Linear Coding of Non-linear Hierarchies: Revitalization of an Ancient Classification Method

Wiebke Petersen

Abstract The article treats the problem of forcing entities into a linear order which could be more naturally organized in a non-linear hierarchy (e.g., books in a library, products in a warehouse or store, ...). The key idea is to apply a technique for the linear coding of non-linear hierarchies which has been developed by the ancient grammarian Pāṇini for the concise representation of sound classes. The article introduces briefly Pāṇini's technique and discusses a general theorem stating under which condition his technique can be applied.

Keywords Classification · Hierarchy · Indian grammar theory · Pāṇini.

1 Introduction

1.1 Why Are Linear Codings Desirable?

There are several situations in daily life where one is confronted with the problem of being forced to order things linearly although they could be organized in a non-linear hierarchy more naturally. For example, due to the one-dimensional nature of book shelves, books in a library or a bookstore have to be placed next to each other in a linear order. One of the simplest solutions to this problem would be to order the books strictly with respect to their authors' names and to ignore their thematic relationships. Such an arrangement forces a user of a library, who is usually interested in literature on a special subject, to cover long distances while collecting the required books. Therefore, librarians normally choose a mixed strategy: books are classified according to their thematic subject and within each class they are ordered alphabetically with respect to their authors' names. International standards for subject classification usually claim that the classification has a tree structures, i.e., each thematic field is partitioned into disjoint subfields which are further subdivided into

Wiebke Petersen
Institute of Language and Information, University of Düsseldorf, Düsseldorf, Germany,
e-mail: petersew@uni-duesseldorf.de

A. Fink et al., (eds.), *Advances in Data Analysis, Data Handling and Business Intelligence*, Studies in Classification, Data Analysis, and Knowledge Organization,
DOI 10.1007/978-3-642-01044-6_28,

disjoint sub-subfields and so forth (e.g., UDC: "Universal Decimal Classification"[1]). However, there will always be books which do not fit neatly into one single most specific class.

Similar problems occur in many types of stores, too: food or toys are, like books, commonly presented on shelves, and clothes are hanging next to each other on racks. The question of whether honey should be placed in the breakfast or in the baking department, or whether clothes should be arranged by their color, by the season in which they are typically worn or their type (trousers, jackets, ...) is comparable to the book-placing problem. In warehouses it is beneficial to place the products such that those which are often ordered together are within striking distance in order to minimize the length of the course which has to be covered while carrying out an order. In this setting, the thematic classification becomes secondary since the "classes" are dynamically defined by the consumers' orders.

1.2 Pāṇini's Śivasūtra-Technique

The problem of linearly ordering entities which bear a complex non-linear relationship to each other is old and predates common product organizing problems. In spoken language everything has to be expressed linearly since language is linear by nature. In ancient India, people were very aware of this problem since their culture based on an oral tradition where script was mainly reserved for profane tasks like trading or administration. Since any text which was considered worth to be preserved was taught via endless recitations, keeping texts as concise as possible was desirable. The aspiration after conciseness is especially noticeable in grammar, for which many techniques to improve the compactness of the grammatical descriptions were invented. Grammar was regarded as the *śāstrānāṃ śāstram* "science of sciences" since it aimed at the preservation of the Vedas, the holy scriptures, of which the oldest parts date around 1200 BC (cf. Staal, 1982).

The culmination point of ancient Indian grammar was Pāṇini's Sanskrit grammar (Böhtlingk, 1887) which dates circa 350 BC. Its main part consists of about 4,000 rules, many of them phonological rules which describe the complex system of Sandhi. Sandhi processes are regular phonological processes which are triggered by the junction of two words or morphemes;[2] they are very common in Sanskrit. Phonological rules are typically of the form "sounds of class A are replaced by sounds of class B if they are preceded by sounds of class C and followed by sounds of class D".[3] Since it is not economical to enumerate for each single rule all sounds which

[1] http://www.udcc.org/, http://www.udc-online.com/.

[2] E.g., the regular alternation between *a* and *an* of the indefinite article in English is a Sandhi phenomenon.

[3] In modern phonology such a rule is denoted as

$$A \rightarrow B/_{C\_\_D} \; . \qquad (1)$$

BookID 187687_ChapID 28_Proof# 1 - 25/07/09

अइउण् ऋऌक् एओङ् ऐऔच् हयवरट् लण् ञमङणनम् झभञ् (I)
घढधष् जबगडदश् खफछठथचटतव् कपय् शषसर् हल्

a·i·uṇ ṛ·ḷk e·oṅ ai·auc hayavaraṭ laṇ ñamaṅaṇanam jhabhañ (II)
ghaḍhadhaṣ jabagaḍadaś khaphachaṭhathacaṭatav kapay śaṣasar hal

a i u M_1 ṛ ḷ M_2 e o M_3 ai au M_4 h y v r M_5 l M_6 ñ m ṅ ṇ n M_7 jh bh M_8 (III)
gh ḍh dh M_9 j b g ḍ d M_{10} kh ph ch ṭh th c ṭ t M_{11} k p M_{12} ś ṣ s M_{13} h M_{14}

Fig. 1 Pāṇini's Śivasūtras (I: Devanāgarī script; II: Latin transcription; III: Analysis – the syllable-building vowels are left out and the meta-linguistical consonants marking the end of a sūtra are replaced by neutral markers M_i)

are involved in it, an appropriate phonological description must include a method to denote sound classes. The method should be such that addressing a natural phonological class becomes easier than addressing an arbitrary set of sounds. In modern phonology one often chooses a set of binary phonetic features like [±consonantal] or [±voiced] in order to define the relevant sound classes. This approach necessarily involves the problem of choosing and naming features and the danger of defining ad-hoc features. However, Pāṇini's method of addressing the relevant sound classes evades this problem.

The first 14 sūtras of Pāṇini's Sanskrit grammar are called Śivasūtras and quoted in Fig. 1. Each sūtra consists of a sequence of sounds ending in a consonant. This last consonant of each sūtra is used meta-linguistically as a marker to indicate the end of a sūtra. As the system behind the naming of the markers is unknown (cf. Misra, 1966), we have replaced them in Fig. 1 (III) by neutral markers M_1 up to M_{14}. Together the Śivasūtras define a linear order on the sounds of Sanskrit. The order is such that each class of sounds on which a phonological rule of Pāṇini's grammar operates forms an interval which ends immediately before a marker element. As a result, Pāṇini could use pairs consisting of a sound and a marker element in order to designate the sound classes in his grammar. Such a pair denotes the continuous sequence of sounds in the interval between the sound and the marker. E.g., the pair $i\,M_2$ denotes the class {i, u, ṛ, ḷ}.

The question whether Pāṇini arranged the sounds in the Śivasūtras in an optimal way and especially whether the double occurrence of the sound h (in the 5th and in the 14th sūtra) is necessary has been widely discussed (cf. Böhtlingk, 1887; Staal, 1962; Cardona, 1969; Kiparsky, 1991). In Petersen (2004) it could be proven that

Since Pāṇini's grammar was designed for oral tradition, he could not make use of visual symbols (like arrows, slashes, ...) to indicate the role of the sound classes in a rule. He takes case suffixes instead which he uses meta-linguistically in order to mark the role a class plays in a rule. In Pāṇinian style rule (1) becomes

$$A + \text{genitive},\ B + \text{nominative},\ C + \text{ablative},\ D + \text{locative}. \qquad (2)$$

there is no shorter solution than the Śivasūtras to the problem of ordering the sounds of Sanskrit in a linear, by markers interrupted list with as few repeated sounds as possible such that each phonological class which is denoted by a sound-marker pair in Pāṇini's grammar can be represented by such a pair with respect to the list.[4] This shows that the double occurrence of *h* is not superfluous and that Pāṇini used a minimal number of markers in the Śivasūtras.

2 Linear Coding of Non-linear Hierarchies: Generalizing Pāṇini's Śivasūtra-Technique

In Sect. 1.1 we have argued that there are several situations in which it is required to force entities in a linear order although it would be more natural to organize them in a non-linear hierarchy. The aim of the present section is to show that Pāṇini's Śivasūtra-technique, which has been introduced in Sect. 1.2, may offer a solution to the mentioned problem in many situations.

2.1 S-Orders and S-Sortability: Formal Foundations

All ordering problems mentioned in Sect. 1.1 are based on a common problem:

Problem 1. Given a set of classes of entities (no matter on what aspects the classification is based) order the entities in a linear order such that each single class forms a continuous interval with respect to that order.

Take for example the problem of ordering books in a library. It would be favorable to order the books linearly on the bookshelves such that all the books belonging to one thematic subfield are placed next to each other on the shelves without having to add additional copies of a book into the order.

Pāṇini solved Problem 1 with his Śivasūtras in a concrete case: The Śivasūtras define a linear order on the set of sounds in Sanskrit (with one sound occurring twice) in which each class of sounds required in his grammar forms a continuous interval.[5] In order to solve concrete instances of Problem 1, one can do without Pāṇini's special technique of interrupting the order by marker elements such that each class interval ends immediately before a marker. In tribute to Pāṇini's Śivasūtras we call a linear order which solves an instance of Problem 1 a *Śivasūtra-order* or short *S-order*. A set of classes is said to be *S-sortable* without duplications

[4] Actually, the Śivasūtras are one of nearly 12,000,000 arrangements which are equal in length (Petersen, 2008).

[5] Actually, for the denotation of some sound classes Pāṇini used different techniques in his grammar (for details see Petersen, 2008, 2009).

if it forms a solvable instance of Problem 1, i.e., if a corresponding S-order exists. Obviously, not every set of classes is S-sortable without duplications. For instance, even the Śivasūtras do not define an S-order as they contain one sound twice. By proving that the double occurrence of *h* is unavoidable it has been shown that no S-order exists for the set of sound classes required by Pāṇini's grammar (cf. Petersen, 2004). However, it should be clear that by a clever duplication of enough elements each set of classes can be S-sorted.

The following definition summarizes formally the terminology which we will use hereinafter in order to generalize and apply Pāṇini's Śivasūtra-technique:

Definition 1. Given a base set $\mathcal{A}$ and a set of subsets Φ with $\bigcup \Phi = \mathcal{A}$, a linear order $<$ on $\mathcal{A}$ is called an *S-order* of $(\mathcal{A}, \Phi)$ if and only if the elements of each set $\phi \in \Phi$ form an interval in $(\mathcal{A}, <)$.

Furthermore, $(\mathcal{A}, \Phi)$ is said to be *S-sortable without duplications* if and only if there exists an S-order $(\mathcal{A}, <)$ of $(\mathcal{A}, \Phi)$.

Two simple examples serve us through the rest of the paper as illustrations:

Example 1. Given the base set $\mathcal{A} = \{a, b, c, d, e, f, g, h, i\}$ and the set of classes $\Phi = \{\{d, e\}, \{a, b\}, \{b, c, d, f, g, h, i\}, \{f, i\}, \{c, d, e, f, g, h, i\}, \{g, h\}\}$, $(\mathcal{A}, \Phi)$ is S-sortable without duplications and $a \prec b \prec c \prec g \prec h \prec f \prec i \prec d \prec e$ is an S-order of $(\mathcal{A}, \Phi)$.

Example 2. Given the base set $\mathcal{A} = \{a, b, c, d, e, f\}$ and the set of classes $\Phi = \{\{d, e\}, \{a, b\}, \{b, c, d\}, \{b, c, d, f\}\}$, $(\mathcal{A}, \Phi)$ is not S-sortable without duplications.

Example 2 is not S-sortable without duplications since $\{b, c, d\} \in \Phi$ demands that no element of $\mathcal{A} \backslash \{b, c, d\}$ may stand between any two elements of $\{b, c, d\}$. Furthermore, from $\{d, e\} \in \Phi$ and $\{a, b\} \in \Phi$ it follows that either $a < b < c < d < e$ or $e < d < c < b < a$ is true. But this is impossible since it contradicts $\{b, c, d, f\} \in \Phi$.

In the following, we will show how an S-order for a set of classes which is S-sortable without duplications can be constructed. Out of the construction process a condition for S-sortability can be derived. This condition is such that it can also help to identify those elements which must be duplicated in the case of a set of classes which is not S-sortable without duplications.

2.2 Constructing S-Orders

In the case of a set of classes which is S-sortable without duplications its S-orders can be read off from its enlarged concept lattice. The term concept lattice is taken from Formal Concept Analysis (FCA), i.e., a mathematical theory for the analysis of data (cf. Ganter, & Wille, 1999). We do not need to evolve the whole apparatus of FCA; it is sufficient to illustrate what we understand by the concept lattice of a set

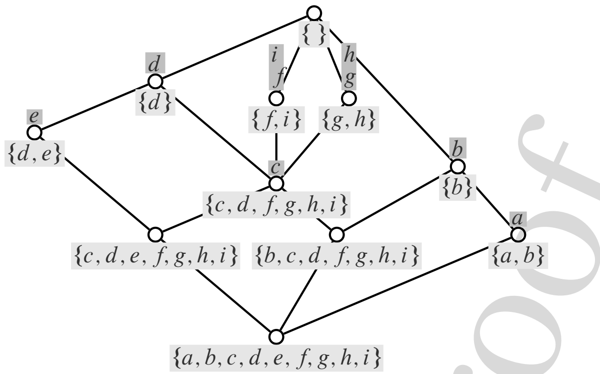
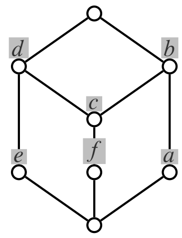


Fig. 2 Concept lattice of $(\mathcal{A}, \Phi)$ from Example 1

Fig. 3 Concept lattice of $(\mathcal{A}, \Phi)$ from Example 2



of classes by an example: Given the base set $\mathcal{A}$ and the set of classes Φ from Example 1, the concept lattice of $(\mathcal{A}, \Phi)$ is given in Fig. 2. It is constructed as follows: All elements of Φ and all possible intersections of elements of Φ are ordered by the set-inclusion relation such that subsets are placed above their supersets. Formally, Fig. 2 shows the Hasse-diagram of the ordered set $(\mathcal{A} \cup \{\phi \mid \phi = \bigcap \Psi \text{ with } \Psi \subseteq \Phi\}, \supseteq)$.[6]

In Fig. 2, you find below each node its corresponding set written. However, it is not necessary to label each node by its corresponding set since it is sufficient to write each element of the base set $\mathcal{A}$ to that node which corresponds to the smallest set which contains the element. The result of this more economical labeling method is shown in Fig. 2 by the labels above the nodes. The set corresponding to a node can be regained from the sparing labels by collecting all labels attached to nodes which can be reached by moving upwards in the graph. Hence, from now on, solely the sparing labels will be shown in figures of concept lattices like in Fig. 3 which shows the concept lattice for Example 2.

Although it would be possible to read off the S-orders for Example 1 from the concept lattice in Fig. 2 (cf. Petersen, 2004, 2008), it is easier to switch to the concept lattice of the enlarged set of classes. Enlarging the set of classes means adding each element of the base set as a singleton set to the set of classes, e.g., in the case of

[6] The Hasse-diagram of a partially ordered set is the directed graph whose vertices are the elements of the set and whose edges correspond to the upper neighbor relation determined by the partial order.

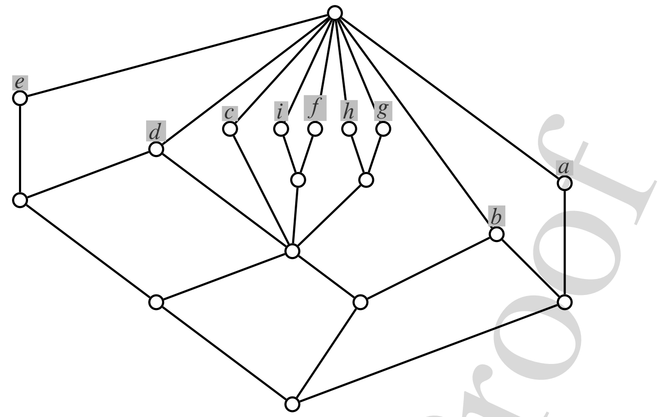


Fig. 4 Enlarged concept lattice for Example 1

Example 1 the set of classes has to be enlarged by the classes $\{a\}, \{b\}, \{c\}, \ldots, \{i\}$. The enlarged concept lattice corresponding to Example 1 is shown in Fig. 4.

The following theorem states the connection between S-orders and concept lattices of enlarged sets of classes:

Theorem 1. *A set of classes* $(\mathcal{A}, \Phi)$ *is S-sortable without duplications if and only if a plane drawing of the Hasse-diagram of the concept lattice of the enlarged set of classes* $(\mathcal{A}, \tilde{\Phi})$ *exists* ($\tilde{\Phi} = \Phi \cup \{\{a\} \mid a \in \mathcal{A}\}$).[7]

The full proof of this theorem is given in Petersen (2009) and sketched in Petersen (2008). It follows immediately from the definition of our concept lattices that whenever the Hasse-diagram of an enlarged concept lattice can be drawn without intersecting edges then an S-order without duplications exists: Concept lattices order sets by set inclusion; this ensures that the labels belonging to the elements of one class out of a set of classes form an interval in the sequence defined by the left-to-right order of the labels in a plane drawing of the Hasse-diagram of the enlarged concept lattice. It follows that this left-to-right defines an S-order without duplications of the set of classes. For example, the plane Hasse-diagram in Fig. 4 defines the S-order $e < d < c < i < f < h < g < b < a$ for the set of classes from Example 1.

The proof of the reversed statement, i.e., that the existence of an S-order implies the existence of a plane drawing of the Hasse-diagram, was first given in Petersen (2004). The proof involves an explicit construction of a plane drawing of the Hasse-diagram of the enlarged concept lattice for any S-order of any S-sortable set of classes. The construction guarantees that the left-to-right order of the labels equals the original S-order.

[7] A drawing of a Hasse-diagram is said to be plane if it shows no intersecting edges.

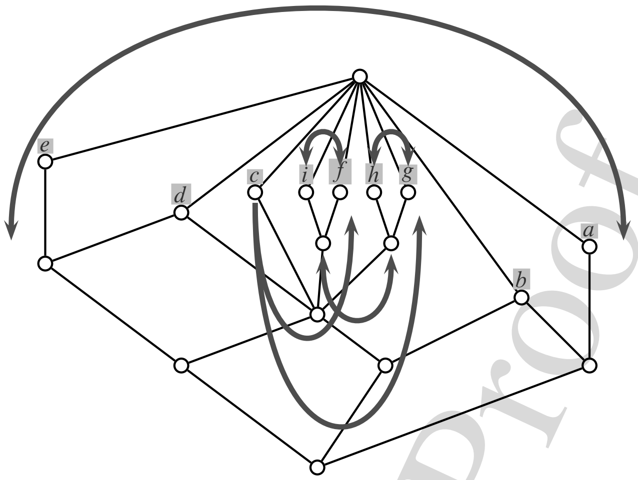


Fig. 5 Illustration of the distinct plane drawings of the enlarged concept lattice for Example 1 resulting in different S-orders

However, the construction method does not deterministically result in one S-order, since usually several plane drawings exist for the concept lattice of an enlarged set of classes. In fact, the proof of the theorem above implies that for every S-order there exists a plane drawing of the concept lattice from which it can be read off. Figure 5 illustrates the distinct plane drawings of the enlarged concept lattice for Example 1 which result in different S-orders, as for example:

$$e < d < c < i < f < h < g < b < a$$
$$e < d < i < f < c < h < g < b < a \qquad \text{(3 variations)}$$
$$e < d < c < f < i < h < g < b < a \qquad \text{(2 variations)}$$
$$e < d < c < i < f < g < h < b < a \qquad \text{(2 variations)}$$
$$e < d < c < h < g < i < f < b < a \qquad \text{(2 variations)}$$
$$a < b < g < h < f < i < c < d < e \qquad \text{(2 variations)}$$

Altogether, Example 1 has 48 ($= 3 \times 2 \times 2 \times 2 \times 2$) distinct solutions, i.e., distinct S-orders.

2.3 *The Problem of Identifying Elements for Duplication*

Theoretically, Theorem 1 enables us to specify for each set of classes whether it is S-sortable without duplications or not. Since in the case of an S-sortable set of classes the proof of the theorem even establishes a method to construct a concrete S-order Theorem 1 solves Problem 1 in theory. Though in practise, deciding whether the

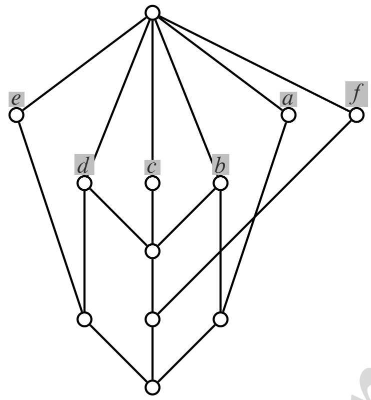


Fig. 6 Enlarged concept lattices for Example 2

Hasse-diagram of a concept lattice can be drawn without intersecting edges is not trivial. For smaller examples like Example 2 a close inspection of the Hasse-diagram of the enlarged concept lattice given in Fig. 6 is sufficient to see that it is impossible to draw this Hasse-diagram without intersecting edges. This proves that the set of classes in Example 2 is not S-sortable. However, for more complex sets of classes like the one given by the sound classes used in Pāṇini's Sanskrit grammar the investigation of the Hasse-diagrams become more awkward. Other necessary as well as sufficient conditions for S-sortability have been developed in Petersen (2008) which are easier to verify, but due to space limits they cannot be evolved here in detail. The most useful condition is based on the property of being bipartite of so-called Ferrers-graphs (cf. Ganter, & Wille, 1999; Petersen, 2008, 2009; Zschalig, 2007). Whether a graph is bipartite can be checked algorithmically; hence, this conditions opens up a way of investigating more complex sets of classes automatically.

The problem of identifying the best candidates for duplication is intricate, too. In order to construct an optimal S-order for a set of classes which is not S-sortable one has to identify those elements whose duplication leads to the "shortest" S-order, i.e., the aim is to duplicate as few elements as possible. In the case of Example 2 it is sufficient to duplicate one element, namely for example b. Duplicating element b means adding a copy b' to the base set $\mathcal{A}$ and changing some instances of b in the set of subsets Φ to b'. One of the optimal solutions to Example 2 is to duplicate b such that the new base set becomes $\{a, b, b', c, d, e, f\}$ and the new set of subsets becomes $\{\{d, e\}, \{a, b'\}, \{b, c, d\}, \{b, c, d, f\}\}$. An S-order of the new set of classes with one duplication is for instance

$$f < b < c < d < e < a < b'.$$

In Petersen (2008) a whole battery of methods for the identification of elements which are good candidates for duplication is developed. Although it can still be hard to identify good candidate elements for duplication, the problem becomes much less challenging if one does not ask for absolutely minimal but just quite minimal S-orders.

For instance, in situations as described in Sect. 1.1, where books or products have to be forced into a linear order, adding additional copies is expensive and space consuming, but not impossible. It can be assumed that by applying S-orders less books or products have to be placed at two distinct regions than by applying standard mono-hierarchical classification methods.

Acknowledgements Part of the research presented in this paper was made possible by the Deutsche Forschungsgemeinschaft within the Forschergruppe FOR 600.

References

Böhtlingk, O. (1887). *Pāṇinis grammatik*. Leipzig: H. Haessel. (Reprinted in 1964, Hildesheim: Georg Olms).

Cardona, G. (1969). Studies in Indian grammarians I: The method of description reflected in the Śiva-Sūtras. *Transactions of the American Philosophical Society, 59*(1), 3–48.

Ganter, B., & Wille, R. (1999). *Formal concept analysis. Mathematical foundations*. Berlin: Springer.

Kiparsky, P. (1991). Economy and the construction of the Śivasūtras. In M. M. Deshpande, & S. Bhate (Eds.), *Pāṇinian Studies*. Ann Arbor, MI: University of Michigan.

Misra, V. N. (1966). *The descriptive technique of Pāṇini. An introduction*. The Hague: Mouton.

Petersen, W. (2004). A mathematical analysis of Pāṇini's Śivasūtras. *Journal of Logic, Language, and Information, 13*(4), 471–489.

Petersen, W. (2008). *Zur Minimalität von Pāṇinis Śivasūtras – Eine Untersuchung mit Methoden der Formalen Begriffsanalyse*. PhD thesis, University of Düsseldorf.

Petersen, W. (2009). On the construction of Pāṇini's Śivasūtras. In A. Kulkarni, & G. Huet (Eds.), *Sanskrit computational linguistics*, *LNAI* (Vol. 5406, pp. 78–97). Berlin: Springer.

Staal, F. J. (1962). A method of linguistic description. *Language, 38*, 1–10.

Staal, F. J. (1982). Ritual, grammar and the origins of science in India. *Journal of Indian Philosophy, 10*(1), 3–35.

Zschalig, C. (2007). Bipartite Ferrers-graphs and planar concept lattices. In S. O. Kuznetsov, & S. Schmidt (Eds.), *Proceedings of the 5th ICFCA*, *LNCS* (Vol. 4390, pp. 313–327). Berlin: Springer.