UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA
CADERNOS DE MATEMÁTICA E ESTATÍSTICA
SÉRIE A: TRABALHO DE PESQUISA

FOUNDATIONS AND APPLICATIONS OF QUALITATIVE REASONING AND MODEL-BASED DIAGNOSIS

PETER STRUSS
WALDIR LEITE ROQUE

SÉRIE A, Nº 35
PORTO ALEGRE, OUTUBRO DE 1993

Foundations and Applications of Qualitative Reasoning and Model-Based Diagnosis*

Peter Struss
Computer Science Department
Technical University of Munich
Orleansstr. 34
D-81677 Munich
Germany
struss@informatik.tu-muenchen.de

Waldir L. Roque
Instituto de Matemática
Universidade Federal do Rio Grande do Sul
Av. Bento Gonçalves, 9500
91501-970 Porto Alegre, RS
Brazil
roque@mat.ufrgs.br

Abstract

In this paper we summarize the foundation for a model-based approach to diagnosis of technical systems and discuss the importance of qualitative reasoning within this framework. Based on this, we propose to apply these diagnostic techniques to model building. Thus, theories about physical systems are checked for consistency and against empirical observations. This leads to a tool that supports the construction and maintenance of model libraries as well as the process of theory formation in research.

1 The Consistency-based Approach to Diagnosis

Knowledge-based diagnosis systems of the *first generation* are crucially based on establishing more or less direct links between symptoms that can be observed and faults (or diseases) that have been known to cause the symptoms (with a certain probability) (Fig. 1). Obviously, this approach heavily depends on the completeness of knowledge about all three elements: the symptoms, the faults and the associations between them. The resulting restriction of the system to what has been encountered and widely experienced before, is prohibitive for most industrial applications; handling newly designed systems and new kinds of failures are a must.

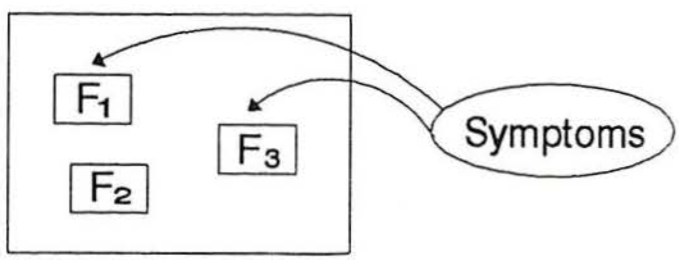

Figure 1: First generation: Linking symptoms to faults.

*Presented at the *Workshop on Information Technology: Cooperative Research with Industrial Partners between Germany and Brazil*, Berlin, July 1993.

An even more significant impediment to industrial applications of this technology lies in the fact that each diagnostic system is dedicated to a particular type of device and has to be developed individually, even though the engineering knowledge required may be essentially the same for a much broader class of devices. Intolerably high costs in development and maintenance of such systems are a consequence. Altogether, these aspects demand for a different approach to diagnosis.

Diagnostic systems of the *second generation* are based on representing and exploiting primarily a different kind of knowledge: principled knowledge about the elements of the system to be diagnosed and about the way they are designed to interact in order to achieve the intended function of the entire system. This is why such knowledge-based systems are called *model-based*. The model of a particular system is constructed from two different sources: models of the constituent parts of the system taken from a *domain-specific library* and the *structure description* which captures the specific knowledge about the respective artifact.

The constituents are normally thought of as active physical components, but they may also be abstract entities, such as processes and characteristic parameters, or passive instruments like sensors. If new kinds of components are introduced, their models have to be added to the library, and after their inclusion in the structure, the changed systems can be diagnosed. In both cases, the diagnostic engine remains unchanged. Obviously, such a system is able to exhibit the flexibility and adaptability required in many diagnostic tasks in technical domains.

As explained above, the system model comprises knowledge about the structure of the device and the behavior of its constituents ("components"). This model is exploited by the so-called *consistency-based* approach as follows: for the situation(s) to be analyzed, the normal, or intended, behavior is predicted from the model. Inconsistencies of the set of predictions and real observations are determined (so-called *discrepancies*), and also their potential origins are identified, which represent diagnostic candidates (Fig. 2). If an inconsistency arises from a prediction based on, say, components C_1 and C_2, then there must be a fault in at least one of C_1 or C_2, whereas there is no reason for suspecting C_3, if it is not involved in the respective prediction.

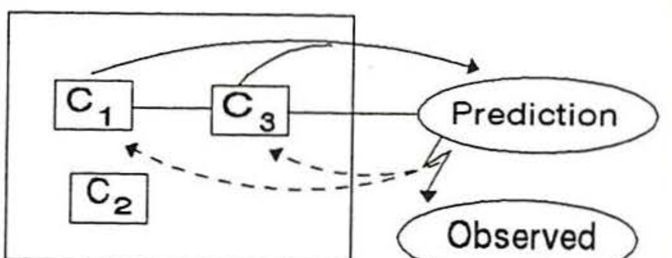

Figure 2: Second generation: Model-based prediction of observable behavior.

Consistency-based diagnosis has a formal description in terms of logic (see e.g. [de Kleer et al. 90]). Given the *MODEL* of the device and the set *OBS* of observations about its actual behavior, the diagnostic task is to find the set Δ of broken components. It is solved by determining (complete) sets of *assumptions* about the correctness of the system's components that are *consistent* with the model and the observations, i.e.

$$MODEL \cup OBS \cup \{FAULTY(C) \mid C \in \Delta\} \cup \{CORRECT(C) \mid C \in COMPS \setminus \Delta\}$$

must be consistent (*COMPS* is the set of all components). This formalization serves as a basis for rigorous analyzing properties of different algorithms and implemented systems, such as the General Diagnostic Engine (GDE) ([de Kleer-Williams 87]) or GDE+ ([Struss-Dressler 89]).

This model-based approach to diagnosis of technical systems promises significant progress in this domain. Its relevance for applications to industrial automation has several aspects: First, it will extend the scope of potential diagnostic applications, because it does not require previous experience with the device to be diagnosed. Second, it is a qualitative advance in competence of knowledge-based diagnosis, since it enables the diagnosis of new and unanticipated kinds of faults. Third, it contributes to the reuse of software (in terms of the component models and the diagnostic algorithms), thus enabling the fast and inexpensive generation of specialized diagnosis systems for new devices. Finally, and more strategically, it opens the perspective of a coherent representation of technological and scientific knowledge that can be exploited for different task-oriented knowledge-based systems, supporting, for instance, design, simulation, monitoring, maintenance etc.

Today, the technology is *mature enough to be applicable* to diagnosing selected domains (see e.g. [Guida-Stefanini 92], [Beschta et al. 93]) and, hence, attracts industrial interest.

Nevertheless, still *much more research* in computer science, mathematics and engineering needs to be done in order to enhance the systems and broaden the application domains.

2 Qualitative Modeling and Diagnosis

The diagnostic approach outlined above does not imply any particular form and content of the model, except that this model and its use *preserves the structure* of the real device.

There are several reasons for incorporating *qualitative modeling* techniques for a model-based system:

- Providing a *structural description* of the device is fundamental for a model-based system. This has been a major research topic of qualitative reasoning.
- There are domains in which *precise models and values* are *not available* or situations in which they are not applicable. Recall the limited precision of measurements in many real cases.
- Much of the naturalness of qualitative arguments is due to the way humans try to *handle complexity* of systems: concentrating on the essential distinctions (e.g. considering only the directions of influences and changes), performing abstractions, working with approximate models and making simplifying assumptions requires qualitative methods. We believe this aspect is of the most important but least developed features of model-based diagnosis.

Qualitative reasoning, or more precisely, qualitative physics ([Weld-de Kleer 89], [Faltings-Struss 90]) has emerged as a special branch of Artificial Intelligence research from two roots: modeling common sense knowledge about the physical world and modeling engineering problem solving skills.

It is obvious that reasoning about systems without precisely knowing their laws and parameters is an essential feature of human interaction with every-day life environment. As a matter of fact, also scientific and engineering work involves analysis on a more abstract level than the solution of differential equations and computation of numerical values. Quite often, these structural and qualitative levels are crucial for solving a problem (and for being able to perform the appropriate calculations).

Sometimes, qualitative reasoning is considered from the narrow perspective of being "calculation with qualitative values". In opposition to this point of view, we want to emphasize that qualitative reasoning, in pursuing the goal to identify and describe the essential features and mechanisms underlying a (physical) system, comprises both:

- Modeling *structure*, i.e., representation of the constituents of a system and of their paths of potential interactions, and
- Modeling *behavior*, i.e., characterizing the system's state and its development over time qualitatively.

To achieve this, qualitative reasoning operates on behavior descriptions that reflect only the *essential* distinctions in the behavior. It derives statements about whole classes of physical systems that share basic behavioral properties. Thus, it avoids having to sample large parameter spaces and sets of initial conditions which is what numerical simulation would require in order to achieve the same results. For this purpose, formal methods and systems have been developed that perform, for instance, inferences about orders of magnitude ([Raiman 86]) and inequality reasoning ([Forbus 84]).

Another technique that derives qualitative statements about interdependencies among variables is regime analysis ([Roque 91, Roque 93]). Qualitative reasoning with *regime analysis* has its foundations in the theory of dimensional analysis, which is quite well-knwon in physics but rather new to the qualitative physics community. The main idea behind regime analysis is to find out dimensionless quantities that describe the core of the physical system and perform a qualitative analysis about the composite behavior of the system forming the so-called regimes.

Regime analysis seems applicable to a number of situations where

- the process is defined by a large number of variables and parameters, and the effect of a change in one of these variables on other variables has to be analyzed,
- the actual physical law that underlies the process or device is only incompletely known,
- the computational costs of a full numerical (quantitative) analysis is high,
- the actual qualitative analysis of the physical process or device satisfies the requirements.

Examples can be found in monitoring of engineering plants, in diagnosing faulty devices, robotics, etc. Actually, this approach can be extended to other non-physical areas where the notion of dimensional representation can be established. Potential candidates are ecology and economy.

Clearly, in many situations, qualitative reasoning is not able to provide an unambiguous and sufficiently detailed behavior of a process, and so, the association with other approaches becomes necessary. Often, different methods of analysis have to be combined.

Quite fundamental too, are techniques that use distinctive "landmark values" and intervals between them as qualitative values ([Kuipers 86], [Struss 88]), with the sign algebra ([de Kleer-Brown 84]) given in Tables 1 and 2 as a special case.

$\oplus$	$-$	0	$+$	?
$-$	$-$	$-$	?	?
0	$-$	0	$+$	?
$+$	?	$+$	$+$	?
?	?	?	?	?

Table 1: Addition of signs. ? denotes unrestricted values.

$\otimes$	$-$	0	$+$	?
$-$	$+$	0	$-$	?
0	0	0	0	?
$+$	$-$	0	$+$	?
?	?	?	?	?

Table 2: Multiplication of signs. ? denotes unrestricted values.

Although this may appear to be rather weak, it often suffices to generate useful information about a system's behavior and diagnostic hypotheses. Qualitative models become even a necessity, if numerical information is not available. In order to illustrate this, consider the simple circuit consisting of a voltage source, three resistors and a light bulb shown in Fig. 3. Assume that the only observation is that the light bulb is dimmed compared to its normal mode, thus indicating a reduced current through R_2.

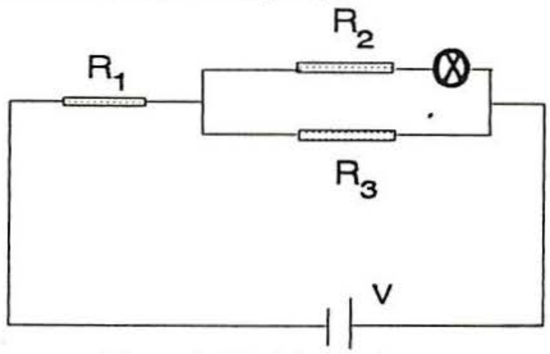

Figure 3: Electric circuit

Let us denote the deviation of the actual value of a variable, x, from its nominal one by $\Delta x = x_{nom} - x_{act}$. From the equation

$$i_2 = \frac{v}{(R_1 + \frac{R_1 * R_2}{R_3} + R_2)}$$

we obtain the "qualitative equation"

$$\Delta i_2 = \Delta v \oplus (\Delta R_3 \ominus \Delta R_1 \ominus \Delta R_2) \otimes v$$

for the deviations from the nominal values. If $v = +$, the dim light bulb indicates that $\Delta i_2 = -$, and the equation implies

$$\Delta i_2 = - \land v = + \Rightarrow \Delta v = - \lor \Delta R_3 = - \lor \Delta R_1 = + \lor \Delta R_2 = +,$$

i.e. a set of deviations of parameter values as possible explanations for the observation.

This, admittedly simple, example illustrates that the use of qualitative models can reflect the nature of the available information about a system and is able to derive useful results even if observations are imprecise and incomplete. As the diagnosis obtained will often be ambiguous, further refinement may be required based on additional observations and/or the use of more detailed models. Hence, one requirement that is frequentely raised by application-oriented work is the necessity of *combinations* of modeling techniques instead of a single one ("Multiple Modeling", see e.g. [Struss 92]).

Under this aspect, it should be noted that the model we introduced in the example is not only qualitative but also simplified, because it neglects, for instance, the resistance of the connecting wires. Such modeling assumptions, although justified in general, may turn out to be inadequate in certain situations, in particular when a fault is present. This is why multiple modeling has to include techniques for handling simplified and approximate models and for reasoning about modeling assumptions, as shown in ([Struss 92]).

3 Diagnosis as Support for Model Formation

Modeling, i.e. the formation of a theory about a physical system in order to perform model-based reasoning for a particular task, turns out to be difficult in practice. In fact, it is, in general, more difficult than collecting a sample of experiential rules because of the much stronger requirements on the generality and robustness of the knowledge captured. Having built a library comprising context-independent behavior models of a set of constituent elements, the system developer has to check for a number of properties. Are these models consistent with each other? Are they together adequate for behavior prediction in the intended contexts? Do they suffice to explain the known phenomena of composite systems? Are they compliant with the "accepted wisdom"? Currently, there is almost no computer support for this task.

At this point, the consistency-based approach to diagnosis explained in section 1 provides the foundation for some help. Remember that this approach can be described as checking a composite model for consistency with a set of given data and determining the possible causes of contradictions to these data in terms of constituent models to be revised. This abstract characterization matches very well with the task of model validation discussed above. Hence, model-based diagnosis techniques can indeed be applied to create a tool that supports the model formation in the following way. The developer of a model library has to

- formulate the "accepted wisdom", i.e. the theory he or she considers undoubted,
- add model fragments that require validation, i.e. an extension to the existing theory, and, using these two elements of the model library,
- construct a composite model of a known system scenario and enter empirical data about its (normal) behavior.

Consistency-based diagnosis is then applied to this model using the empirical data as observations and will either confirm consistency or return a list of model fragments as possible origins of contradictions which are candidates for a revision of the theory. Two issues are worth noticing.

Such a system cannot only point to the model fragments that may need inspection. It can often *suggest certain specific modifications* to the model that would remove the contradiction. This happens because the system proposes suspected model fragments because they contradict predictions of other models (or the given data). Hence, the latter predictions indicate what the suspect should predict in order to be appropriate.

On the other hand, this approach does not suffice to discover all deficiencies. If a model is simply too weak to make particular predictions about the expected behavior, it would be consistent with the given data and, hence, pass consistency-based diagnosis. So, what we want to make sure is that the known behavior is actually *entailed* by the composite model (rather than being merely consistent). The appropriate type of diagnosis for this task is *abductive* diagnosis. ([Dressler-Struss 92]) describes an integration of abductive diagnosis into the consistency-based approach which is needed in the proposed model-building tool.

Such a tool would be a significant support for the development of model-based systems and help to reduce development time and efforts spent on maintenance of model libraries. While this indicates the importance for the industrial production and application of knowledge-based systems, it should be noted that this kind of tool may as well be used in research activities in areas where the formation or extension of a theory can be supported. Here, the models diagnosed would not necessarily be component models, but fragments or hypotheses in a theory which are checked against other confirmed theories, empirical data, or expected predictions.

For instance, this might support the formation of models of biological and ecological systems, in particular in conjunction with qualitative modeling techniques, since often data in these domains are inherently qualitative. Furthermore, as the consistency-based diagnostic techniques provide facilities for proposing useful measurements ([de Kleer-Williams 87]) and for generating tests ([Struss 93]), they can also be used for the design of experiments to check and promote the evolution of the theory.

4 Summary and Outlook

We have shown that model-based diagnosis in conjunction with reasoning with qualitative and multiple models make an important contribution to the application of knowledge-based systems to industrial automation. Furthermore, the techniques can be applied to the process of building such systems, in supporting the development of model libraries or, more generally, of theories of physical systems.

The authors intend to continue collaborating on model-based diagnosis, qualitative modeling and model formation. For the validation of the research results, discussions are in progress to jointly develop

- with Riosoft company an application to diagnosis of the pulp production process and
- with ELETROSUL/UFSC and SIEMENS Munich an application in the domain of power transportation networks.

These proposals fit in the context of the Brazilian project on "Integrated Automation Technology". Furthermore, we intend to evaluate the facilities for qualitative modeling and scientific theory formation in other areas, such as ecology.

References

Beschta et al. 93 Beschta, A., Dressler, O., Freitag, H. and Struss, P., *A Model-based Approach to Fault Localization in Power Transmission Networks.* Intelligent Systems Engineering, **3**, 1993.

de Kleer-Brown 84 de Kleer, J. and Brown, J.S., *A Qualitative Physics Based on Confluences.* Artificial Intelligence **24**(1-3), 1984.

de Kleer et al. 90 de Kleer, J., Mackworth, A. and Reiter, R., *Characterizing Diagnosis.* In Proceedings AAAI-90.

de Kleer-Williams 87 de Kleer, J. and Williams, B. C., *Diagnosing Multiple Faults.* Artificial Intelligence, **32**, 97-130, 1987.

Dressler-Struss 92 Dressler, O. and Struss, P., *Back to Defaults: Characterizing and Computing Diagnoses as Coherent Assumption Sets.* In Proceedings of ECAI-92; corrected version in Working Papers of the Third International Workshop on Principles of Diagnosis DX-92.

Faltings-Struss 92 Faltings, B. and Struss, P. (eds.), *Recent Advances in Qualitative Physics*, MIT Press, 1992.

Forbus 84 Forbus, K., *Qualitative Process Theory.* Artificial Intelligence **24**(1-3), 1984.

Guida-Stefanini 92 Guida, G. and Stefanini, A. (eds.), *Industrial Applications of Knowledge-Based Diagnosis.* Elsevier, Amsterdam 1992.

Kuipers 86 Kuipers, B., *Qualitative Simulation*, Artificial Intelligence **29**(3), 1986.

Raiman 86 Raiman, O., *Order of Magnitude Reasoning.* In Proceedings of AAAI-86.

Roque 91 Roque, W. L., *Qualitative Physics Reasoning about Systems and Devices.* In Brazilian Symposium on Artificial Intelligence, ed. A. Arcela, pp. 137-141, (Editora da UnB, 1991).

Roque 93 Roque, W. L., *Learning Qualitative Physics Reasoning with Regime Analysis.* In Proceedings of the Artificial Intelligence and Symbolic Mathematical Computation, Springer-Verlag, 1993 (to appear in LNCS).

Struss 88 Struss, P., *Mathematical Aspects of Qualitative Reasoning.* International Journal for Artificial Intelligence in Engineering, **3**(3), 1988.

Struss 92 Struss, P., *What's in SD? Towards a Theory of Modeling for Diagnosis.* In Readings in Model-based Diagnosis, Hamscher, W., Console, L. and de Kleer, J. (eds.), San Mateo, 1992.

Struss 93 Struss, P., *Multiple Models of Physical Systems - Modeling Intermittent Faults, Inaccuracy, and Tests in Diagnosis*, to appear in: Annals of Mathematics and Artificial Intelligence, 1993.

Struss-Dressler 89 Struss, P. and Dressler, O., *Physical Negation - Integrating Fault Models into the General Diagnostic Engine.* In Proceedings of IJCAI-89.

Weld-de Kleer 90 Weld, D. and de Kleer, J. (eds), *Qualitative Reasoning about Physical Systems*, Morgan Kaufmann, San Mateo, 1990.

Publicações do Instituto de Matemática da UFRGS

Cadernos de Matemática e Estatística

Serie A: Trabalho de Pesquisa

.. Marcos Sebastiani Artecona - Transformation des Singularités - MAR/89

2. Jaime Bruck Ripoll - On a Theorem of R. Langevin about Curvature and Complex Singularities - MAR/89

3. Eduardo Cisneros, Miguel Ferrero e Maria Inés Gonzales - Prime Ideals of Skew Polynomial Rings and Skew Laurent Polynomial Rings - ABR/89

4. Oclide José Dotto - ε-Dilations - JUN/89

5. Jaime Bruck Ripoll - A Characterization of Helicoids - JUN/89

6. Mark Thompson e V. B. Moscatelli - Asymptotic Distribution of Liusternik-Schnirelman Eigenvalues for Elliptic Nonlinear Operators - JUL/89

7. Mark Thompson - The Formula of Weyl for Regions with a Self-Similar Fractal Boundary - JUL/89

8. Jaime Bruck Ripoll - A Note on Compact Surfaces with Non Zero Constant Mean Curvature - OUT/89

9. Jaime Bruck Ripoll - Compact ε-Convex Hypersurfaces - NOV/89

0. Jandyra Maria G. Fachel - Coeficientes de Correlação Tipo-Contigência - JAN/90

1. Jandyra Maria G. Fachel - The Probality of Ocurrence of Heywood Cases - JAN/90

12. Jandyra Maria G. Fachel - Heywood Cases in Unrestricted Factor Analysis - JAN/90

13. Julio Cesar R. Claeyssen e Tereza Tsukazan de Ruiz - Dynamical Solutions of Linear Matrix Diferential Equations - JAN/90

14. Maria T. Albanese - Behaviour of de Likelihood in Latent Analysis of Binary Data - ABR/91

15. Maria T. Albanese - Measuremente of the Latent Trait Analysis of Binary Data - ABR/91

16. Maria Teresinha Albanese - Adequacy of the Asymptotic Variance-Covariance Matrix Using Bootstrap Jackknife Techniques in Latent Trait Analysis of Binary Data - ABR/91

17. Maria Teresinha Albanese - Latent Variable Models for Binary Response - ABR/91

18. Mark Thompson - Kinematic Dynamo in Random Flows - DEZ/90

19. Jaime Bruck Ripoll e Marcos Sebastiani Artecona- The Generalized Map and Applications - AGO/91

20. Jaime Bruck Ripoll, Suzana Fornari e Katia Frensel - Hypersurfaces with Constant Mean Curvature in the Complex Hyperbolic Space - AGO/91

21. Suzana Fornari e Jaime Bruck Ripoll - Stability of Compact Hypersurfaces with Constant Mean Curvature - JAN/92

22. Marcos Sebastiani Artecona - Une Généralisation de L'Inveriant de Malgrange - FEV/92

23. Cornelis Kraaikamp e Artur Lopes - The Theta Group and the Continued Fraction with Even Partial Quotients - MAR/92

4. Silvia Lopes - Amplitude Estimation in Multiple Frequency Spectrum - MAR/92

5. Silvia Lopes e Benjamin Kedem - Sinusoidal Frequency Modulated Spectrum Analysis - MAR/92

6. Silvia Lopes e Benjamin Kedem - Iteration of Mappings and Fixed Spectrum Analysis - MAR/92

7. Miguel Ferrero, Eduardo Cisneros e Maria Ines Gonzales - Ore Extensions and Jacobson Rings - MAI/92

8. Sara C. Carmona - An Asymptotic Problem for a Reaction-Diffusion Component - JUL/92

9. Luiz Fernando Carvalho da Rocha - Unique Ergodicity of Interval Exchange Maps - JUL/92

). Sara C. Carmona - Wave Front Propagation for a Cauchy Problem With a Fast Component - OUT/92

. Marcos Sebastiani Artecona e Iván Pan Pérez - Intersections Transverses dans l'Espace Projectif - OUT/92

. Miguel Ferrero - Closed Bimodules over Prime Rings: Closed Submodules and Applications to Rings Extensions - DEZ/92

. Dinara W. X. Fernandez - Método da Maxima Verossimilhança Restrita para Estimação de Componentes de Variância - SET/93

. Martin Knott e M. Teresa Albanese - Polymiss: A Computer Program for Fitting a One- or Two-Factor Logit-Probit Latent Variable Model to Polynomous Data when Observations may be Missing - OUT/93

35. Peter Struss e Waldir L. Roque - Foundations and Applications of Qualitative Reasoning and Model-Based Diagnosis - OUT/93

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA
NÚCLEO DE ATIVIDADES EXTRA CURRICULARES

Os Cadernos de Matemática e Estatística publicam as eguintes séries:

érie A: Trabalho de Pesquisa

érie B: Trabalho de Apoio Didático

érie C: Colóquio de Matemática SBM/UFRGS

érie D: Trabalho de Graduação

érie F: Trabalho de Divulgação

érie G: Textos para Discussão

Toda correspondência com solicitação de números ublicados e demais informações deverá ser enviada para:

NAEC - NÚCLEO DE ATIVIDADES EXTRA CURRICULARES
INSTITUTO DE MATEMATICA - UFRGS
AV. BENTO GONÇALVES, 9500 - PRÉDIO 43111
CEP 91509 - 900 AGRONOMIA - POA/RS
FONE: 336 92 22 OU 339 13 55 OU 228 16 33
RAMAL 6197
FAX: 336 15 12