



What a
RegExp!



charset
≠
character set



RegExp
≠
regular expression



RegExp
 \supseteq
regular expression

Reguläre Ausdrücke

- \emptyset ist ein regulärer Ausdruck.
- ε ist ein regulärer Ausdruck.
- Für jedes $a \in \Sigma$ ist a ein regulärer Ausdruck.
- Wenn α und β reguläre Ausdrücke sind, dann auch $\alpha\beta$, $(\alpha \mid \beta)$ und $(\alpha)^*$.
- Das sind alle regulären Ausdrücke.

$$(1) \quad A? = \varepsilon \mid A$$

$$(2) \quad A+ = A A^*$$

$$(3) \quad A\{0\} = \varepsilon$$

$$(4) \quad A\{n\} = A A\{n-1\} \quad n \in \mathbb{N}; \quad n \geq 1$$

$$(5) \quad A\{n, \} = A\{n\} A^* \quad n \in \mathbb{N}$$

$$(6) \quad A\{, n\} = (A?)\{n\} \quad n \in \mathbb{N}$$

$$(7) \quad A\{m, n\} = A\{m\} A\{, n-m\} \quad m, n \in \mathbb{N}; \quad m \leq n$$

$$(8) \quad [] = \varepsilon$$

$$(9) \quad [a] = a$$

$$(10) \quad [a\varphi] = a \mid [\varphi]$$

$$(11) \quad [^{\wedge}\varphi] = [\psi]$$

$$(12) \quad \cdot = [\omega]$$

$[abc]\{2,3\}$

$[abc]\{2\} [abc]\{,1\}$

$[abc] [abc]\{1\} [abc]? \{1\}$

$[abc] [abc] [abc]\{0\} [abc]? [abc]? \{0\}$

$[abc] [abc] [abc]?$

$[abc] [abc] (\epsilon \mid [abc])$

$(a \mid [bc]) (a \mid [bc]) (\epsilon \mid (a \mid [bc]))$

$(a \mid b \mid [c]) (a \mid b \mid [c]) (\epsilon \mid (a \mid b \mid [c]))$

$(a \mid b \mid c) (a \mid b \mid c) (\epsilon \mid (a \mid b \mid c))$



Noam Chomsky

Photo: Andrew Rusk, CC BY 2.0

Grammatik

$$G = (V, \Sigma, P, S)$$

Variablen

Startsymbol $S \in V$

Symbole (Alphabet)

Produktionen (Ableitungsregeln)

$$V \cap \Sigma = \emptyset$$

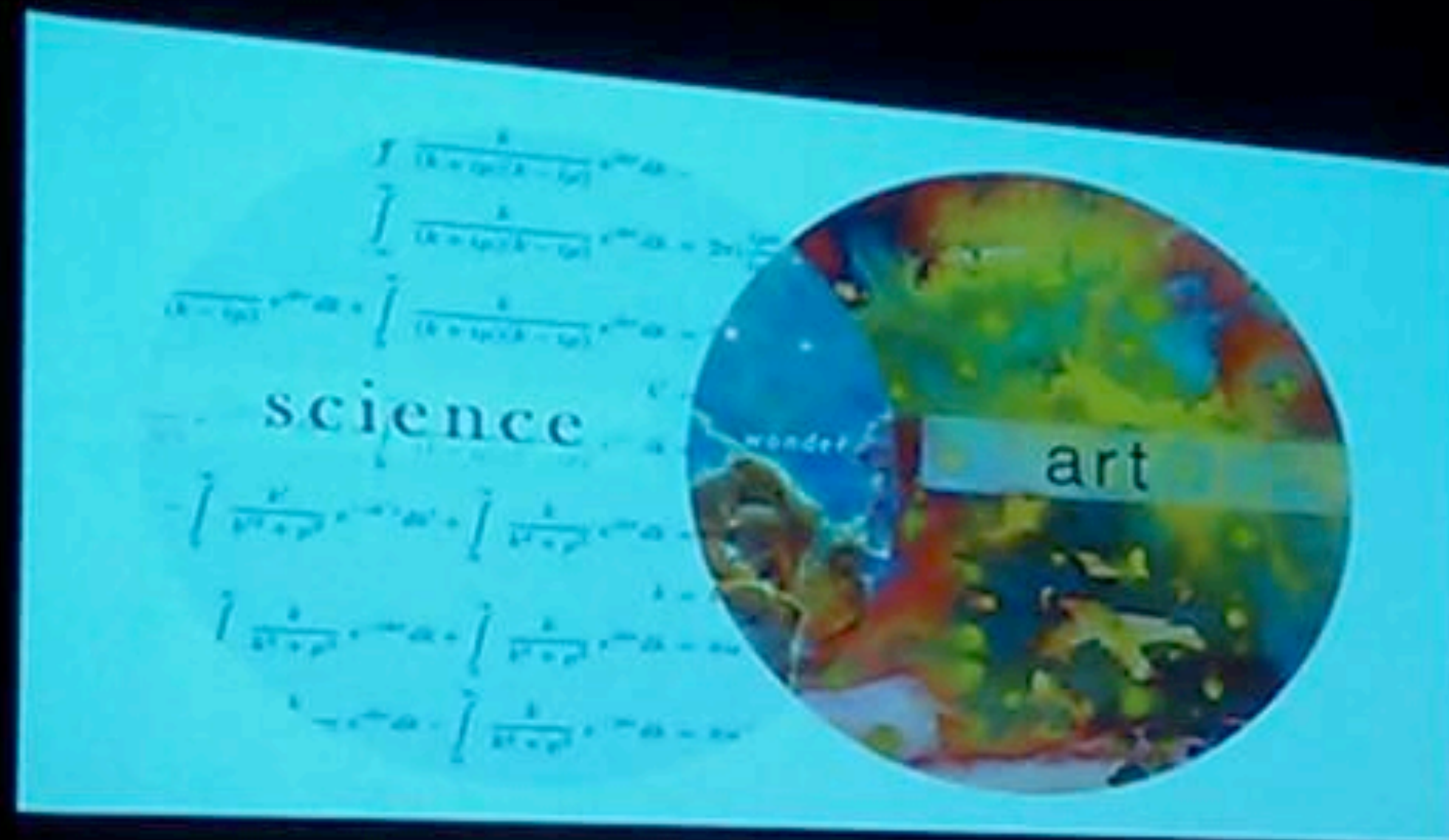
Sprache

$$L(G) = \{ w \in \Sigma^* \mid S \xRightarrow{*}_G w \}$$

Wörter

Chomsky-Hierarchie

Sprache	Automat	Ableitungsregeln	Beispiel
Typ 0: rekursiv aufzählbar	Turingmaschine	$\gamma \rightarrow a$	
Typ 1: kontextsensitiv	linear platzbeschränkte nichtdeterministische Turingmaschine	$aA\beta \rightarrow a\gamma\beta$	$\{a^n b^n c^n, n > 0\}$
Typ 2: kontextfrei	nichtdeterministischer Kellerautomat	$A \rightarrow a$	$\{a^n b^n, n > 0\}$
Typ 3: regulär	endlicher Automat	$A \rightarrow a$ $A \rightarrow aB$	$\{a^n, n \geq 0\}$



science \cap art = wonder

– Brian Suda

Chomsky-Hierarchie

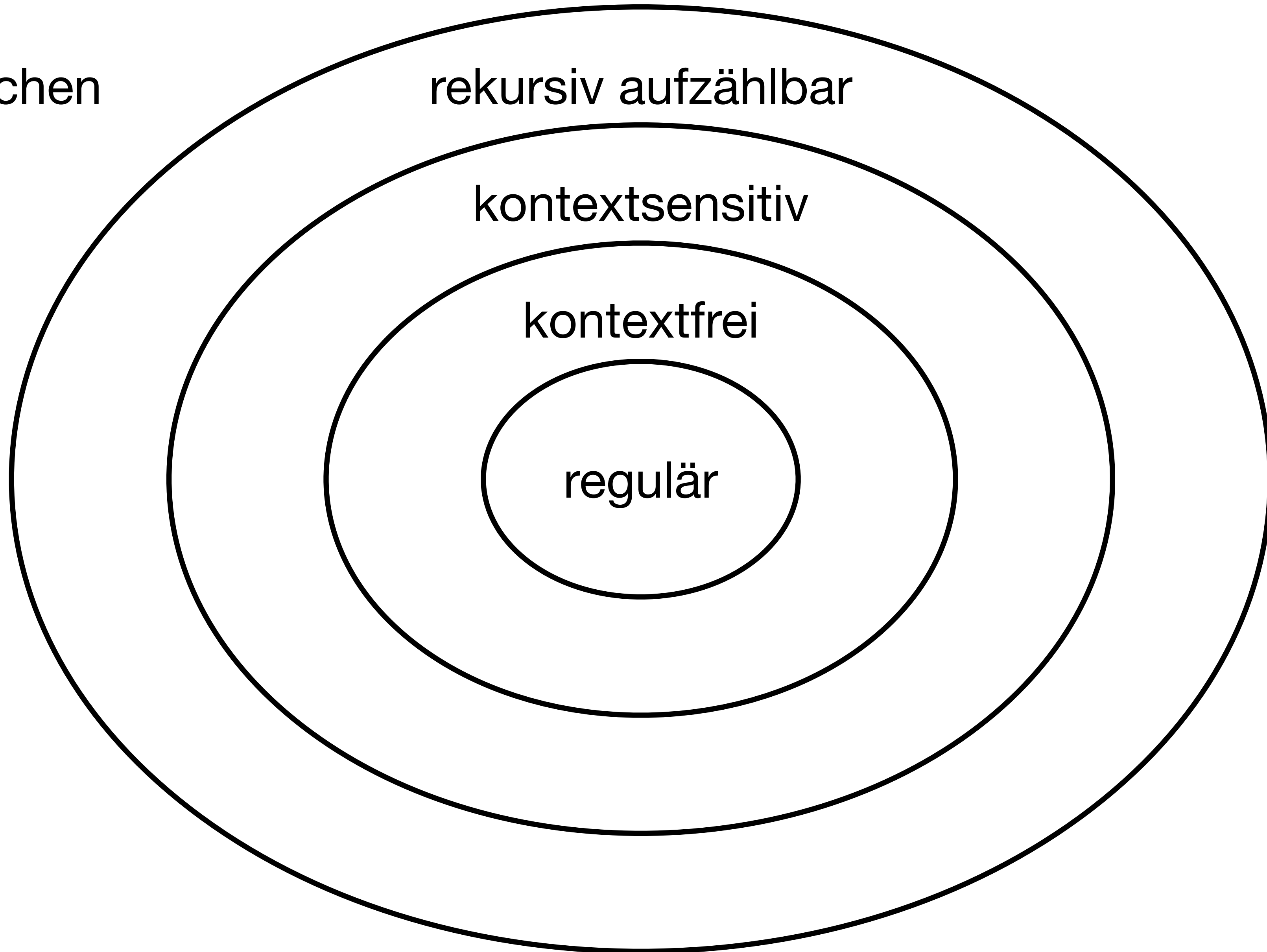
alle Sprachen

rekursiv aufzählbar

kontextsensitiv

kontextfrei

regulär



Chomsky-Hierarchie

Sprache	Automat	Ableitungsregeln	Beispiel
Typ 0: rekursiv aufzählbar	Turingmaschine	$\gamma \rightarrow a$	
Typ 1: kontextsensitiv	linear platzbeschränkte nichtdeterministische Turingmaschine	$aA\beta \rightarrow a\gamma\beta$	$\{a^n b^n c^n, n > 0\}$
Typ 2: kontextfrei	RegExp nichtdeterministischer Kellerautomat	$A \rightarrow a$	$\{a^n b^n, n > 0\}$
Typ 3: regulär	endlicher Automat regulärer Ausdruck	$A \rightarrow a$ $A \rightarrow aB$	$\{a^n, n \geq 0\}$

-?\d*(?:(?:?:?:?:[02468][048]|[13579][26])?:
[02468][048]|[13579][26])|(?:[02468][1235679]|
[13579][01345789])(?:0[48]|[2468][048]|[13579]
[26]))-02-(?:0[1-9]|[12]\d)|(?:(?:[02468]
[1235679]|[13579][01345789])00|\d\d(?:[02468]
[1235679]|[13579][01345789]))-02-(?:0[1-9]|
1\d|2[0-8])|\d{4}-(?:(?:0[13578]|1[02])-
(?:0[1-9]|[12]\d|3[01])|(?:0[469]|11)-
(?:0[1-9]|[12]\d|30)))

```

-?\d* # für die Ewigkeit
(?:
  (?:
    (?:
      (?: [02468][048] | [13579][26]) # Jahrhundert durch 4 teilbar
      (?: [02468][048] | [13579][26]) # Jahr durch 4 teilbar
      |
      (?: [02468][1235679] | [13579][01345789]) # Jahrhundert nicht durch 4 teilbar
      (?: 0[48] | [2468][048] | [13579][26]) # Jahr durch 4 teilbar
      # außer volle Jahrhunderte
    )
  )
  - # Trennzeichen
  02 # Februar im Schaltjahr
  - # Trennzeichen
  (?: 0[1-9] | [12]\d) # 01 bis 29

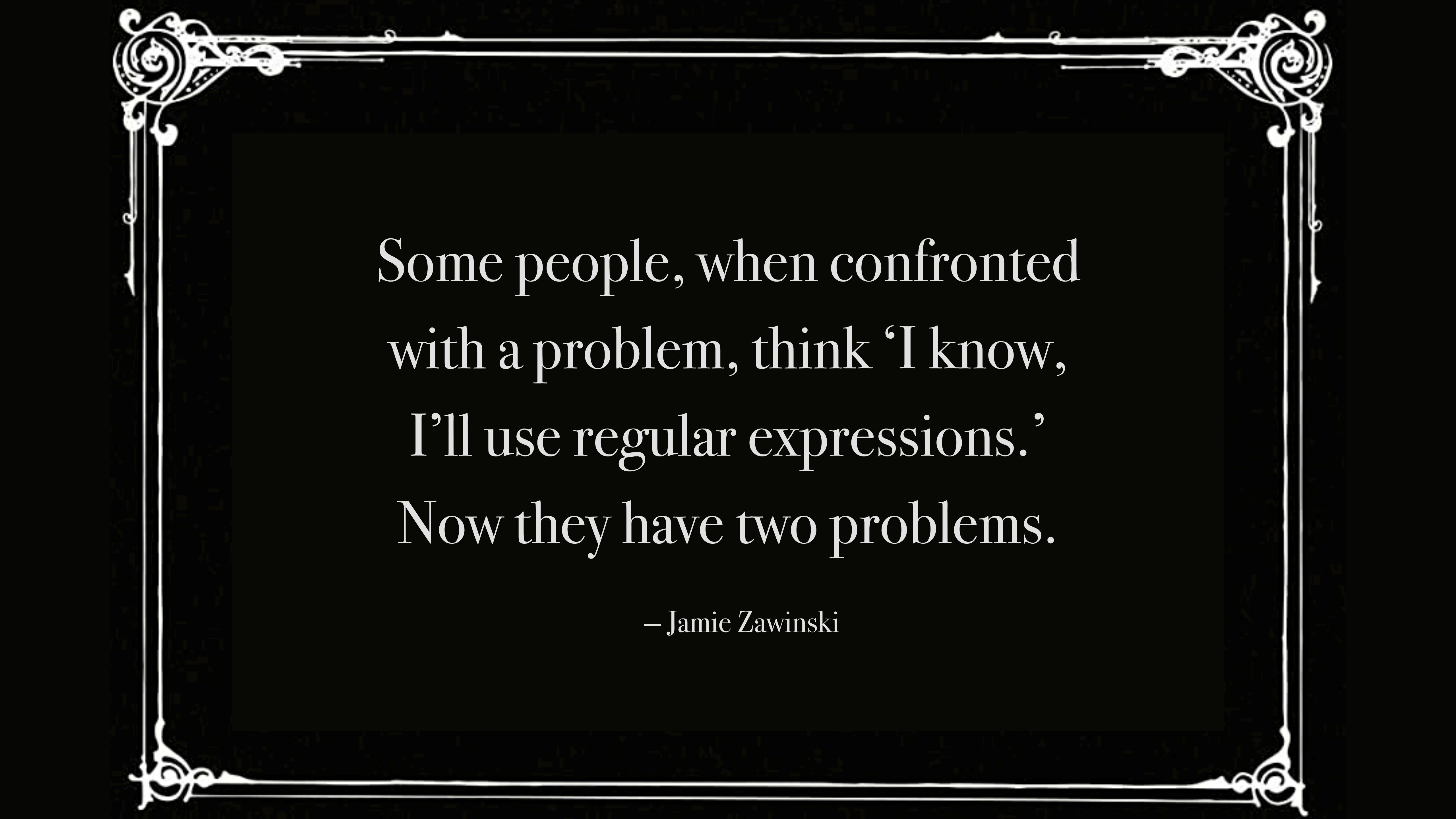
```

```
|
(?:
  (?: [02468][1235679] | [13579][01345789]) # Jahrhundert nicht durch 4 teilbar
  00 # selbsterklärend
  |
  \d\d # beliebiges Jahrhundert
  (?: [02468][1235679] | [13579][01345789]) # Jahr nicht durch 4 teilbar
)
- # Trennzeichen
02 # Februar im Nicht-Schaltjahr
- # Trennzeichen
(?: 0[1-9] | 1\d | 2[0-8]) # 01 bis 28
)
```

```
|
\d{4} # beliebiges Jahr
- # Trennzeichen
(?:
  (?: 0[13578] | 1[02]) # langer Monat
  - # Trennzeichen
  (?: 0[1-9] | [12]\d | 3[01]) # 01 bis 31
  |
  (?: 0[469] | 11) # kurzer Monat
  - # Trennzeichen
  (?: 0[1-9] | [12]\d | 30) # 01 bis 30
)
)
```



```
[./-] # Trennzeichen
[0-9]{4} # Jahr
|
29[./-]02[./-] # 29.02.
( ?:
    [0-9]{2} ( ?: 0[48] | [2468][048] | [13579][26] ) # beliebiges Jahrhundert,
                                                         # durch 4 teilbares Jahr
                                                         # außer volle Jahrhunderte
    |
    ( ?: [02468][048] | [13579][26] ) 00 # durch 4 teilbares Jahrhundert
)
```



Some people, when confronted
with a problem, think ‘I know,
I’ll use regular expressions.’
Now they have two problems.

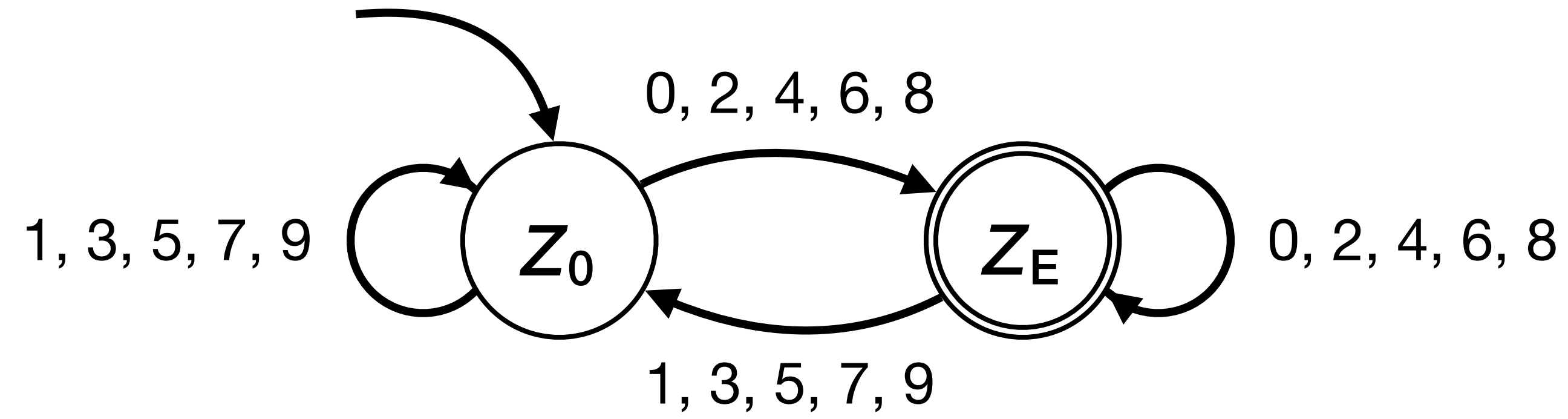
– Jamie Zawinski

Endlicher Automat

$$M = (Z, \Sigma, \delta, z_0, E)$$



Endlicher Automat



Natürliche Zahlen

Grammatik

$V = \{S\}$

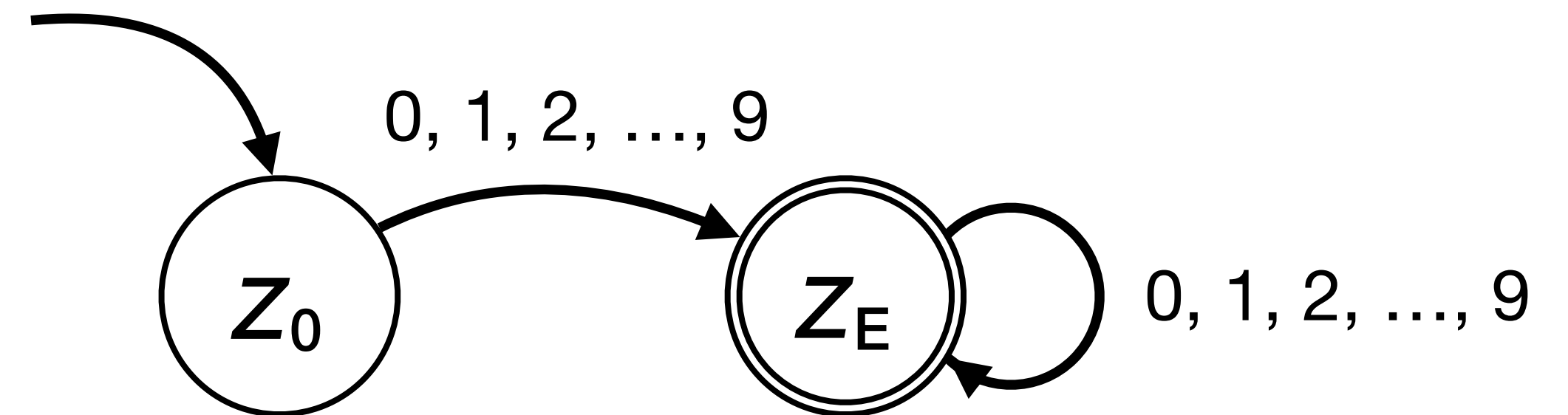
$\Sigma = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$

$S \rightarrow 0$	$S \rightarrow 0S$
$S \rightarrow 1$	$S \rightarrow 1S$
$S \rightarrow 2$	$S \rightarrow 2S$
$S \rightarrow 3$	$S \rightarrow 3S$
$S \rightarrow 4$	$S \rightarrow 4S$
$S \rightarrow 5$	$S \rightarrow 5S$
$S \rightarrow 6$	$S \rightarrow 6S$
$S \rightarrow 7$	$S \rightarrow 7S$
$S \rightarrow 8$	$S \rightarrow 8S$
$S \rightarrow 9$	$S \rightarrow 9S$

regulärer Ausdruck

$[0-9]^+$

endlicher Automat



Gerade Zahlen

Grammatik

$V = \{S\}$

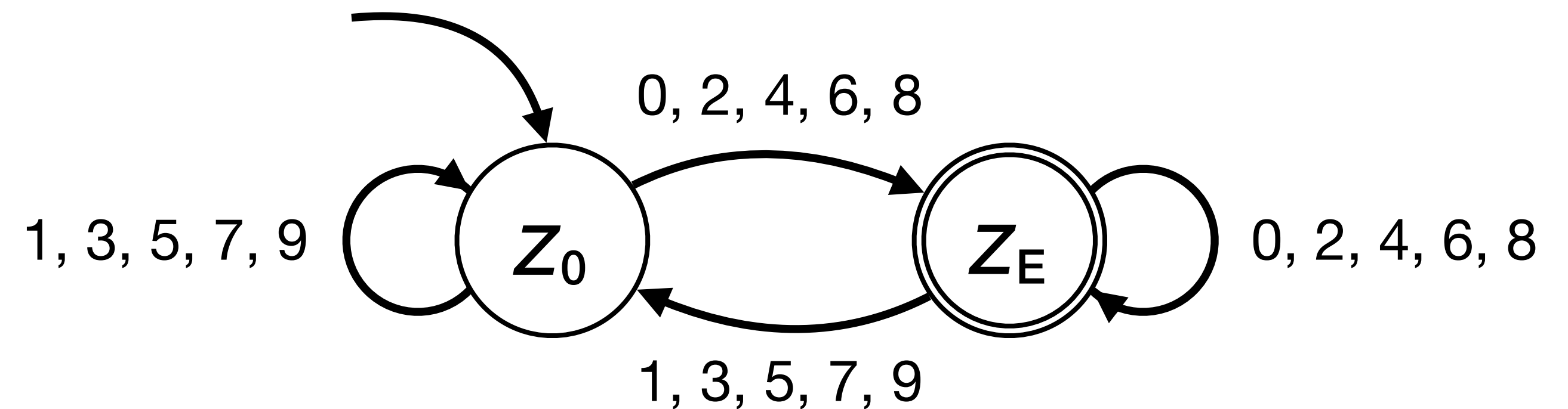
$\Sigma = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$

$S \rightarrow 0$	$S \rightarrow 0S$
	$S \rightarrow 1S$
$S \rightarrow 2$	$S \rightarrow 2S$
	$S \rightarrow 3S$
$S \rightarrow 4$	$S \rightarrow 4S$
	$S \rightarrow 5S$
$S \rightarrow 6$	$S \rightarrow 6S$
	$S \rightarrow 7S$
$S \rightarrow 8$	$S \rightarrow 8S$
	$S \rightarrow 9S$

regulärer Ausdruck

$[0-9]^*[02468]$

endlicher Automat



Teilbarkeit durch 5

Grammatik

$V = \{S\}$

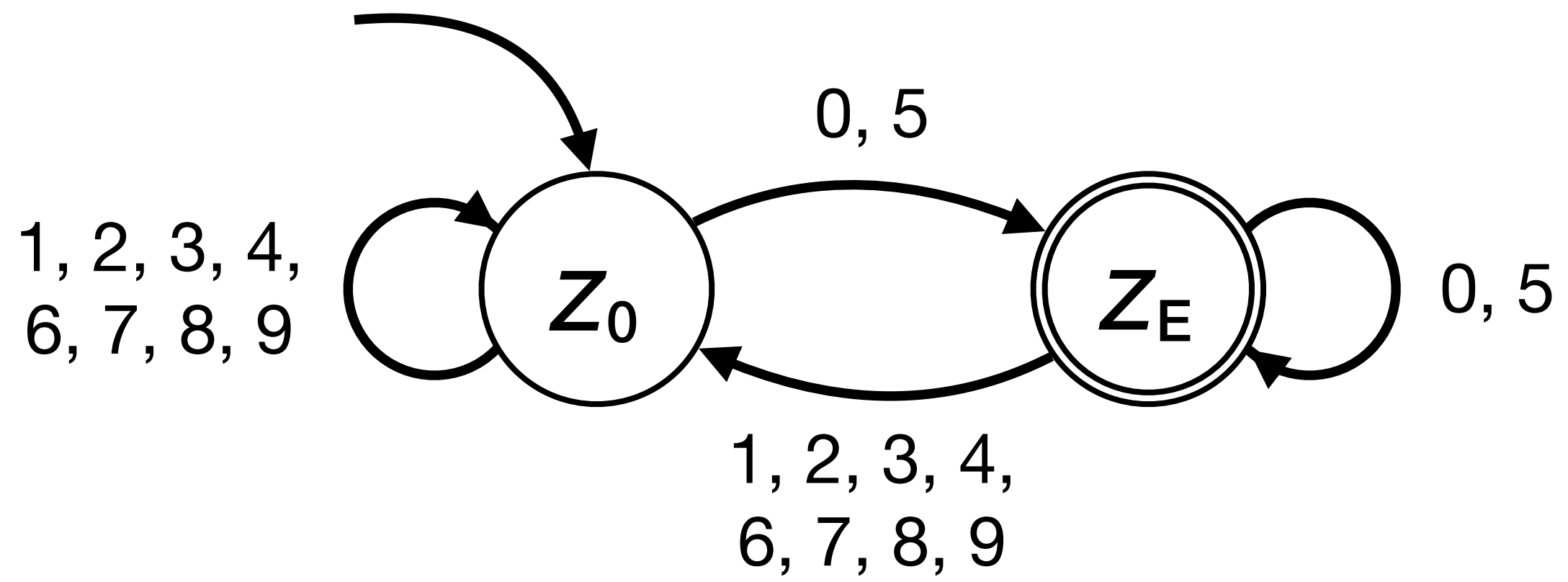
$\Sigma = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$

$S \rightarrow 0$	$S \rightarrow 0S$
	$S \rightarrow 1S$
	$S \rightarrow 2S$
	$S \rightarrow 3S$
	$S \rightarrow 4S$
$S \rightarrow 5$	$S \rightarrow 5S$
	$S \rightarrow 6S$
	$S \rightarrow 7S$
	$S \rightarrow 8S$
	$S \rightarrow 9S$

regulärer Ausdruck

$[0-9]^*[05]$

endlicher Automat



Teilbarkeit durch 10

Grammatik

$V = \{S\}$

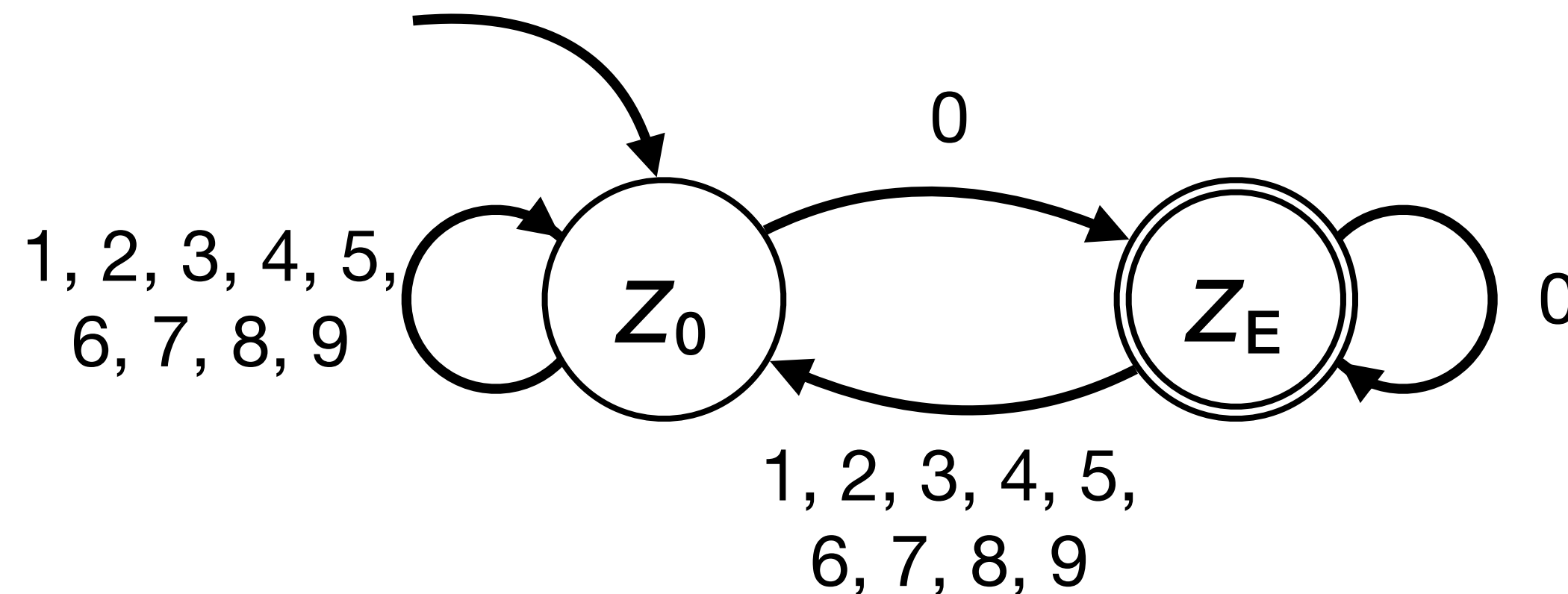
$\Sigma = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$

$S \rightarrow 0$ $S \rightarrow 0S$
 $S \rightarrow 1S$
 $S \rightarrow 2S$
 $S \rightarrow 3S$
 $S \rightarrow 4S$
 $S \rightarrow 5S$
 $S \rightarrow 6S$
 $S \rightarrow 7S$
 $S \rightarrow 8S$
 $S \rightarrow 9S$

regulärer Ausdruck

$[0-9]^*0$

endlicher Automat



Teilbarkeit durch 4

Grammatik

$V = \{S, A\}$

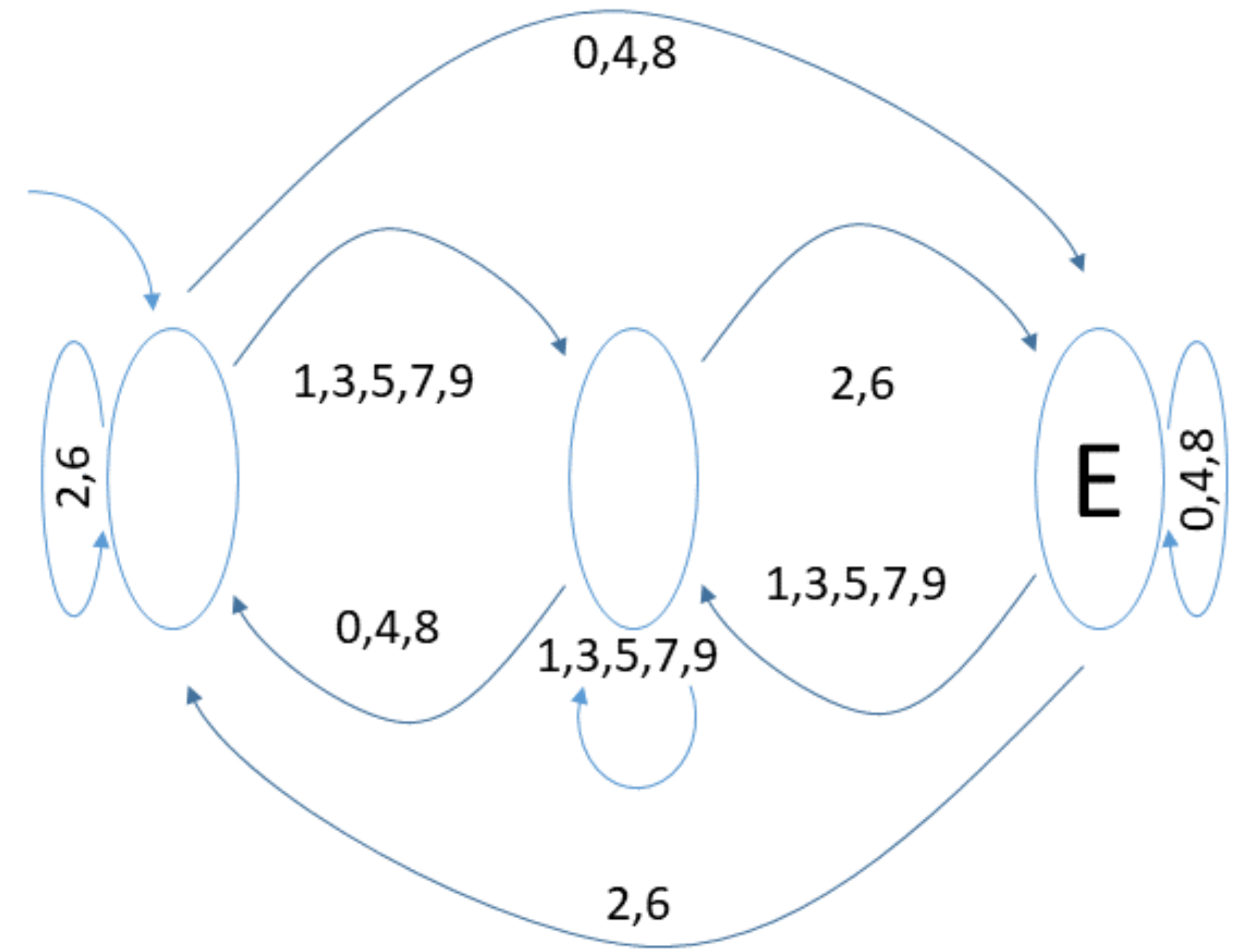
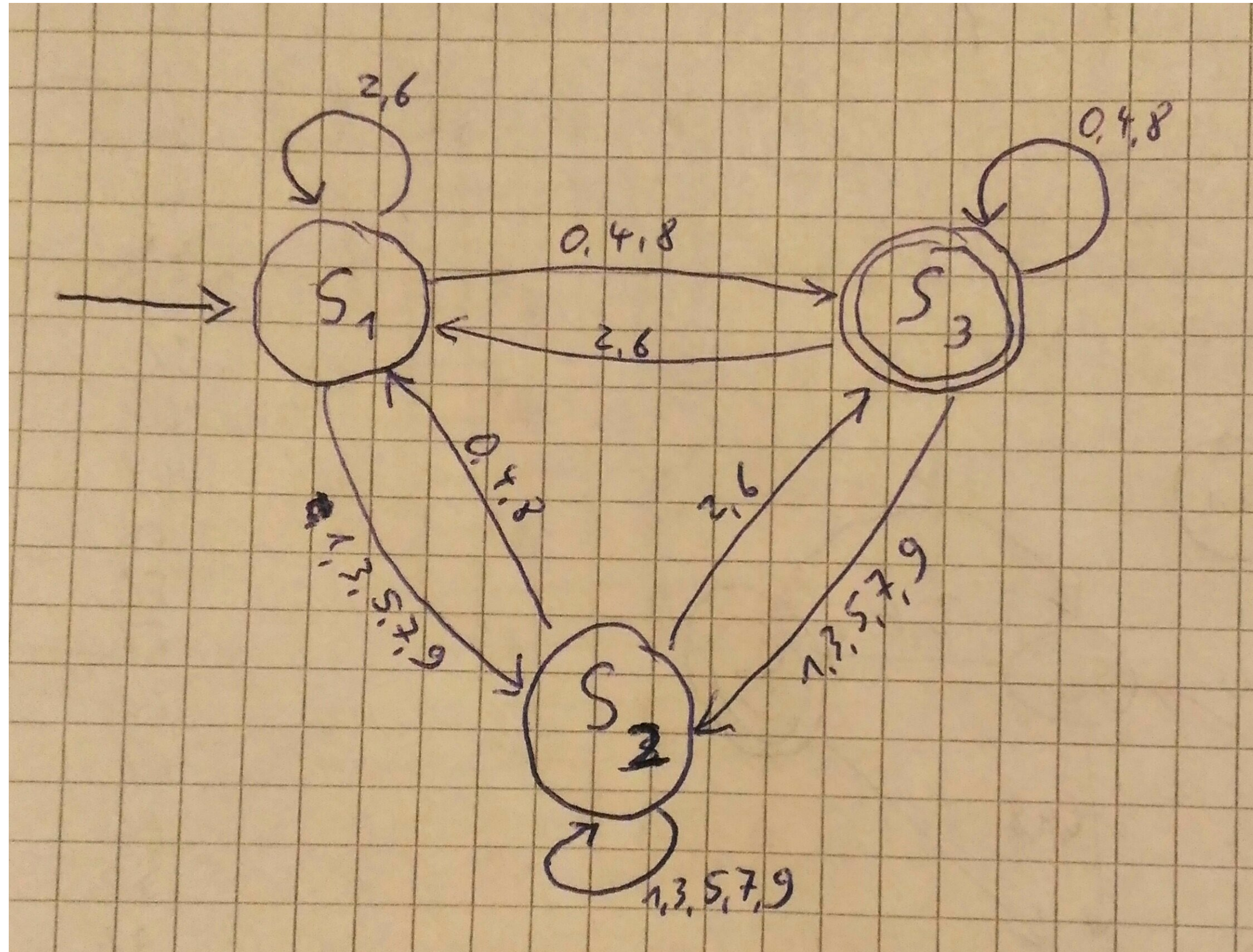
$\Sigma = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$

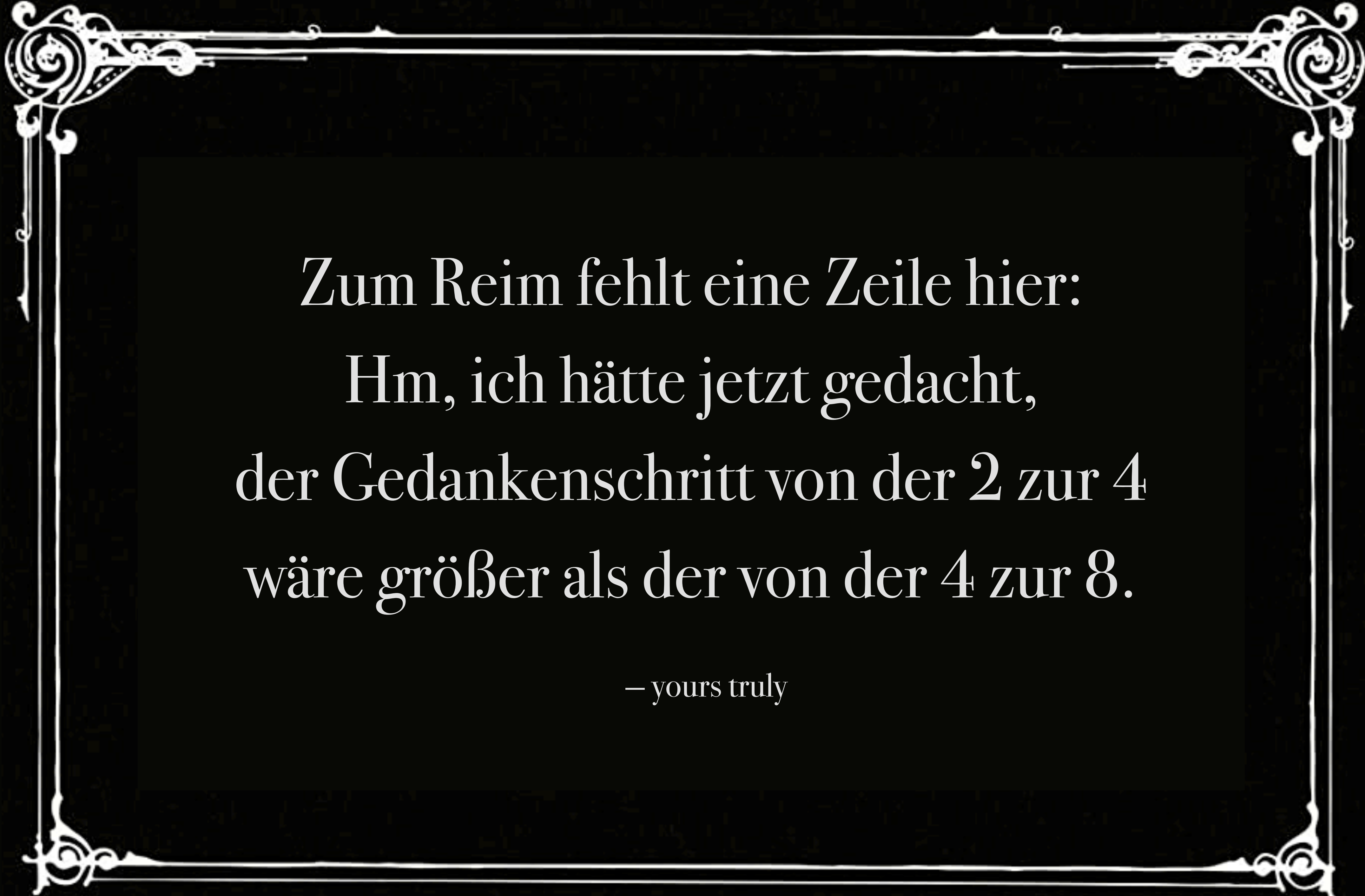
$S \rightarrow 0$	$S \rightarrow 0S$	$A \rightarrow 0S$
	$S \rightarrow 1A$	$A \rightarrow 1A$
$A \rightarrow 2$	$S \rightarrow 2S$	$A \rightarrow 2S$
	$S \rightarrow 3A$	$A \rightarrow 3A$
$S \rightarrow 4$	$S \rightarrow 4S$	$A \rightarrow 4S$
	$S \rightarrow 5A$	$A \rightarrow 5A$
$A \rightarrow 6$	$S \rightarrow 6S$	$A \rightarrow 6S$
	$S \rightarrow 7A$	$A \rightarrow 7A$
$S \rightarrow 8$	$S \rightarrow 8S$	$A \rightarrow 8S$
	$S \rightarrow 9A$	$A \rightarrow 9A$

regulärer Ausdruck

$([0-9]^*[02468])?[048]$
 $| [0-9]^*[13579][26]$

Teilbarkeit durch 4

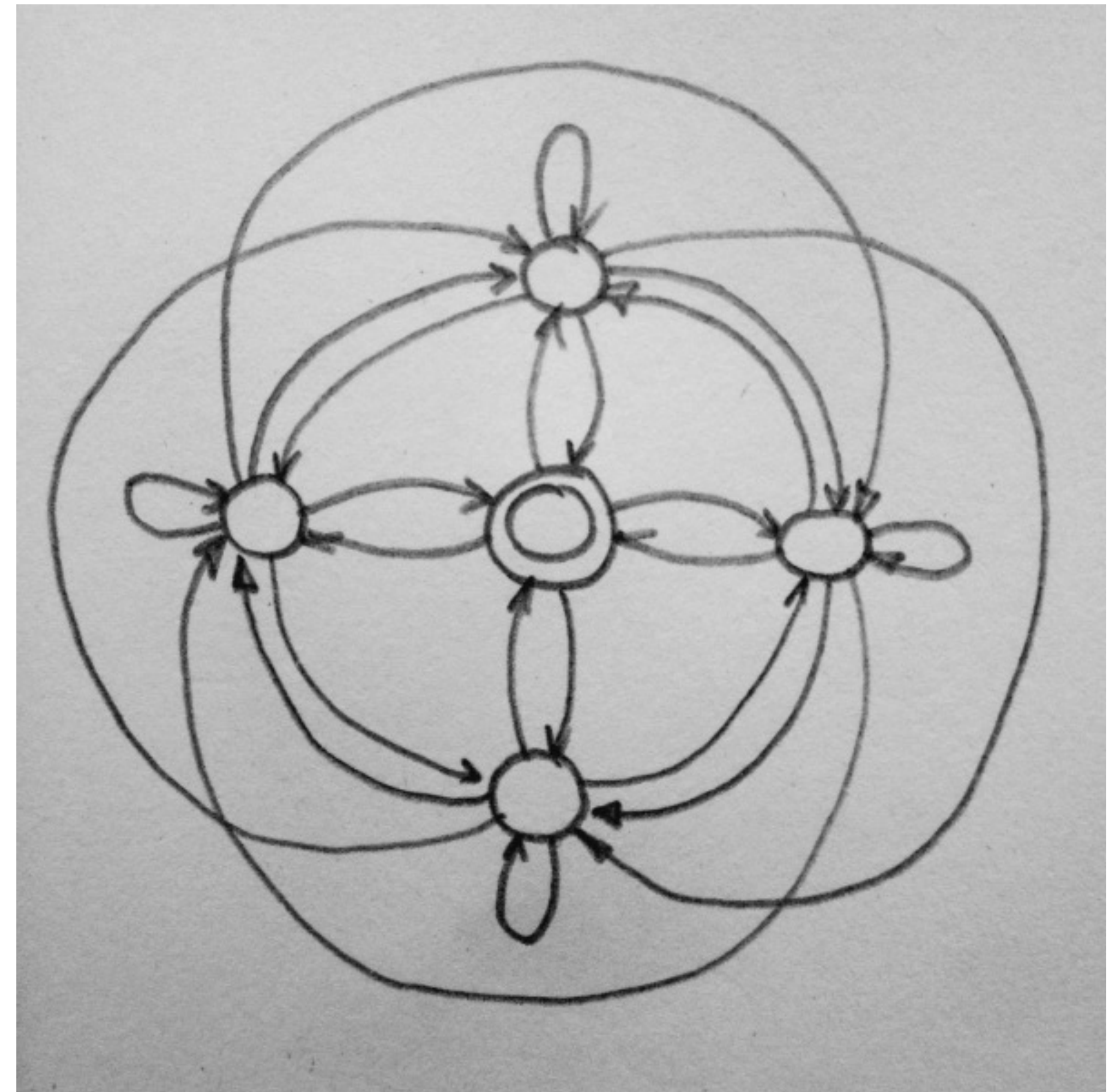
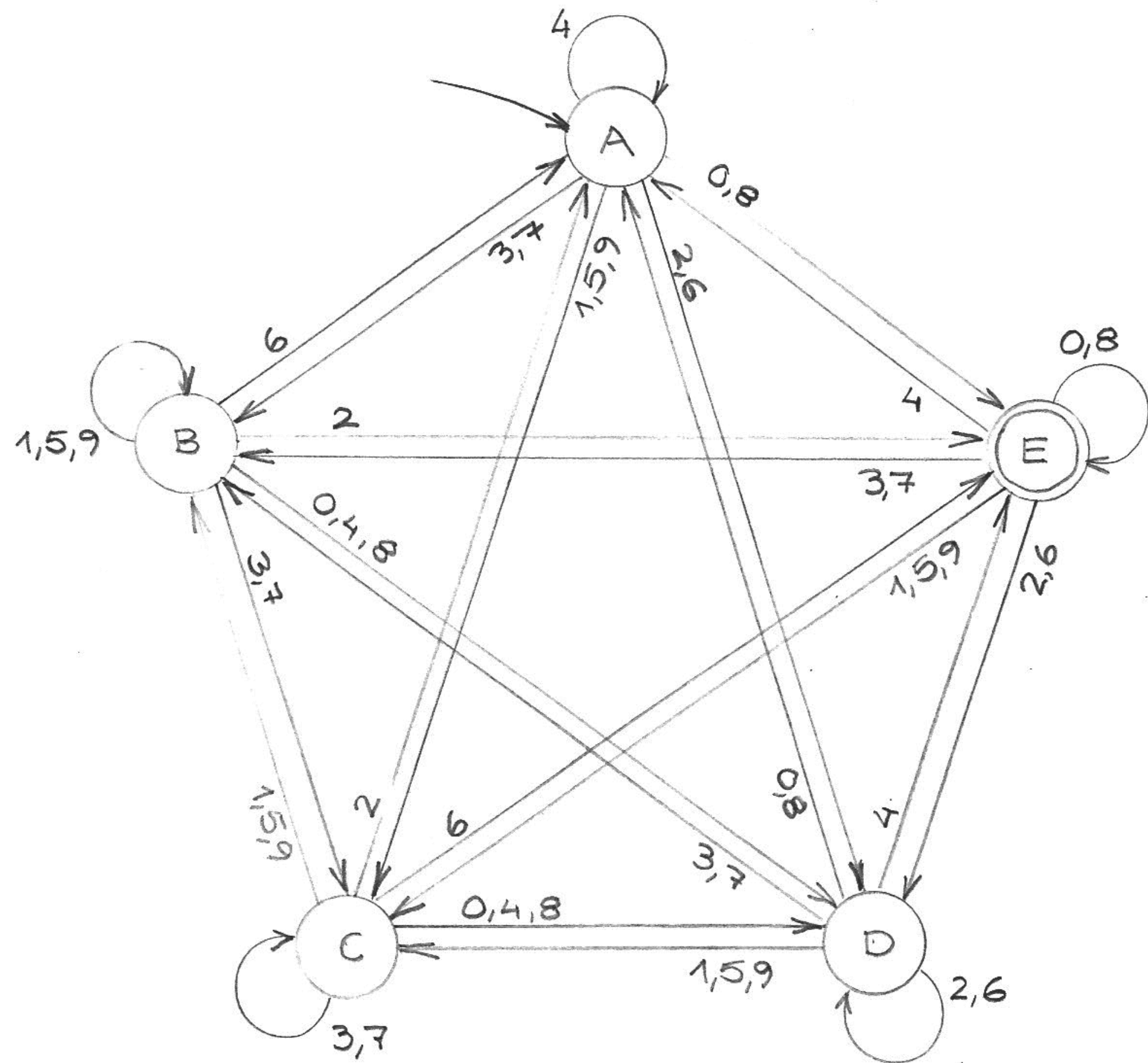


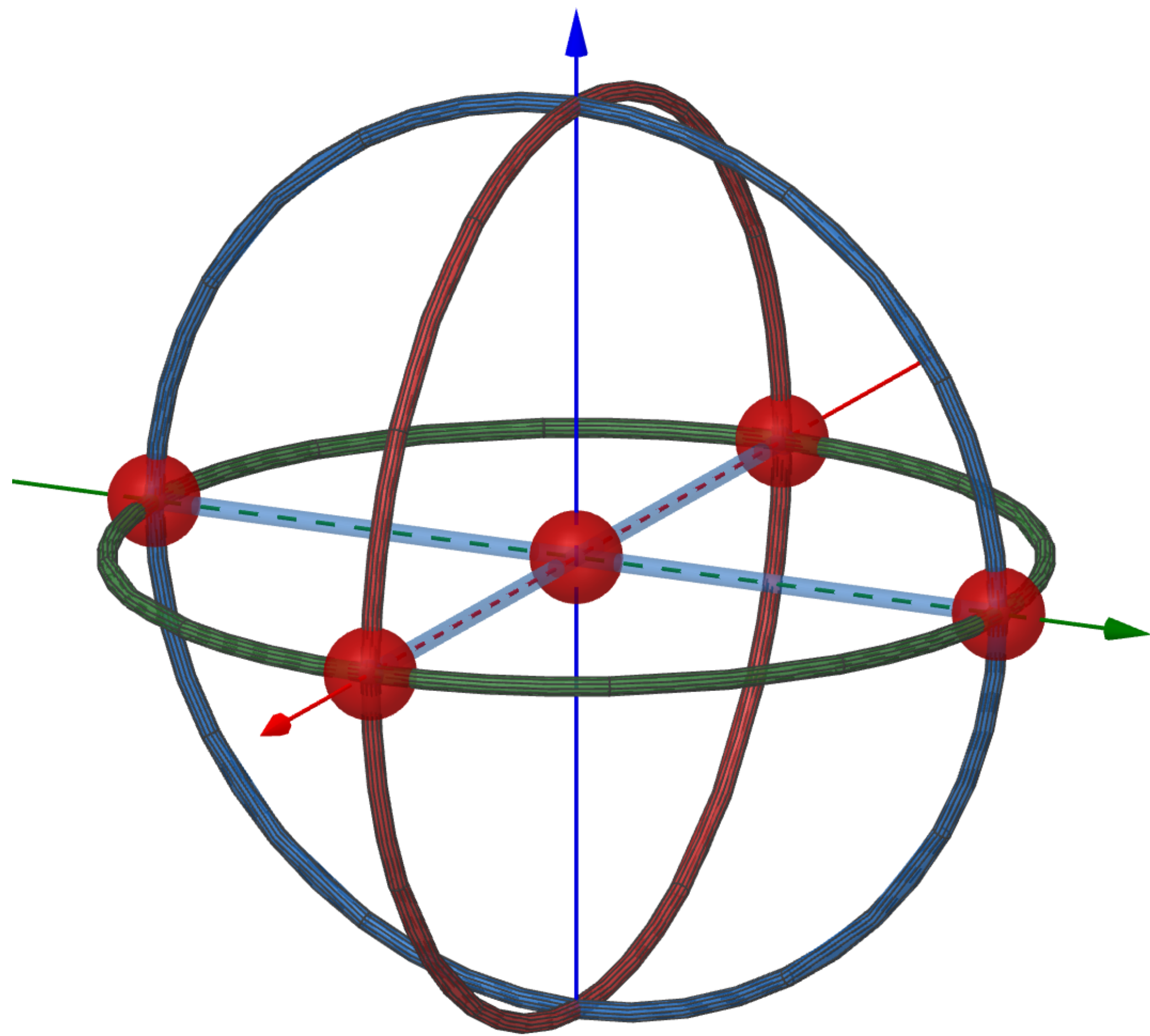
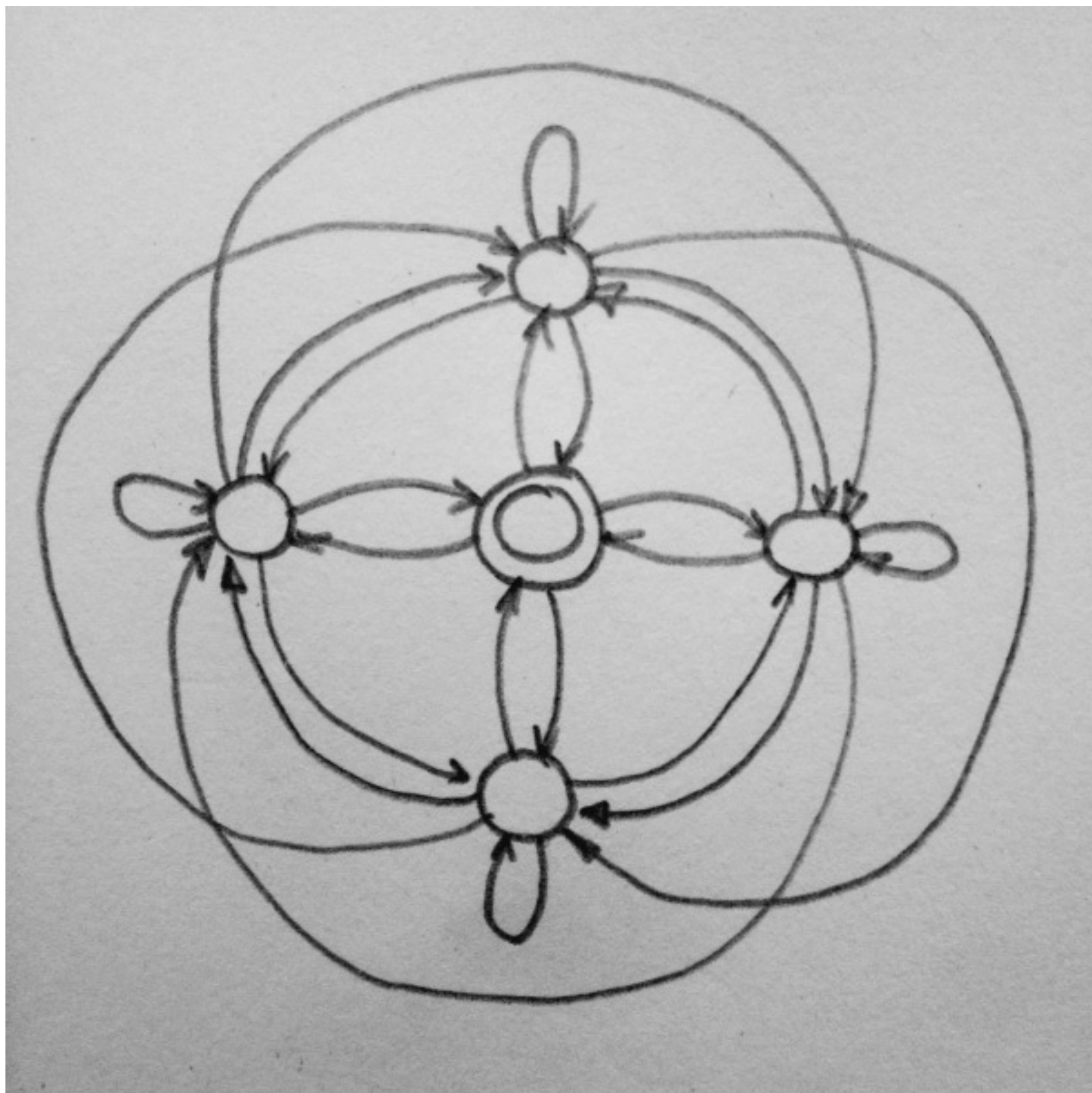


Zum Reim fehlt eine Zeile hier:
Hm, ich hätte jetzt gedacht,
der Gedankenschritt von der 2 zur 4
wäre größer als der von der 4 zur 8.

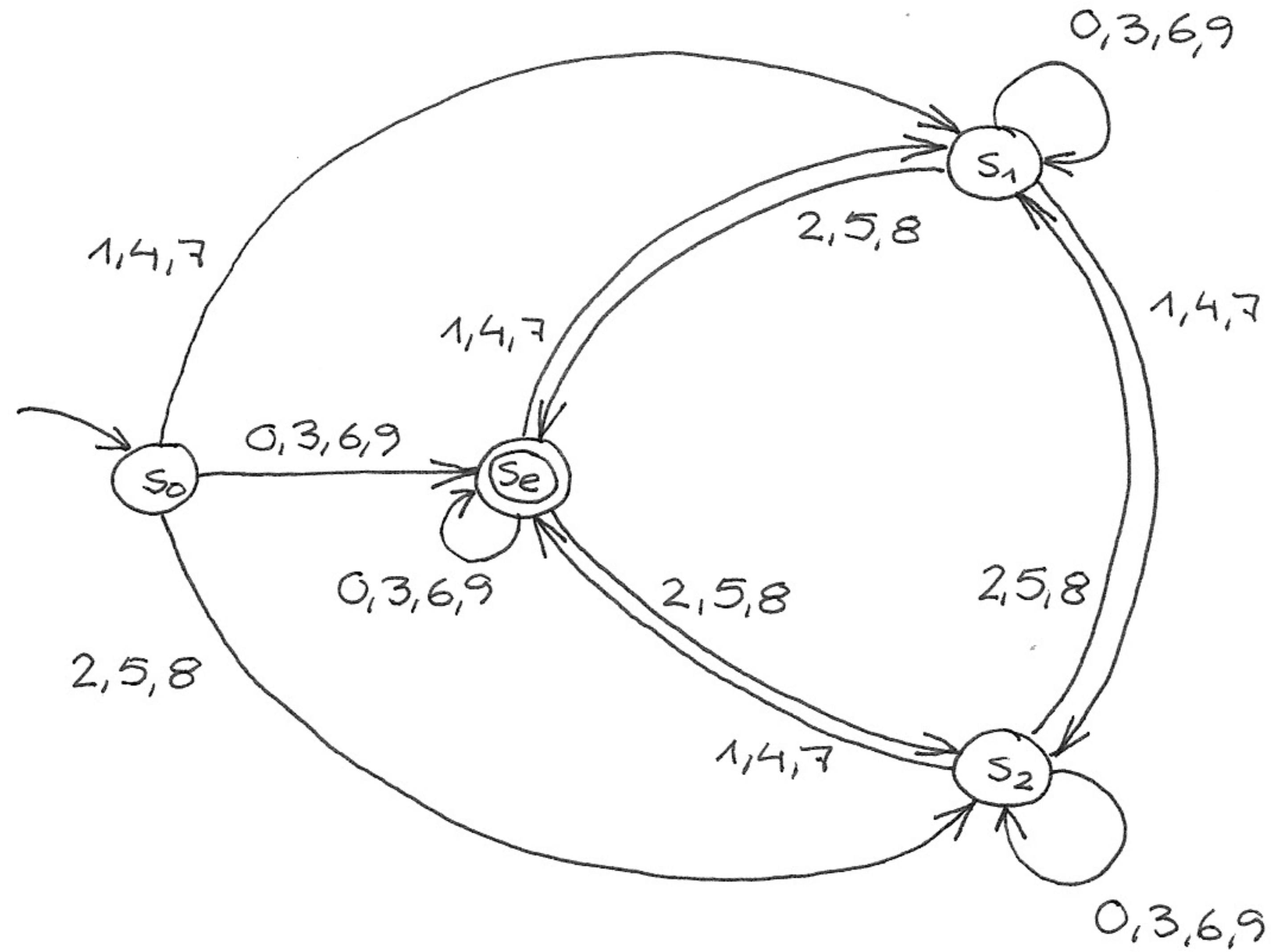
– yours truly

Teilbarkeit durch 8

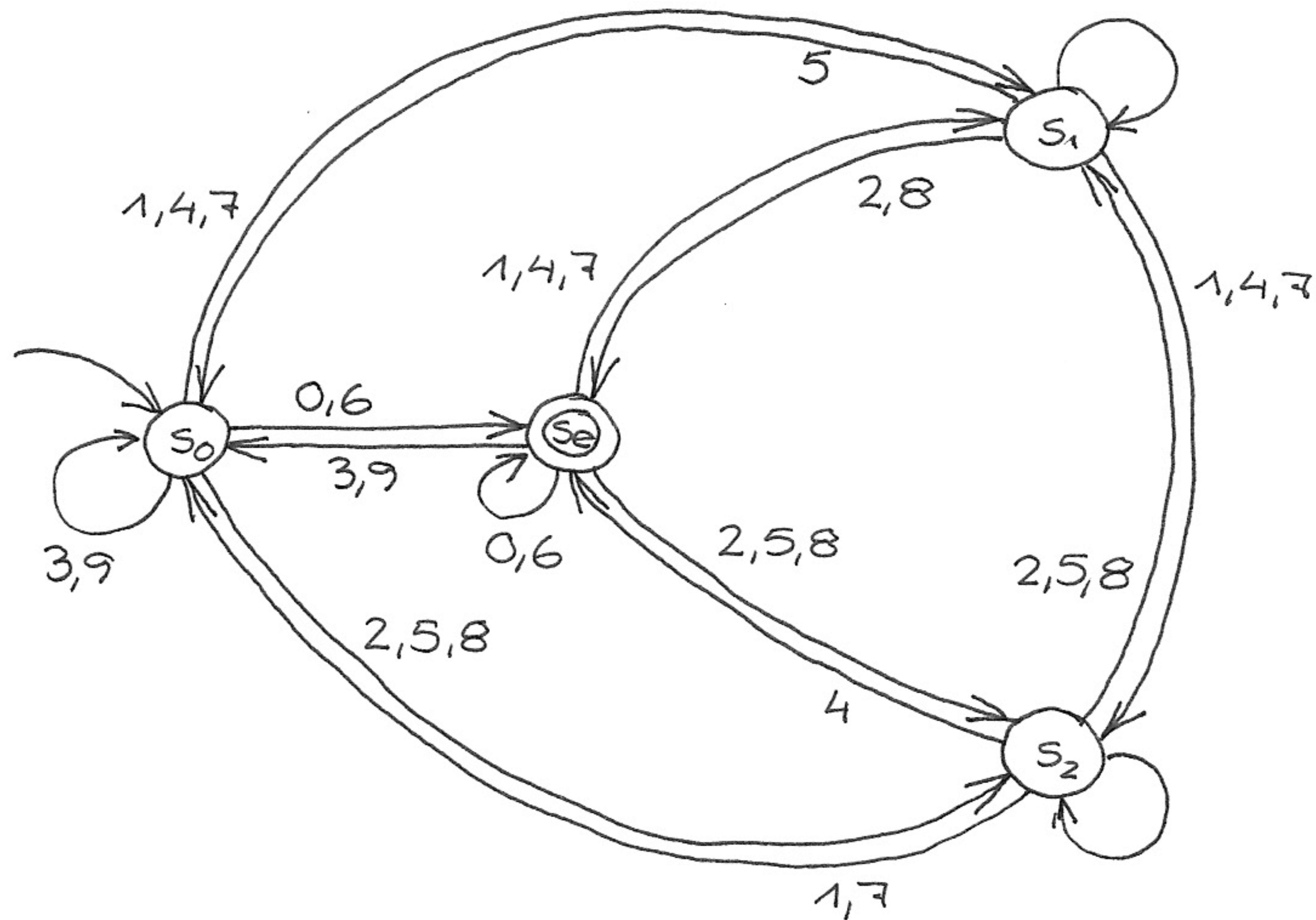




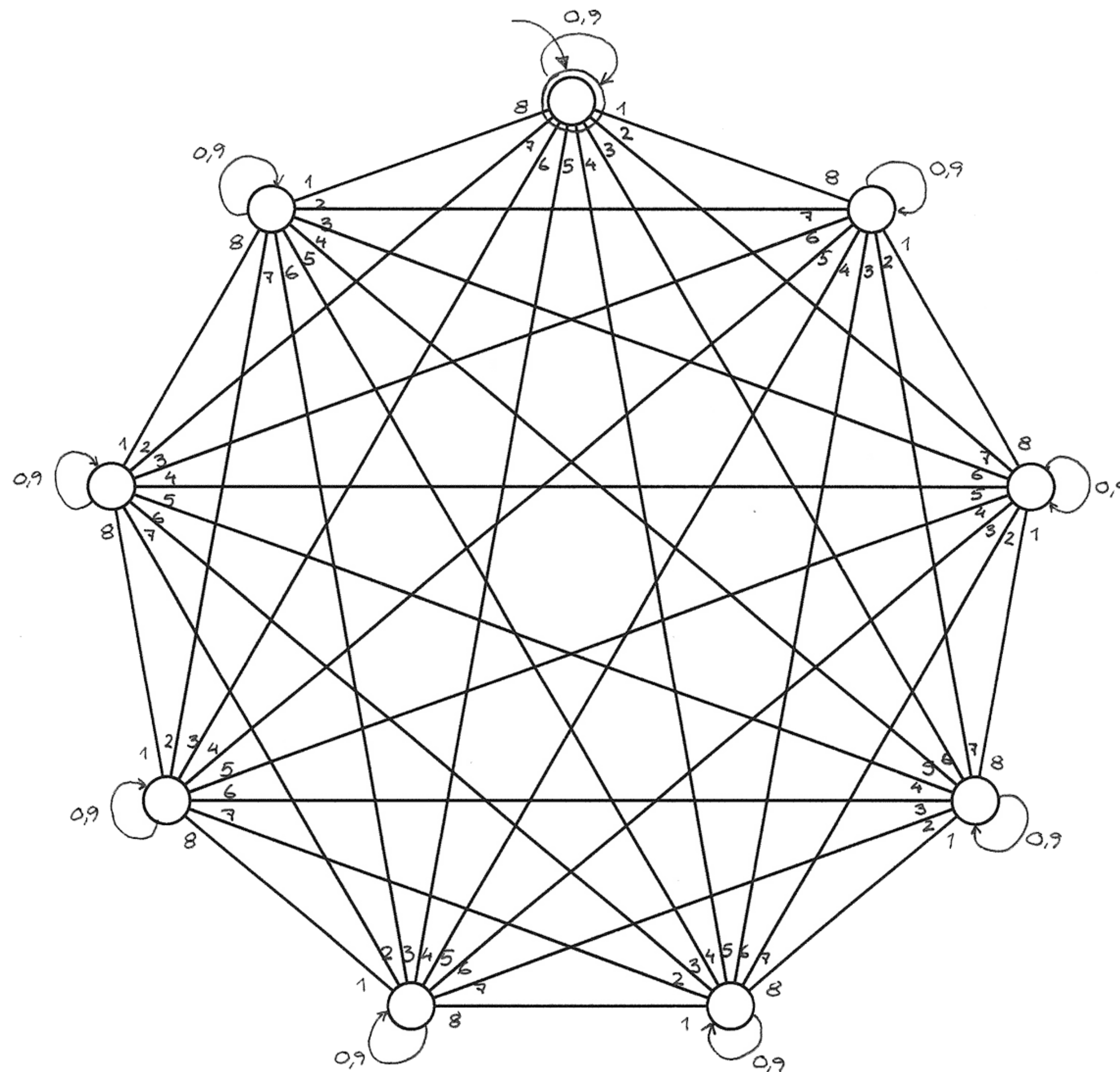
Teilbarkeit durch 3



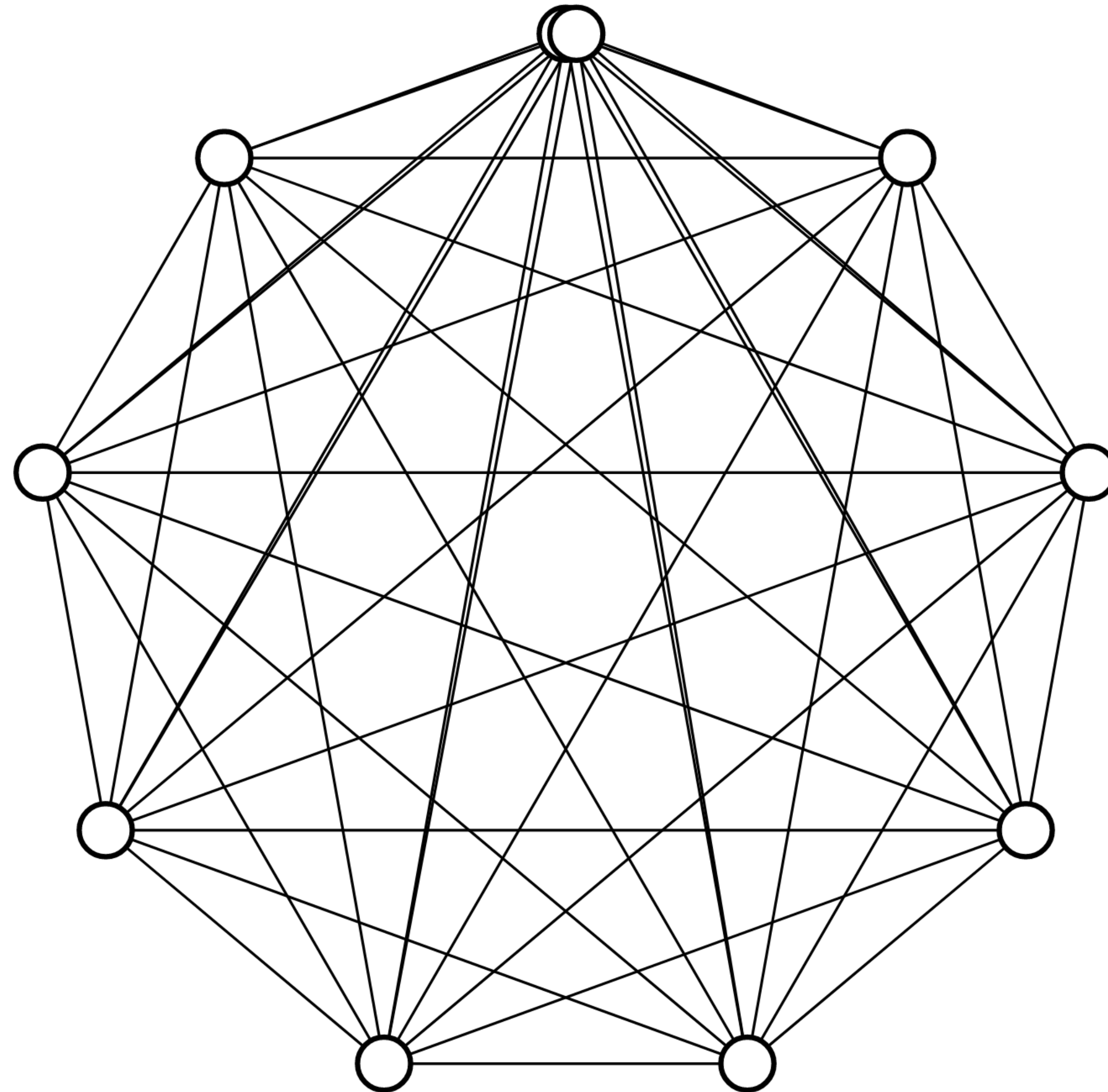
Teilbarkeit durch 6



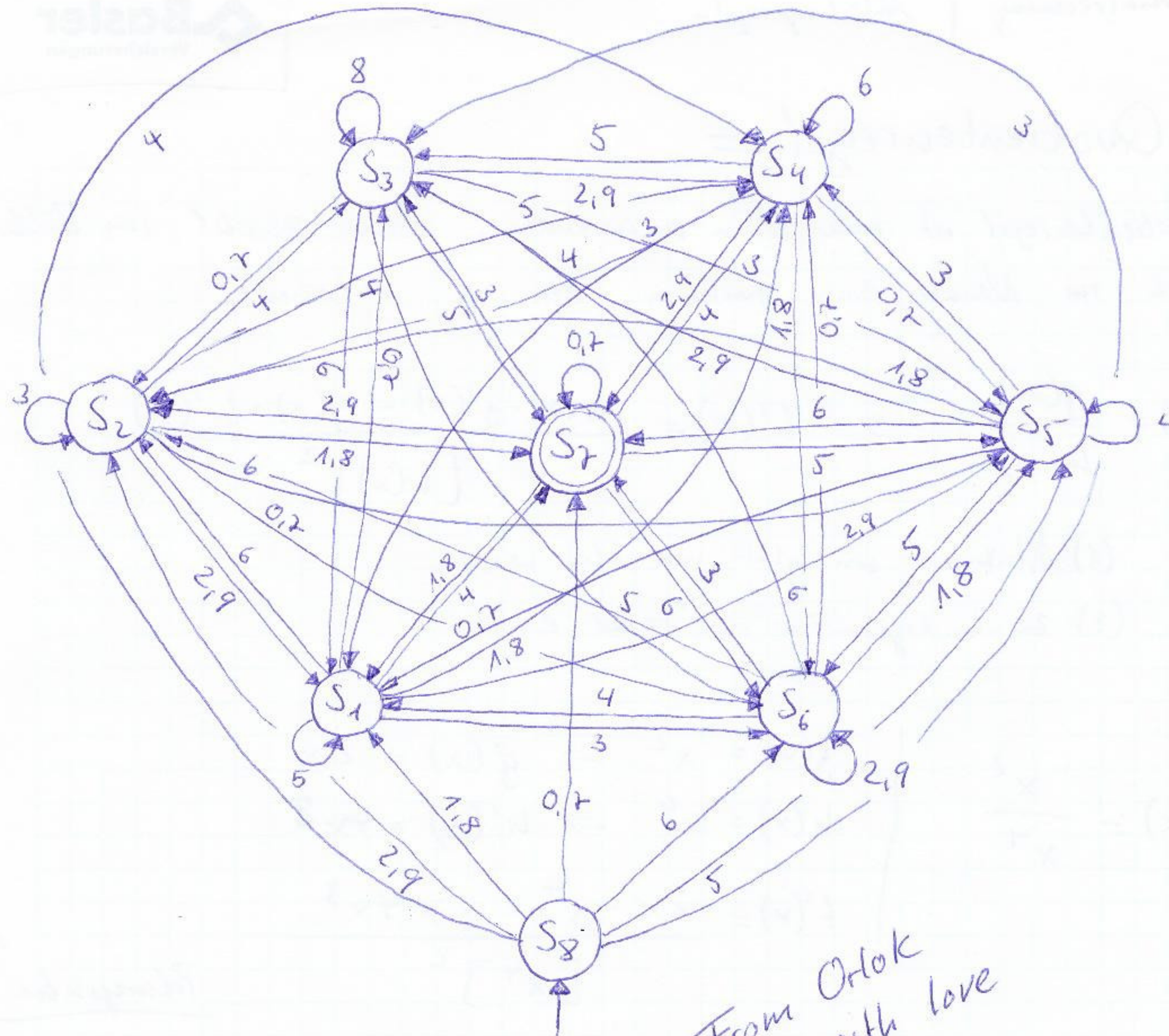
Teilbarkeit durch 9



Teilbarkeit durch 9



Teilbarkeit durch 7



Teilbarkeit durch beliebige Zahl n

$$M = (Z, \Sigma, \delta, z_{-0}, E)$$

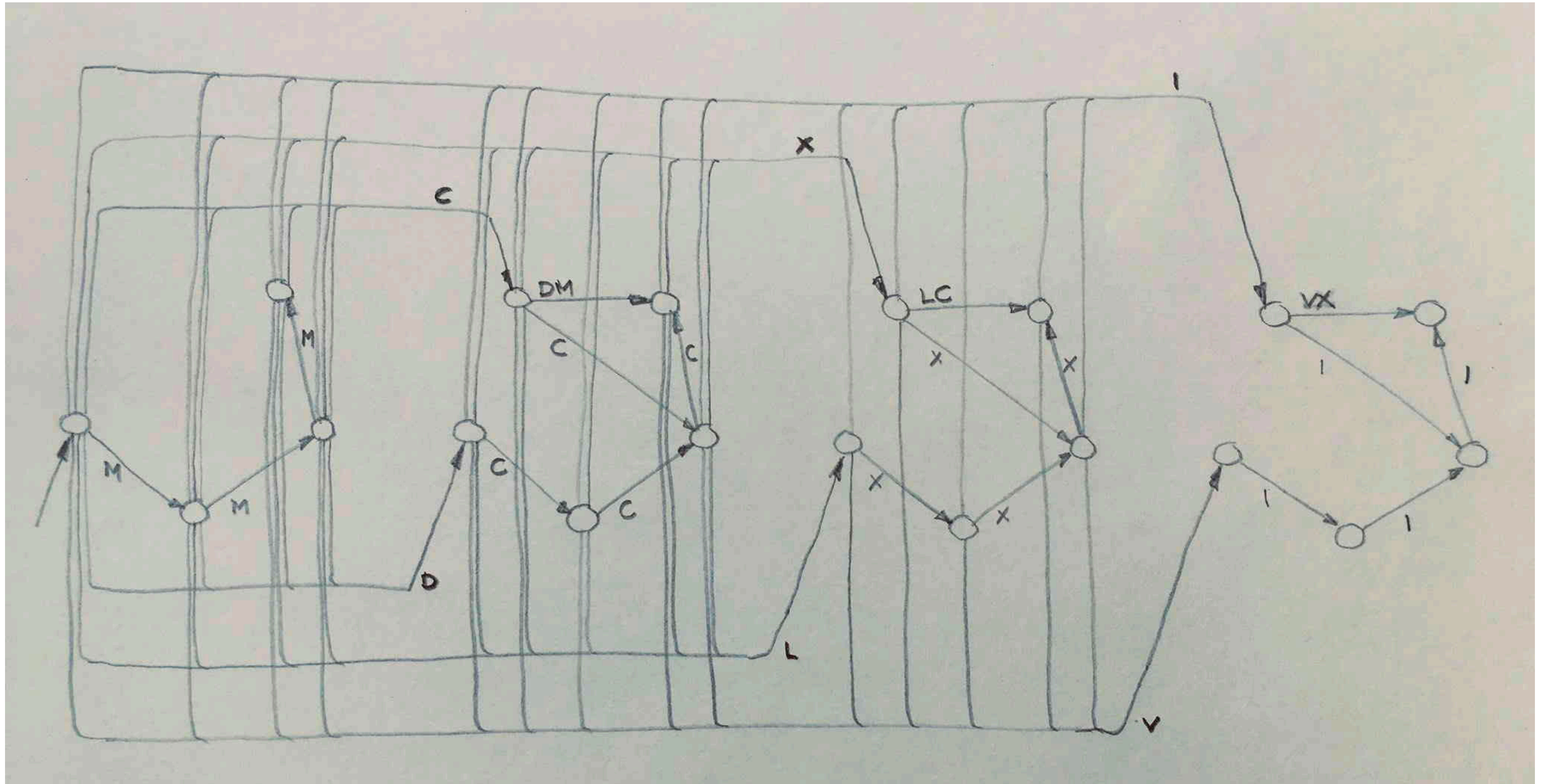
$$Z = \{z_{-0}, z_0, z_1, z_2, \dots, z_{n-1}\}$$

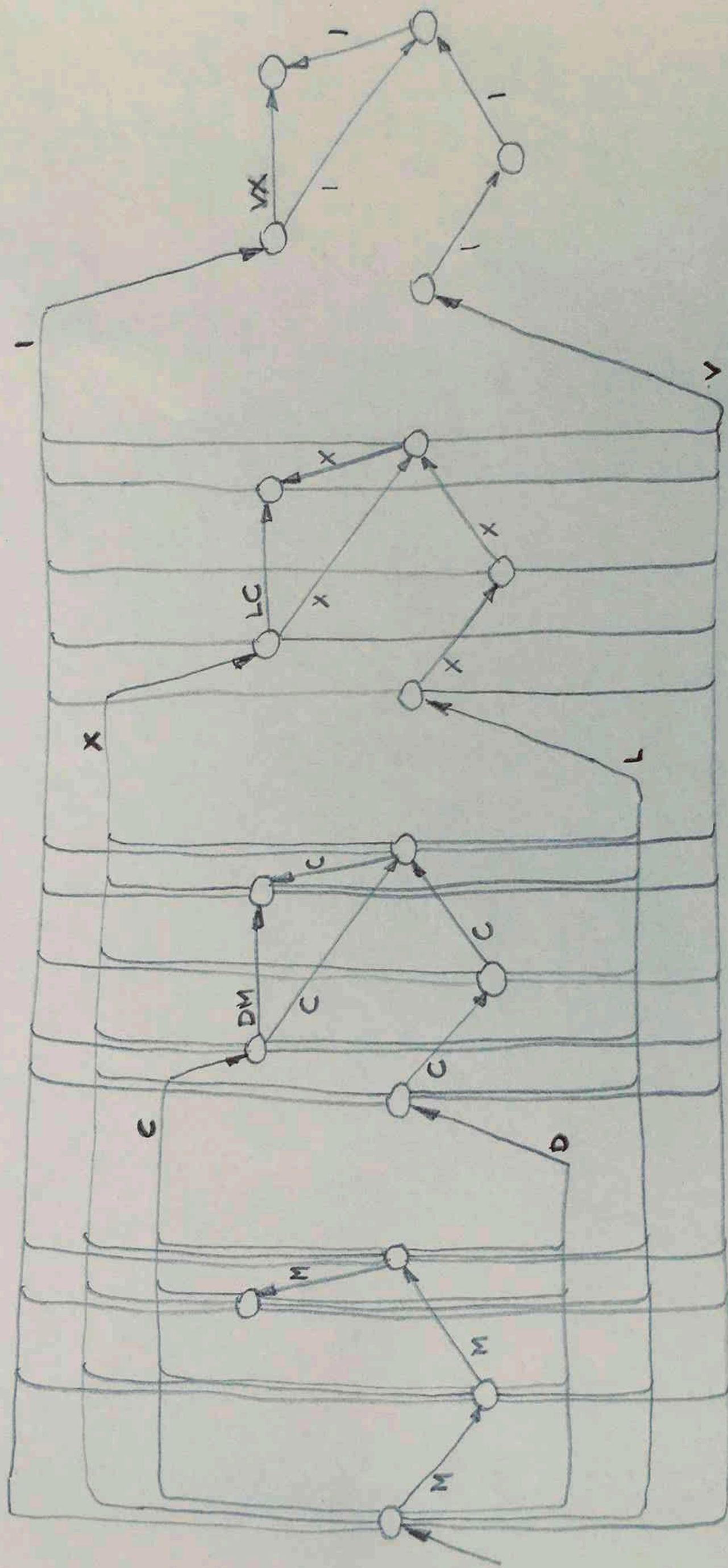
$$\Sigma = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

$$\delta(z_u, d) = z_v \text{ mit } v = (10u + d) \bmod n$$

$$E = \{z_0\}$$

römische Zahlen



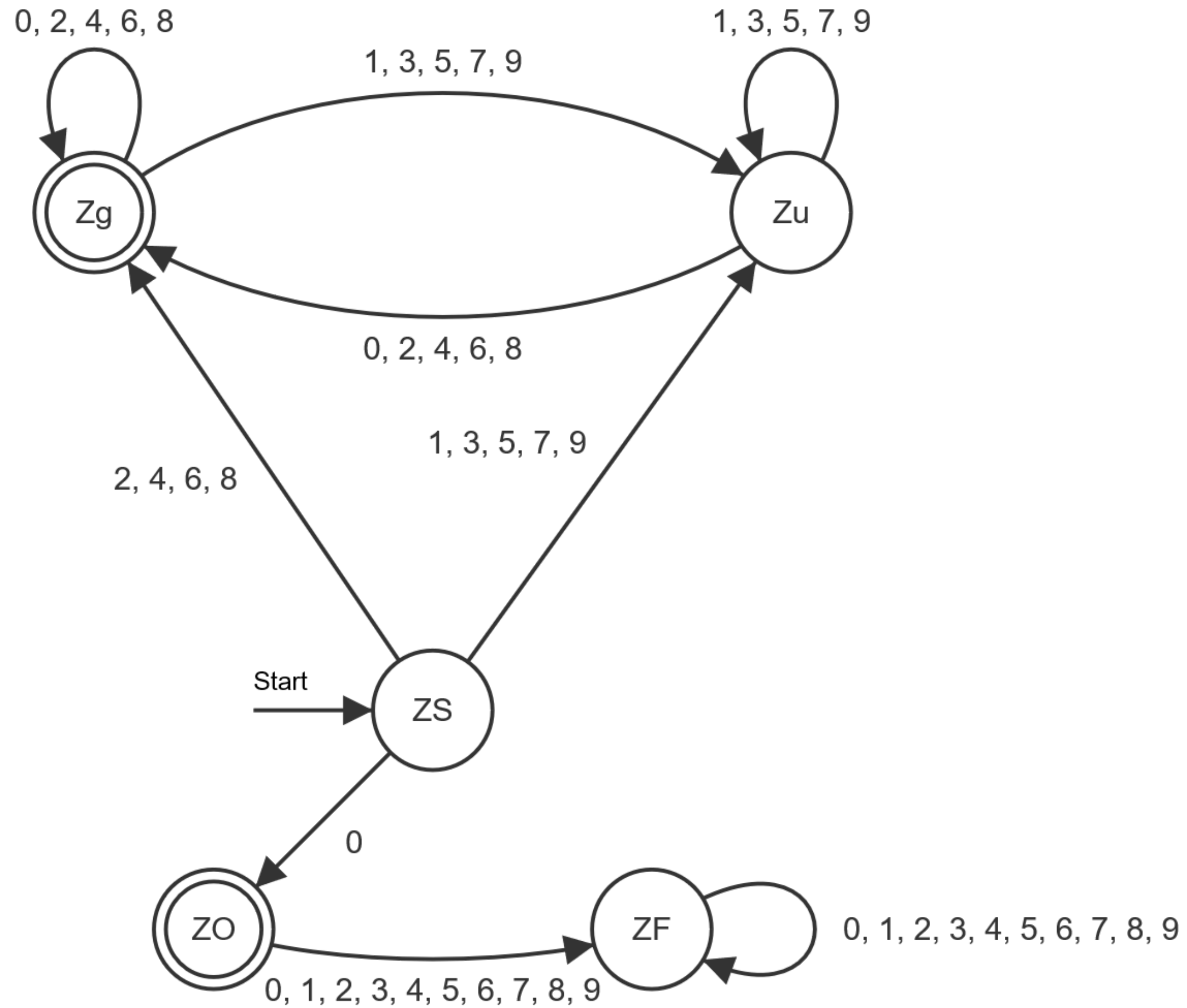


The background is a solid dark red color. It features a central illustration of a girl with a surprised expression sitting in a bathtub. Surrounding her are several circular and oval shapes, each containing a different facial expression or hairstyle, creating a collage-like effect. The text 'TIL' and 'about' is centered over the bathtub area.

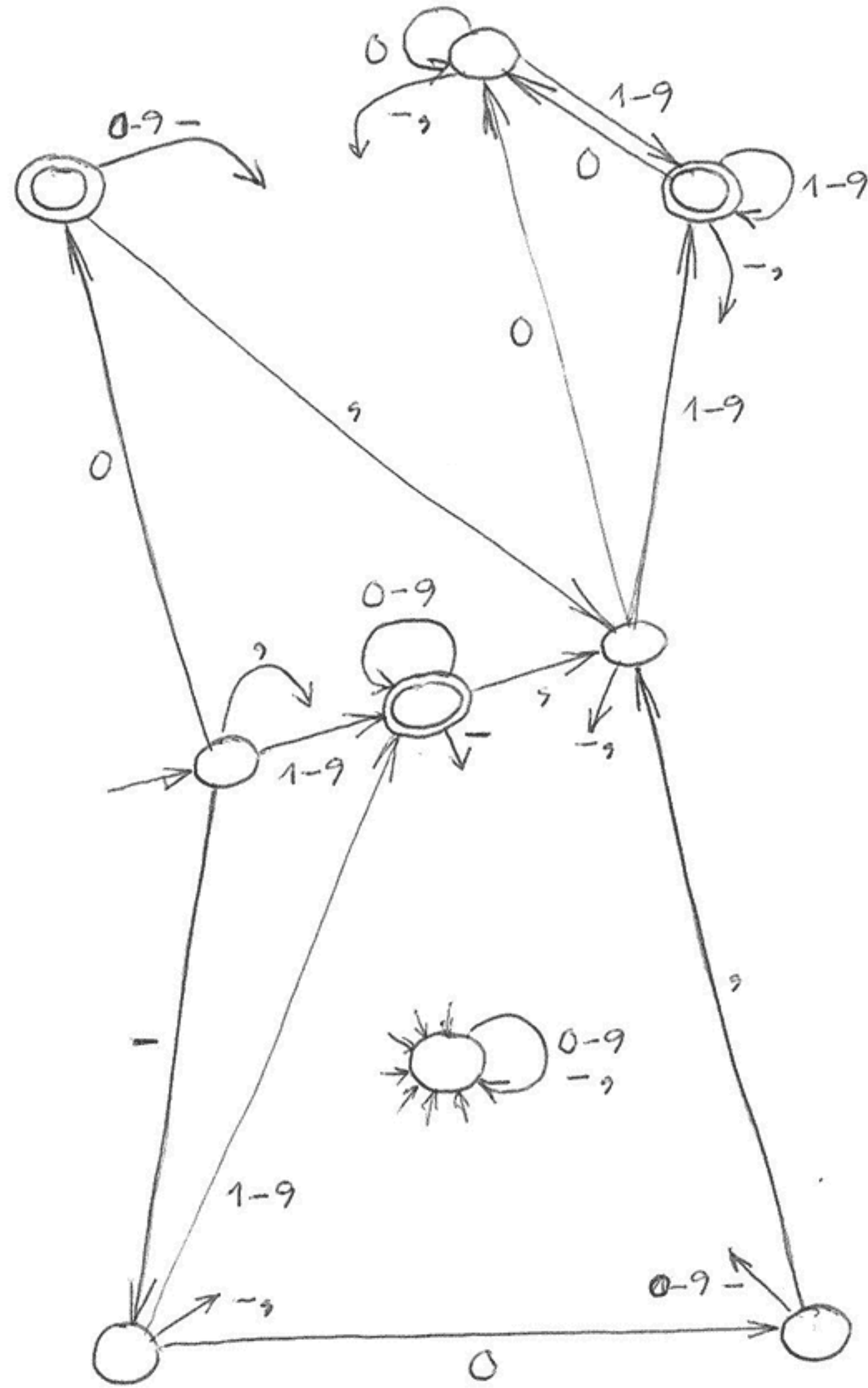
TIL
about

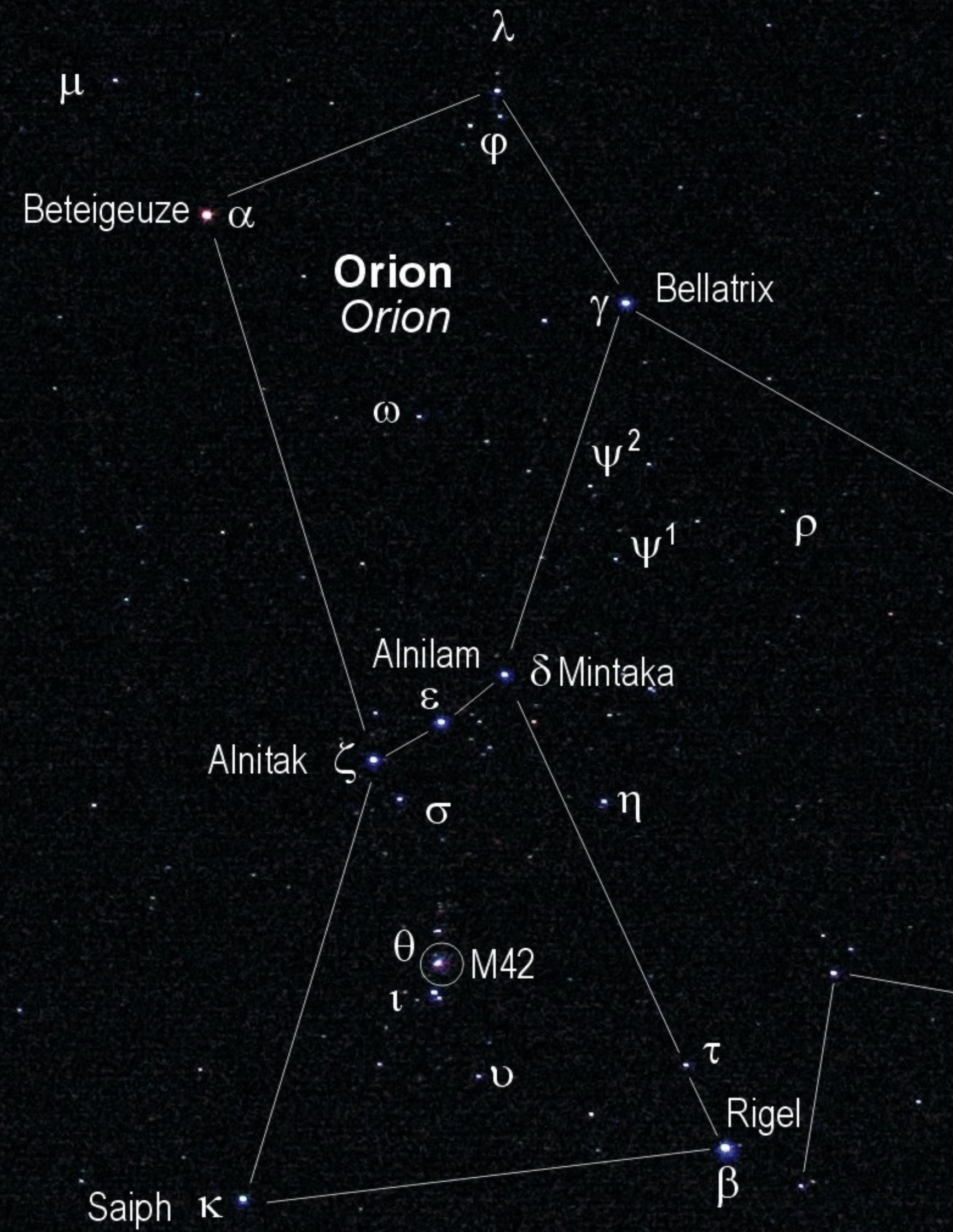
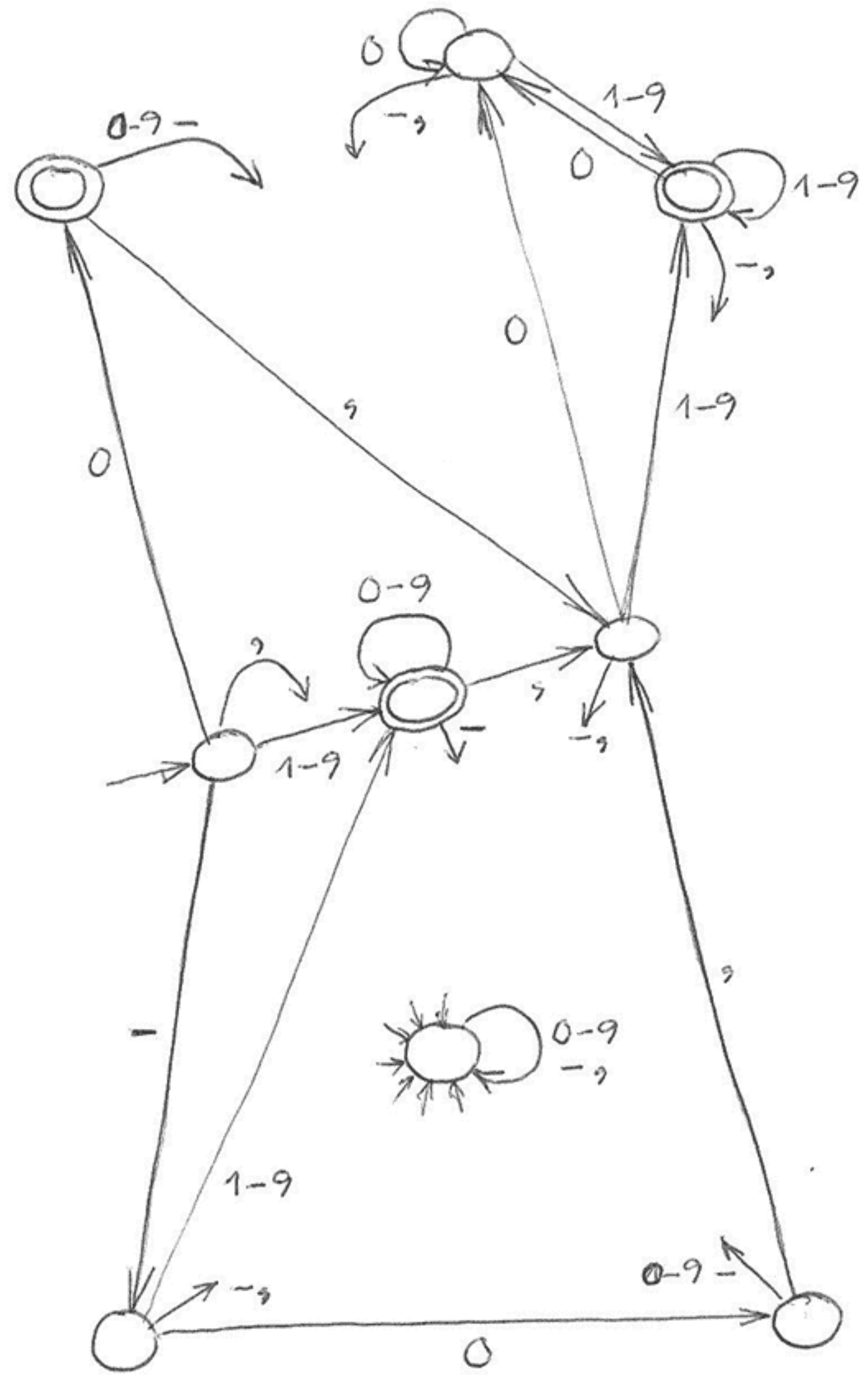
五
言
詩

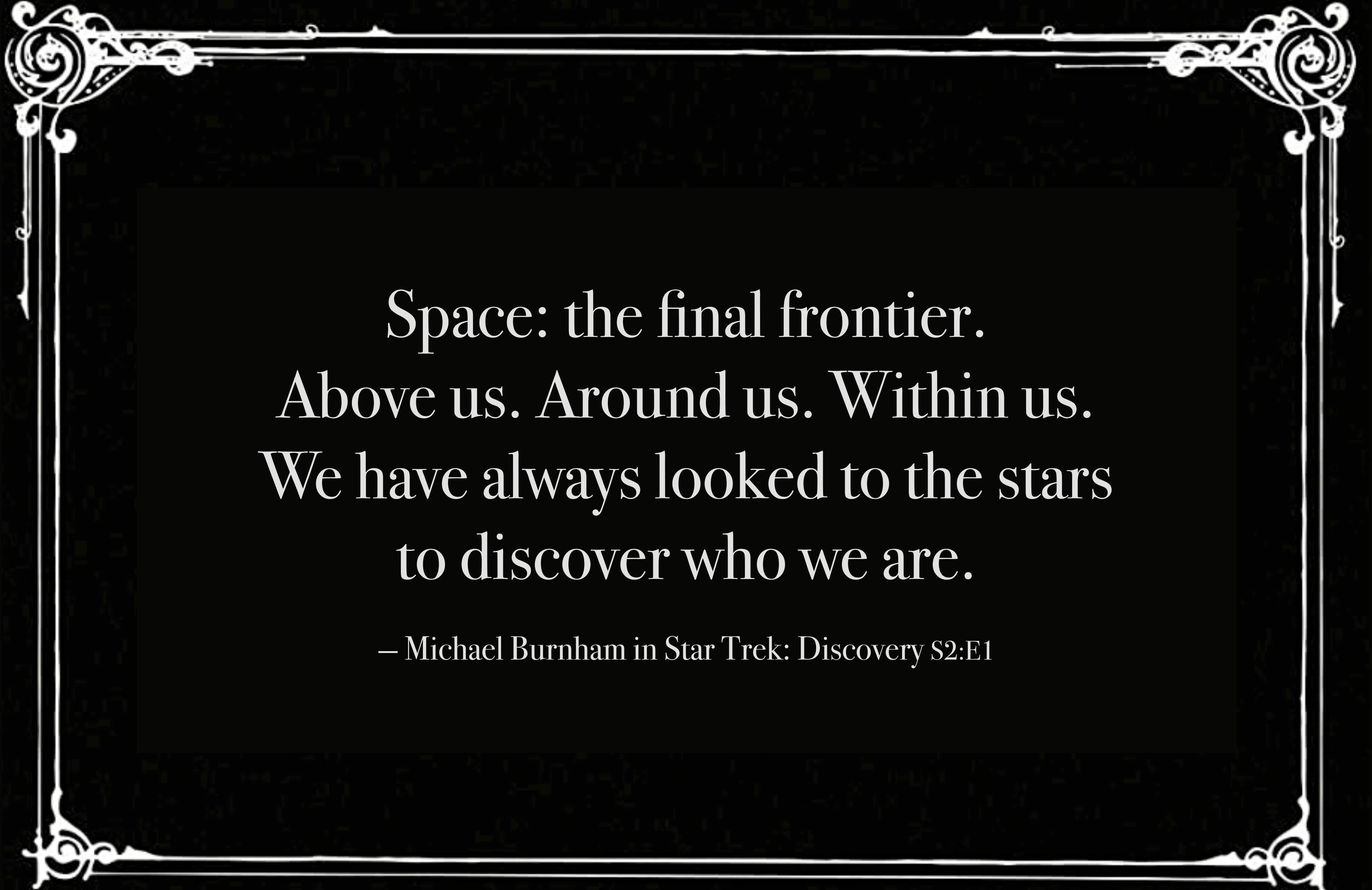
Gerade Zahlen ohne führende Nullen



Kommazahlen







Space: the final frontier.
Above us. Around us. Within us.
We have always looked to the stars
to discover who we are.

– Michael Burnham in *Star Trek: Discovery* S2:E1