How formal concept lattices solve a problem of ancient linguistics

Wiebke Petersen
Department of Computational Linguistics
Institute of Language and Information
University of Düsseldorf

Pāṇini's Śivasūtras

अइउण्॥ऋऌक्॥एओङ्॥ऐऔच्॥हयवरट्॥लण्॥ञमङणनम्॥झभञ्।
।घढधष्॥जबगडदश्॥खफछठथचटतव्॥कपय्॥शषसर्॥हल्॥

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

Phonological rules

A is replaced by **B** if preceded by **C** and followed by **D**

- in modern form: $A \rightarrow B/_{C\_\_D}$
- as context-sensitive rule: $CAD \rightarrow CBD$

Example: final devoicing in German (Hunde - Hund)

[d]→[t] / _#,   [b]→[p] / _#,   [g]→[k] / _#, ...

$$\begin{bmatrix} +consonantal \\ -nasal \\ +voiced \end{bmatrix} \rightarrow \begin{bmatrix} +consonantal \\ -nasal \\ -voiced \end{bmatrix} /\_\#$$

Pāṇini's coding of rules

$$A \rightarrow B/_{C\_\_D}$$

A + genitive, B + nominative, C + ablative, d + locative

6.1.77. *iko yaṇ aci* (इको यण् अचि)

$[\text{ik}]_{\text{genitive}} \quad [\text{yaṇ}]_{\text{nominative}} \quad [\text{ac}]_{\text{locative}}$

$$[\text{iK}] \rightarrow [\text{yṆ}]/\_\_[\text{aC}]$$

Pāṇini's Śivasūtras

अइउण्॥ऋऌक्॥एओङ्॥ऐऔच्॥हयवरट्॥लण्॥ञमङणनम्॥झभञ्।
।घढधष्॥जबगडदश्॥खफछठथचटतव्॥कपय्॥शषसर्॥हल्॥

a·i·uṇ || ṛ·ḷk || e·oṅ || ai·auc || hayavaraṭ || laṇ || ñamaṅaṇanam || jhabhañ ||

ghaḍhadhaṣ || jabagaḍadaś || khaphachaṭhathacaṭatav || kapay || śaṣasar || hal ||

Pāṇini's Śivasūtras

anubandha

sūtras

1.	a	i	u			Ṇ
2.				ṛ	ḷ	K
3.		e	o			Ṅ
4.		ai	au			C
5.	h	y	v	r		Ṭ
6.					l	Ṇ
7.	ñ	m	ṅ	ṇ	n	M
8.	jh	bh				Ñ
9.			gh	ḍh	dh	Ṣ
10.	j	b	g	ḍ	d	Ś
11.	kh	ph	ch	ṭh	th	
			c	ṭ	t	V
12.	k	p				Y
13.		ś	ṣ	s		R
14.	h					L

Phonological classes/ *pratyāhāras*

1.	a	i	u				Ṇ
2.				ṛ	ḷ		K
3.		e	o				Ṅ
4.		ai	au				C
5.	h	y	v	r			Ṭ
6.					l		Ṇ
7.	ñ	m	ṅ	ṇ	n		M
8.	jh	bh					Ñ

iC

Phonological classes are denoted by *pratyāhāras.*
E.g., the *pratyāhāra* *iC* denotes the set of segments in the continuous sequence starting with *i* and ending with *au*, the last element before the *anubandha* *C.*

Minimality criteria

1. The length of the whole list is minimal.
2. The length of the sublist of the anubandhas is minimal and the length of the whole list is as short as possible.
3. The length of the sublist of the sounds is minimal and the length of the whole list is as short as possible.

Minimality criteria

1. The length of the whole list is minimal.
2. The length of the sublist of the anubandhas is minimal and the length of the whole list is as short as possible.
3. The length of the sublist of the sounds is minimal and the length of the whole list is as short as possible.
   - no duplication of *h*
   - less anubandhas

Basic concepts

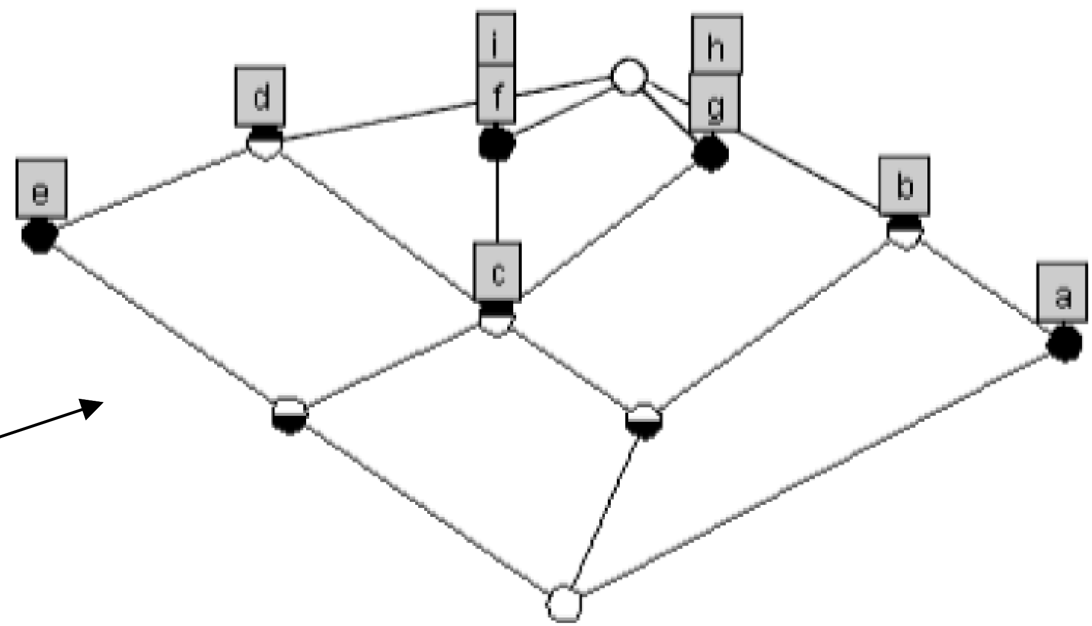
S-encodable set of sets: $\Phi=\{\{d,e\},\{b,c,d,f,g,h,i\},\{a,b\},\{f,i\},\{c,d,e,f,g,h,i\},\{g,h\}\}$

S-alphabet $(\mathcal{A},\Sigma,<)$ of Φ: e d M_1 c i f M_2 g h M_3 b M_4 a M_5

alphabet

marker

total order on $\mathcal{A}\cup\Sigma$

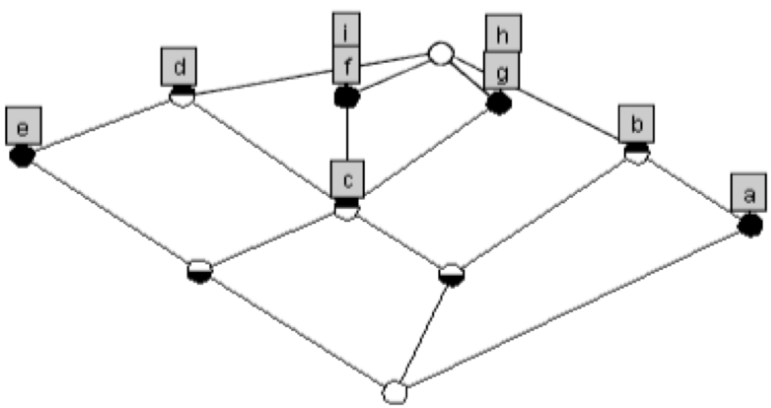


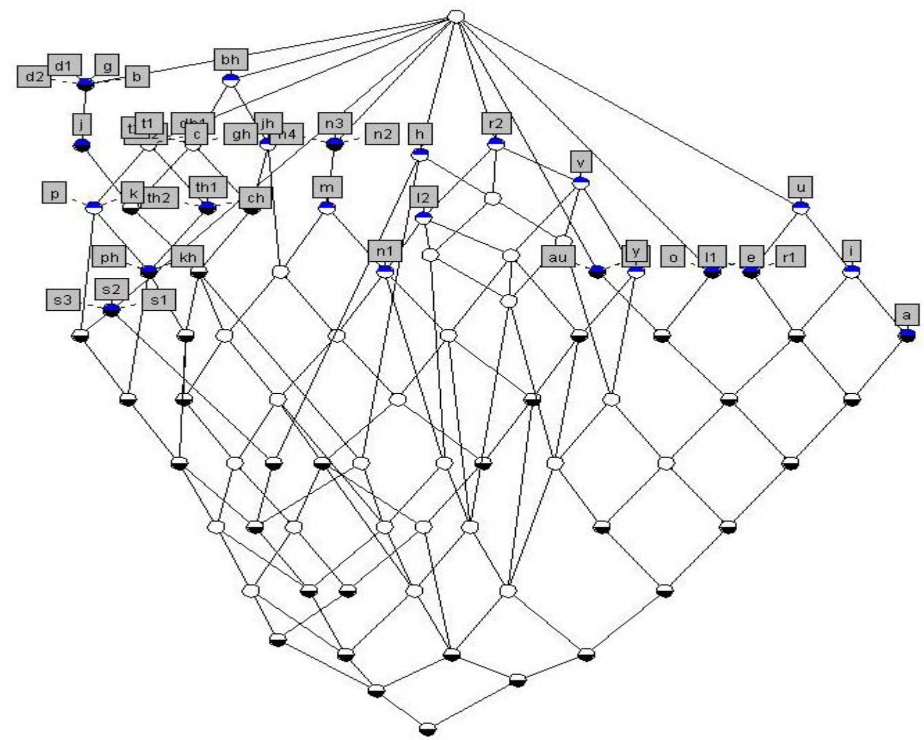
$\underline{\mathcal{B}}(\Phi,\mathcal{A},\ni)$

$$\mathcal{A}=\bigcup_{\phi\in\Phi}\phi$$

S-encodability and planar formal concept lattices

If Φ is S-encodable, then the formal concept lattice $\underline{\mathcal{B}}(\Phi, \mathcal{A}, \ni)$ is planar

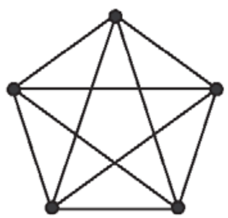




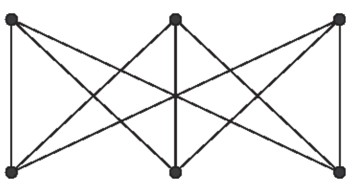
Hasse-diagram for Pāṇini's *pratyāhāras*

S-encodability and planar formal concept lattices

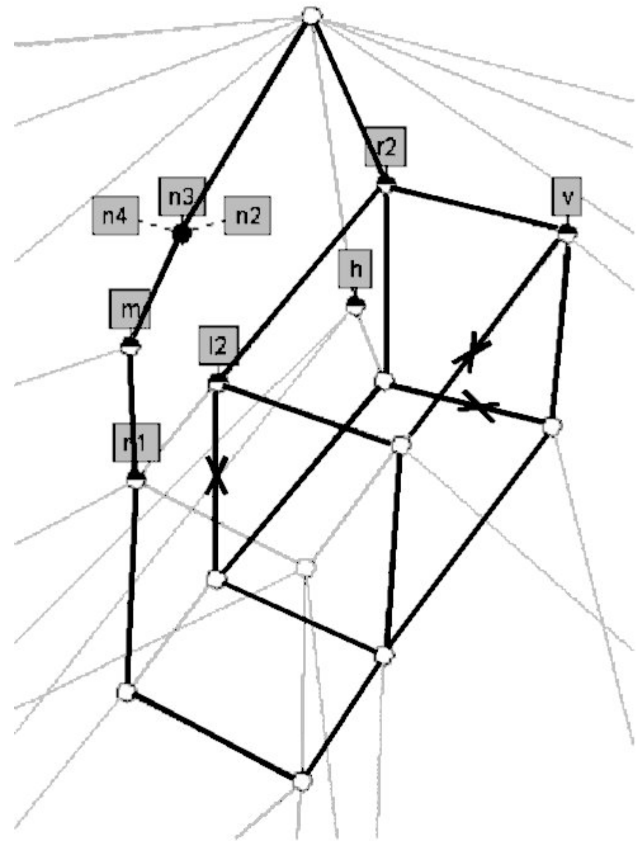
Criterion of Kuratowski:
A graph is planar iff it has neither K^5 nor $K_{3,3}$ as a *minor*.



K^5

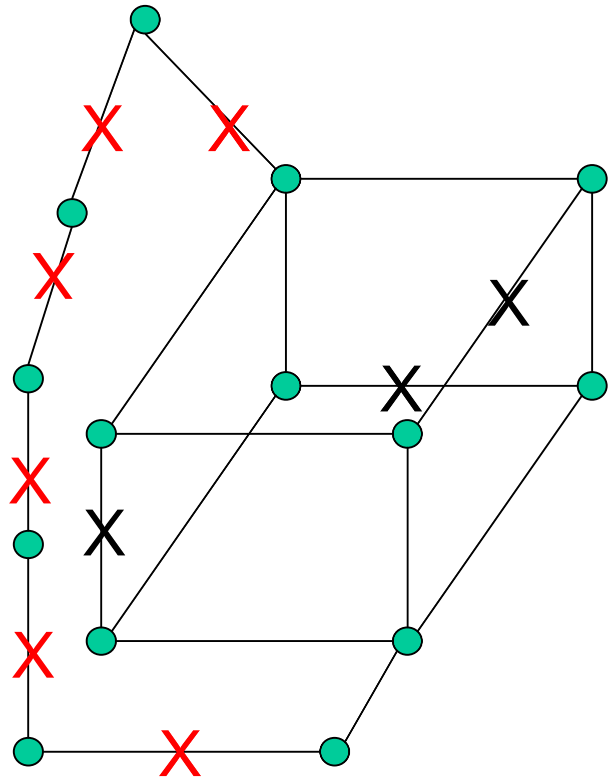


$K_{3,3}$

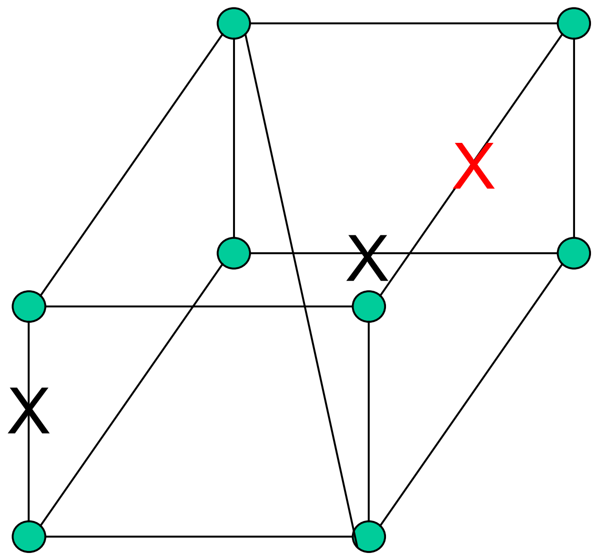


part of the concept lattice for Pāṇini's phonological classes

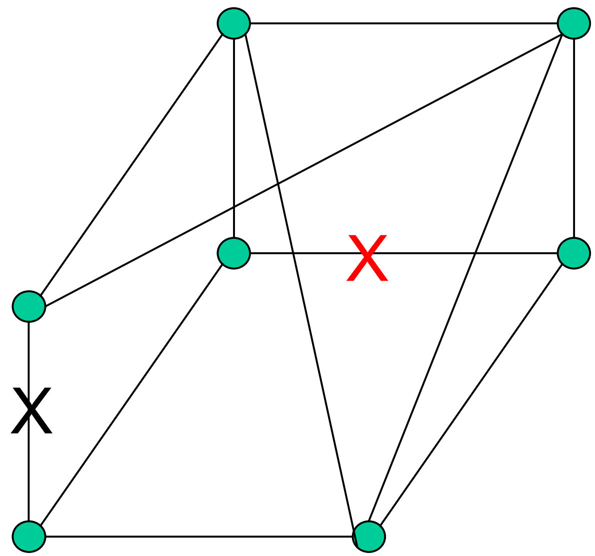
K^5 is a minor of the concept lattice for Pāṇini's phonological classes



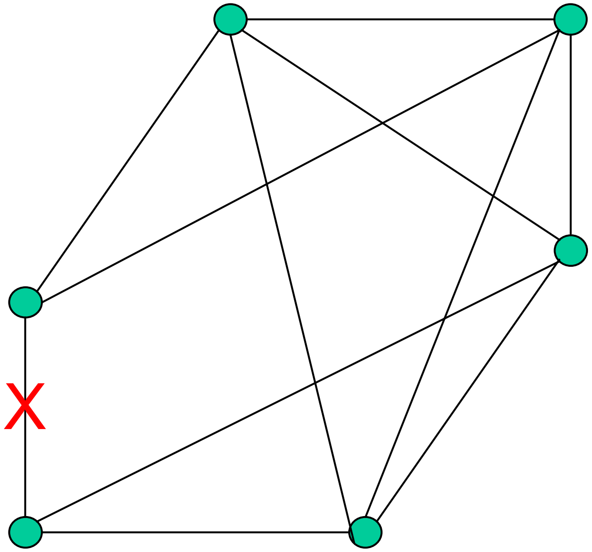
K^5 is a minor of the concept lattice for Pāṇini's phonological classes



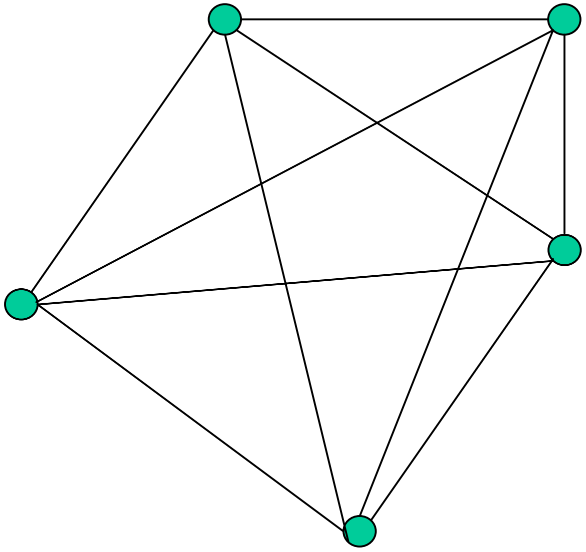
K^5 is a minor of the concept lattice for Pāṇini's phonological classes



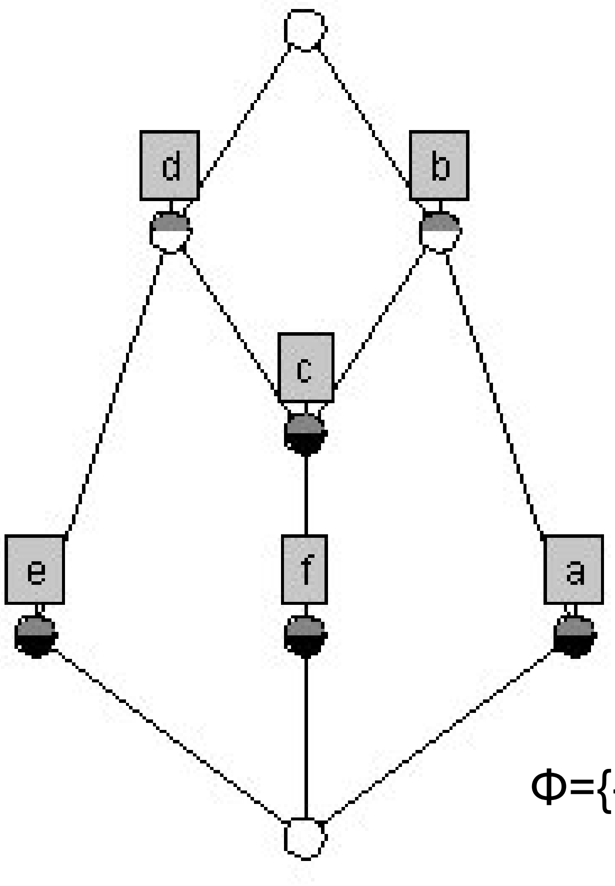
K^5 is a minor of the concept lattice for Pāṇini's phonological classes



K^5 is a minor of the concept lattice for Pāṇini's phonological classes



We are not done yet!



plane but not S-encodable!

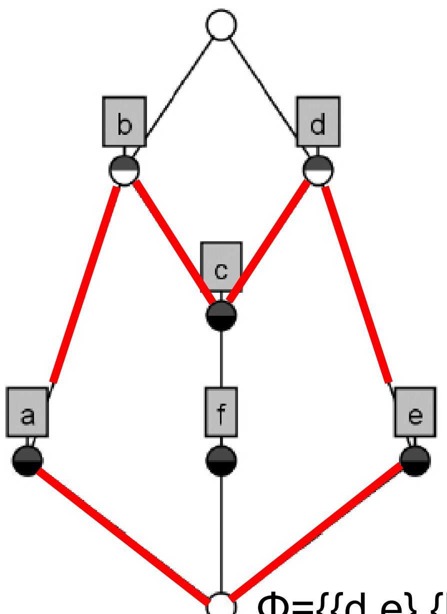
$\Phi=\{\{d,e\},\{b,c,d,f,\},\{a,b\},\{b,c,d\}\}$

Existence of S-alphabets

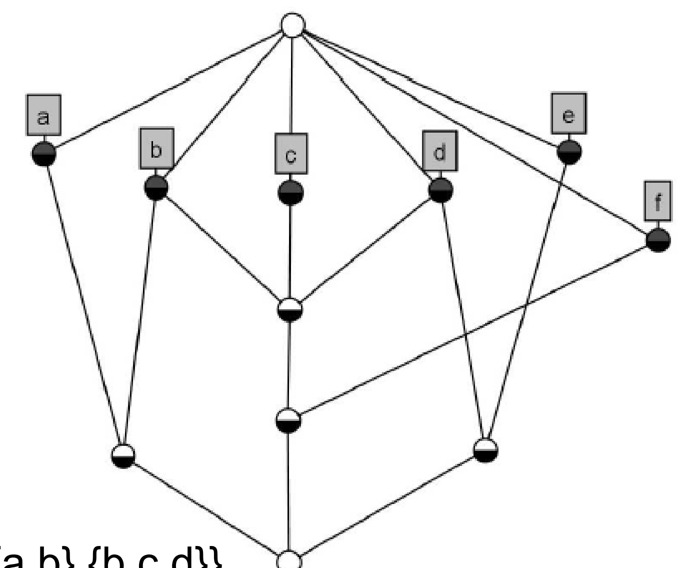
$\bar{\Phi} = \Phi \cup \{\{a\} : a \in \mathcal{A}\}$

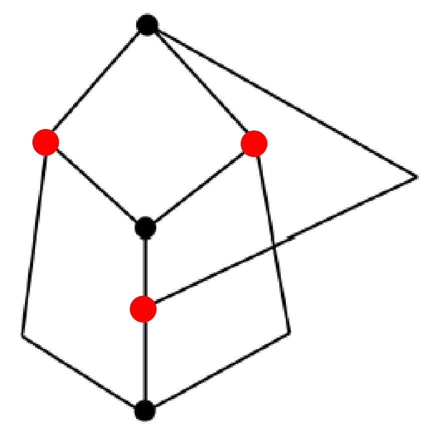
The following statements are equivalent:

1. $(\mathcal{A}, \Phi)$ is S-encodable
2. $\underline{\mathcal{B}}(\bar{\Phi}, \mathcal{A}, \ni)$ is planar



Φ={{d,e},{b,c,d,f,},{a,b},{b,c,d}}



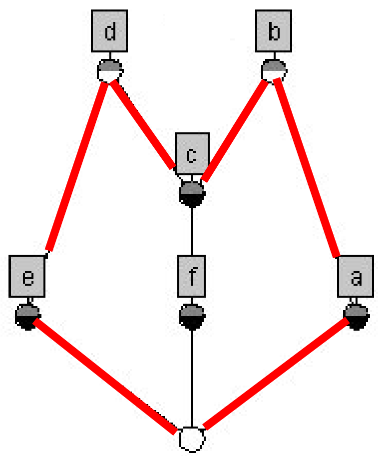


Existence of S-alphabets

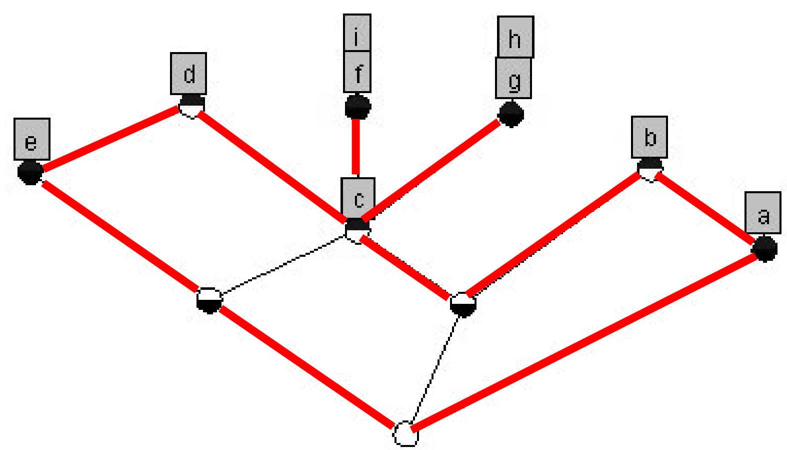
$\bar{\Phi} = \Phi \cup \{\{a\} : a \in \mathcal{A}\}$

The following statements are equivalent:

1. $(\mathcal{A}, \Phi)$ is S-encodable
2. $\underline{\mathcal{B}}(\bar{\Phi}, \mathcal{A}, \ni)$ is planar
3. the S-graph contains all attribute concepts

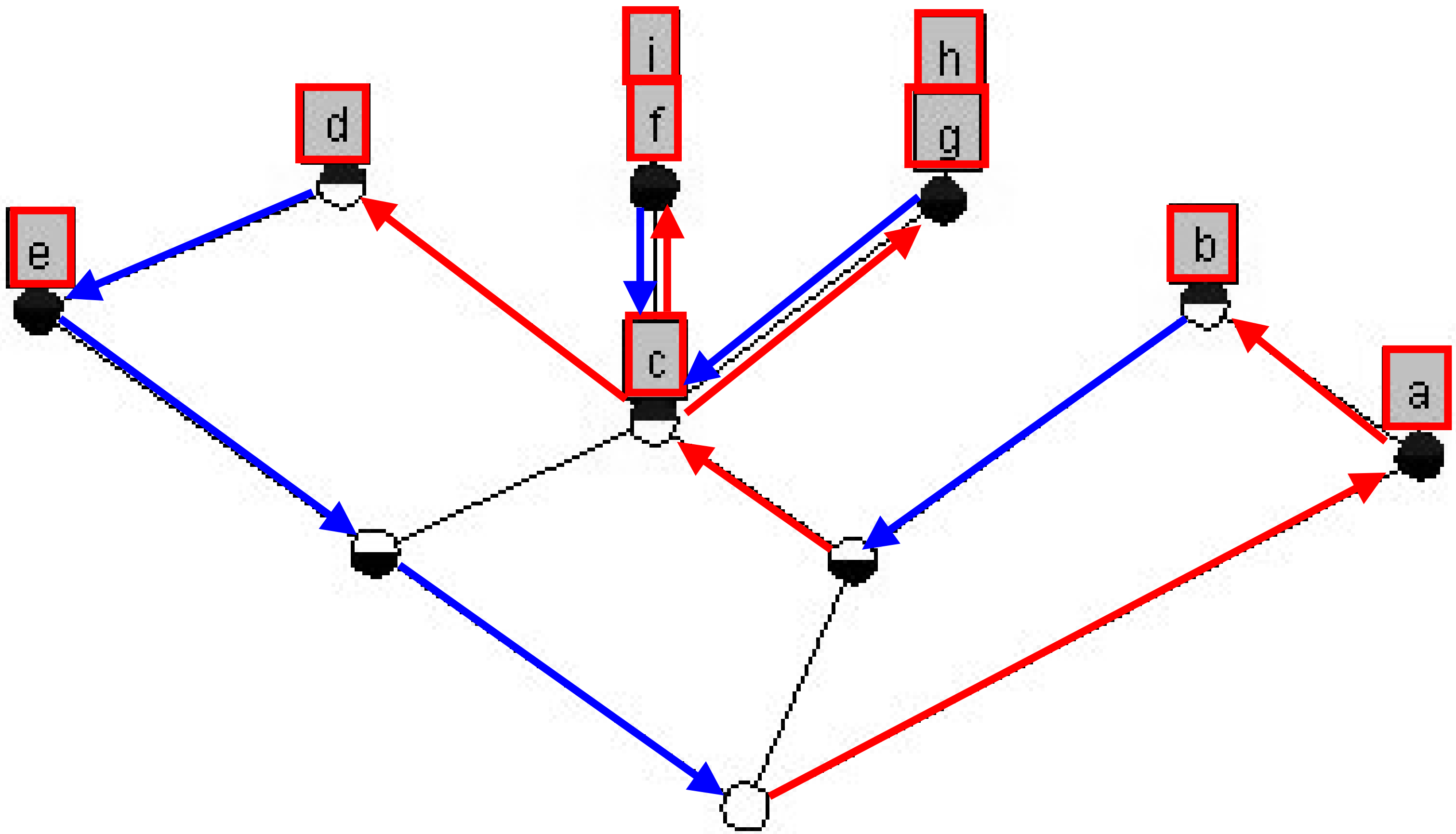


not S-encodable



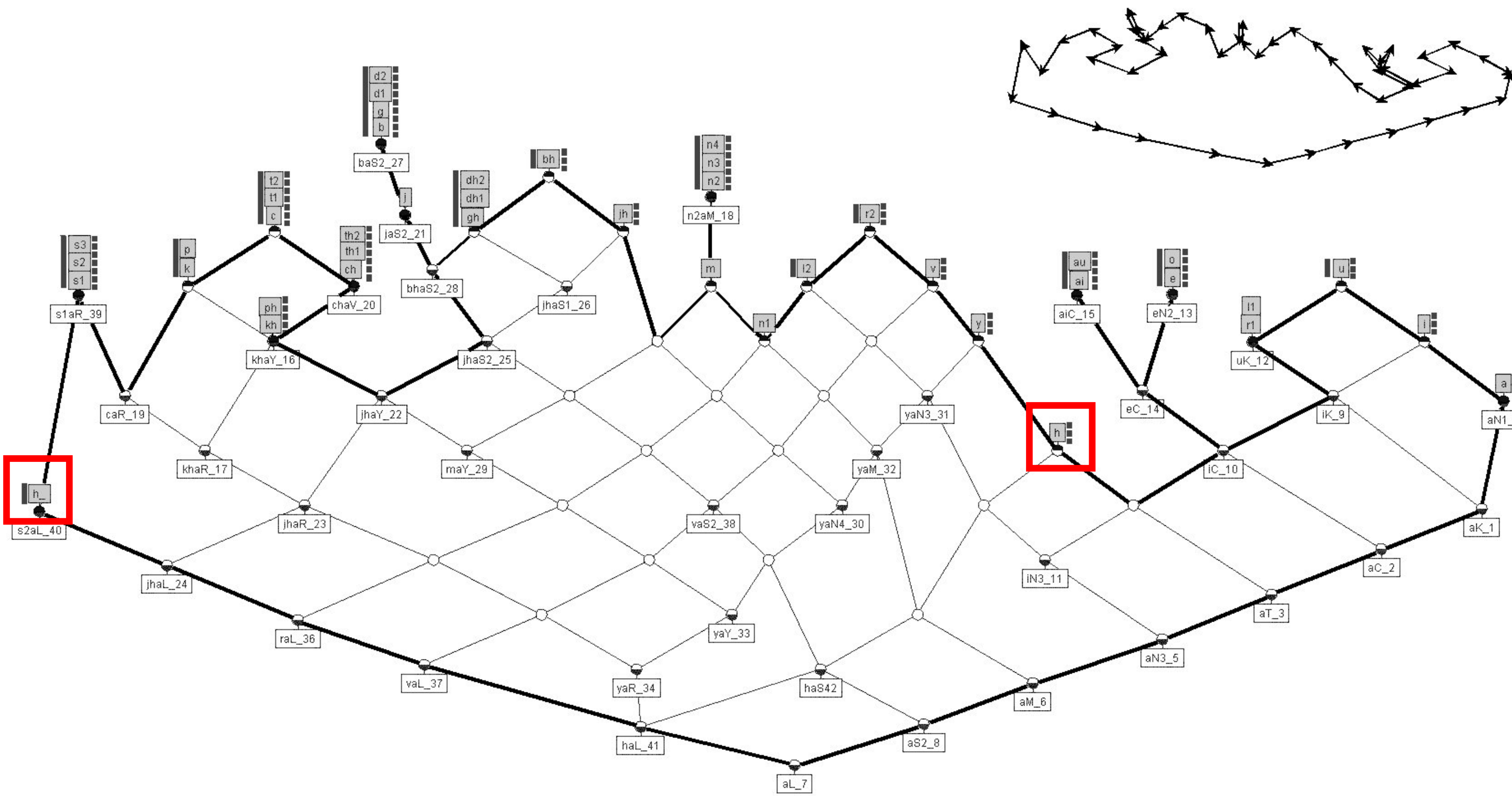
S-encodable

Construction of S-alphabets

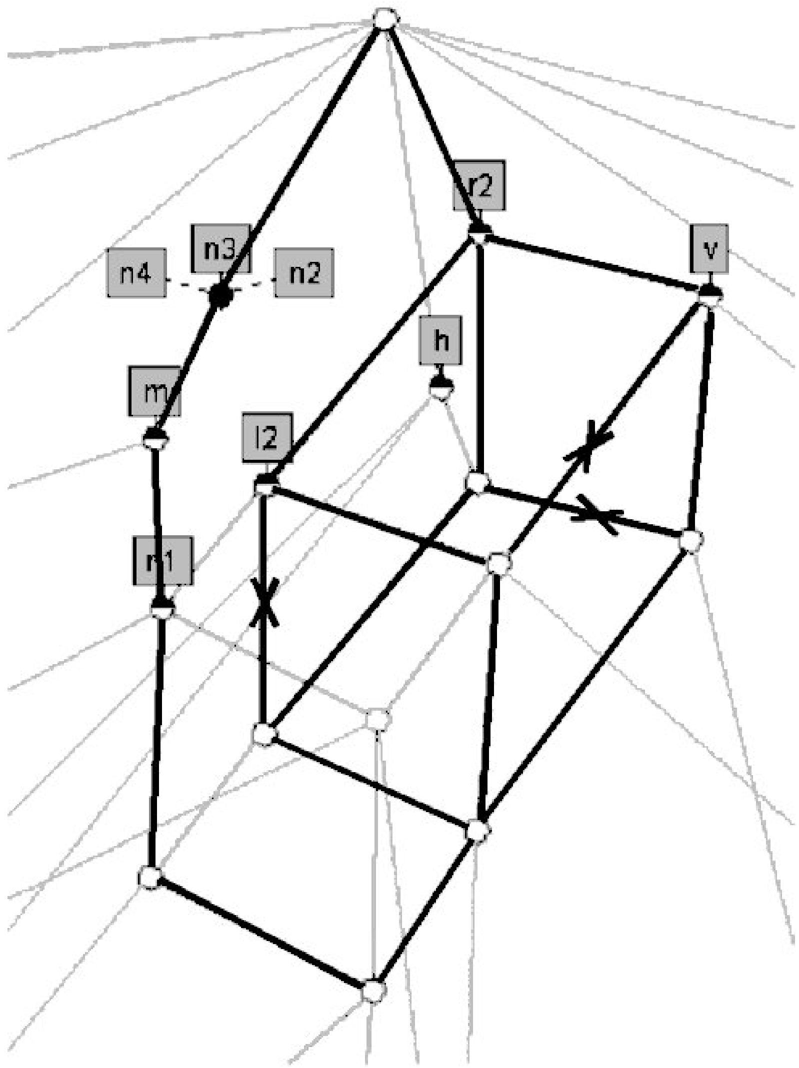


a b $\mathbf{M_1}$ c g h $\mathbf{M_2}$ ~~c~~ i f $\mathbf{M_3}$ ~~c~~ d $\mathbf{M_4}$ e $\mathbf{M_5}$

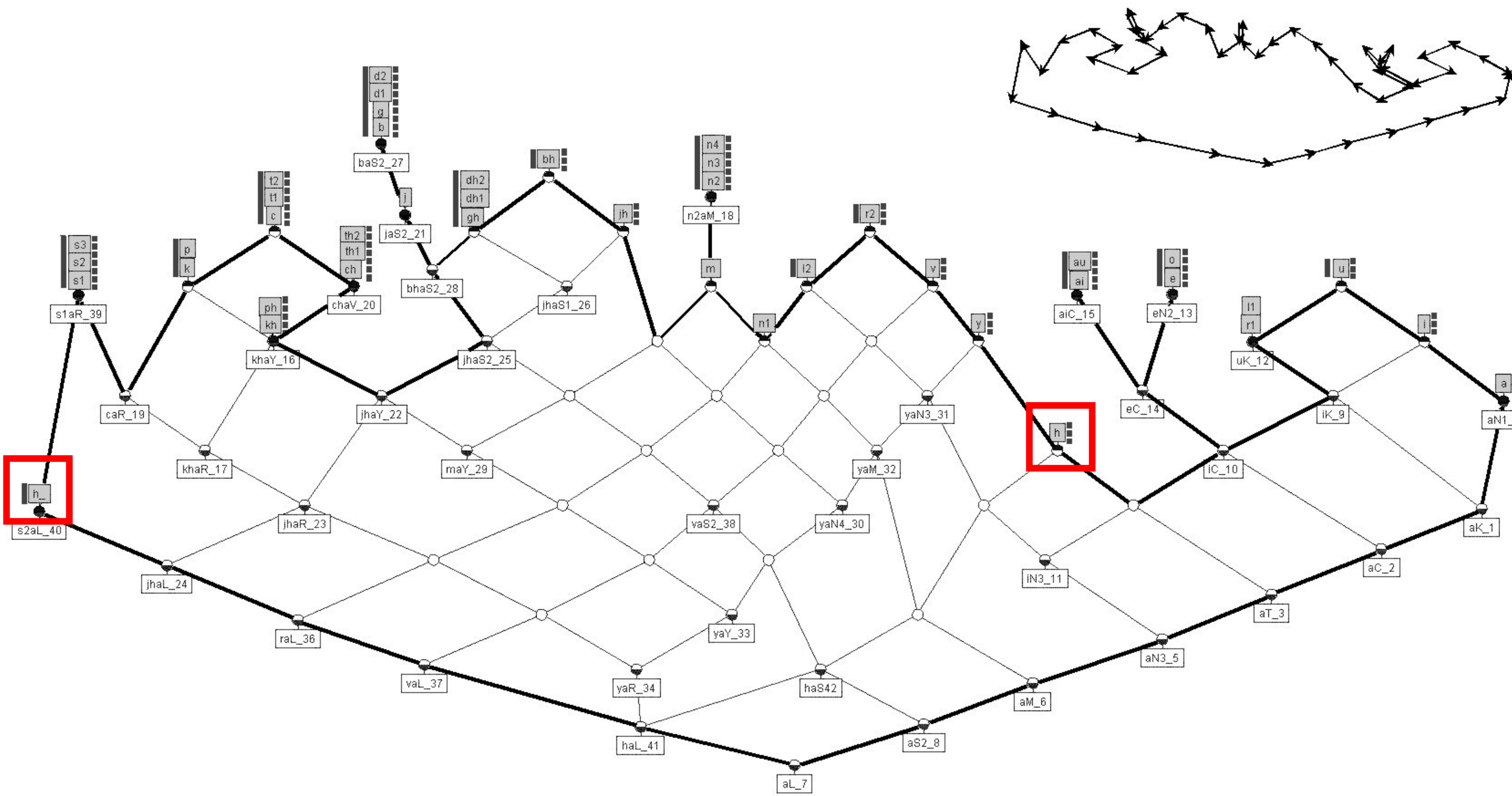
Pāṇini's Śivasūtras are optimal



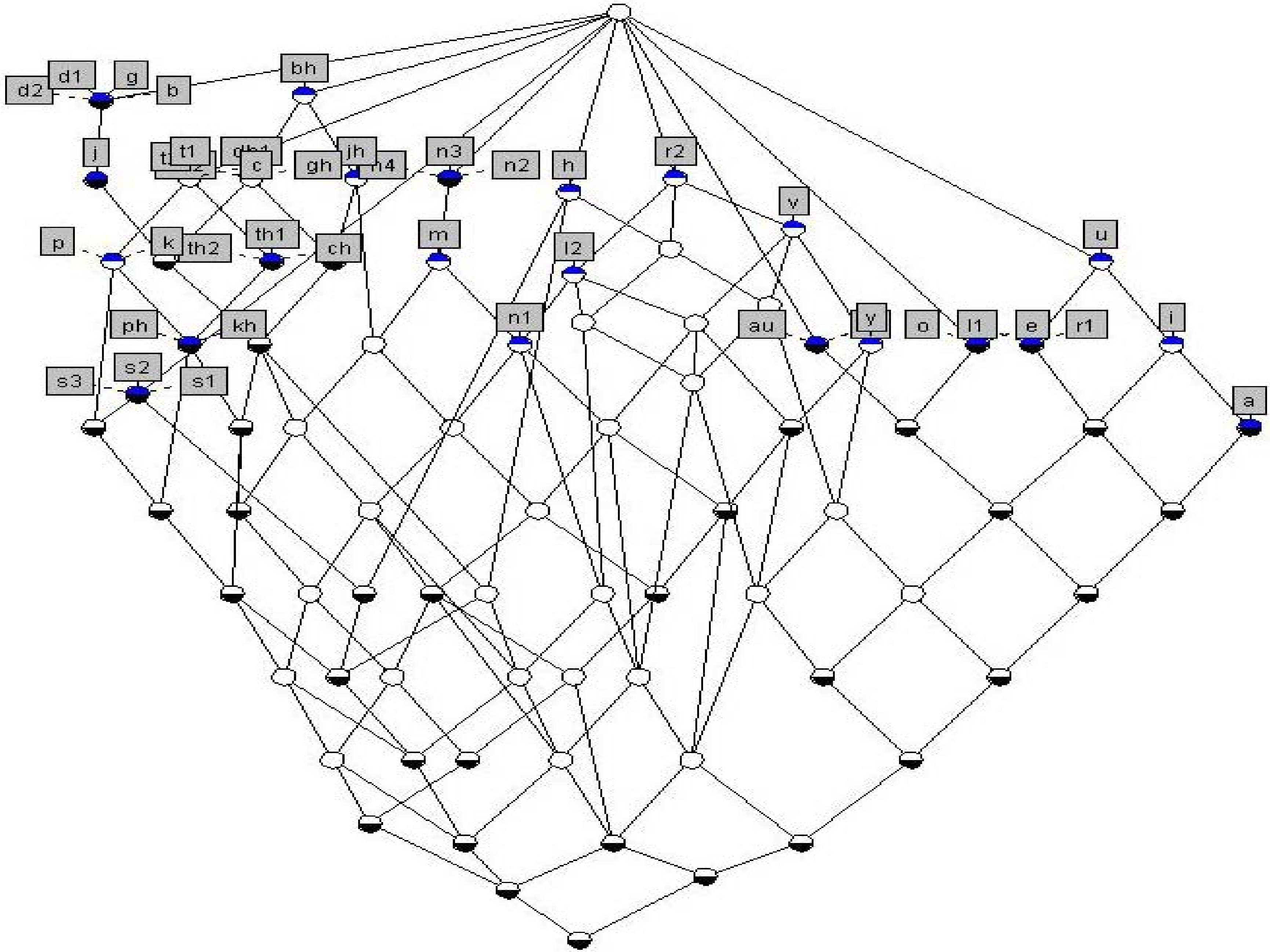
HEINRICH HEINE UNIVERSITÄT DÜSSELDORF


r2
n3
n4
n2
v
h
m
l2
n1

Pāṇini's Śivasūtras are optimal



HEINRICH HEINE
UNIVERSITÄT
DÜSSELDORF



d2
d1
g
b
bh
j
t1
c
gh
jh
n4
n3
n2
h
r2
v
p
k
th2
th1
ch
m
l2
u
ph
kh
n1
au
y
o
l1
e
r1
i
s2
s3
s1
a