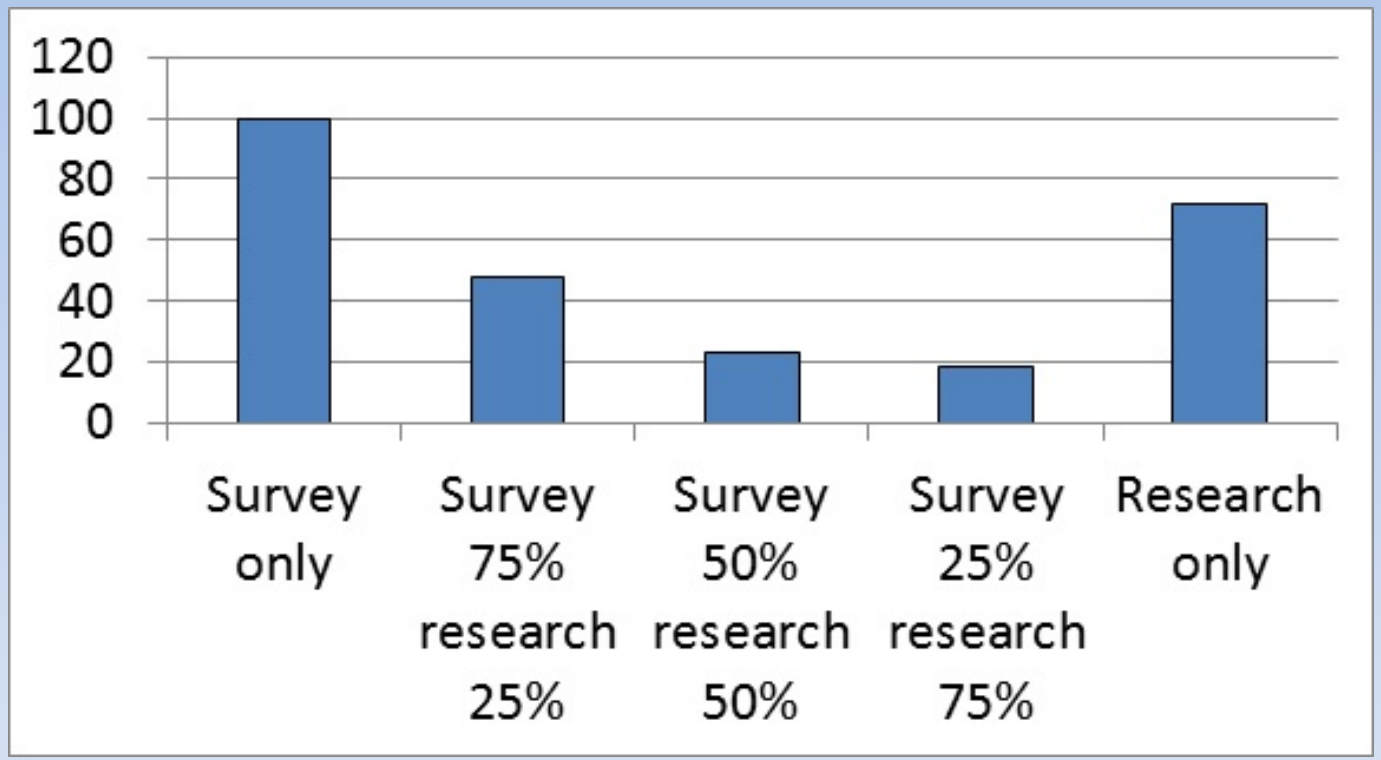


Figure 9: Survey papers versus research papers, what generates more citations?
Explanation: Surveys generate more,
unless an extraordinary research paper is generated in a popular field.

Appendix

One possible research plan based on the methodology presented in this paper is elaborated here,
week by week
(between two logical weeks,
one can have one or more physical weeks)

SURVEY PAPER

1. Read about the general subject, to worm up.

2. Collect 20 to 40 papers,
   on various approaches from the open literature.

3. For each example (covered by one or more papers),
   write the main 7 sentences, as explained in this paper.
   Explain why the chosen template enables easy comparison,
   and therefore represents a contribution to science.

4. Decide about classification criteria generate a classification,
   sort the found examples by classes,
   and form Figure #1, as explained in this paper.
   Explain why the proposed classification represents a contribution to science.

5. For each example, generate two figures
   (for example, one block scheme of the structure,
   and one pseudo code presentation of the algorithm).
   Choose the presentation form which indicates
   the essence of the class that the example belongs to.

6. If the generated classification
   includes a class without examples
   (which is highly desirable,
   since that points to possible new research avenues),
   define the research strategy of interest for those
   who decide to take that avenue.
   Form a section with appropriate discussions.

7. Define the research strategy for those
   who decide to analyze the hybrid approaches
   (those consisting of elements of two different classes).
   Hybrid approaches can be either a symbiosis
   (the two solutions used interchangeably, as the conditions dictate),
   or a synergy (the two solutions combined into one).
   Discuss possible new solutions or both types (symbiosis and synergy).
   Discuss other possible avenues leading to new inventions
   (transdisciplinarization and retrajectorization).

8. Add the preamble and the postambule,
   and create the list of annotated references.
   Form the final text of the paper.
   Generate a pearl of wisdom that sheds light
   on the essence of the paper,
   and increases the probability that the paper be referenced a lot.

9. Ask peers to review your paper,
   while you look for a suitable journal to publish it.

10. Submit the paper to a journal.

RESEARCH PAPER

1. For the best subset of ideas from the position paper,
   make appropriate simulator changes,
   and run the newly generated original solutions,
   comparatively with the best solution from the open literature.
   Create the tables and figures with results.

2. Write the paper.

3. Bounce the paper off the peers, and submit it to a journal.

thanks!

Q&A

vm@etf.rs