Schreier G, Hayn D, Ammenwerth E (Hrsg.). Tagungsband der eHealth2010: Health Informatics meets eHealth.
6.-7. Mai 2010, Wien. OCG Books Nr. 264. Österreichische Computer Gesellschaft. 2010

# CHAOS OR CONTENT: TOWARDS GENERIC STRUCTURING OF MEDICAL DOCUMENTS BY UTILISATION OF ARCHETYPE-CLASSES

Borovicka M[1], Laschitz L[1], Reiter D[2], Gächter E[3], Specht G[1]

***Abstract***

*Interoperability is the key concept for approaches on building future-proof EHR systems. Two-level modelling and archetypes methodologies bring this goal one step closer. The challenge lies in fully structuring medical documents as to not leave valuable information hidden when searching. A study for laboratory processes from the IHE was taken and refined along the nature of archetypes and how they can best be applied to medical concepts used by medics during the whole EHR data cycle. We defined classes of archetypes and developed a model for bridging the gap between document creation and examination ordering, and tested it with a proof-of concept implementation usable in the Austrian ELGA HL7 CDA scenario.*

***Keywords – Archetypes, CDA, Two-level dual modelling approach, Interoperability***

## 1. Introduction

Health information systems are facing a vast amount of society and associated paradigm changes. Emerging Trends in health care show that the domain is getting more and more multidisciplinary [10], collaborative, therefore spanning geographic regions. Several E-Health-Initiatives like Infoway, Canada, NHS CRS, United Kingdom, or AORTA from The Netherlands are each implementing an infrastructure for sharing patient information on a national level. In Austria, the ELGA E-Health-Initiative has taken the lead to provide this national infrastructure.

From the point of modelling and representing intended meaning [12] the health domain as such is very complex [23], which has deterred many interoperability approaches. To help aid and manage this complexity came the introduction of two-level modelling [13] for EHRs and archetype methodologies. Although by definition they can be used for any domain, they where specifically invented for the medical domain with the intention of been shared by health institutions on a broader, ideally global scale [10], [12].

For ELGA, the HL7v3 Clinical Document Architecture (CDAr2) [11] was chosen to serve as the common document exchange format. Its advantage is that all medical documents can be shared and

[1] Leopold Franzens University of Innsbruck, Institute for Computer Science. Databases and Information Systems
[2] TILAK University Hospital Innsbruck, Department for Information Technology
[3] Zentralinstitut für Bluttransfusion und Immunologische Abteilung (Blutbank) TILAK,Universitätskliniken

read among different systems. But as the content within does not have to offer a defined structure, no further support is possible. Information may still be hidden within the vast amount of documents belonging to one patient. Therefore, electronical exchange of documents can only be the first step. It becomes important to also structure them in a benefiting way, leveraging their full potential [1]. Archetypes provide a container for storing this structuring knowledge and therewith offer a concept to build future-proof EHR systems.

Introducing archetypes requires significant additional effort. This effort can best be justified when they allow for more than directions on how to structure a document. When considering a view of the whole process, as seen in *Figure 1*, from examination ordering to visualisation of results, archetypes must therefore cope with the data transformation problem [20] and also proof beneficial for data visualisation.

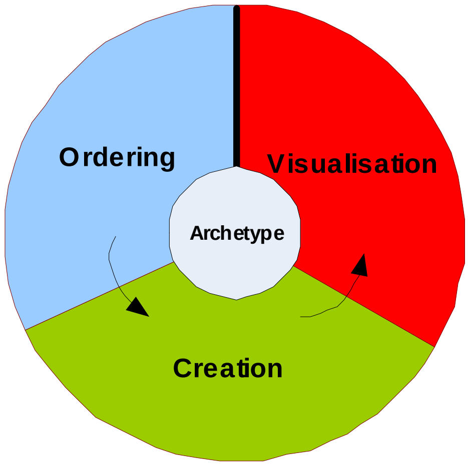


**Figure 1: EHR data cycle**

This paper examines how archetypes can best support a generic process to create documents using the CDA under the scope of laboratory examination results. It tries to identify the capabilities of archetypes as well as their boundaries, and looks at how the challenge to make deployed legacy systems compatible to an EHR infrastructure using archetypes is met best.

## 2. Related Work

National implementation guidelines [16], [7] form the basis for valid, well structured, and sound documents. This general structure is further refined by the implementation guides for the document kind, e.g., laboratory [22] report or diagnostic imaging [25] report. As the implementation guides are provided in human readable form, no automated procedure can be invoked to instantiate/create a document. Integrating the knowledge contained in these guidelines, systems based on the dual model or two-level modelling approach [14] support the separation between information and knowledge.

In general, many approaches exist to structure a document by means of a computer interpretable form, mostly of proprietary format. Standardized candidates suitable for constructing future-proof systems are the HL7 Template [5] model and the openEHR [14] (and closely related EN/ISO 13606) archetype model. As Heard [15] points out, the terms they introduce are used among themselves in different meaning. In [5], Bointner and Duftschmid offer a comparison for the different approaches. OpenEHR also provides a knowledge description language called Archetype Definition Language [3] (ADL) to create a formal definition for an archetype. Thereby, archetypes can be seen as formal structured models covering one specific medical concept [13]. The ADL also supports the binding with medical coding systems (SNOMED CT, LOINC, ICD-10) to gain semantic grounding [6]. Such bindings support the possibility to structure e.g., “level 3” CDA documents, which makes them computer-interpretable. Furthermore, they allow for more flexibility and evolvement, building highly adaptive systems to account for the inevitable changes taking place in clinical requirements or medical knowledge [8], [12].

Both Kohl et al. [18] and Chen et al. [8] argue that there should exist an instance which validates available archetypes. OpenEHR proposed one method to reach this common sense by establishing an open repository called Clinical Knowledge Manager (CKM). Others try to grasp it by defining archetypes as maximum-archetypes [4] or maximum dataset [9] for the clinical concept being modelled, evidence based or based on best practise [10]. All agree that an archetype for a medical concept must be comprised of the union of items of all possible instances of that concept [4], [9], [10], [8], [18]. Still, according to Chen [9], no existing EHR product is using or is built on the archetype methodology beyond the prototype stage.

Together with the concept of openEHR templates as means to construct compositions of archetypes, using the CDA R-MIM the two-level modelling approach from openEHR [13] can be applied to the handling of CDA documents [23]. Maldonaldo et al. [20] developed an ADL-Editor called LinkEHR, supporting the creation of what they call integration archetypes for openEHR, ISO 13606, and CDA, which also allow mappings to data sources.

## 3. Analysis and Methodology

A recently published study [1] conducted by the IHE Laboratory Committee analysed laboratory processes and defined several daily routine high-level use-cases (UC1-UC3). According to Beale [3], archetypes offer a promising, future-proof concept for managing complexity in the health care domain. It is now examined if they may be deployed in a benefitting way along the whole data cycle, as was seen in *Figure 1*. Adopting a design science approach [17], we combined field reports

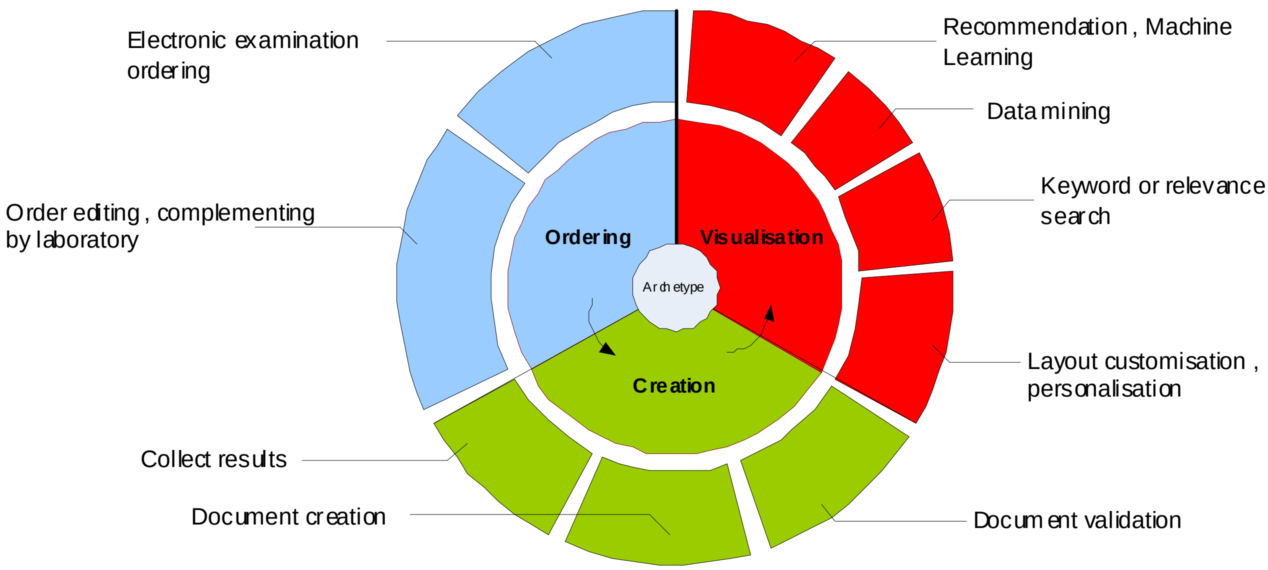


**Figure 2: Extended EHR data cycle**

with expert opinions and refined them through interviews conducted with the biggest regional hospital in Tyrol (IT Department, Central Laboratory Department, Blood Transfusion and Immunologic Department, Department of Internal Medicine), as well as with general practitioners, and an external laboratory to formulate a model and implement a proof-of concept prototype.

From the study [1] use-case UC1 (external order management) is chosen as it best fits our general scenario of independent participants. UC1 describes the process from examination ordering to the actual document instantiation (examination report). The term "external" is defined as the setting where there exists an external origin (e.g., general practitioner, clinician at hospital) which creates the order with respect to the laboratory where the order is being processed, implying that they both share no common information system. *Figure 2* shows the refined data cycle according to UC1. The first step from the use-case UC1 is examination ordering. Depending on the grade of IT support, the order is accomplished either electronically or paper-based. In the above observed medical departments both variants are currently being used. In both cases, a unique number (the IHE "order placer number" [19]) is generated in the laboratory information system (LIS) to represent the order.

When looking at the daily routine of general practitioners/clinicians, depending on their actually scope they either work with single patient parameters, e.g., haemoglobin or medical concepts, e.g., full blood count[1], which represent groups of parameters. Although A4 paper sheets only offer limited space, both single parameters and groups are depicted, among a clinician may choose from. This suggests, that doctors are also thinking in both terms when executing the examination order.
The next step in UC1 is order processing, which is done at the laboratory the order is sent to. An analysis of the observed laboratory departments revealed that they use different terms ("Analyt", "Meßwert", "Verfahren"), and ("Gruppe", "Serie") respectively, but still they all imply the same meaning. These terms are either built into or induced by their legacy software systems. What they all have in common is the usage of single parameters and the ability for grouping/combining them, similar to the LOINC Order Panels (Batteries) [21]. From a technical perspective, these findings are in correspondence with the concept of archetypes. This suggests that these can be utilised to construct archetype instances for examination ordering, which are build around such a group of parameters, confining a medical concept, useable in the examination ordering process.

The final step in UC1 is order completion, which is comprised of collecting results and document creating/validating, resulting in an electronic document exchange. The archetypes can thereby be directly passed on to document creation, as they are identical to the ones needed for an approach to structure medical documents in a uniform way representing a generic solution. To complete the cycle in *Figure 2*, also the visualisation part will benefit when it has information about the structure of the document, as to enhance overview and interpretation [24], [10], [6]. This way, archetypes become the basis instrument for a future-proof, "level3" semantic interoperable electronic health infrastructure, allowing for automated evaluation, arrangement, and recommendation of information.

## 4. Model and Framework

As examined in the previous chapter, utilisation of archetypes can be beneficial for every part of the whole EHR data cycle. An important prerequisite for sharing on a national level is that medi-

[1] In German: Blutbild

cine and medical concepts are equal among the participants so that there can be no misinterpretation, i.e. a doctor practising conventional medicine would have no use for archetypes from traditional Chinese medicine.

As Kohl [18] put it, building quality archetypes which are accepted among domain experts requires appropriate governed development processes [18]. To circumvent this problem, it is argued here that the granularity [2] is of utmost importance when trying to formulate a generic model for this use-case. The right granularity can be compared to a common denominator between medics at the clinic, general practitioners, and laboratory doctors and workers, possessing the right level of shareable concepts.

This approach views archetypes as kind of a building bricks covering a clinical concept, which aligns with the perception of openEHR [14]. We go along Eccher [12] when we say that the focus of archetypes lies in covering one method of a medical examination concept of a specific part of the human body (e.g., thyroid etc) or for certain kind of decease (Rubella[1] etc) which result in actual measurable parameters. It must be noted that in this matter of sense an x-ray examination result is also seen as a parameter (as the result is a picture and has to be interpreted in the same way like a parameter-value). *Table 1a* holds these definitions.

**Table 1: Definitions (a) and examples for clinical concepts (b) with corresponding parameters**

a)

| Term | Definition |
|---|---|
| *Parameter* | *Single measured value* |
| *Method* | *Group of parameters, which from a medical view (medical concept) makes sense* |

b)

| Method | Parameter |
|---|---|
| *Rubella* | *IgM-ELISA, IgG-ELISA* |
| *Thyroid* | *TSH, fT3, fT4* |
| *Blood count* | *Leukocytes, erythrocytes, …* |
| *X-Ray* | *Thorax x-ray picture* |

Literature and discussions with medical personnel confirm that those methods are generally accepted, understood and already shared in a way that they serve as this common denominator. The more elementary or atomic [14] a concept is, the higher its shareability, thereby shifting the problem of acceptance [18] to the next higher level in the hierarchy, i.e. to the compositional (openEHR template) level. As the archetypes now entail a proper meaning, they are reusable and can be shared nationally. An example list of sharable concepts is given in *Table 1b*.

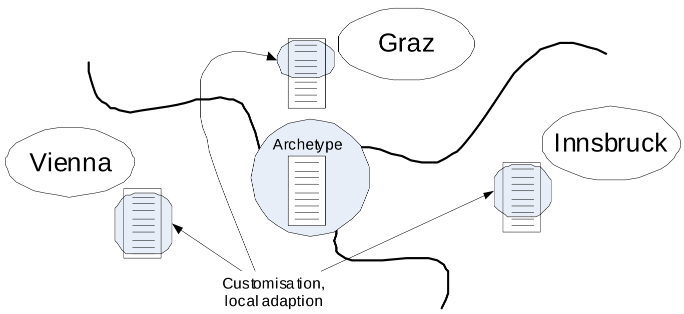


**Figure 3: Maximised minimum-archetype**

[1] In German: Röteln

Schreier G, Hayn D, Ammenwerth E (Hrsg.). Tagungsband der eHealth2010: Health Informatics meets eHealth.
6.-7. Mai 2010, Wien. OCG Books Nr. 264. Österreichische Computer Gesellschaft. 2010

Granularity is seen as the key concept. Still, sharing a generic, common archetype requires adopting it locally, as seen in *Figure 3*, to specific situations and restrictions. This is necessary, as not every laboratory has the same possibilities, i.e. equipment, to measure all parameters. Archetypes must therefore be minimised, to serve as common denominator, and maximised to cover all aspects in this minimal world view. According to [9], usage can be guaranteed through the compositional pattern they allow. As explained in Chapter 2, documents are built [7] according to national implementation guidelines. These for one part provide means to structure the national defined format and are comprised of the jurisdictional requirements necessary for a medical document (meta-data, e.g., author, custodian etc). In the Austrian scenario it defines the header and body part of the CDA, but without structuring the actual medical content in detail (measured or examined parts of the patient).

Similar to [12], archetypes within this model are classified according to their kind of function. A guideline-archetype is based on a guideline defining the structure of the examination document, while a method-archetype represents a medical concept and thereby measured data from a patient. Templates in this model represent the link between archetypes and their local hierarchal combination/visualisation to medics to select/choose from during the examination ordering procedure. Depending on the nature of the examination order, a valid CDA document requires one guideline-archetype as well as a combination of method-archetypes. As seen in *Figure 4*, archetypes are the key element. They act as the missing link between examination ordering, processing and creating the medical document. The laboratory receives its order in the form of archetypes, together with the information which parameters where selected.

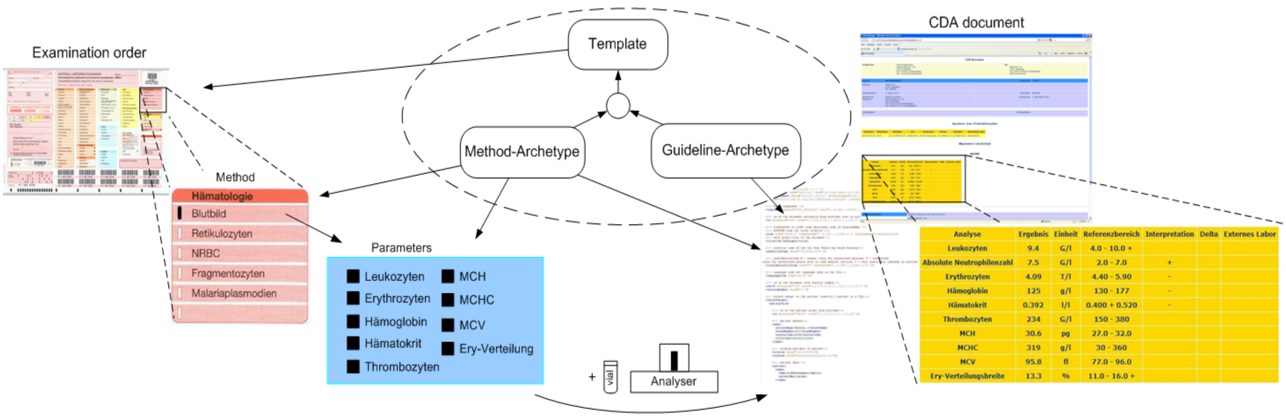


**Figure 4: Document order and creation model based on archetype classes**

## 5. Implementation and Discussion

This paper is part of the early stages of the HERMED project[1] undertaken at the Leopold-Franzens-University of Innsbruck, where a semantic interoperable EHR system using a problem-oriented medical record (POMR) is currently being established. As proof of concept, according to the model, based on a TILAK examination order sheet, one guideline-archetype (ELGA Laboratory) and one method-archetype (full blood count, containing nine single parameters) were created using ADL and the LinkEHR tool [23]. A Java prototype was developed which is able to handle/process a generic examination order request with archetypes, and using them to generate a corresponding ELGA CDA document structured up to level 3. As the HL7 templates lack an obligate definition

[1] http://dbis-informatik.uibk.ac.at/147-1-HERMED-Health-Information-Systems.html

for the computer interpretable part [5], CDA R-MIM based archetypes where chosen as structuring method.

An intended outcome of the implementation was to improve and further refine the model. ADL can be used to restrict and validate data [3], but it does not contain a technique how to access the data it needs to fill the document, which is a necessary requirement for an automated creation process. The archetype per se cannot contain this information, as it is shared nationally, and every laboratory host database has its own but different data management. For this purpose, additional archetype adjustments (so called "AAA"-XML files) are introduced as local extensions.

## 6. Conclusion and Outlook

We successfully demonstrated how archetypes can be integrated into the EHR data cycle, starting from laboratory examination ordering for a general practitioner/clinician. The implemented prototype offers the possibility to use the proposed model of fine-grained archetypes in a composition to generate well-structured and semantic interoperable CDA documents usable for the ELGA infrastructure. The main contribution of this paper is the analysis of laboratory ordering processes and the approach of linking these with archetype concepts. In order to do this, some assumptions where made. Archetype nodes are seen as entities, and are not allowed to change their meaning, even when a new version of the archetype is constructed. Second, the medical concepts used for the archetypes are derived only for the Tyrolean area. A more exhaustive study has to be conducted to reach more certainty. Also, meta-data which are an integral part of the clinical content are not yet considered in our model. These in combination with the possibility of local adaption (AAA) pose an important topic for further considerations.

## 7. References

[1] ANDRIES, H. (2009) External Laboratory Ordering Whitepaper.
URL:http://www.ihe.net/Technical_Framework/upload/Whitepaper-External-Lab-Orders_v1-0_PC_20090422.pdf

[2] BEALE, T. and S. HEARD (2005) Archetypes Definitions and Principles.
URL:http://www.openehr.org/releases/1.0.2/architecture/am/archetype_principles.pdf

[3] BEALE, T. and S. HEARD, The openEHR Archetype Model Archetype Definition Language.
URL:http://www.openehr.org/releases/1.0.1/architecture/am/adl.pdf, 2007.

[4] BERNSTEIN, K., et al., Can openEHR Archetypes Be Used in a National Context? The Danish Archetype Proof-of-Concept Project. Medical Informatics, 2009.

[5] BOINTNER, K. and G. DUFTSCHMID, HL7 Template Modell und EN/ISO 13606 Archetyp Objekt Modell - Eine Gegenüberstellung. Tagungsband der eHealth2009 und eHealth Benchmarking 2009, 2009.

[6] BOROVICKA, M., DMIS: design and prototype of a future clinical eHealth system, in Proc. of the 6th Intl. Conference on Advances in Mobile Computing and Multimedia. 2008, ACM: Linz, Austria. p. 485-489.

[7] BRANDSTÄTTER, J. (2009) ELGA: CDA Dokumente für das österreichische Gesundheitswesen.
URL:http://www.arge-elga.at/index.php?id=32

[8] CHEN, R., et al., An archetype-based testing framework. EHealth Beyond the Horizon: Get It There, 2008: p. 401-406.

[9] CHEN, R., et al., Archetype-based conversion of EHR content models: pilot experience with a regional EHR system. BMC Medical Informatics and Decision Making, 2009. 9(1).

[10] DIPAK, K., et al. (2008) Management and maintenance policies for EHR interoperability resources.

[11] DOLIN, R.H., et al. (2006) HL7 Clinical Document Architecture, Release 2. URL:http://jamia.bmj.com/content/13/1/30.full.pdf

[12] ECCHER, C., et al., Ontologies supporting continuity of care: The case of heart failure. Computers in biology and medicine, 2006. 36(7): p. 789-801.

[13] GARDE, S., et al., Expressing clinical data sets with openEHR archetypes: A solid basis for ubiquitous computing. International Journal of Medical Informatics, 2007. 76: p. 334-341.

[14] GARDE, S., et al., Towards Semantic Interoperability for Electronic Health Records Domain Knowledge Governance for openEHR Archetypes. Methods Inf Med, 2007.

[15] HEARD, S., et al., Templates and Archetypes: how do we know what we are talking about? URL:http://www.openehr.org/publications/archetypes/templates_and_archetypes_heard_et_al.pdf, 2003.

[16] HEITMANN, K.U., et al. (2006) VHitG: Arztbrief auf Basis der HL7 Clinical Document Architecture Release 2 für das deutsche Gesundheitswesen. URL:http://download.vhitg.de/Leitfaden-VHitG-Arztbrief-v150.pdf

[17] HEVNER, A.R., et al., Design Science in Information Systems Research. MIS Quarterly, 2004. 28(1): p. 75-105.

[18] KOHL, C.D., S. GARDE, and P. KNAUP, Facilitating the openEHR approach organizational structures for defining high quality archetypes. EHealth Beyond the Horizon: Get It There, 2008: p. 437-442.

[19] MACARY, F. and N. CHIBA, IHE: Laboratory Technical Framework. URL:http://www.ihe.net/, 2008.

[20] MALDONADO, J.A., et al., LinkEHR-Ed: A multi-reference model archetype editor based on formal semantics. International Journal of Medical Informatics, 2009. 78(8): p. 559-570.

[21] MCDONALD, C., et al., Logical Observation Identifiers Names and Codes: Benutzerhandbuch URL:http://www.dimdi.de/dynamic/en/ehealth/loinc/downloads/loinc/german_loinc_user_guide.pdf, 2009.

[22] MENSE, A., S.W. SABUTSCH, and S. SAUERMANN (2009) ELGA: CDA Laborbefund für das österreichische Gesundheitswesen. URL:http://www.arge-elga.at/index.php?id=29

[23] MONER, D., et al., Archetype-based semantic integration and standardization of clinical data. Proceedings of the 28th IEEE EMBS Annual International conference, 2006: p. 5141-5144.

[24 VAN DER LINDEN, H., T. AUSTIN, and J. TALMON, Generic screen representations for future-proof systems, is it possible?: There is more to a GUI than meets the eye. Computer methods and programs in biomedicine, 2009. 95(3): p. 213-226.

[25] WEIGL, M. and A.O. LINDNER (2009) ELGA: CDA Befund Radiologie für das österreichische Gesundheitswesen. URL:http://www.arge-elga.at/index.php?id=30

**Corresponding Author**
Michael Borovicka
Leopold Franzens University of Innsbruck, Databases and Information Systems
Technikerstraße 21a, A-6020 Innsbruck
Email: michael.borovicka@uibk.ac.at