

Introspective Pushdown Analysis

Chris Earl
University of Utah
cwearl.com

Ilya Sergey
KU Leuven
[@ilyasergey](https://twitter.com/ilyasergey)

Matt Might
University of Utah
matt.might.net
[@mattmight](https://twitter.com/mattmight)

Dave Van Horn
Northeastern University
lambda-calculus.us
[@lambda_calculus](https://twitter.com/lambda_calculus)

“grad-school Vietnam”

“charred remains”

“would-be Ph.D.s”

-Olin Shivers

“academic War on Terror”

“unwinnable”

“never-ending”

-Me, to my grad students

What is flow analysis?

What is flow analysis?

What is wrong with it?

What is flow analysis?

What is wrong with it?

How do we fix it?

How do we fix the fixes?

We use fixed points.

What is flow analysis?

What is control-flow analysis?

(f x)

What is f ?

Why not run the program?

CEK

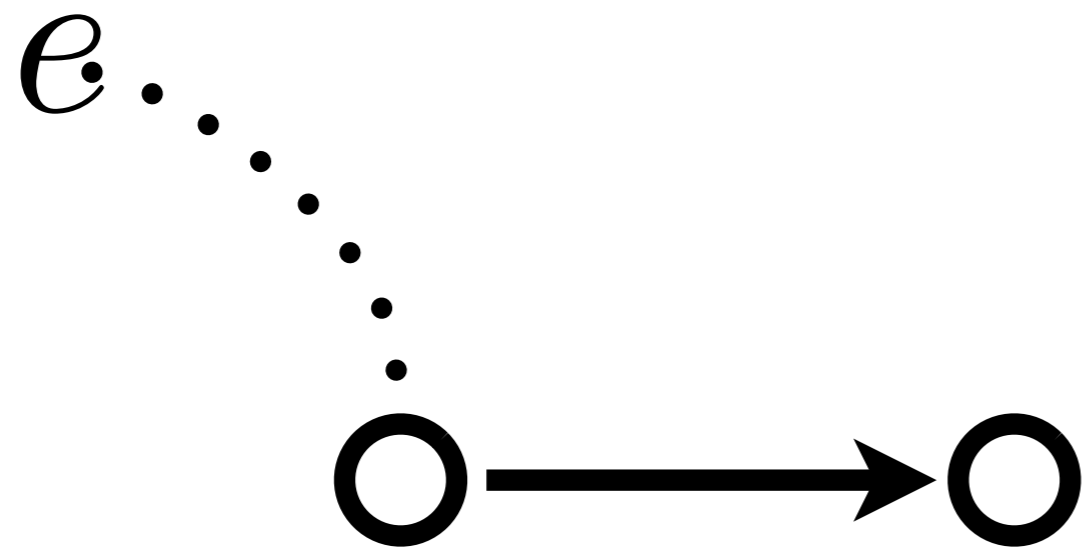
CEK

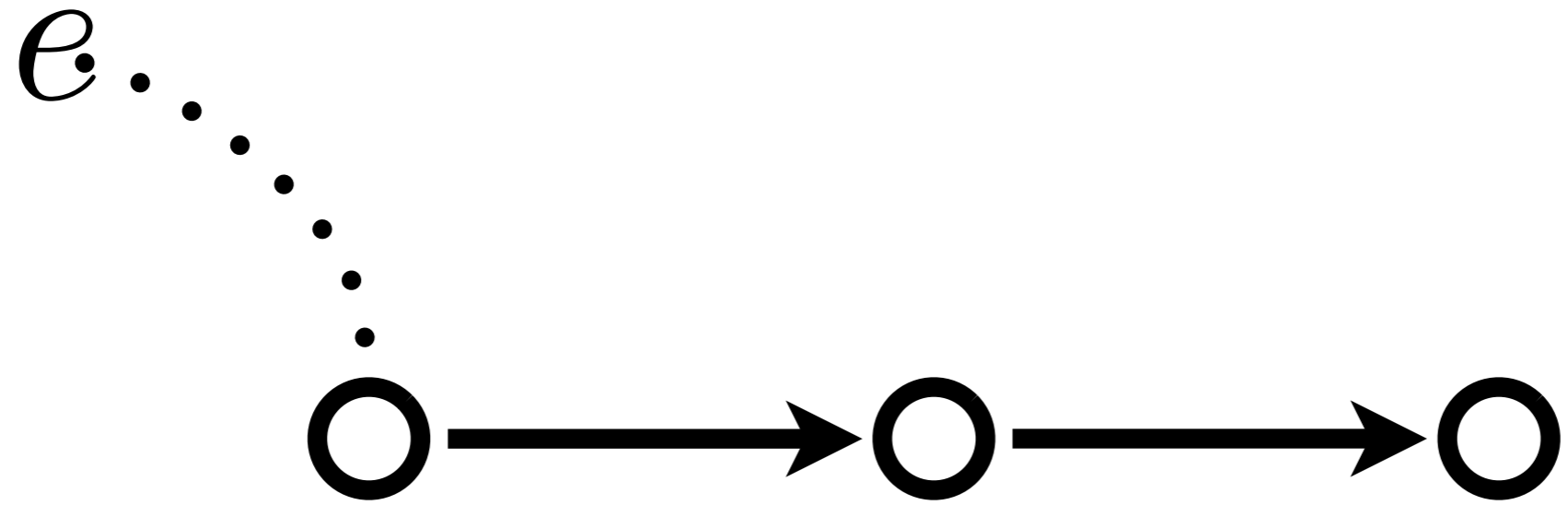
(Felleisen and Friedman, 1986)

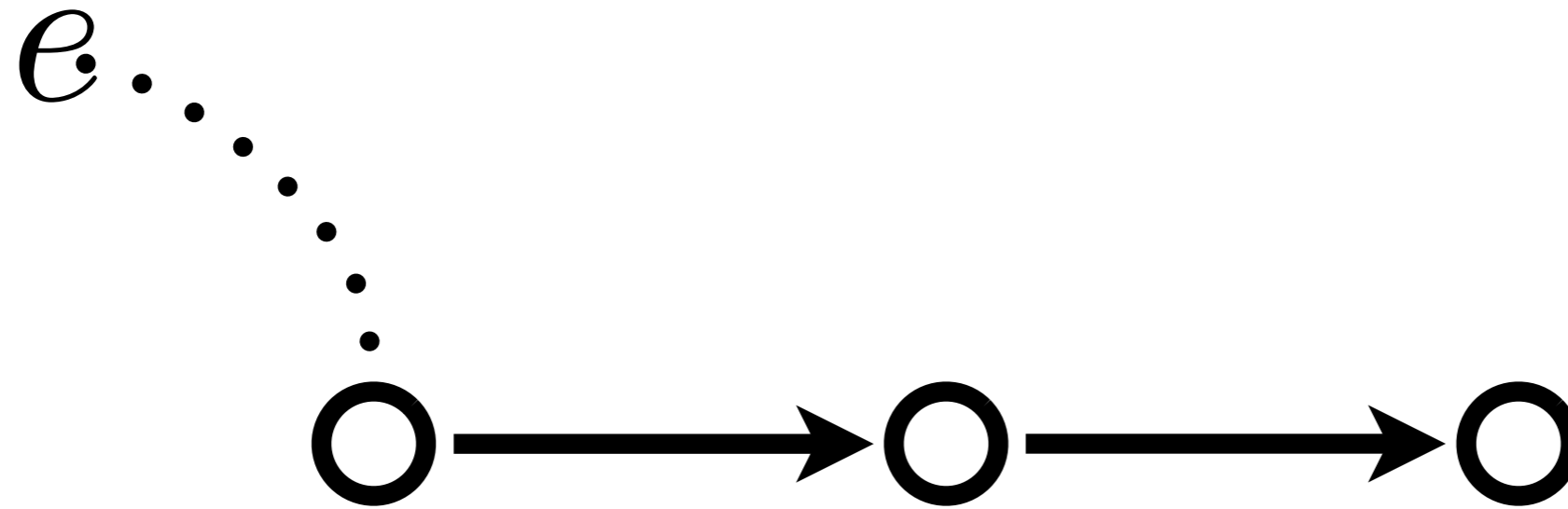
e

e

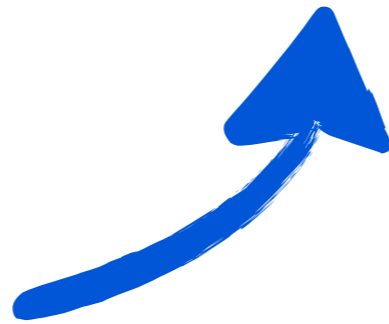
e. . . .
O

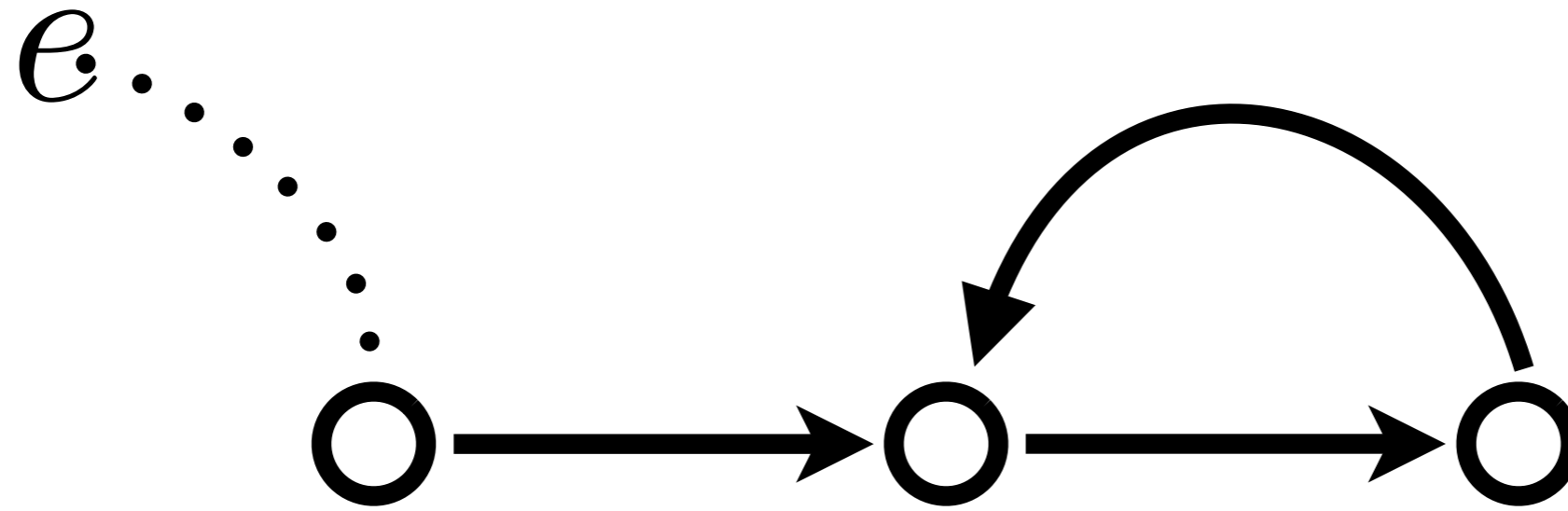




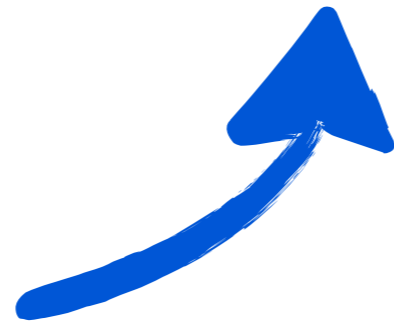


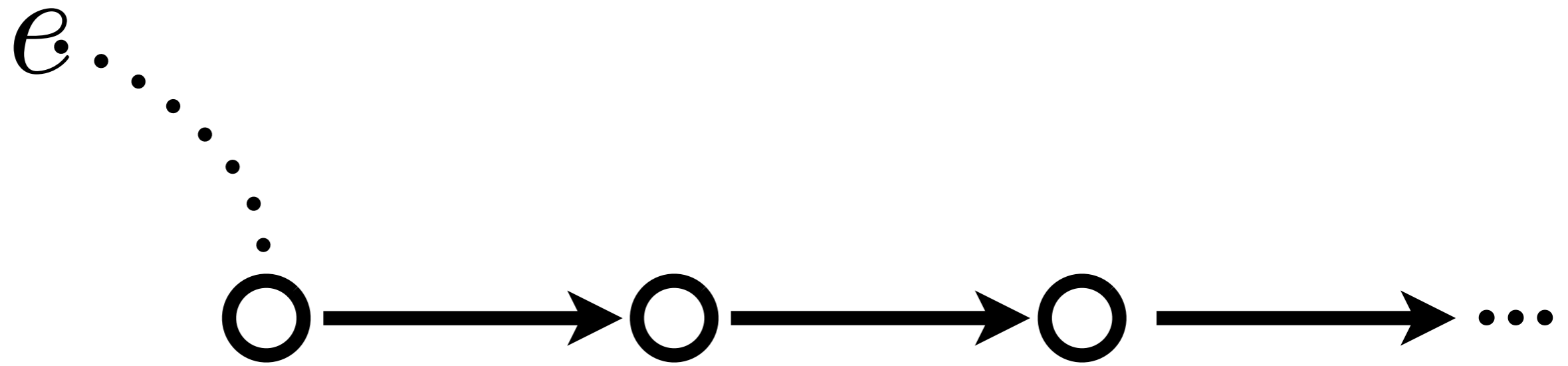
What is f , here?





What is f , here?





Make it terminate?

Make it finite.

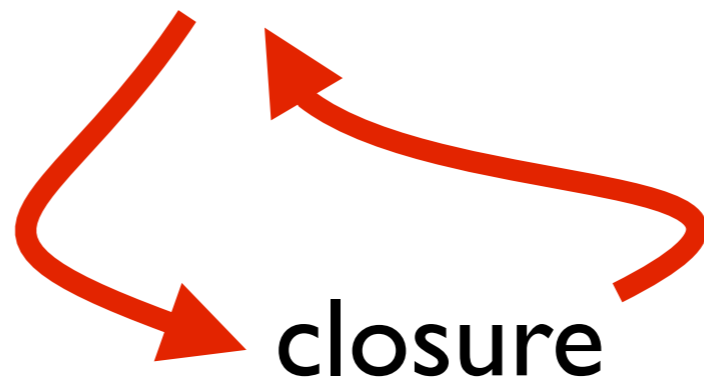
Make it finite.

(Might, SAS 2010)

(Van Horn and Might, ICFP 2010)

CEK

CEK



closure

CEK

$E = V \rightarrow \lambda x E$

CEK

$E = V \rightarrow \lambda x E$



CEK

E = V → A

CEK



CEK

CEK

CESK

CESK


(Felleisen and Friedman, 1987)

CESK

S = A → λ × E

CESK

$S = A \rightarrow \lambda \times E$



CESK

CESK



CESK*

CESK*

$K^* \subset A$

CESK*

$S = A \rightarrow \lambda \times E + K$

CESK*

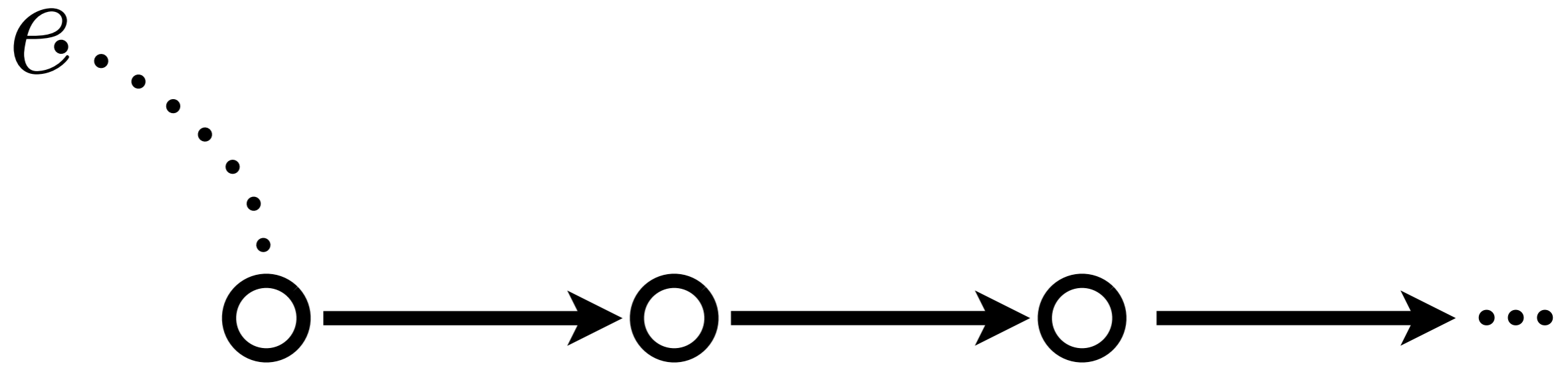
CESK*

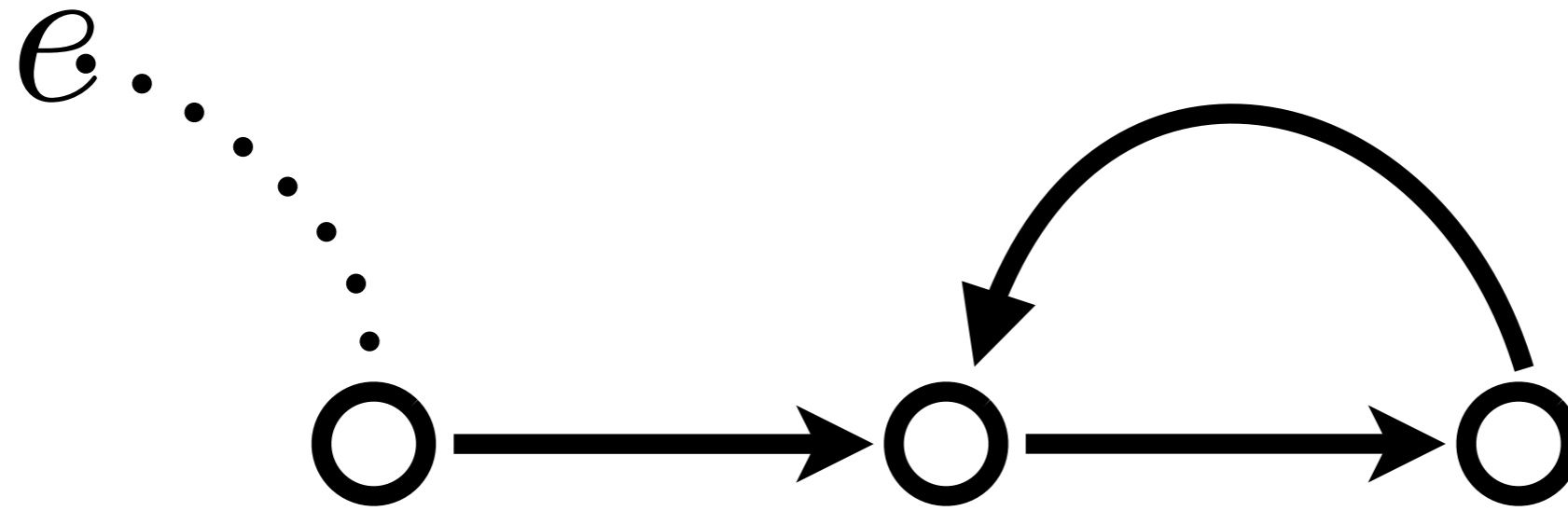
CĚŠK*



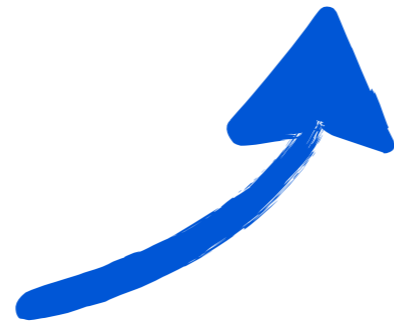
CESK*

e





What is f , here?



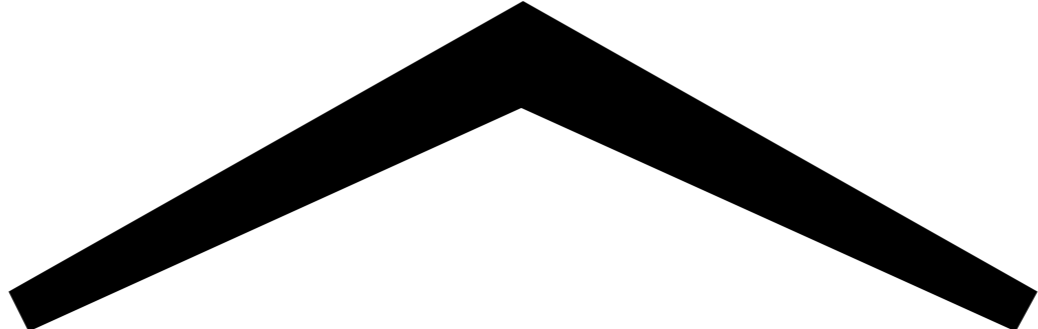
What's wrong with flow analysis?

Control-flow forks.

Data-flow merges.

why?

S = A → D



S = A → D

S = \hat{A} → D

$$S = \hat{A} \rightarrow D$$

$$S = \hat{A} \rightarrow \hat{D}$$

$$\hat{S} = \hat{A} \rightarrow \mathcal{P}(\hat{D})$$

$(f \ x)$

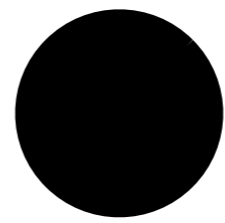
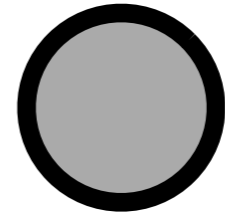
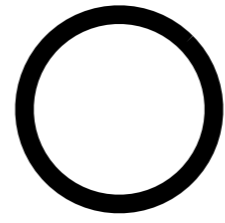
What is f ?

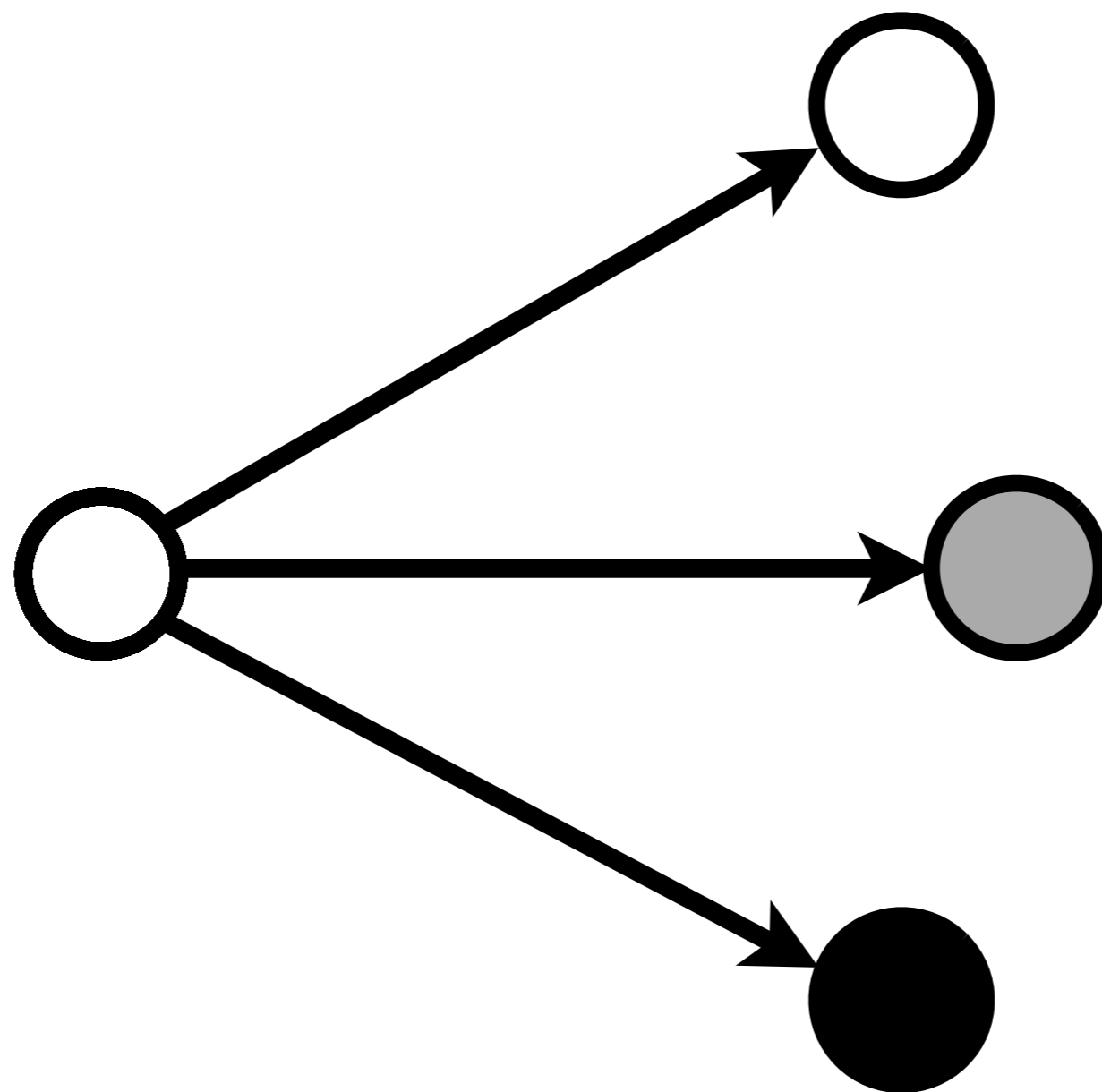
f is

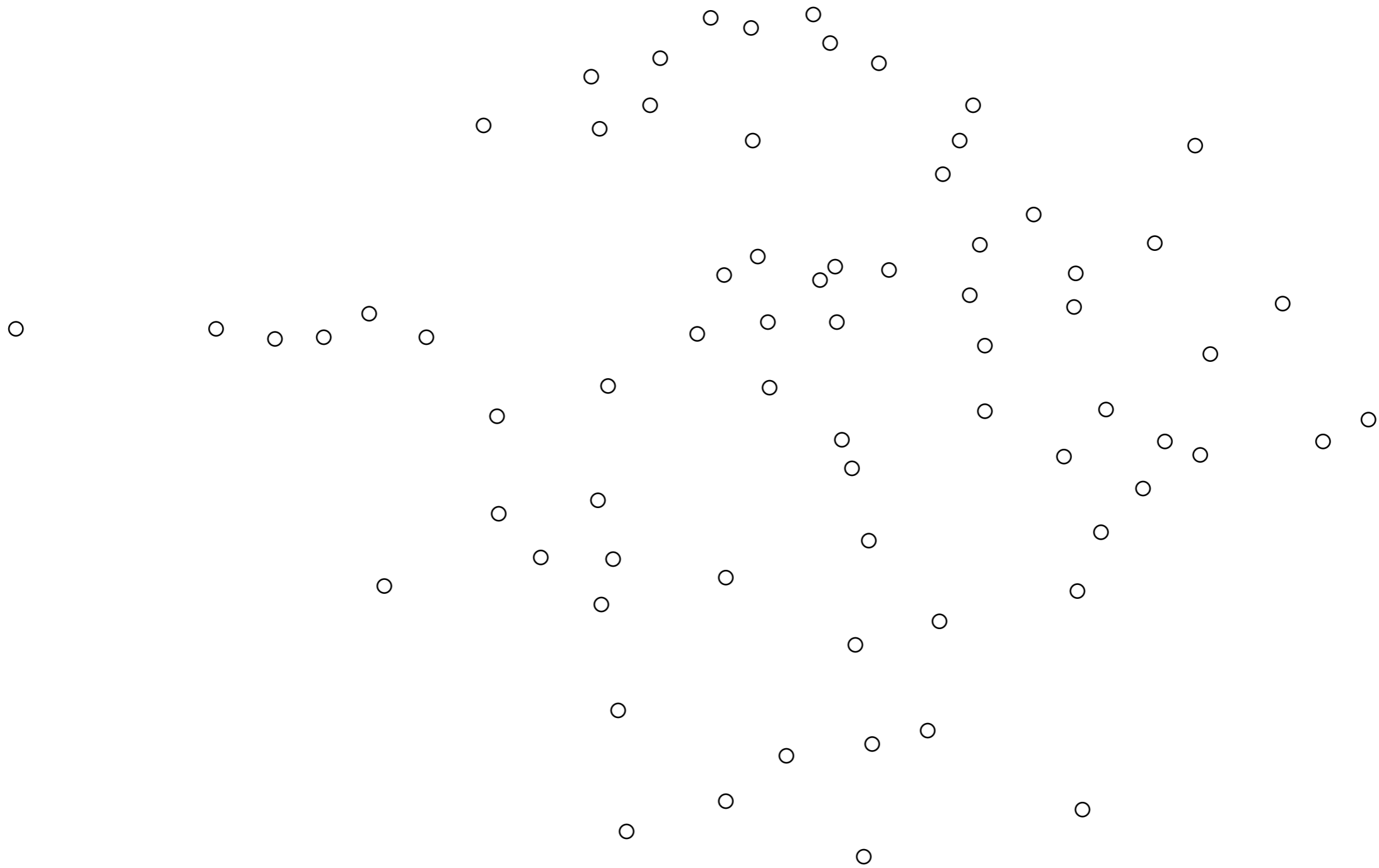
or

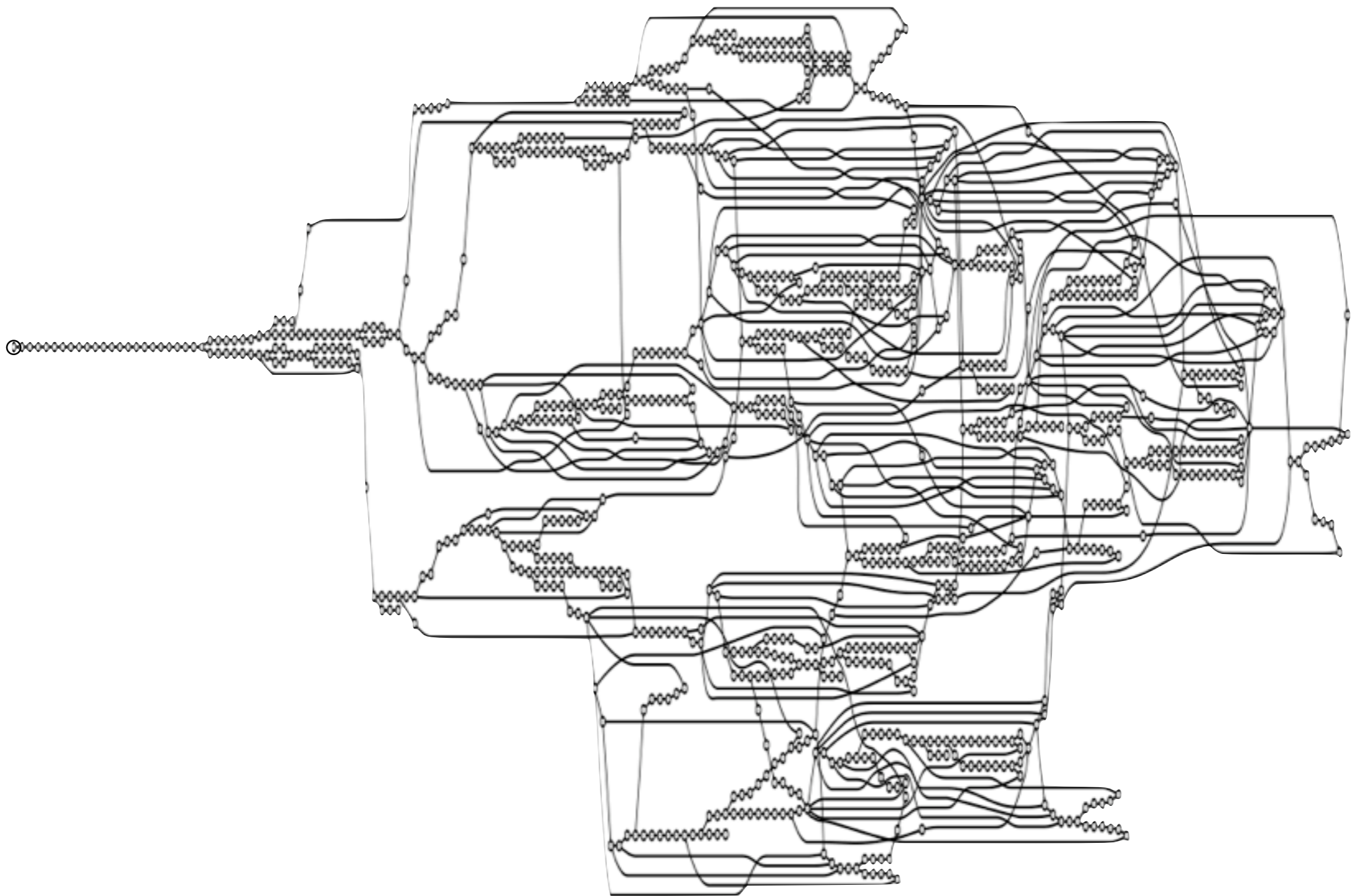
or

f is  or  or 









Problem?

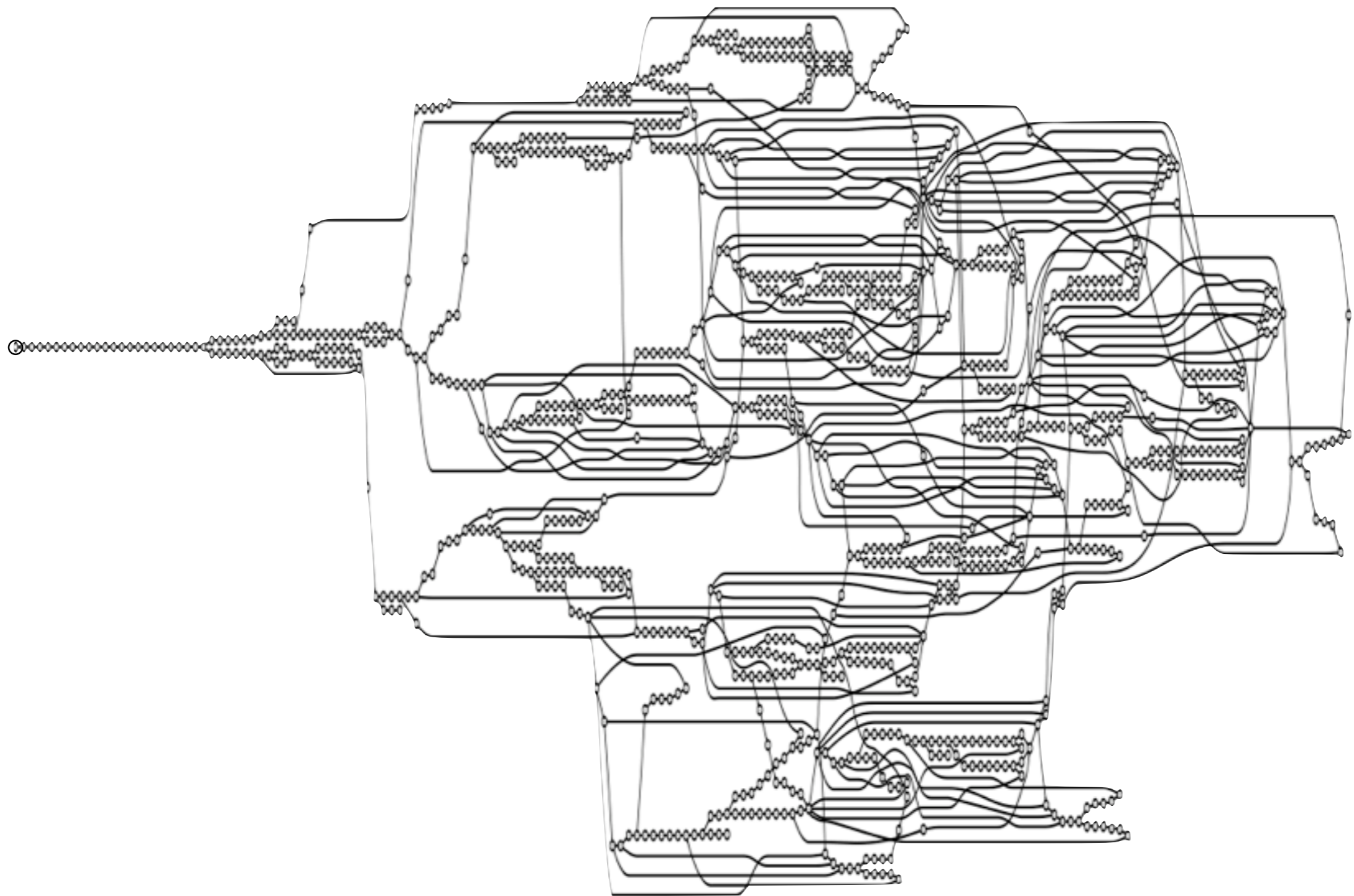
Finite store.


Solution?

Toss garbage.

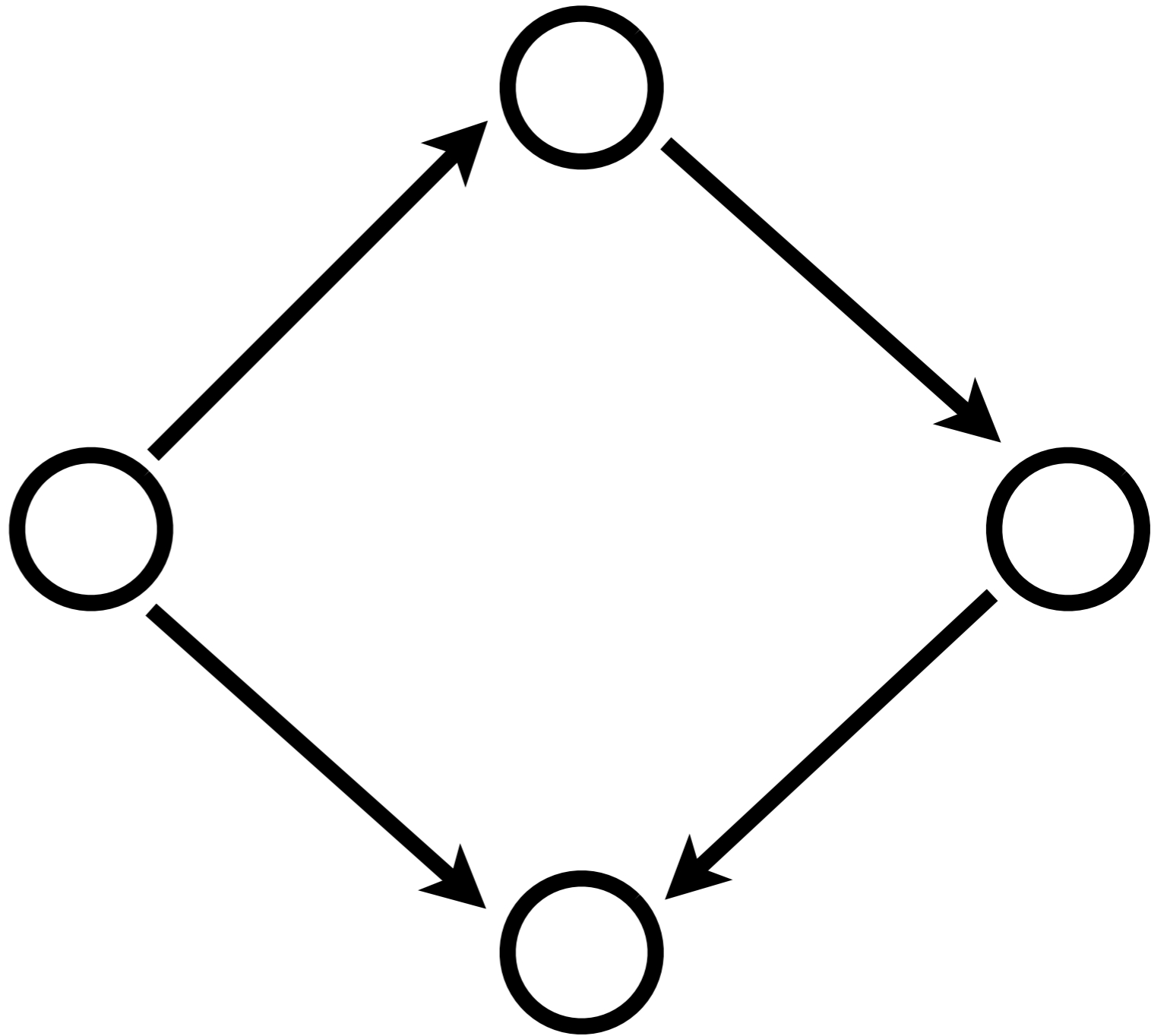
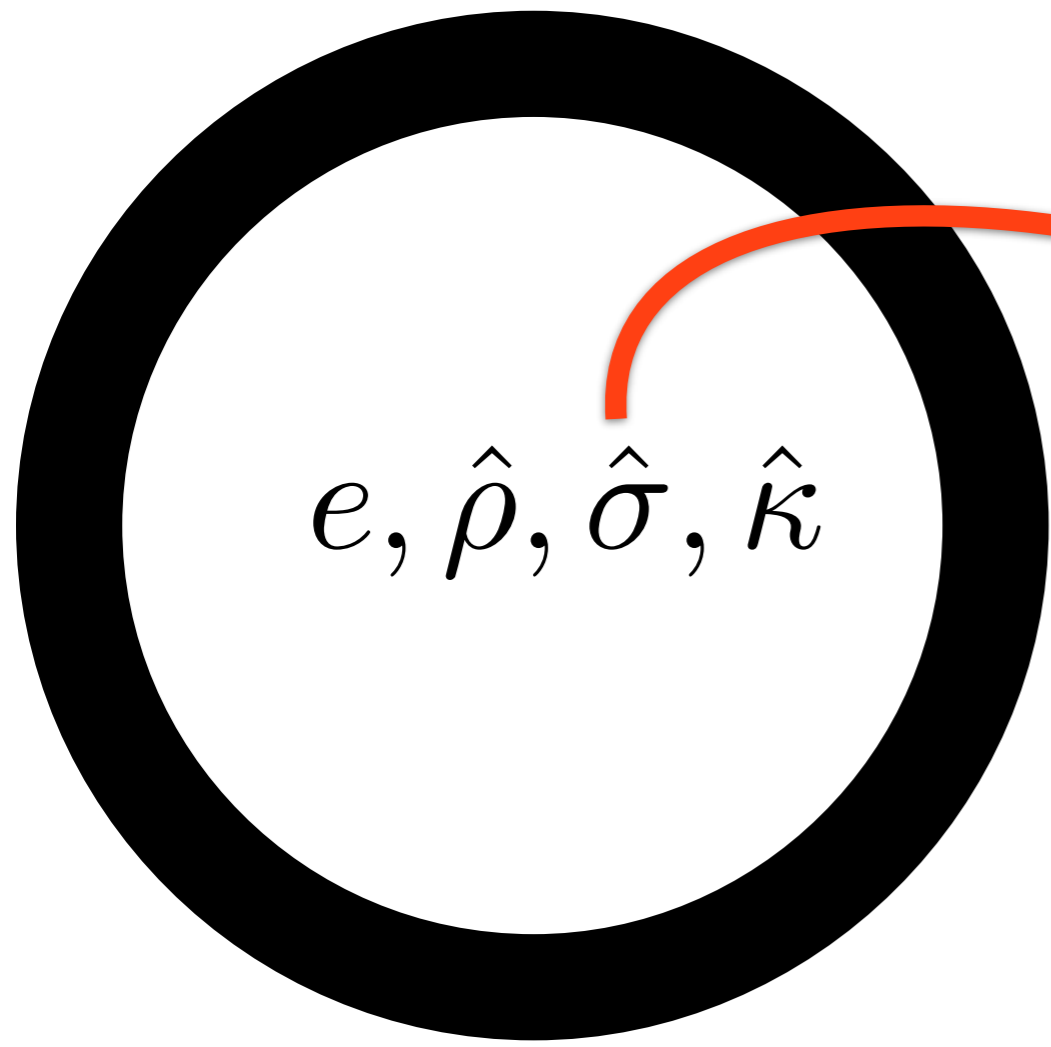
Toss garbage.

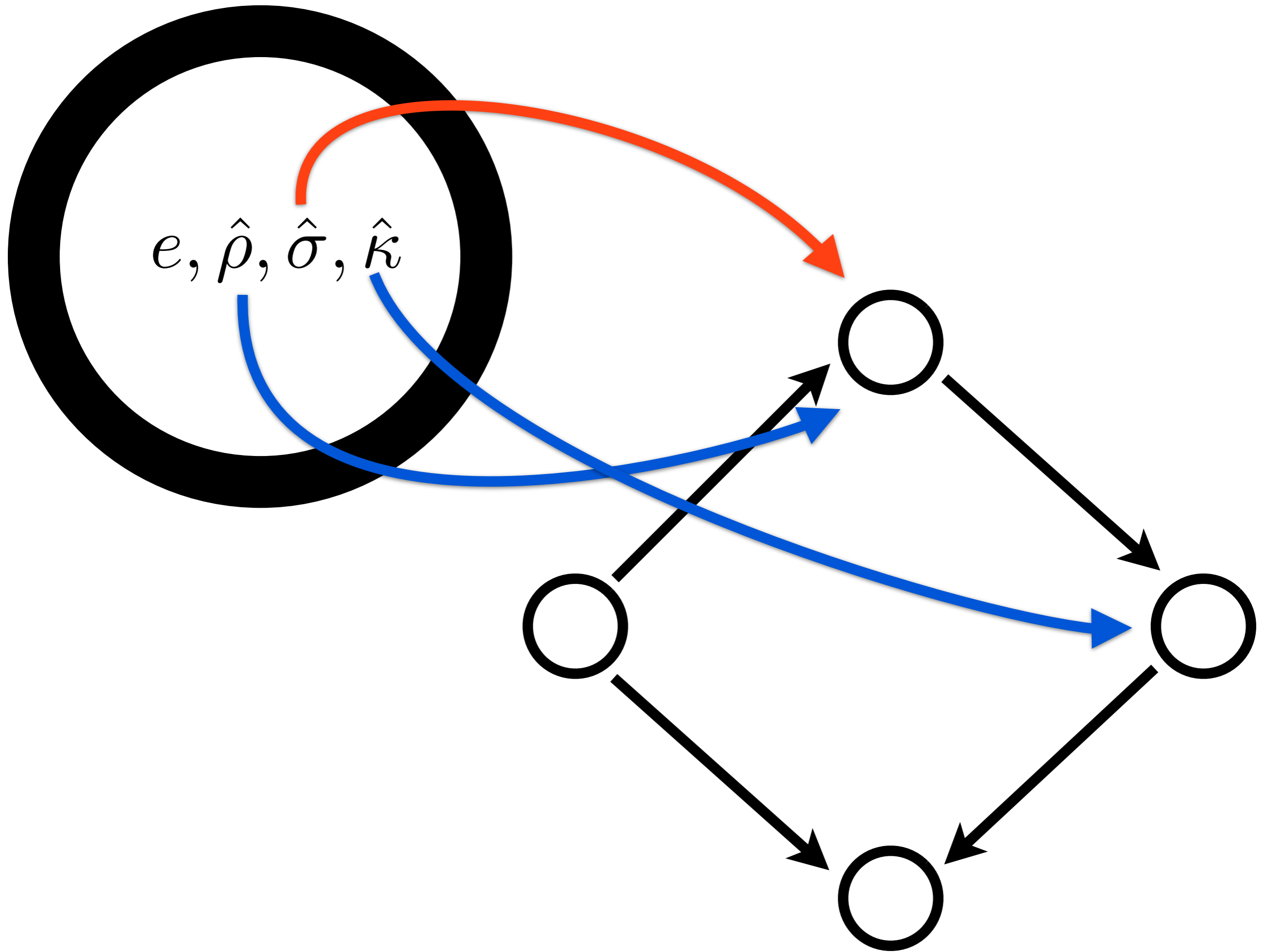
(Might & Shivers, ICFP 2006)

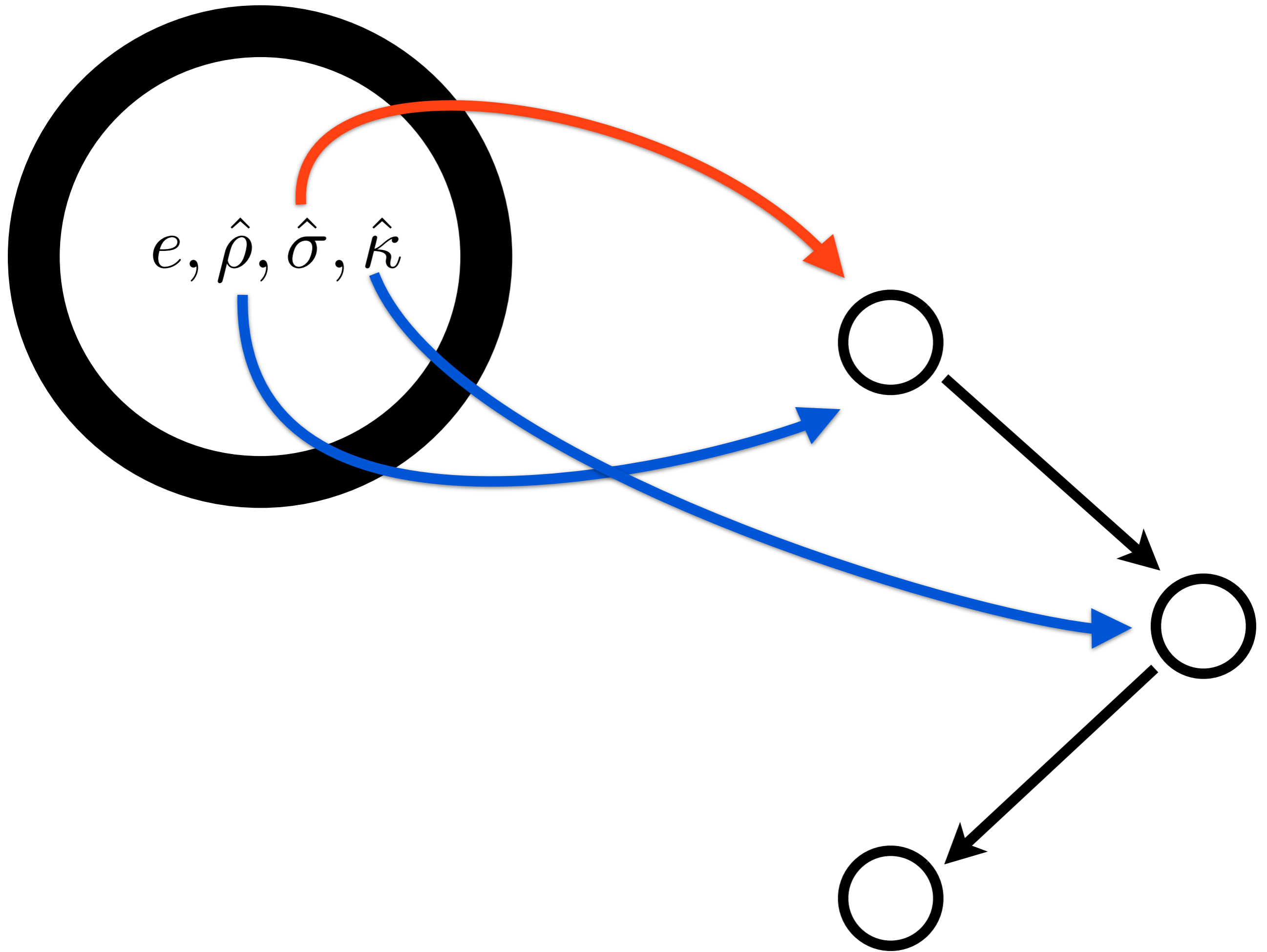





$e, \hat{\rho}, \hat{\sigma}, \hat{K}$

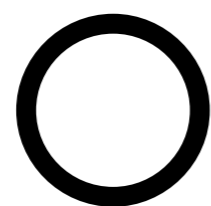


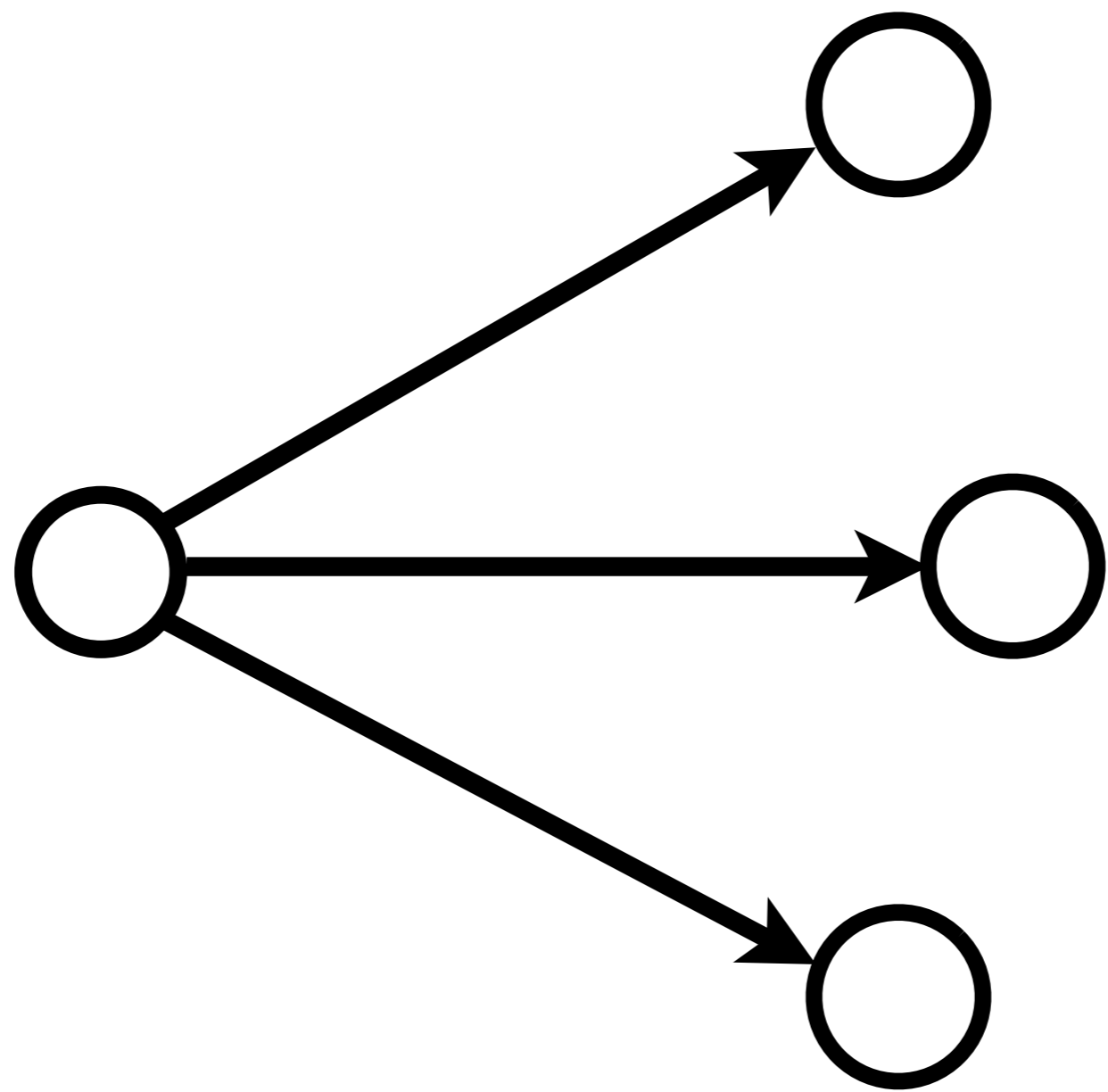




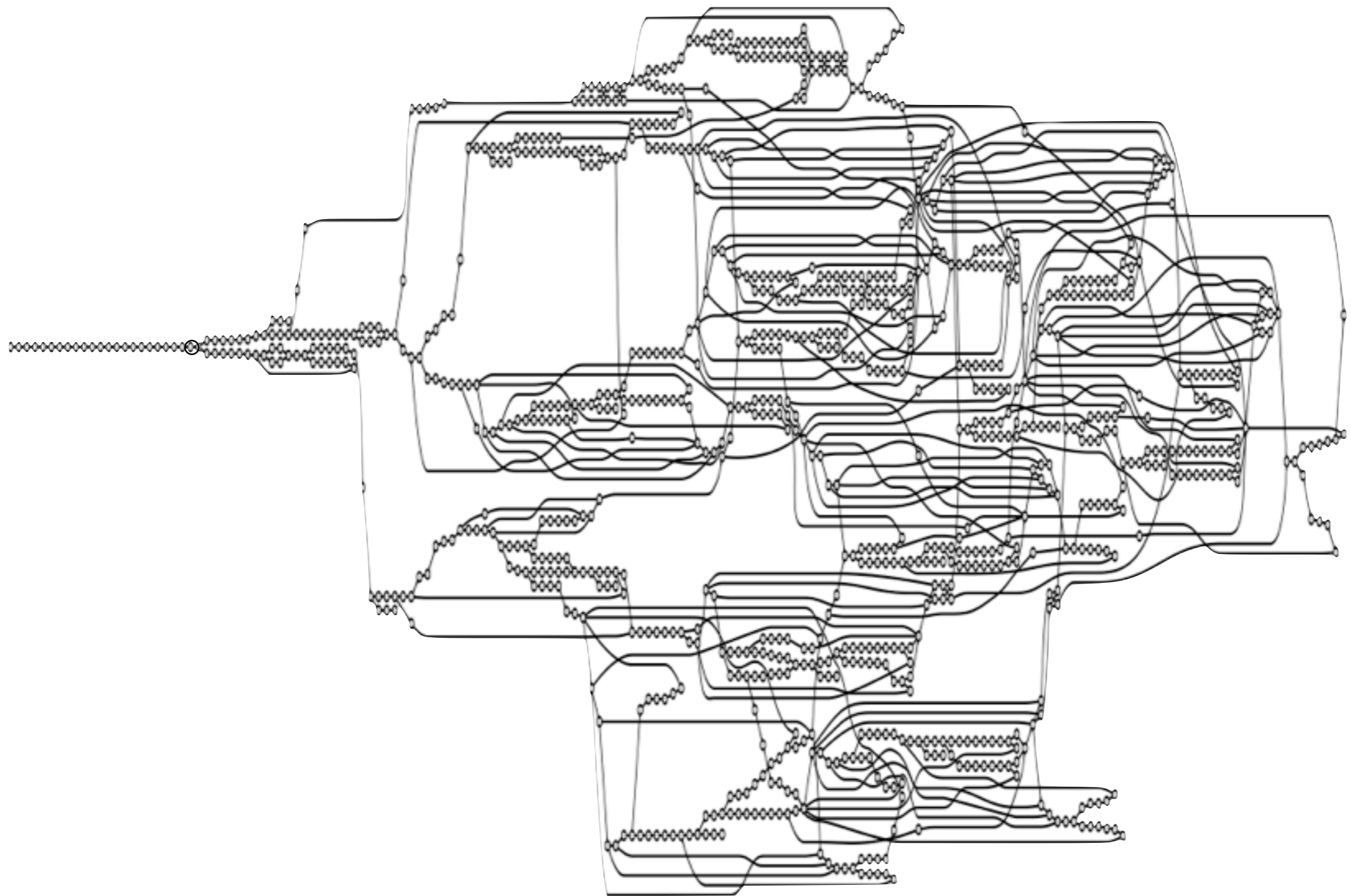


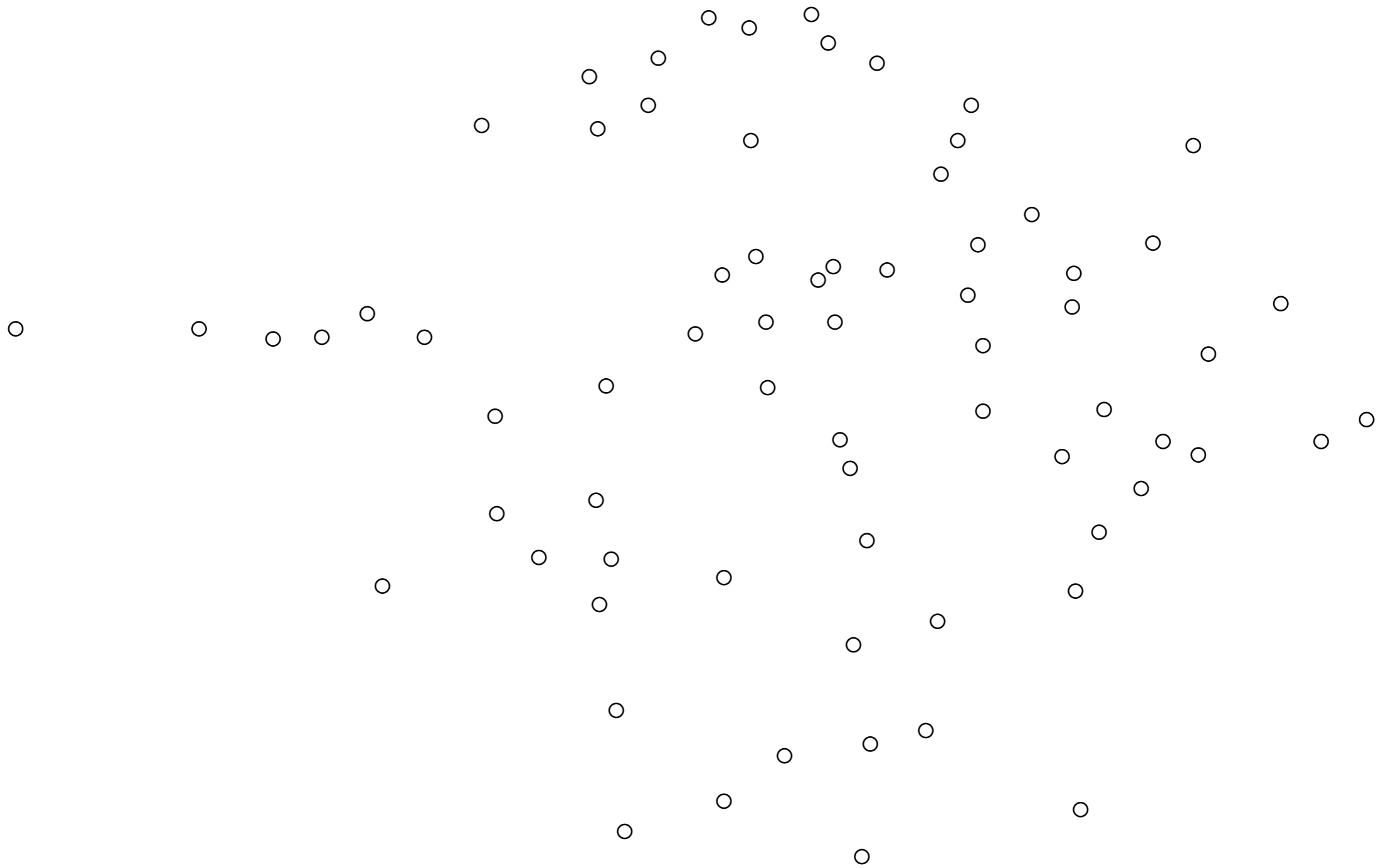
$e, \hat{\rho}, \hat{\sigma}', \hat{k}$

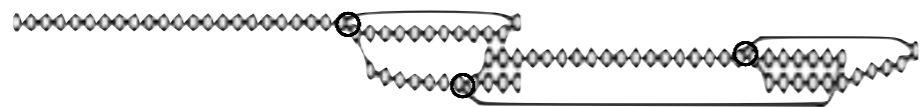












Control-flow forks.

Return-flow forks.

(foo)

(define (foo)

...)

(foo)

(foo)



(define (foo)

...)

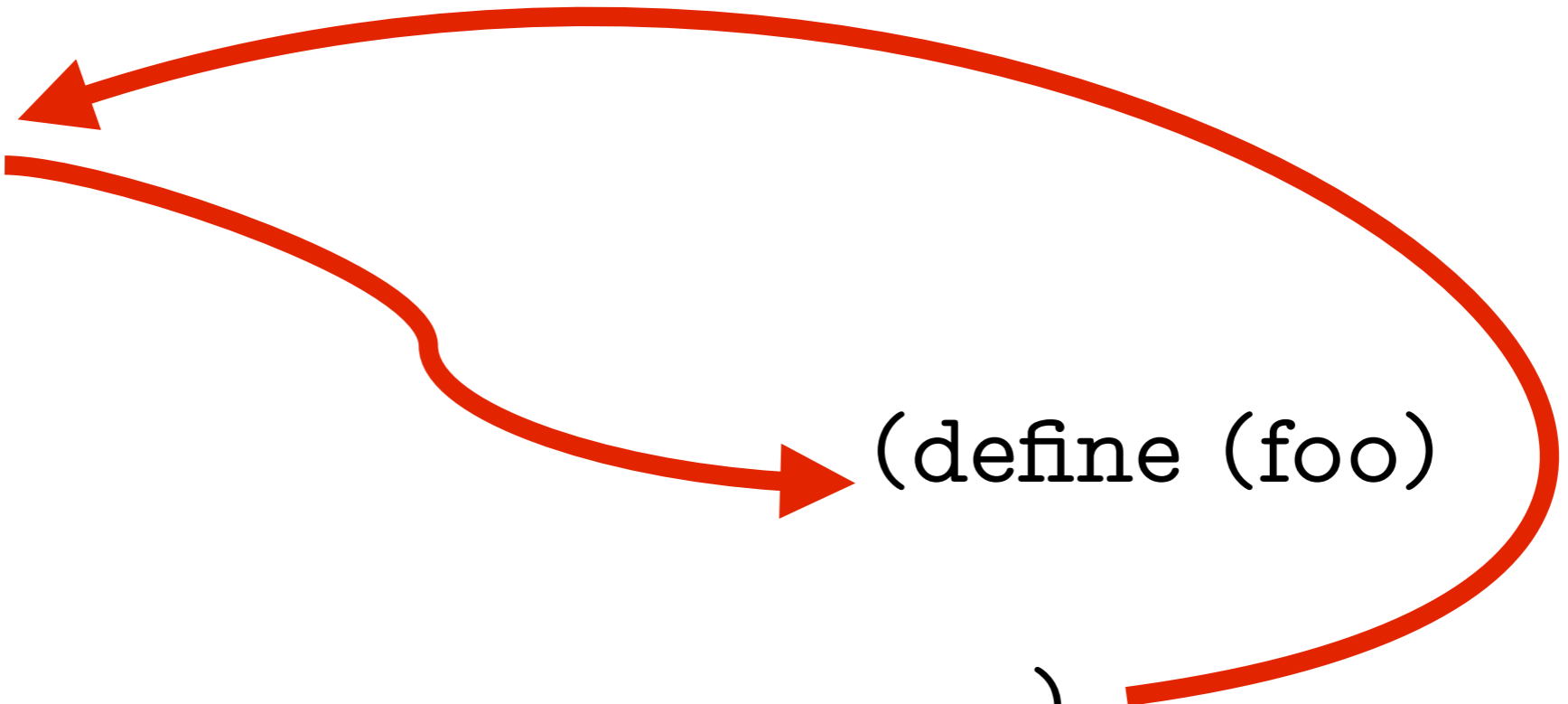
(foo)

(foo)

(define (foo)

...)

(foo)

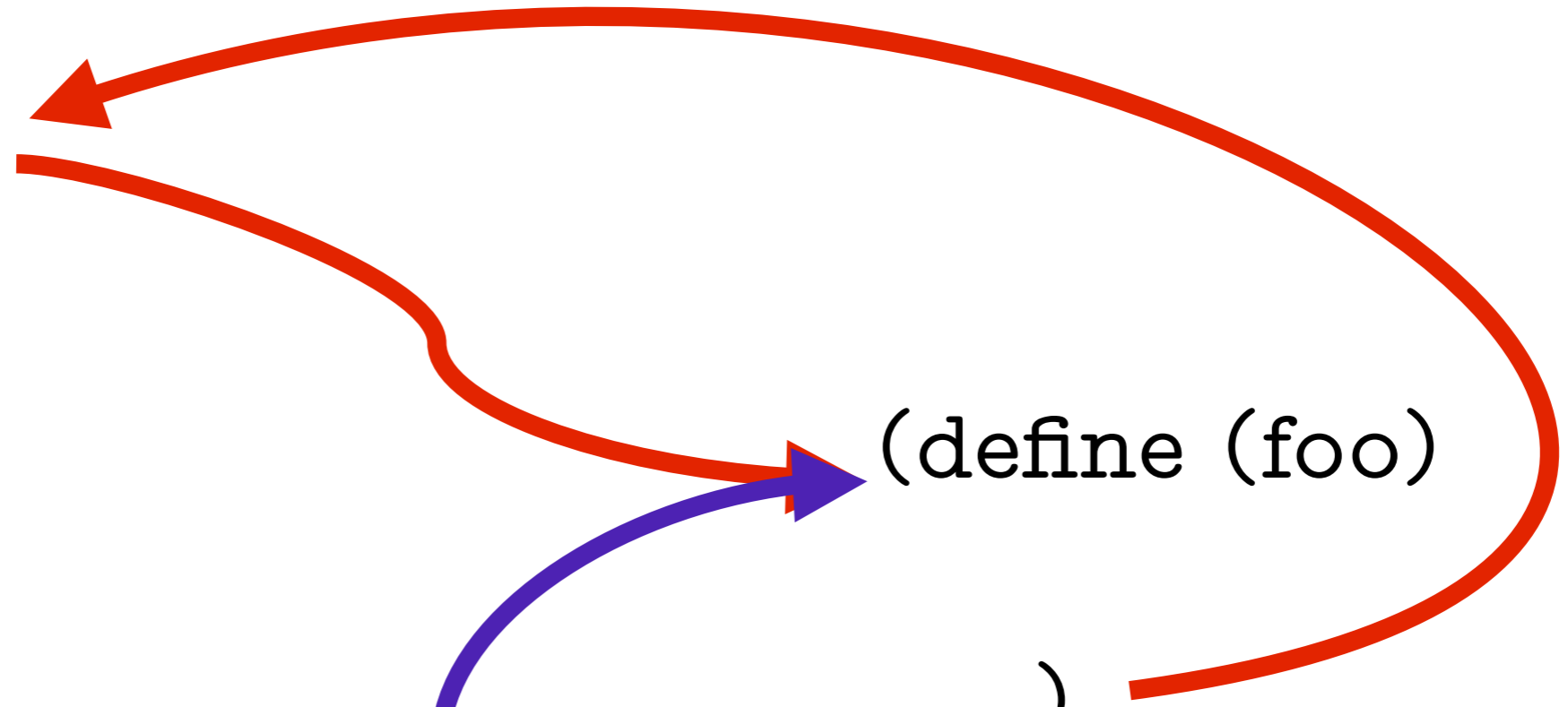


(foo)

(define (foo)

...)

(foo)

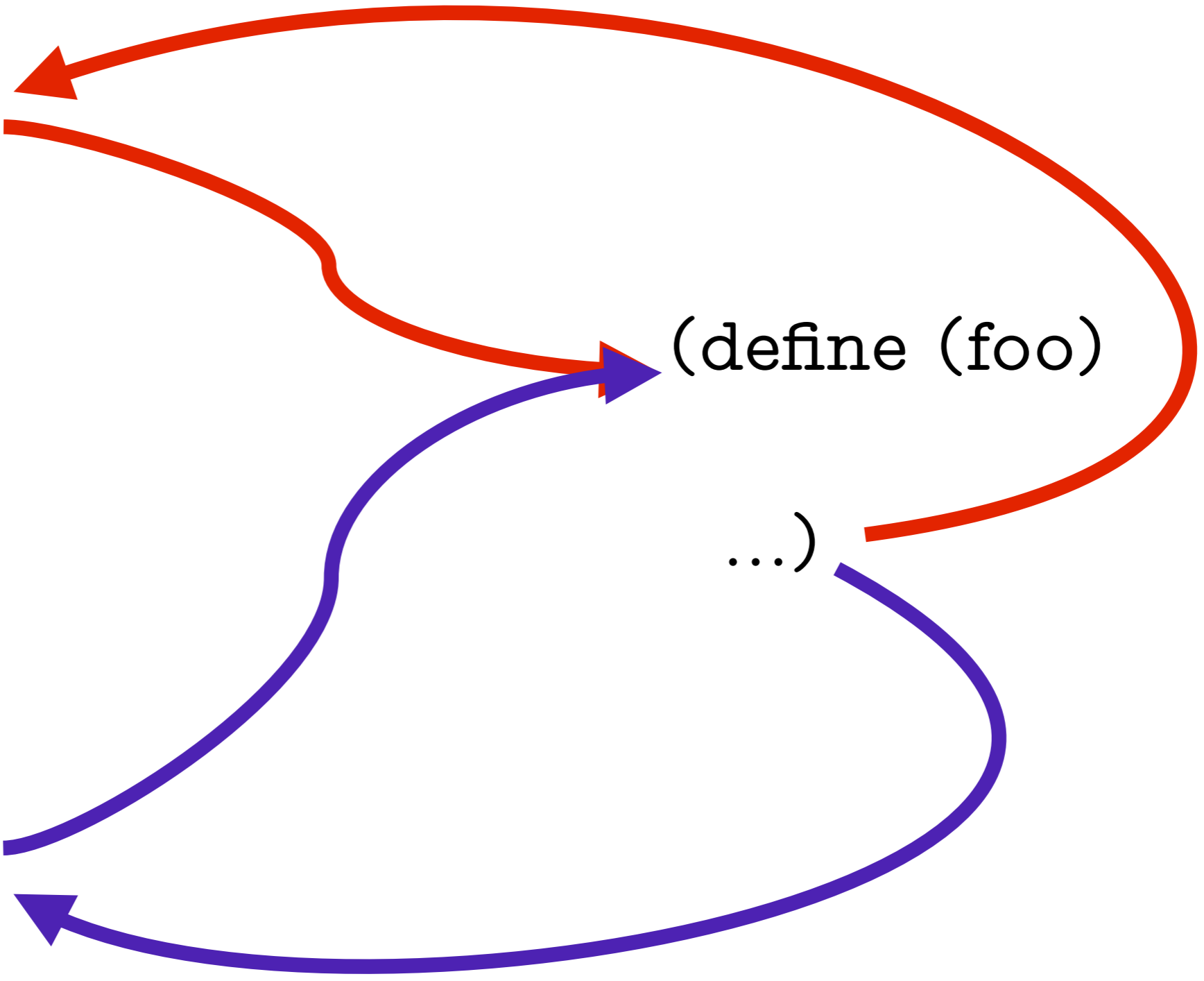


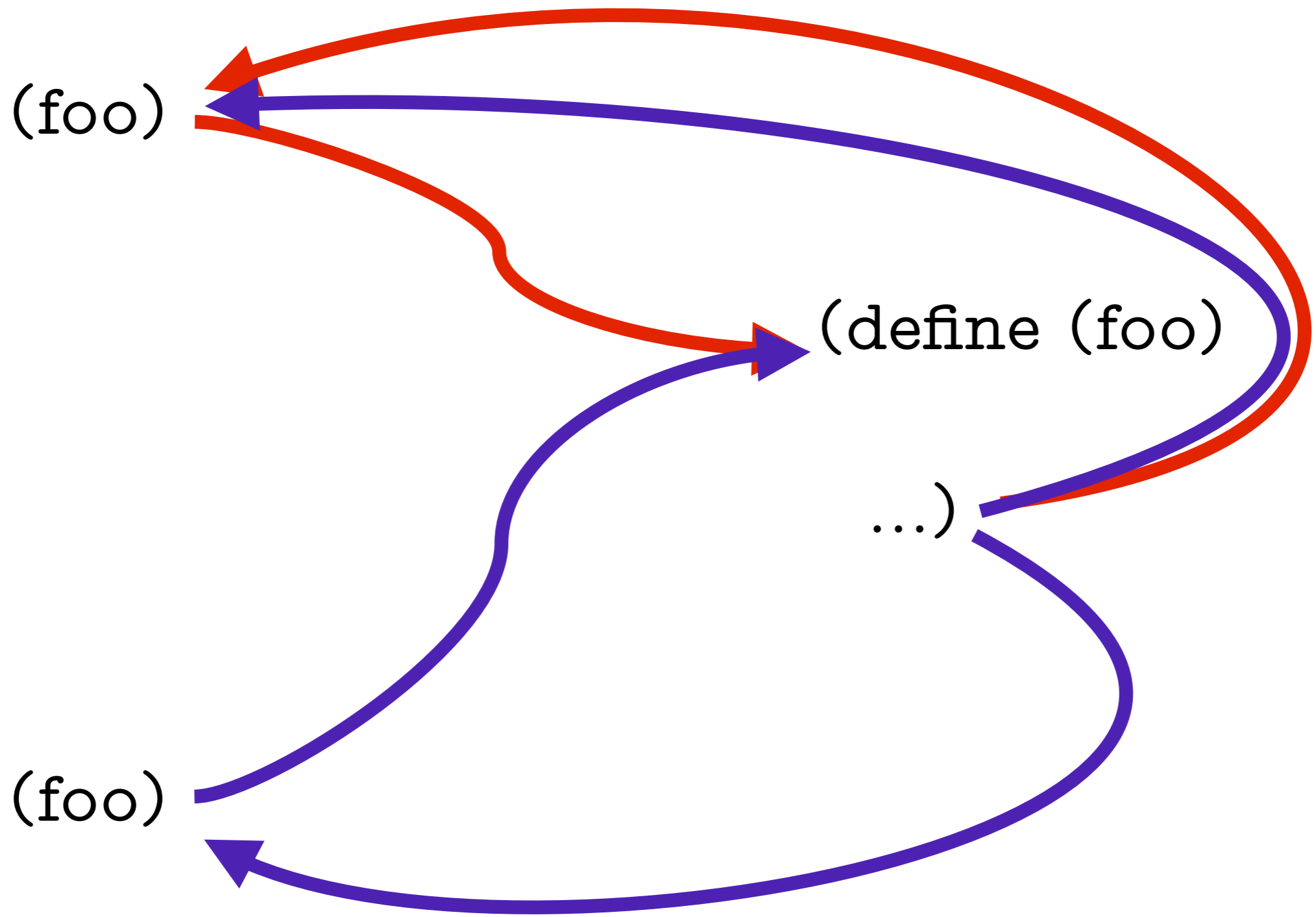
(foo)

(define (foo)

...)

(foo)





$$\hat{S} = \hat{A} \rightarrow \mathcal{P}(\hat{D})$$

$$\hat{S} = \hat{A} \rightarrow \mathcal{P}(\lambda \times \hat{E} + \hat{K})$$

CESK

CESK*



CESK*

CESK

CES K

$\widehat{CES} \widehat{K}$

$\widehat{CES} \widehat{K}$

control state

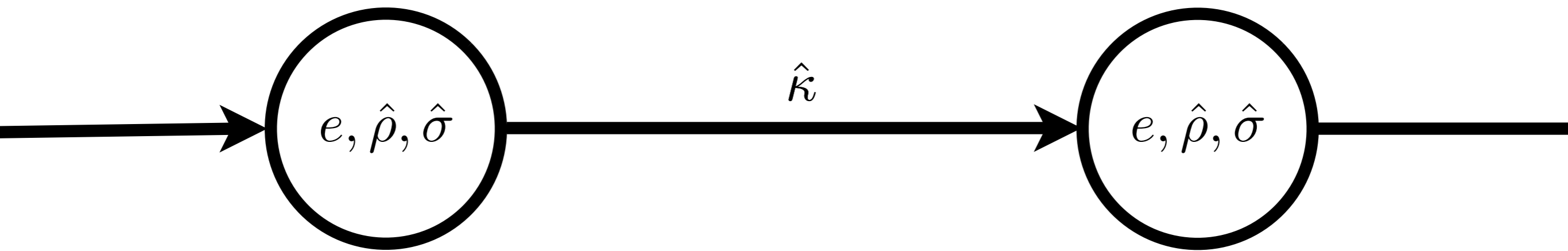
\widehat{CES} \widehat{K}
control state stack

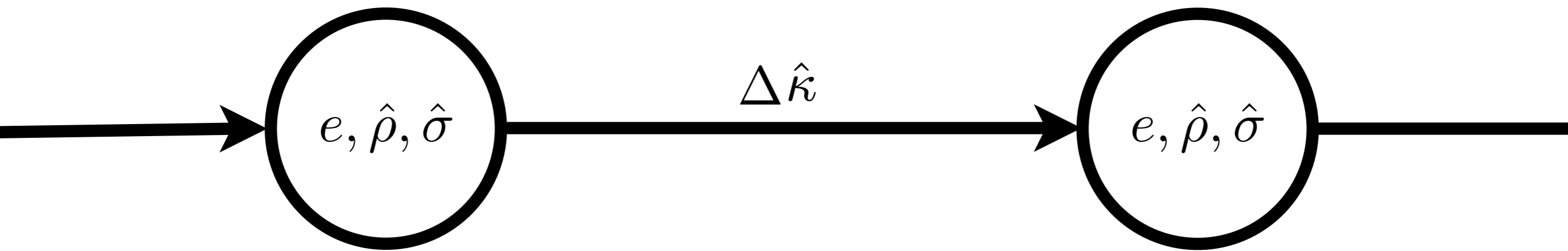
control state + stack = pushdown

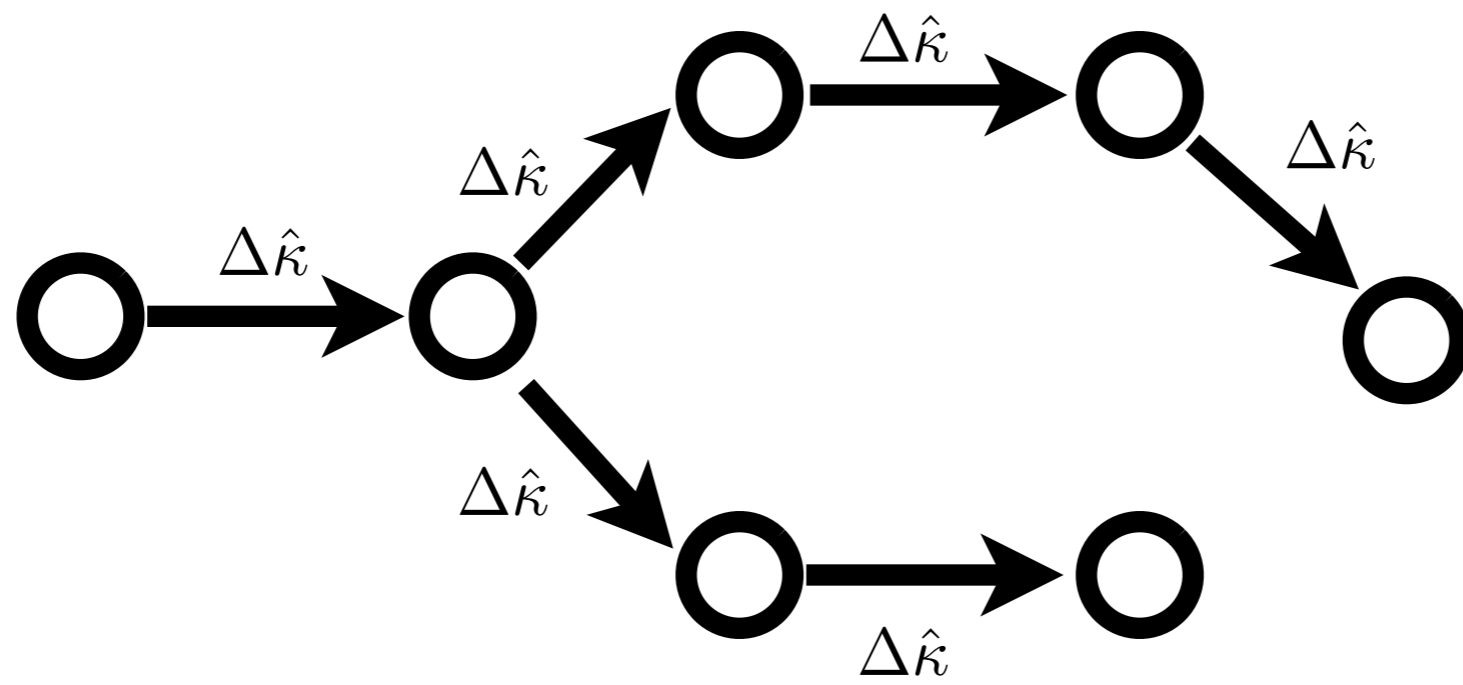
control state + stack = pushdown

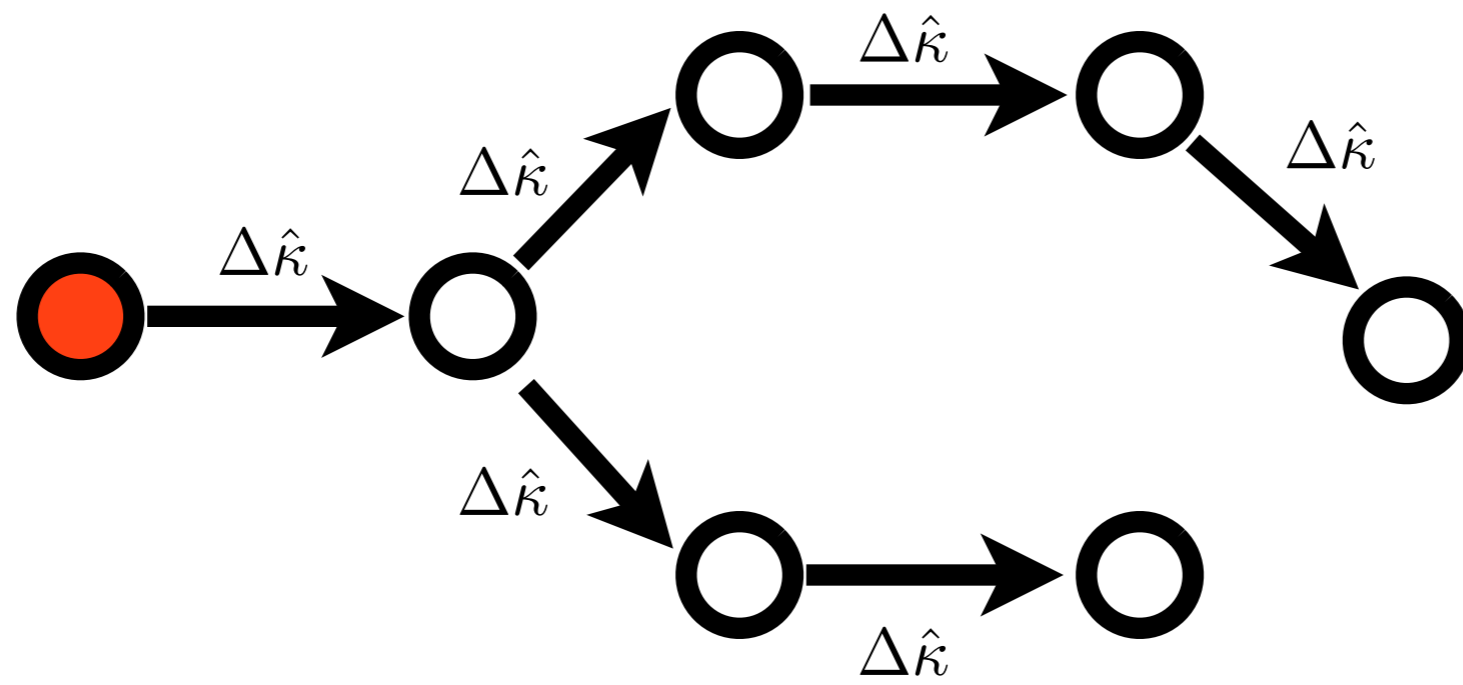
(Vardoulakis and Shivers, CFA2)

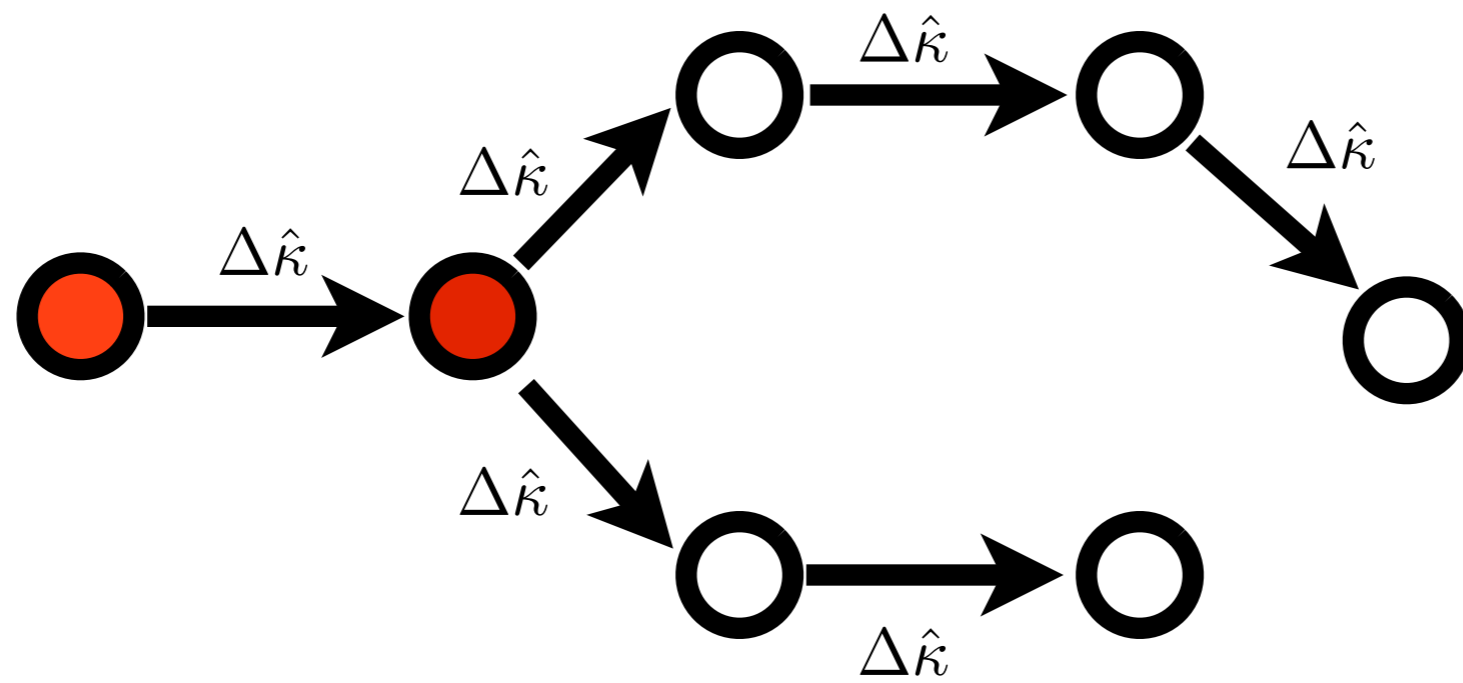


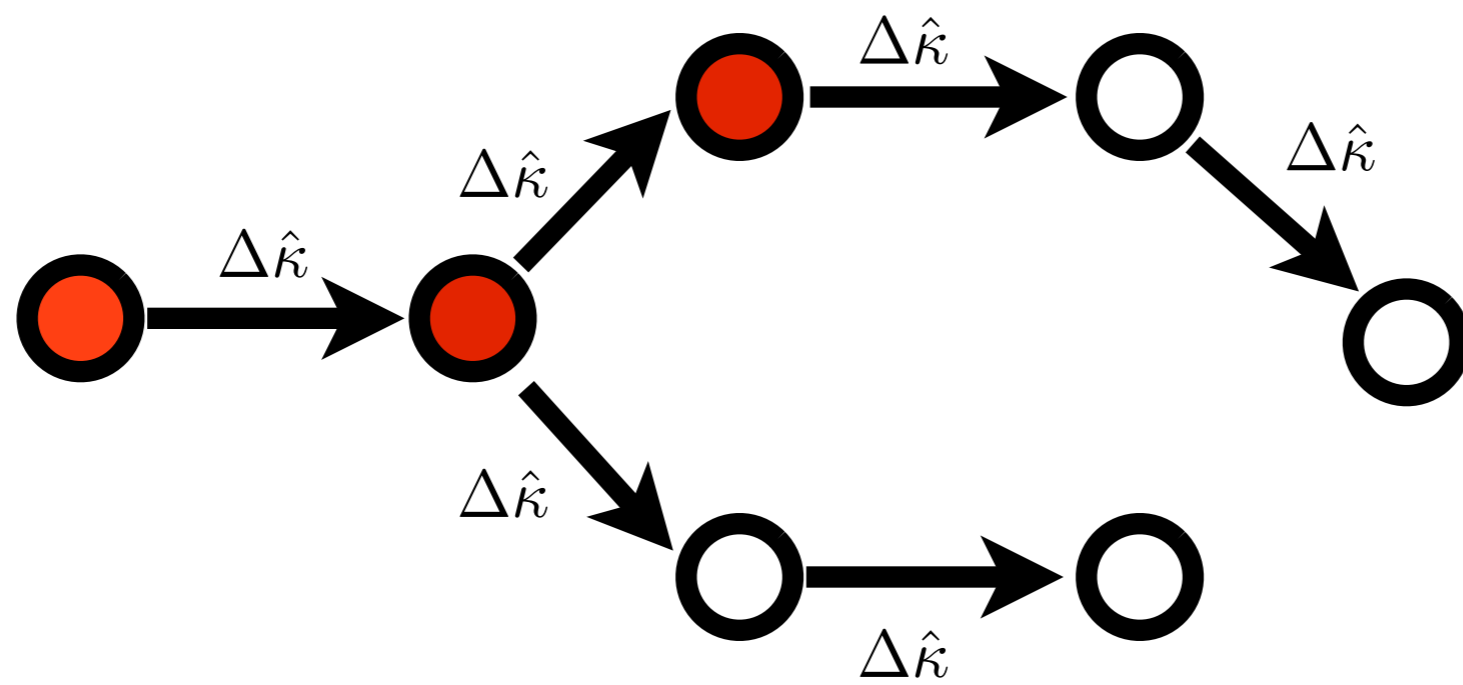


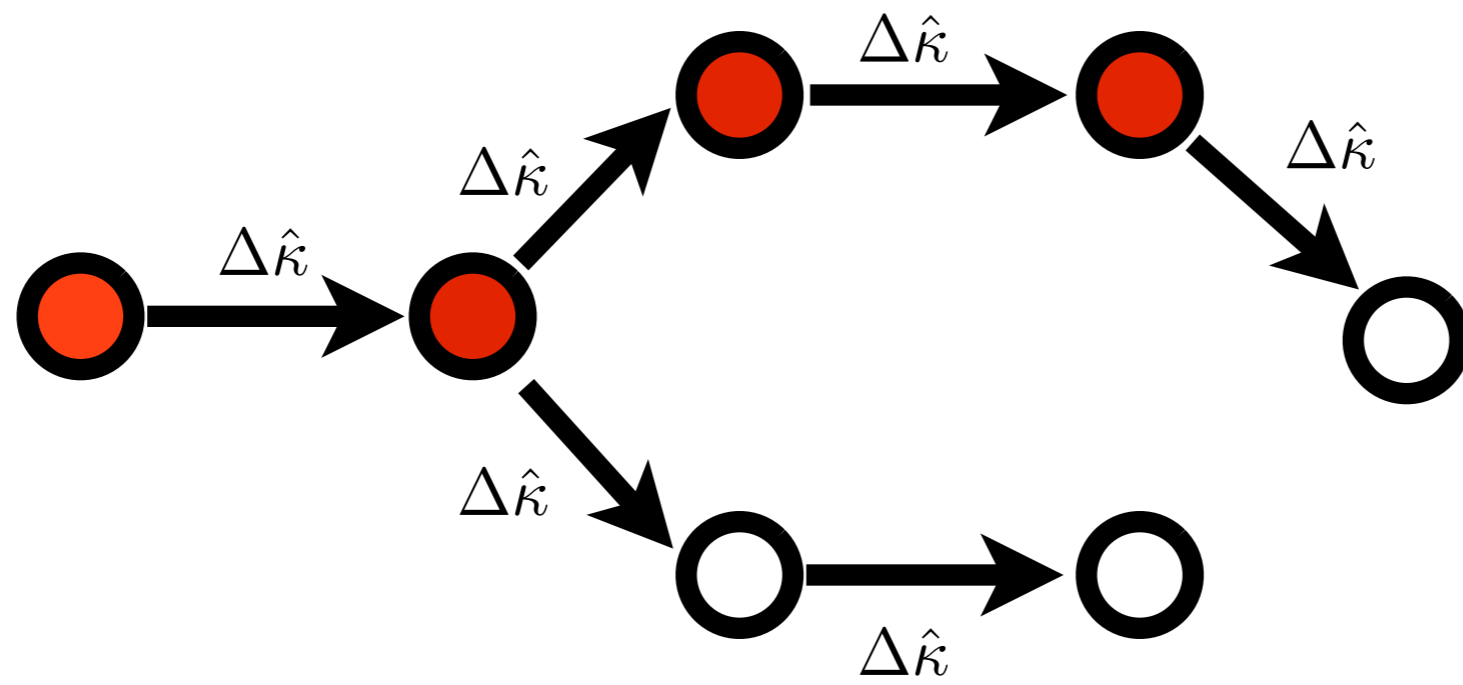


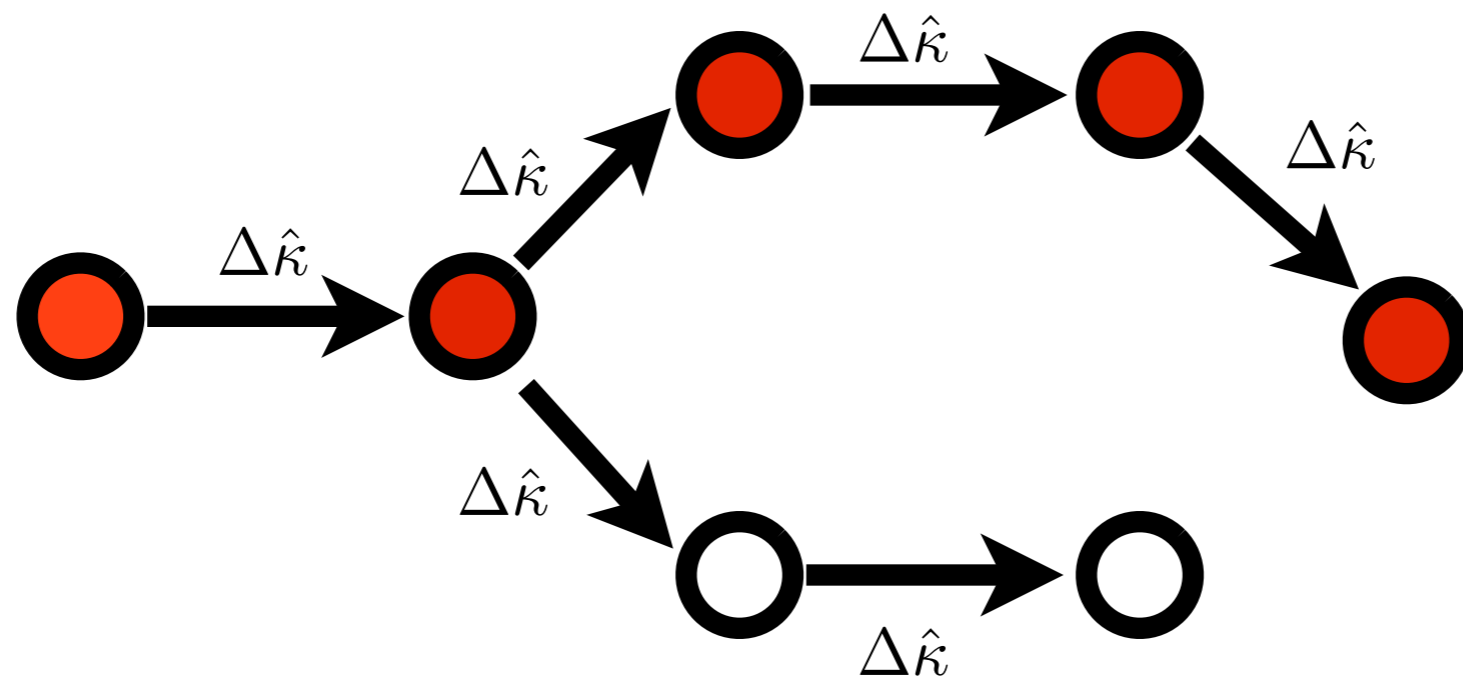


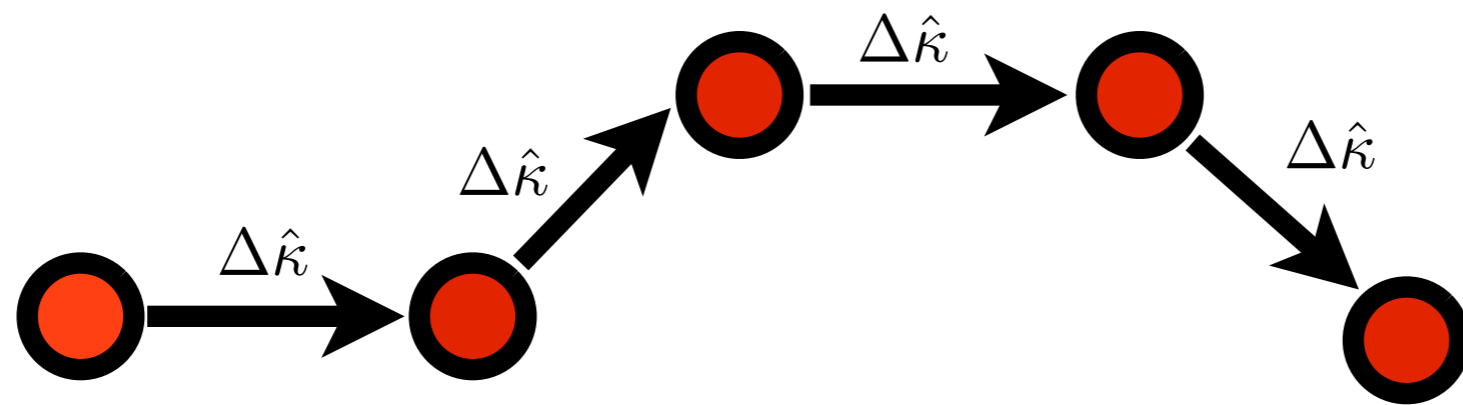


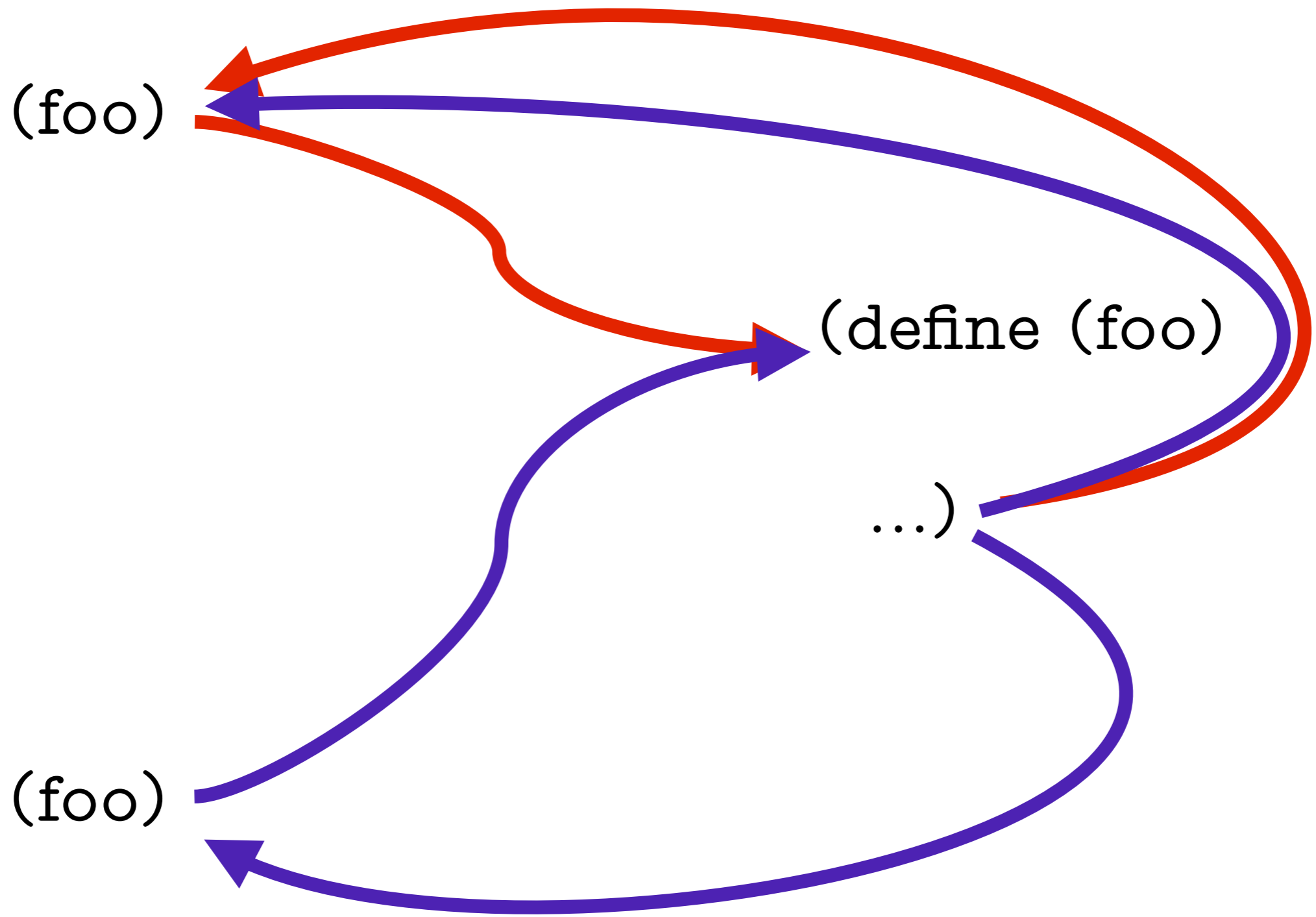










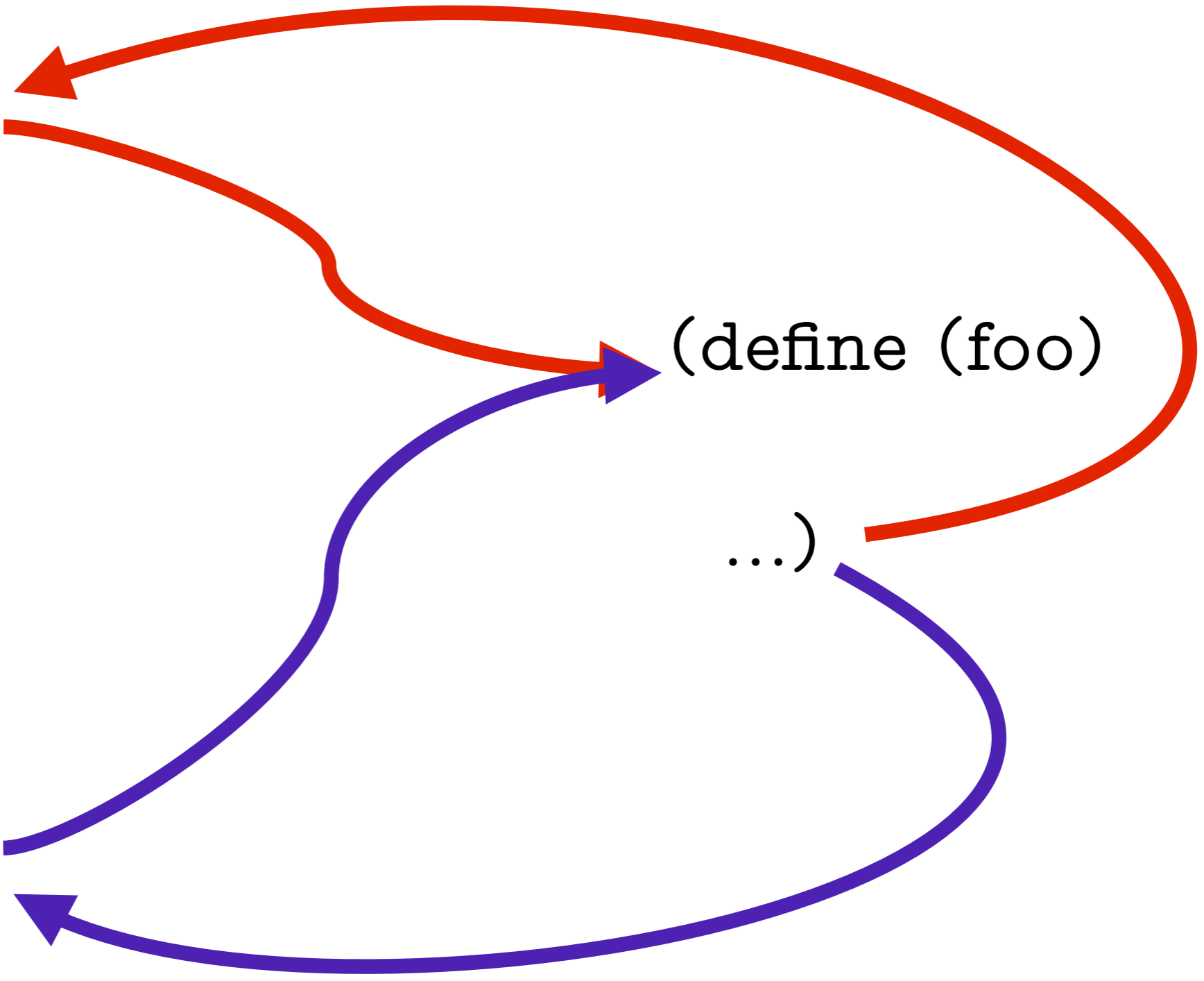


(foo)

(define (foo)

...)

(foo)



GC & pushdown?

Halt!

$$\delta : Q \times \Delta\Gamma \rightarrow \mathcal{P}(Q)$$

Control state



Next states



$$\delta : Q \times \Delta\Gamma \rightarrow \mathcal{P}(Q)$$



Stack change

$$\delta : Q \times \Delta\Gamma \times \Gamma^* \rightarrow \mathcal{P}(Q)$$

Entire stack



$$\delta : Q \times \Delta\Gamma \times \Gamma^* \rightarrow \mathcal{P}(Q)$$

$$\delta : Q \times \Delta\Gamma \times \mathcal{P}(\Gamma^*) \rightarrow \mathcal{P}(Q)$$

Infinite stacks

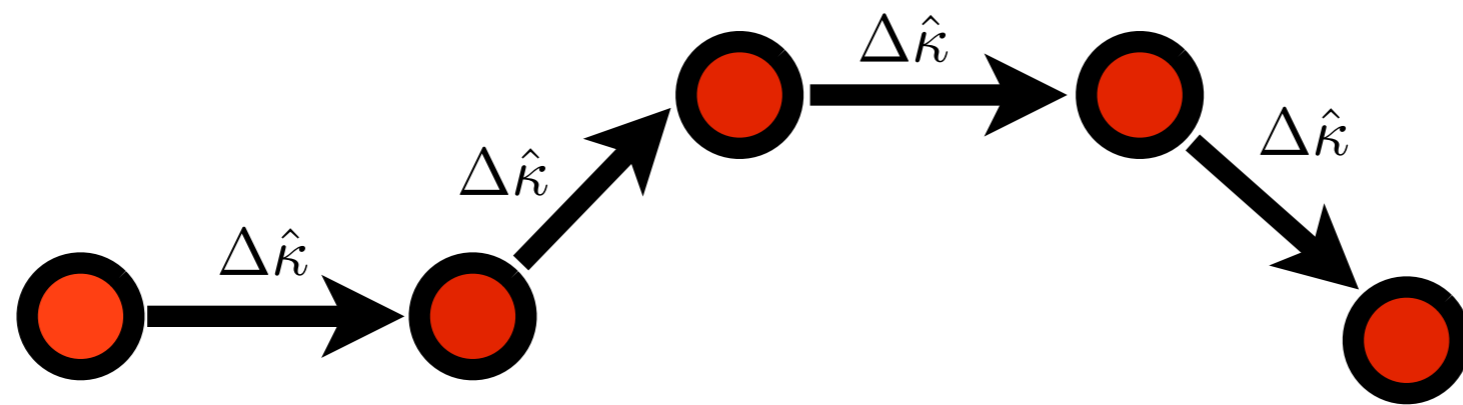


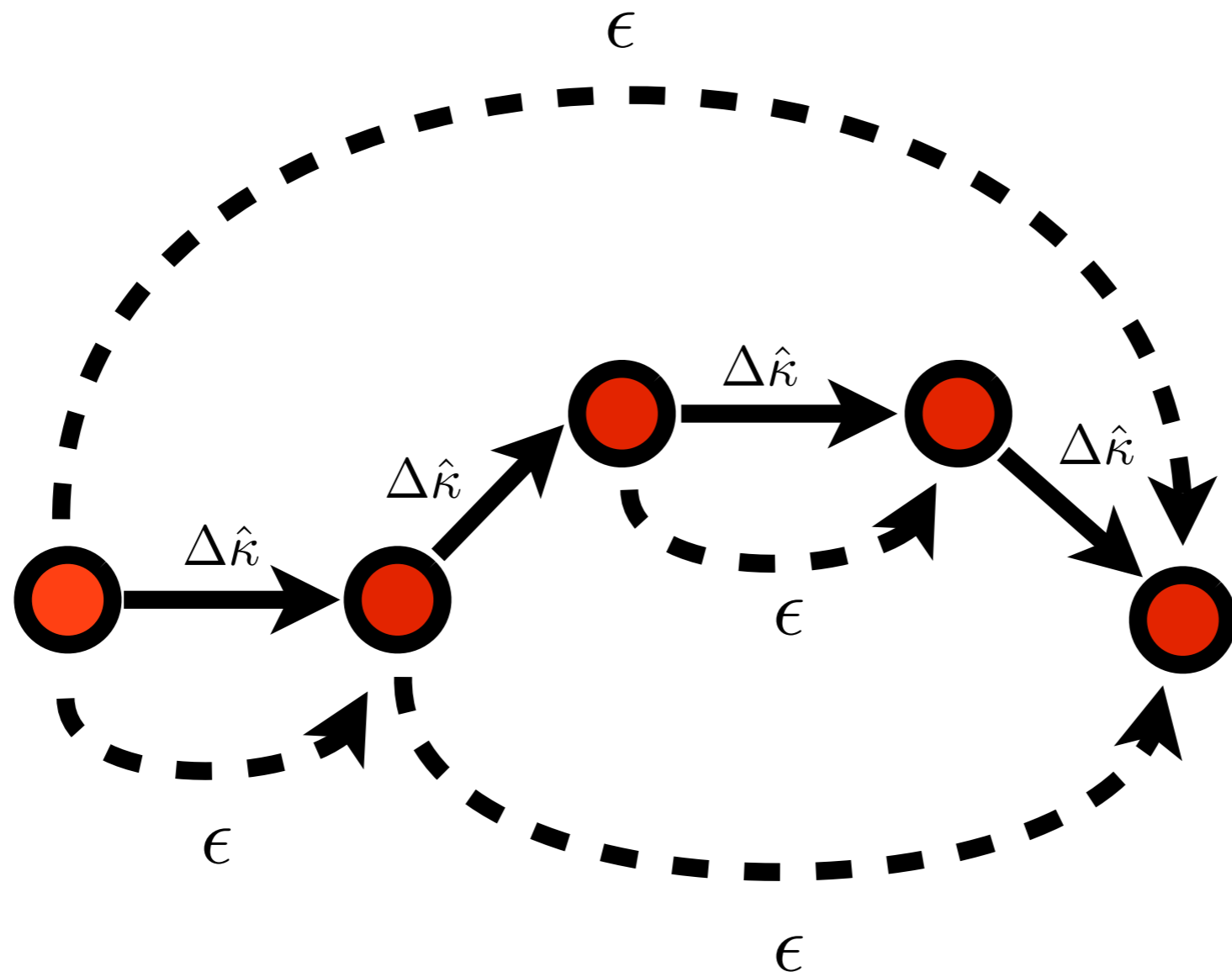
$$\delta : Q \times \Delta\Gamma \times \mathcal{P}(\Gamma^*) \rightarrow \mathcal{P}(Q)$$

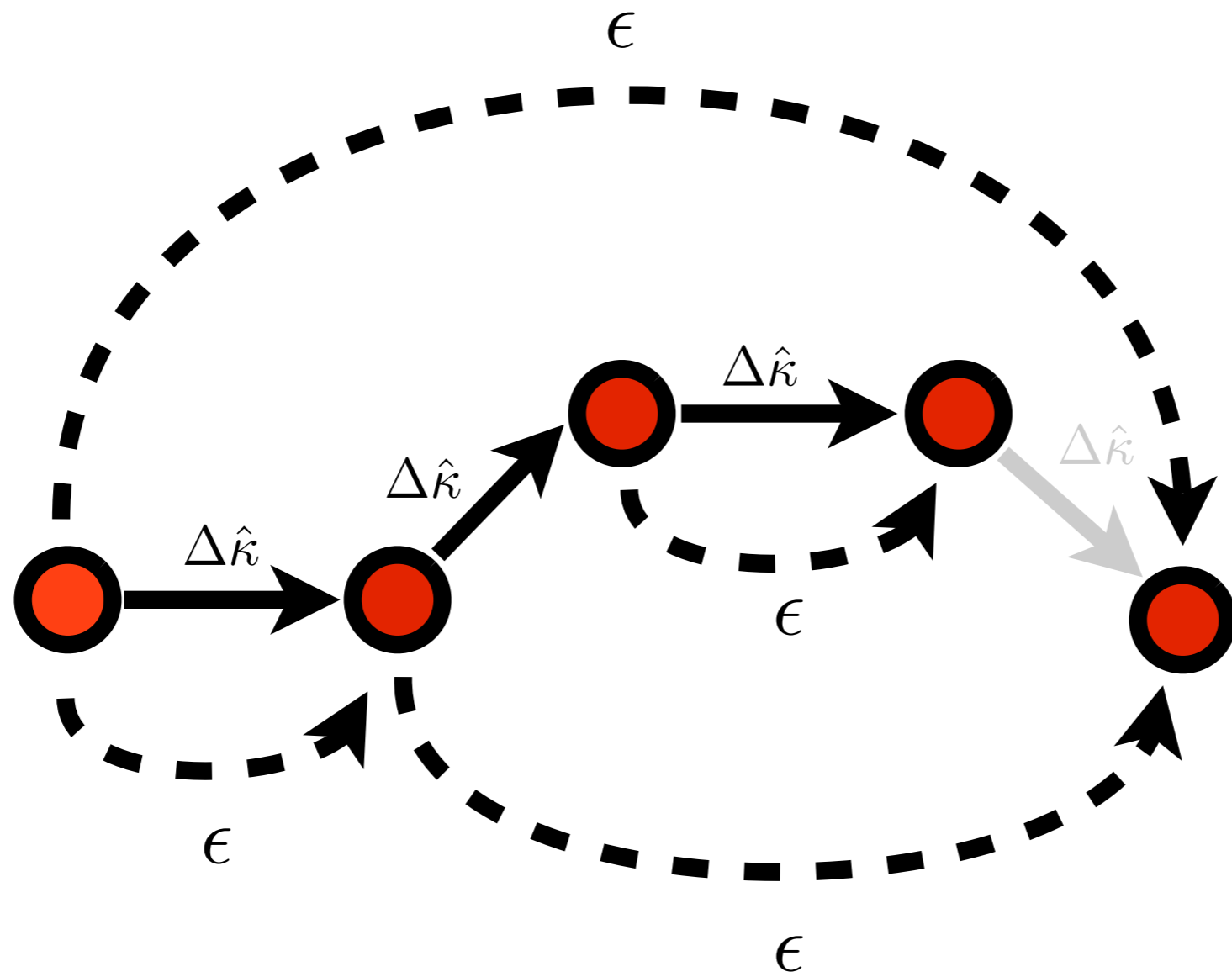
How do we fix the fixes?

$$\delta : Q \times \Delta\Gamma \times \mathcal{P}(\Gamma^*) \rightarrow \mathcal{P}(Q)$$

$$\mathcal{P}(\Gamma^*)$$

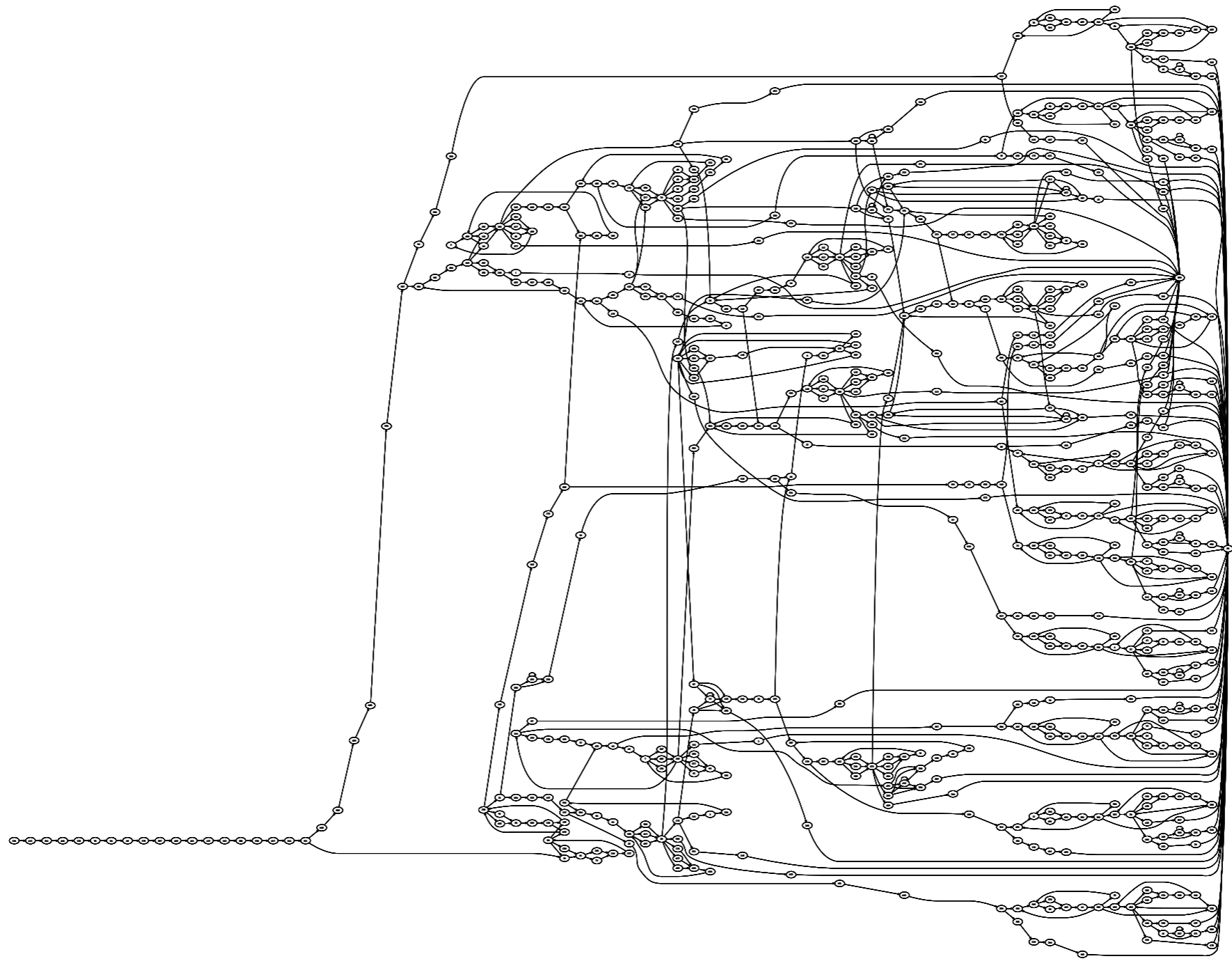


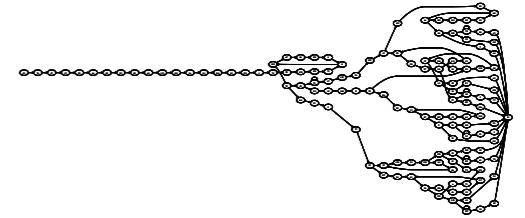
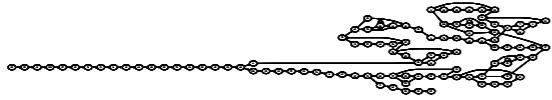
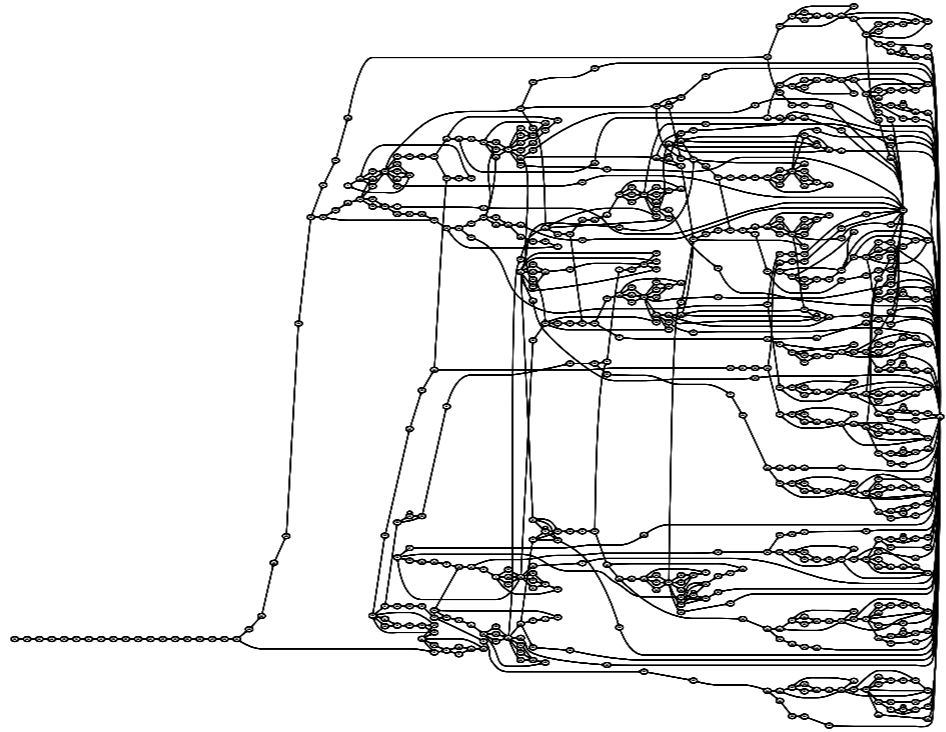


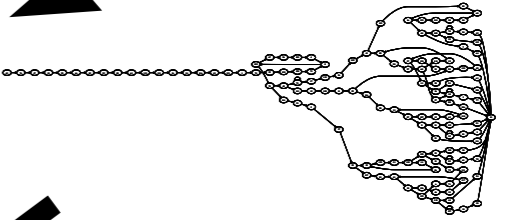
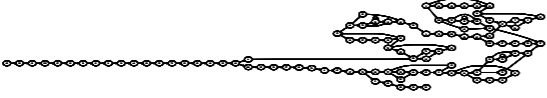
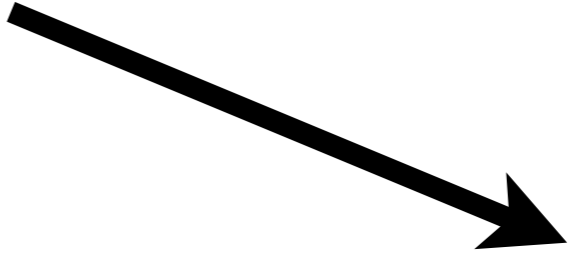
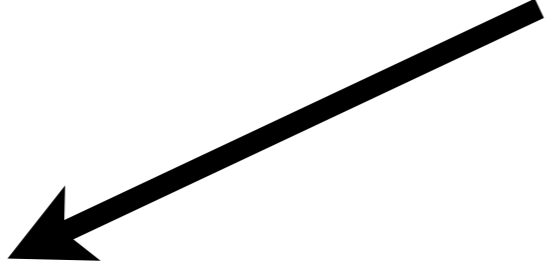
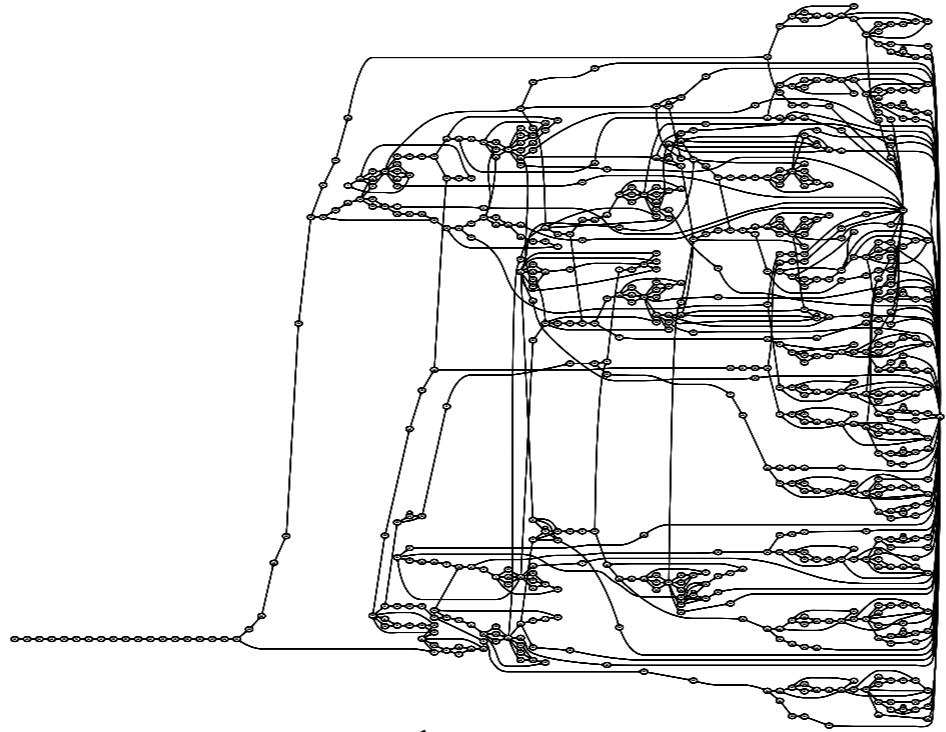


$\mathcal{P}(\Gamma^*)$

Result?







$$O((|\Gamma|m^2) \times (|\Gamma_{\pm}|^3 m^3)) = O(|\Gamma|^4 m^5).$$

$$g \in \Gamma_{\pm} ::= \epsilon \quad (q, \vec{\gamma}) \xrightarrow[M]{\epsilon} (q', \vec{\gamma}) \text{ iff } q \xrightarrow{\epsilon} q' \in \delta$$

$$\begin{array}{l} | \gamma_+ \\ | \gamma_- \end{array} \quad (q, \gamma : \vec{\gamma}) \xrightarrow[M]{\gamma_{\pm}} (q', \vec{\gamma}) \text{ iff } q \xrightarrow{\gamma_{\pm}} q' \in \delta$$

$$(q, \vec{\gamma}) \xrightarrow[M]{\gamma_+} (q', \gamma : \vec{\gamma}) \text{ iff } q \xrightarrow{\gamma_+} q' \in \delta$$

$$c \xrightarrow[M]{g} c' \text{ iff } (q_0, \langle \rangle) \xrightarrow[M]^* c \text{ and } c \xrightarrow[M]{g} c'.$$

$$[\vec{g} \gamma_+ \gamma_- \vec{g}'] = [\vec{g} \vec{g}'] \quad [\vec{g} \epsilon \vec{g}'] = [\vec{g} \vec{g}'].$$

$$[\gamma_+ \gamma'_+ \dots \gamma_+^{(n)}] = \langle \gamma^{(n)}, \dots, \gamma', \gamma \rangle$$

$$q \xrightarrow[M]{g} q' \text{ iff } (q, \vec{\gamma}) \xrightarrow[M]{g} (q', \vec{\gamma}') \text{ for some stacks } \vec{\gamma}, \vec{\gamma}'$$

$$c_0 \xrightarrow[M]{g_1 \dots g_n} c_n \text{ iff } c_0 \xrightarrow[M]{g_1} c_1 \xrightarrow[M]{g_2} \dots \xrightarrow[M]{g_{n-1}} c_{n-1} \xrightarrow[M]{g_n} c_n$$

$$\{q : q_0 \xrightarrow[M]{} q\}$$

$$c \in \text{Conf} = \text{Exp} \times \text{Env} \times \text{Store} \times \text{Kont}$$

$$\rho \in \text{Env} = \text{Var} \rightarrow \text{Addr}$$

$$\sigma \in \text{Store} = \text{Addr} \rightarrow \text{Clo}$$

$$\text{clo} \in \text{Clo} = \text{Lam} \times \text{Env}$$

$$\kappa \in \text{Kont} = \text{Frame}^*$$

$$\phi \in \text{Frame} = \text{Var} \times \text{Exp} \times \text{Env}$$

$$a \in \text{Addr} \text{ is an infinite set of addresses}$$

$$\mathcal{A}(\text{lam}, \rho, \sigma) = (\text{lam}, \rho)$$

$$\mathcal{A}(v, \rho, \sigma) = \sigma(\rho(v))$$

$$\phi \in \text{Frame} = \text{Var} \times \text{Exp} \times \text{Env}$$

$$a \in \text{Addr} \text{ is an infinite set of addresses}$$

$$\mathcal{E}(e) = \{c : \mathcal{I}(e) \Rightarrow^* c\}$$

$$\hat{c}_0 = \hat{\mathcal{I}}(e) = (e, [], [], \langle \rangle).$$

$$e \in \text{Exp} ::= (\text{let } ((v \text{ call})) e)$$

$$| \text{call}$$

$$| \mathfrak{x}$$

$$f, \mathfrak{x} \in \text{Atom} ::= v \mid \text{lam}$$

$$\text{lam} \in \text{Lam} ::= (\lambda (v) e)$$

$$\text{call} \in \text{Call} ::= (f \mathfrak{x})$$

$$v \in \text{Var} \text{ is a set of identifiers}$$

$$\overbrace{([\![\text{let } ((v \text{ call})) e]\!] , \rho, \sigma, \kappa)}^c \Rightarrow \overbrace{(\text{call}, \rho, \sigma, (v, e, \rho) : \kappa)}^{c'}$$

$$\overbrace{([\![f \mathfrak{x}]\!] , \rho, \sigma, \kappa)}^c \Rightarrow \overbrace{(e, \rho'', \sigma', \kappa)}^{c'}, \text{ where}$$

$$([\![\lambda (v) e]\!] , \rho') = \mathcal{A}(f, \rho, \sigma)$$

$$a = \text{alloc}(v, c)$$

$$\rho'' = \rho'[v \mapsto a]$$

$$\sigma' = \sigma[a \mapsto \mathcal{A}(\mathfrak{x}, \rho, \sigma)].$$

$$(\hat{\sigma} \sqcup \hat{\sigma}')(\hat{a}) = \hat{\sigma}(\hat{a}) \cup \hat{\sigma}'(\hat{a}).$$

$$\hat{\mathcal{E}}(e) = \{\hat{c} : \hat{\mathcal{I}}(e) \rightsquigarrow^* \hat{c}\}$$

$$\text{Addr} = \mathbb{N}$$

$$\text{alloc}(v, (e, \rho, \sigma, \kappa)) = 1 + \max(\text{dom}(\sigma)).$$

$$\overbrace{(\mathfrak{x}, \rho, \sigma, (v, e, \rho') : \kappa)}^c \Rightarrow \overbrace{(e, \rho'', \sigma', \kappa)}^{c'}, \text{ where}$$

$$a = \text{alloc}(v, c)$$

$$\rho'' = \rho'[v \mapsto a]$$

$$\sigma' = \sigma[a \mapsto \mathcal{A}(\mathfrak{x}, \rho, \sigma)].$$

$$\hat{c} \in \widehat{\text{Conf}} = \text{Exp} \times \widehat{\text{Env}} \times \widehat{\text{Store}} \times \widehat{\text{Kont}}$$

$$\hat{\rho} \in \widehat{\text{Env}} = \text{Var} \rightarrow \widehat{\text{Addr}}$$

$$\hat{\sigma} \in \widehat{\text{Store}} = \widehat{\text{Addr}} \rightarrow \mathcal{P}(\widehat{\text{Clo}})$$

$$\widehat{\text{clo}} \in \widehat{\text{Clo}} = \text{Lam} \times \widehat{\text{Env}}$$

$$\hat{\kappa} \in \widehat{\text{Kont}} = \widehat{\text{Frame}}^*$$

$$\hat{\phi} \in \widehat{\text{Frame}} = \text{Var} \times \text{Exp} \times \widehat{\text{Env}}$$

$$\hat{a} \in \widehat{\text{Addr}} \text{ is a finite set of addresses}$$

$$\overbrace{([\![\text{let } ((v \text{ call})) e]\!] , \hat{\rho}, \hat{\sigma}, \hat{\kappa})}^{\hat{c}} \rightsquigarrow \overbrace{(\text{call}, \hat{\rho}, \hat{\sigma}, (v, e, \hat{\rho}) : \hat{\kappa})}^{\hat{c}'}$$

$$\overbrace{([\![f \mathfrak{x}]\!] , \hat{\rho}, \hat{\sigma}, \hat{\kappa})}^{\hat{c}} \rightsquigarrow \overbrace{(e, \hat{\rho}'', \hat{\sigma}', \hat{\kappa})}^{\hat{c}'}, \text{ where}$$

$$([\![\lambda (v) e]\!] , \hat{\rho}') \in \hat{\mathcal{A}}(f, \hat{\rho}, \hat{\sigma})$$

$$\hat{a} = \widehat{\text{alloc}}(v, \hat{c})$$

$$\hat{\rho}'' = \hat{\rho}'[v \mapsto \hat{a}]$$

$$\hat{\sigma}' = \hat{\sigma} \sqcup [\hat{a} \mapsto \hat{\mathcal{A}}(\mathfrak{x}, \hat{\rho}, \hat{\sigma})].$$

$$\widehat{\text{Addr}} = \text{Var}$$

$$\text{alloc}(v, \hat{c}) = v.$$

$$\widehat{\text{Addr}} = \text{Var} \times \text{Exp}$$

$$\text{alloc}(v, (e, \hat{\rho}, \hat{\sigma}, \hat{\kappa})) = (v, e).$$

$$\widehat{\text{Addr}} = \text{Var} + \text{Var} \times \text{Exp}$$

$$\text{alloc}(v, ([[f \mathfrak{x}]] , \hat{\rho}, \hat{\sigma}, \hat{\kappa})) = \begin{cases} (v, [[f \mathfrak{x}]]) & f \text{ is let-bound} \\ v & \text{otherwise.} \end{cases}$$

$$(\text{lam}, \hat{\rho}) \sqsubseteq (\text{lam}, \hat{\rho}') \text{ iff } \hat{\rho} \sqsubseteq \hat{\rho}';$$

$$(v, e, \hat{\rho}) \sqsubseteq (v, e, \hat{\rho}') \text{ iff } \hat{\rho} \sqsubseteq \hat{\rho}';$$

$$\widehat{\text{Addr}} = \text{Var} \times \text{Exp}^k$$

$$\widehat{\text{alloc}}(v, \langle (e_1, \hat{\rho}_1, \hat{\sigma}_1, \hat{\kappa}_1), \dots \rangle) = (v, \langle e_1, \dots, e_k \rangle).$$

$$(e, \hat{\rho}, \hat{\sigma}, \hat{\kappa}) \sqsubseteq (e, \hat{\rho}', \hat{\sigma}', \hat{\kappa}') \text{ iff } \hat{\rho} \sqsubseteq \hat{\rho}' \text{ and } \hat{\sigma} \sqsubseteq \hat{\sigma}' \text{ and } \hat{\kappa} \sqsubseteq \hat{\kappa}'.$$

$$\alpha(e, \rho, \sigma, \kappa) = (e, \alpha(\rho), \alpha(\sigma), \alpha(\kappa))$$

$$\alpha(\rho) = \lambda v. \alpha(\rho(v))$$

$$\alpha(\sigma) = \lambda \hat{a}. \bigsqcup_{\alpha(a) = \hat{a}} \{\alpha(\sigma(a))\}$$

$$\alpha\langle \phi_1, \dots, \phi_n \rangle = \langle \alpha(\phi_1), \dots, \alpha(\phi_n) \rangle$$

$$\alpha(v, e, \rho) = (v, e, \alpha(\rho))$$

$$\alpha(a) \text{ is determined by the allocation functions.}$$

Theorem 4.1. *If:*

$$\alpha(c) \sqsubseteq \hat{c} \text{ and } c \Rightarrow c',$$

then there must exist $\hat{c}' \in \widehat{\text{Conf}}$ such that:

$$\alpha(c') \sqsubseteq \hat{c}' \text{ and } \hat{c} \rightsquigarrow \hat{c}'.$$

$$\hat{G}(e, \hat{\rho}, \hat{\sigma}, \hat{\kappa}) = (e, \hat{\rho}, \hat{\sigma} | \text{Reachable}(\hat{c}), \hat{\kappa}),$$

$$\text{Reachable}(\hat{c}) = \{\hat{a} : \hat{a}_0 \in \text{Root}(\hat{c}) \text{ and } \hat{a}_0 \xrightarrow[\hat{\sigma}]{*} \hat{a}\}$$

$$\widehat{\mathcal{PDS}}(e) = (Q, \Gamma, \delta, q_0), \text{ where}$$

$$Q = \text{Exp} \times \widehat{\text{Env}} \times \widehat{\text{Store}}$$

$$\Gamma = \widehat{\text{Frame}}$$

$$(q, \epsilon, q') \in \delta \text{ iff } (q, \hat{\kappa}) \rightsquigarrow (q', \hat{\kappa}) \text{ for all } \hat{\kappa}$$

$$(q, \hat{\phi}_-, q') \in \delta \text{ iff } (q, \hat{\phi} : \hat{\kappa}) \rightsquigarrow (q', \hat{\kappa}) \text{ for all } \hat{\kappa}$$

$$(q, \hat{\phi}_+, q') \in \delta \text{ iff } (q, \hat{\kappa}) \rightsquigarrow (q', \hat{\phi}' : \hat{\kappa}) \text{ for all } \hat{\kappa}$$

$$(q_0, \langle \rangle) = \hat{\mathcal{I}}(e).$$

$$(\rightsquigarrow_{\text{GC}}) = (\rightsquigarrow) \circ \hat{G}$$

$$\hat{a} \xrightarrow[\hat{\sigma}]{*} \hat{a}' \text{ iff there exists } (\text{lam}, \hat{\rho}) \in \hat{\sigma}(\hat{a}) \text{ such that } \hat{a}' \in \text{range}(\hat{\rho}).$$

$$q \xrightarrow[M]{g} q' \text{ iff there exists } \vec{g} \text{ such that } q_0 \xrightarrow[M]{\vec{g}} q \text{ and } (q, [\vec{g}], g, q') \in \delta,$$

$$\text{where } q \xrightarrow[M]{g_1 \dots g_n} q' \text{ iff } q \xrightarrow[M]{g_1} q_1 \xrightarrow[M]{g_2} \dots \xrightarrow[M]{g_n} q'.$$

$$q_0 \xrightarrow[M]{g} q \text{ iff } (q_0, \langle \rangle, g, q) \in \delta.$$

$$S = \{q : q_0 \xrightarrow[M]{\vec{g}} q \text{ for some stack-action sequence } \vec{g}\}$$

$$E = \{q \xrightarrow[M]{g} q' : q \xrightarrow[M]{g} q'\}$$

$$\text{Stacks}(\overbrace{S, \Gamma, E, s_0}^M)(s) = (S, \Gamma, \delta, s_0, \{s\}), \text{ where}$$

$$(s', \gamma, s'') \in \delta \text{ if } (s', \gamma_+, s'') \in E$$

$$(s', \epsilon, s'') \in \delta \text{ if } s' \xrightarrow[M]{\vec{g}} s'' \text{ and } [\vec{g}] = \epsilon.$$

$$\mathcal{F}(M) = f, \text{ where}$$

$$M = (Q, \Gamma, \delta, q_0)$$

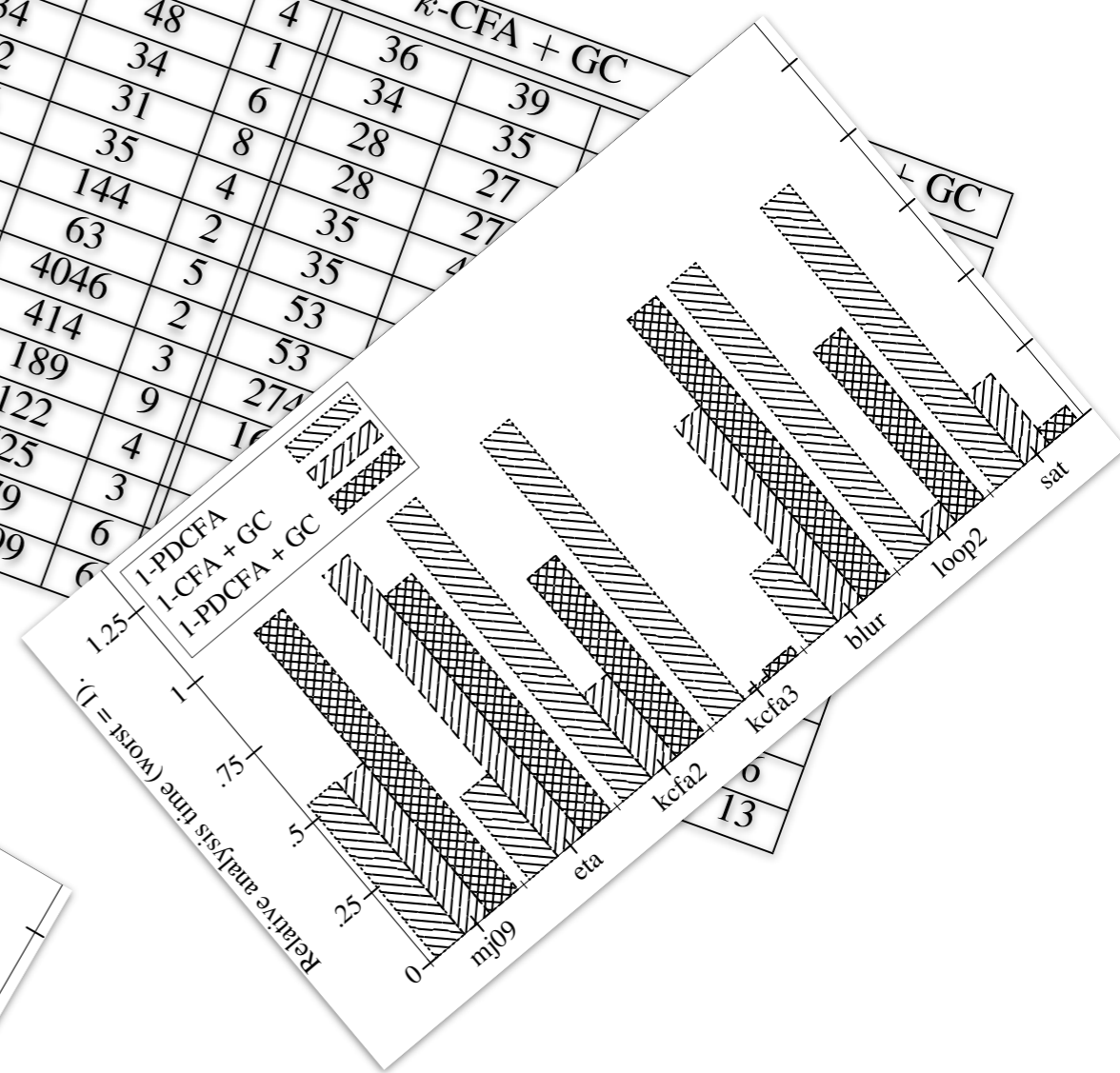
$$f(S, \Gamma, E, s_0) = (S', \Gamma, E', s_0), \text{ where}$$

$$S' = S \cup \{s' : s \in S \text{ and } s \xrightarrow[M]{} s'\} \cup \{s_0\}$$

$$E' = E \cup \{s \xrightarrow[M]{g} s' : s \in S \text{ and } s \xrightarrow[M]{g} s'\}.$$

Theorem 7.1. $\text{DSG}(M) = \text{lfp}(\mathcal{F}(M)).$

Program	Exp	Var	k	k-CFA				k-PDCFA				k-CFA + GC	
				Time	Time	Time	Time	Time	Time	Time	Time	Time	Time
mj09	19	8	0	83	107	4	38	38	4	36	39		
eta	21	13	1	454	812	1	44	48	1	34	35		
kcfa2	20	10	0	63	74	4	34	34	6	28	27		
kcfa3	25	13	1	33	33	8	32	31	8	28	27		
blur	40	20	0	194	236	3	36	35	4	35	35		
loop2	41	14	1	970	1935	1	87	87	2	53	53		
sat	63	31	0	> 7119	> 14201	4	58	144	5	35	35		
			1	261	> 2435	≤ 1	1761	63	2	53	53		
			0	228	340	≤ 3	280	4046	3	274	274		
			1	> 10867	252	9	177	414	4	16	16		
			0	> 5362	> 16040	4	113	189	3	16	16		
			1	> 8395	> 7610	≤ 3	411	122	9	274	274		
			0	> 12391	> 12391	≤ 6	775	525	4	16	16		
			1	> 12391	> 12391	≤ 6	7979	979	3	16	16		
			0	> 10299	> 10299	6	10299	10299	6	16	16		



<http://github.com/ilyasergey/reachability>

<http://github.com/ilyasergey/reachability>



Finitization is double-edged.

Progress attacks finitization.

We can skirt inside decidability.

Tak.

Complexity?

- Monovariant, global store? Polynomial.
- Polyvariant? Exponential.
- Flat environments, global store? Polynomial.

Alternative?

$$\delta : Q \times \Delta\Gamma \times \mathcal{P}(\Gamma^*) \rightarrow \mathcal{P}(Q)$$

$$\delta : Q \times \Delta\Gamma \times \mathcal{P}(\Delta\Gamma) \rightarrow \mathcal{P}(Q)$$

$$\delta : Q \times \mathcal{P}(\Delta\Gamma) \times \Delta\Gamma \rightarrow \mathcal{P}(Q)$$

control states

$$\delta : Q \times \mathcal{P}(\Delta\Gamma) \times \Delta\Gamma \rightarrow \mathcal{P}(Q)$$