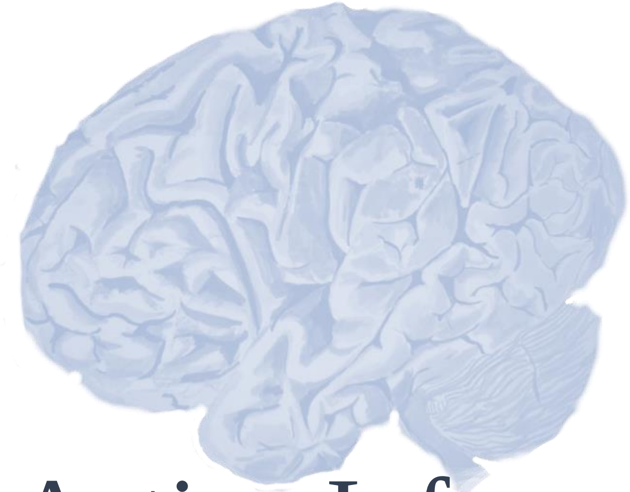
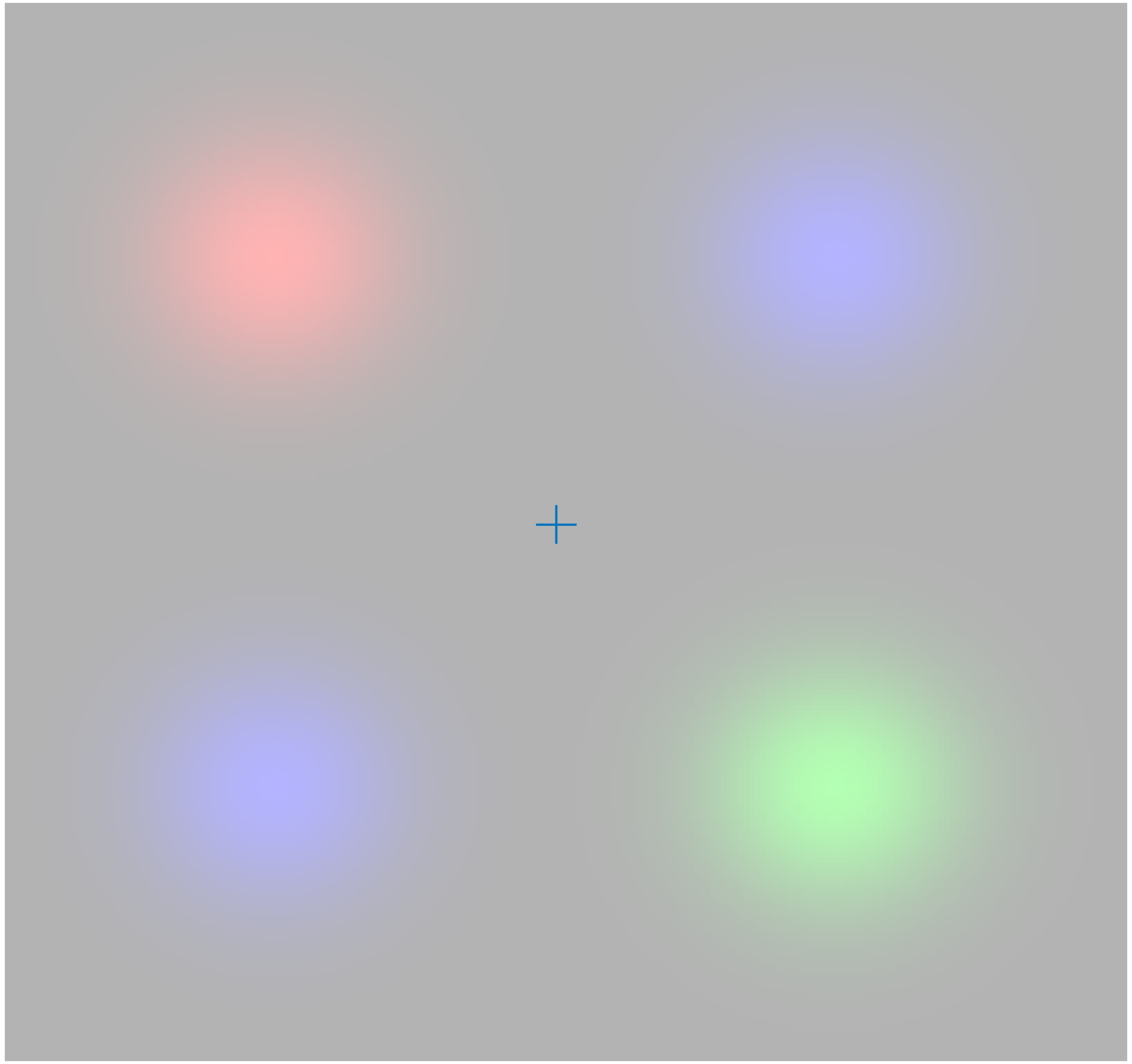
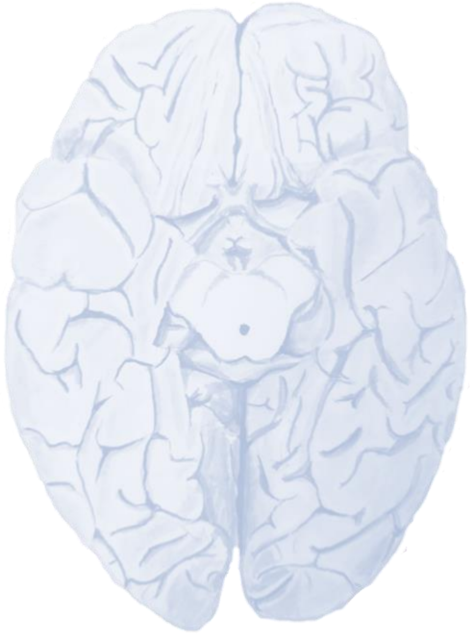


Active Inference

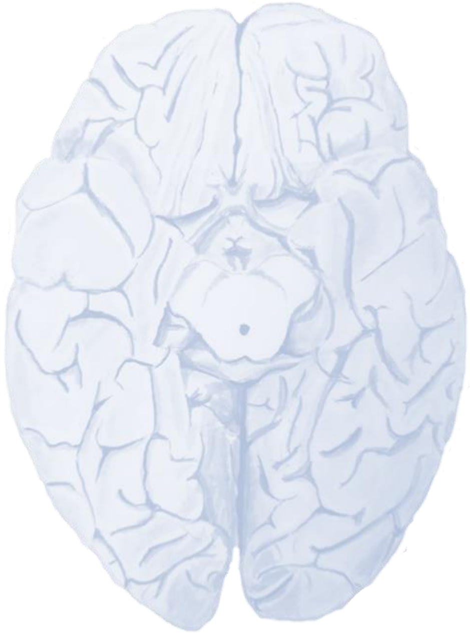
Annual Conference of the Australasian Bayesian Network Modelling Society





Active Inference

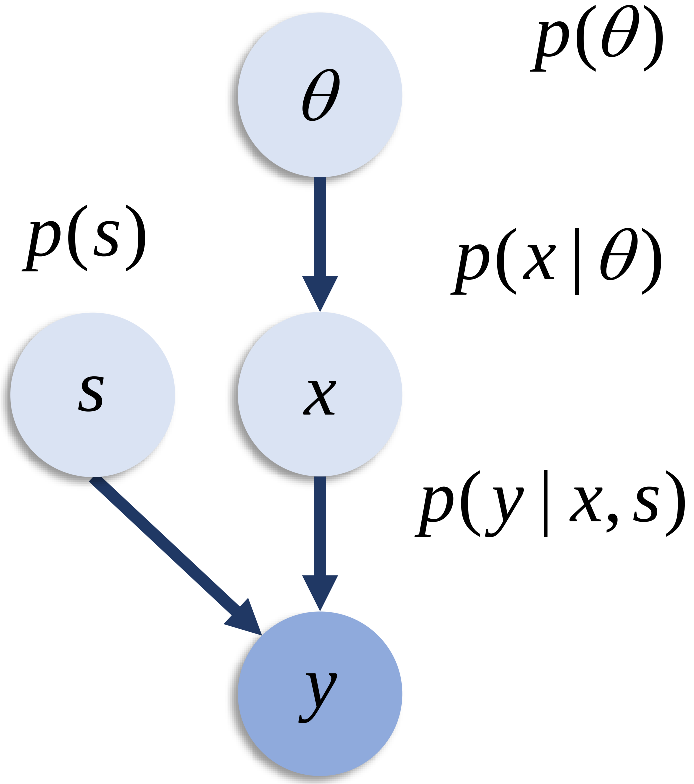
Self-evidencing
Message passing
Discrete time (planning)
Continuous time (movement)
Hierarchical models
Summary



Active Inference

Self-evidencing
Message passing
Discrete time (planning)
Continuous time (movement)
Hierarchical models
Summary

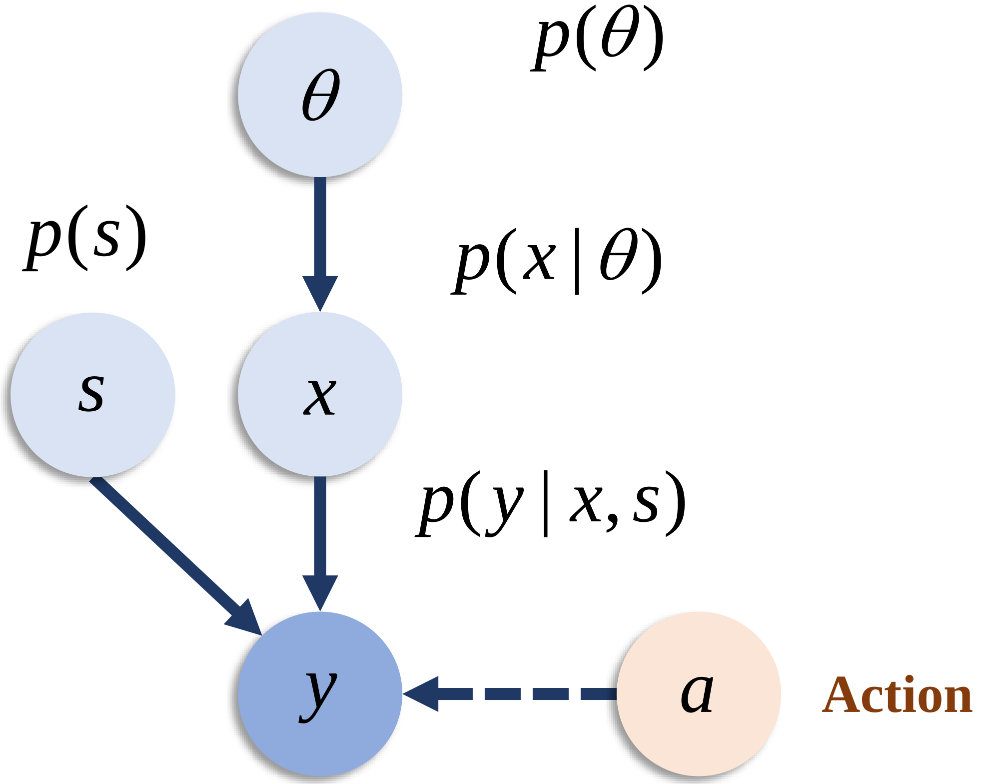
Self-evidencing

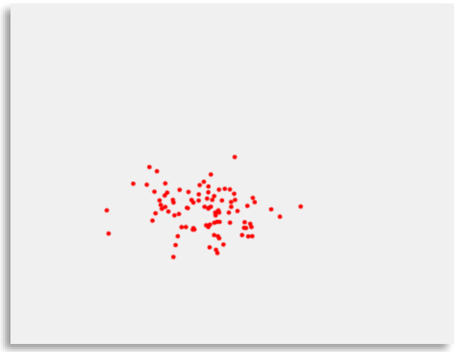


$$p(y) = \sum_s \iint p(y, x, s, \theta) dx d\theta$$ **Evidence**

$$p(y, x, s, \theta) = p(y \mid x, s) p(s) p(x \mid \theta) p(\theta)$$ **Model**

Self-evidencing

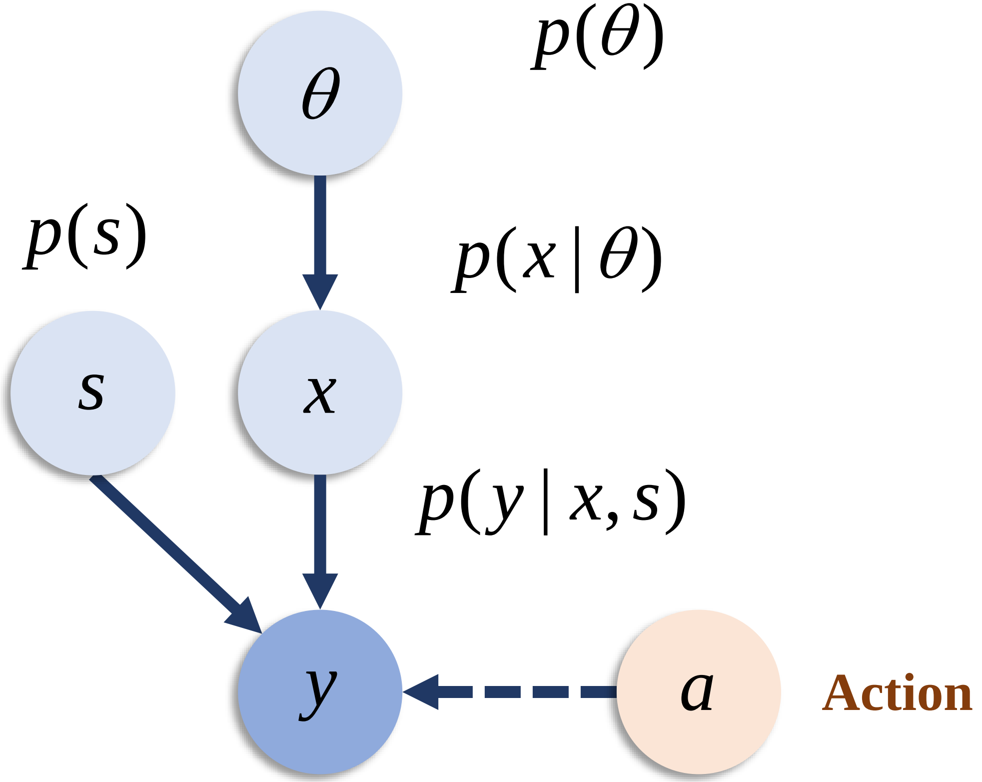


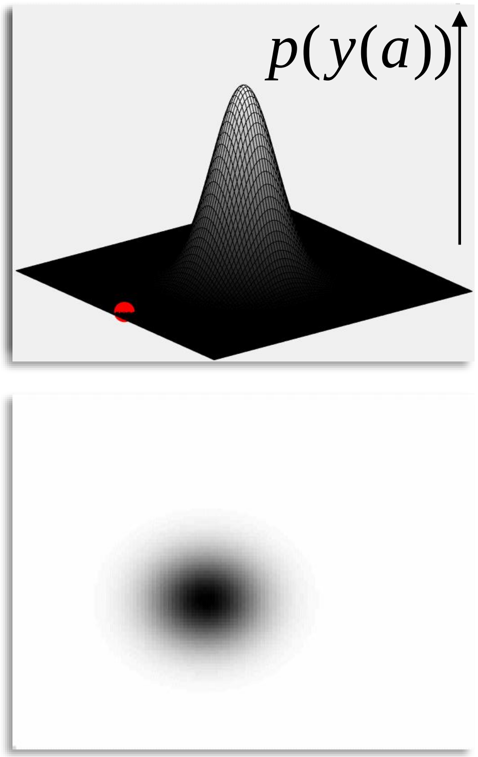


$$p(y(a)) = \sum_s \iint p(y(a), x, s, \theta)\, dx\, d\theta$$ **Evidence**

$$p(y(a), x, s, \theta) = p(y(a)|x, s)\, p(s)\, p(x|\theta)\, p(\theta)$$ **Model**

Self-evidencing





$$p(y(a)) = \sum_s \iint p(y(a), x, s, \theta)\, dx\, d\theta$$ **Evidence**

$$p(y(a), x, s, \theta) = p(y(a)|x,s)\, p(s)\, p(x|\theta)\, p(\theta)$$ **Model**

Free energy

$$F(y(a)) = \underbrace{E_q\left[\ln q(x,s,\theta) - \ln p(x,s,\theta \mid y)\right]}_{=D_{KL}\left[q(x,s,\theta)\,\|\,p(x,s,\theta \mid y)\right]\geq 0} - \ln p(y(a))$$

Variational inference

$$q(x,s,\theta) = \arg\min_q F(y(a)) \approx p(x,s,\theta \mid y)$$

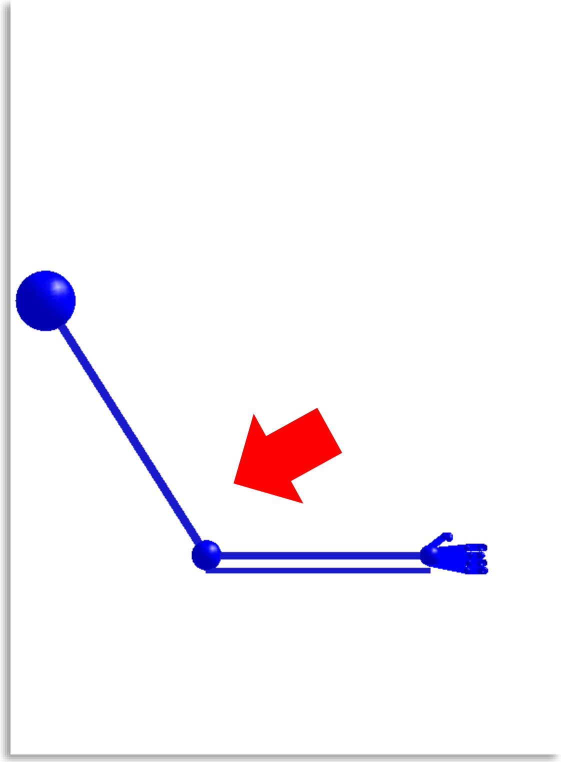
$$\Rightarrow F(y(a)) \approx -\ln p(y(a))$$

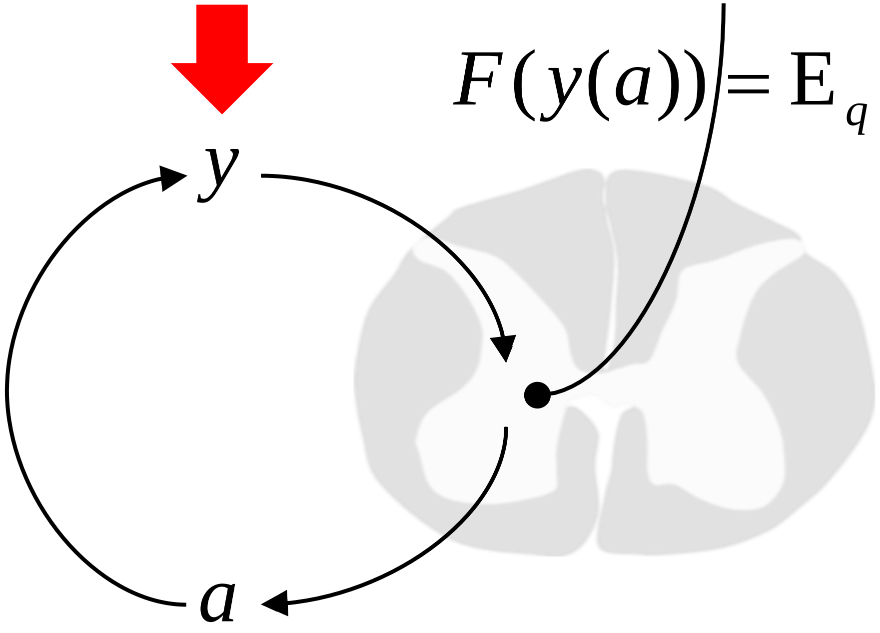
Active inference

$$q(x,s,\theta) = \arg\min_q F(y(a)) \approx p(x,s,\theta \mid y)$$

$$a = \arg\min_a F(y(a)) \approx \arg\max_a p(y(a))$$

Reflexes



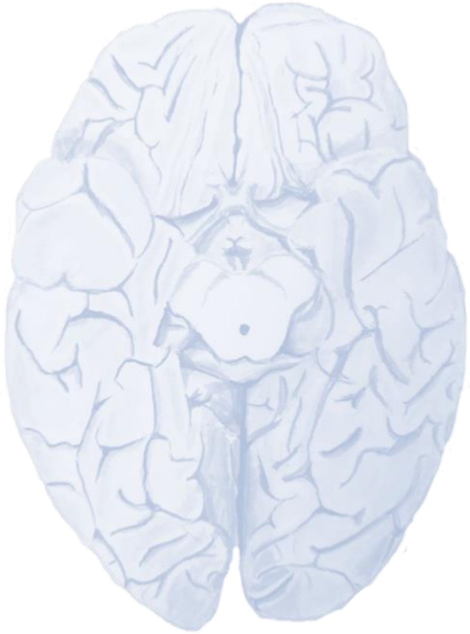


$$F(y(a)) = E_q[\ln p(y(a) \mid x, s, \theta)] + \ldots$$

Active inference

$$q(x, s, \theta) = \underset{q}{\arg\min}\, F(y(a)) \approx p(x, s, \theta \mid y)$$

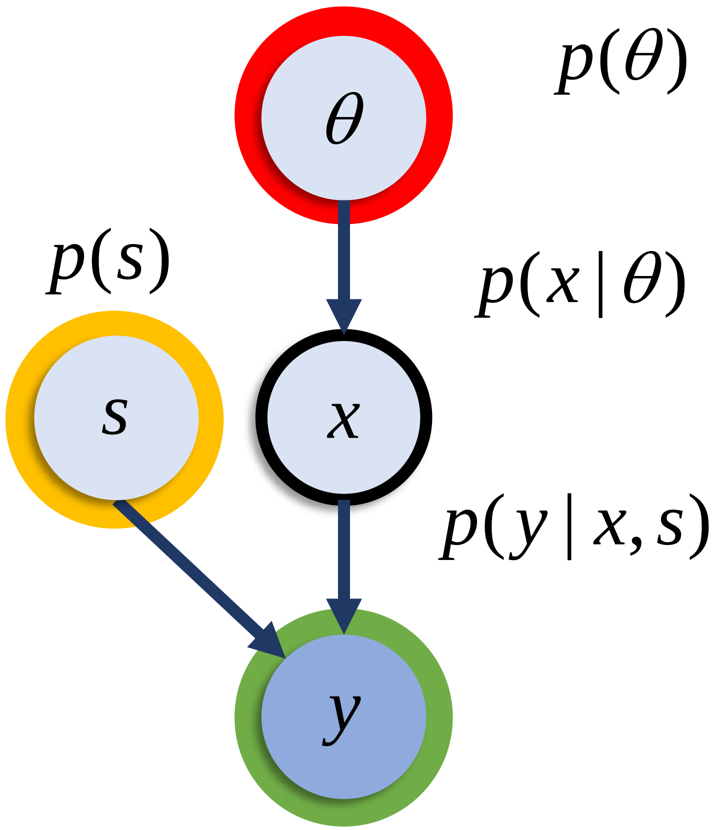
$$a = \underset{a}{\arg\min}\, F(y(a)) \approx \underset{a}{\arg\max}\; p(y(a))$$



Active Inference

Self-evidencing
Message passing
Discrete time (planning)
Sequential time (movement)
Hierarchical models
Summary

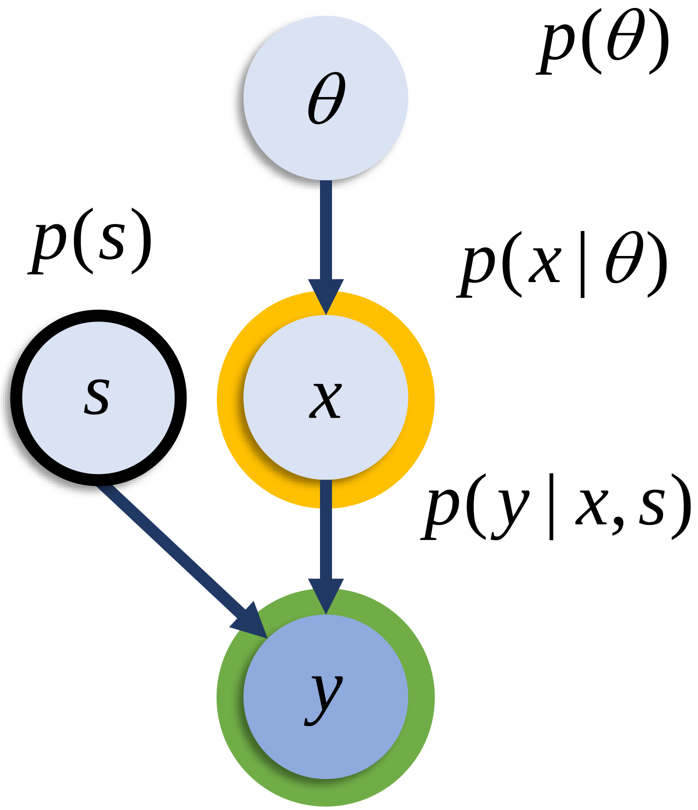
Markov blankets



Parents
Children
Co-parents

$$p(\mu,\eta|b) = p(\mu|b)\,p(\eta|b) \Leftrightarrow \mu \perp \eta | b$$

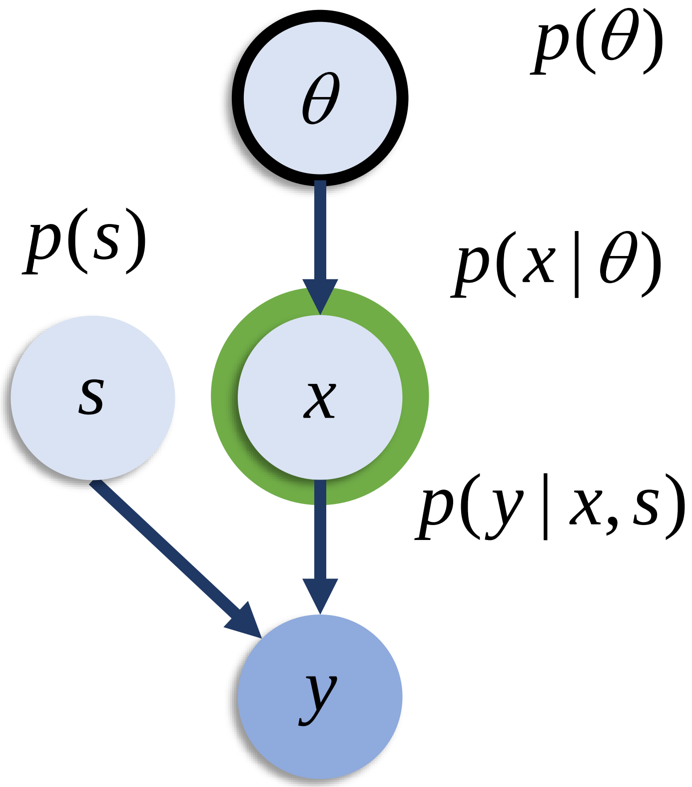
Markov blankets



Parents
Children
Co-parents

$$p(\mu,\eta|b)=p(\mu|b)\,p(\eta|b)\Leftrightarrow\mu\perp\eta|b$$

Markov blankets

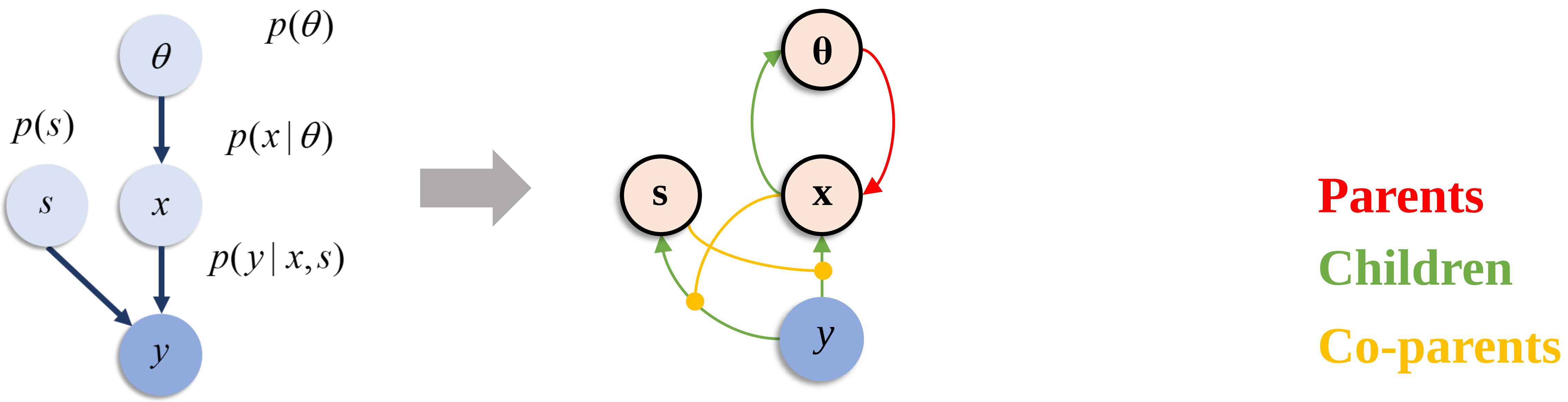


Parents

Children

Co-parents

$$p(\mu,\eta|b) = p(\mu|b)\,p(\eta|b) \Leftrightarrow \mu \perp \eta | b$$

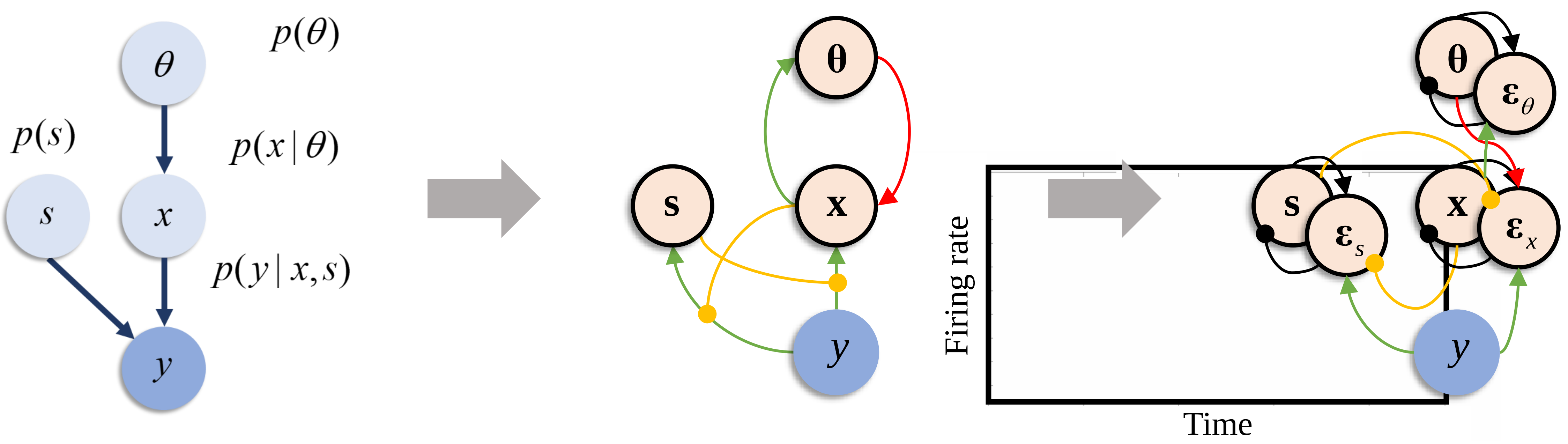


Variational inference

$$q(\theta) = \arg\min_{q(\theta)} \left\{ \ldots \mathrm{E}_{q(\theta)q(x)}\left[\ln p(x|\theta)\right] + \mathrm{E}_{q(\theta)}\left[\ln p(\theta)\right] - \mathrm{E}_{q(\theta)}\left[\ln q(\theta)\right] \right\}$$

$$q(x) = \arg\min_{q(x)} \left\{ \ldots \mathrm{E}_{q(\theta)q(x)}\left[\ln p(x|\theta)\right] + \mathrm{E}_{q(x)q(s)}\left[\ln p(y|x,s)\right] - \mathrm{E}_{q(x)}\left[\ln q(x)\right] \right\}$$

$$q(s) = \arg\min_{q(s)} \left\{ \ldots \mathrm{E}_{q(x)q(s)}\left[\ln p(y|x,s)\right] - \mathrm{E}_{q(s)}\left[\ln q(s)\right] \right\}$$



Dynamic (gradient descent) inference

$$\dot{\boldsymbol{\theta}} = \boldsymbol{\varepsilon}_{\theta}$$

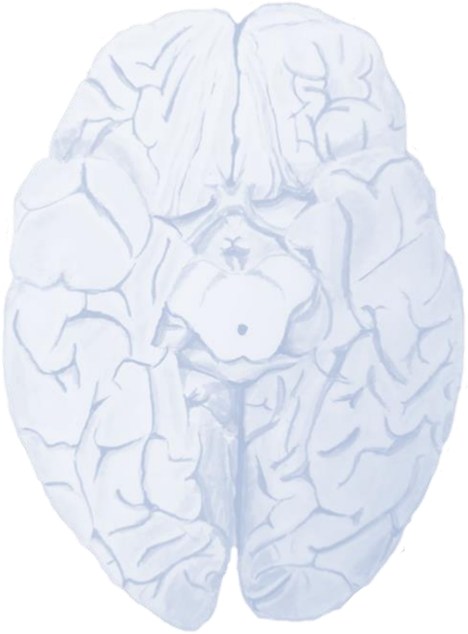
$$\boldsymbol{\varepsilon}_{\theta} = \nabla_{\phi(\boldsymbol{\theta})} \mathrm{E}_{q_{\boldsymbol{\theta}}(\theta)} [\ln q_{\boldsymbol{\theta}}(\theta)] - \nabla_{\phi(\boldsymbol{\theta})} \mathrm{E}_{q_{\boldsymbol{\theta}}(\theta) q_{\mathbf{x}}(x)} [\ln p(x \mid \theta)] - \nabla_{\phi(\boldsymbol{\theta})} \mathrm{E}_{q_{\boldsymbol{\theta}}(\theta)} [\ln p(\theta)]$$

$$\dot{\mathbf{x}} = \boldsymbol{\varepsilon}_{x}$$

$$\boldsymbol{\varepsilon}_{x} = \nabla_{\phi(\mathbf{x})} \mathrm{E}_{q_{\mathbf{x}}(x)} [\ln q_{\mathbf{x}}(x)] - \nabla_{\phi(\mathbf{x})} \mathrm{E}_{q_{\boldsymbol{\theta}}(\theta) q_{\mathbf{x}}(x)} [\ln p(x \mid \theta)] - \nabla_{\phi(\mathbf{x})} \mathrm{E}_{q_{\mathbf{x}}(x) q_{\mathbf{s}}(s)} [\ln p(y \mid x, s)]$$

$$\dot{\mathbf{s}} = \boldsymbol{\varepsilon}_{s}$$

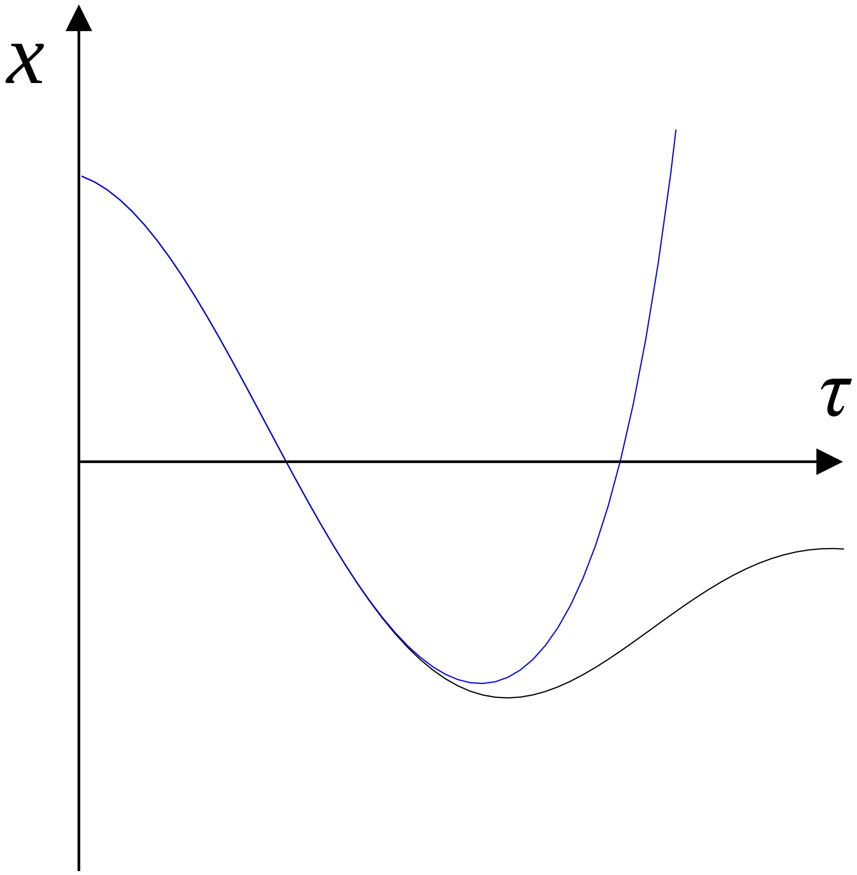
$$\boldsymbol{\varepsilon}_{s} = \nabla_{\phi(\mathbf{s})} \mathrm{E}_{q_{\mathbf{s}}(s)} [\ln q_{\mathbf{s}}(s)] - \nabla_{\phi(\mathbf{s})} \mathrm{E}_{q_{\mathbf{x}}(x) q_{\mathbf{s}}(s)} [\ln p(y \mid x, s)]$$

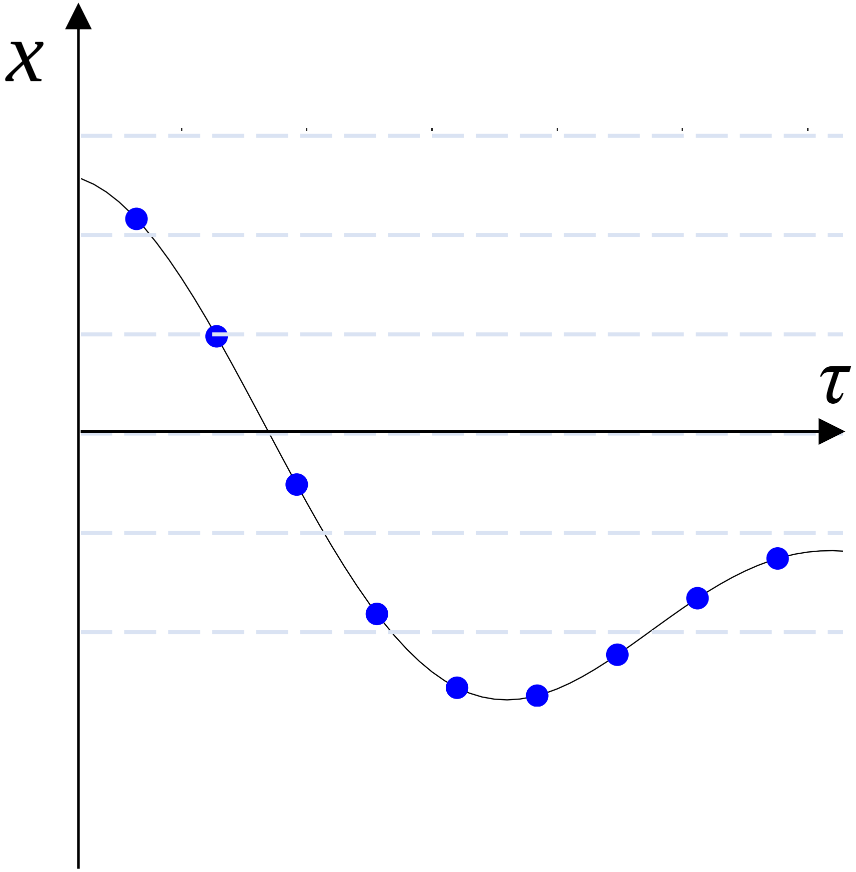


Active Inference

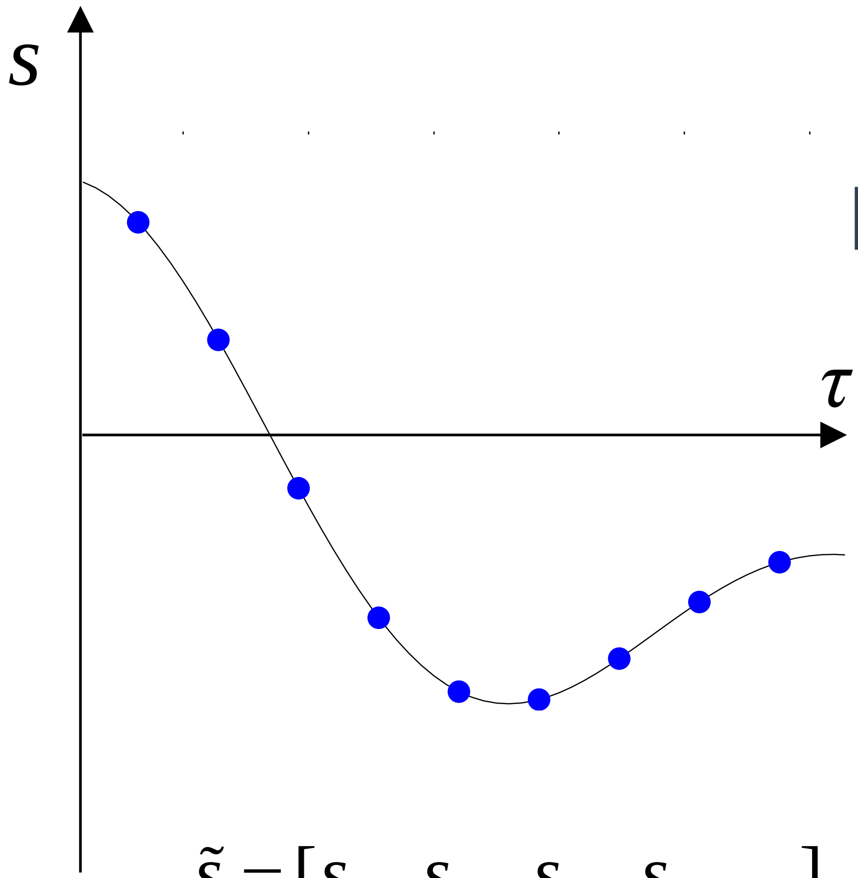
Self-evidencing
Message passing
Discrete time (planning)
Continuous time (movement)
Hierarchical models
Summary

Representing dynamics in generative models

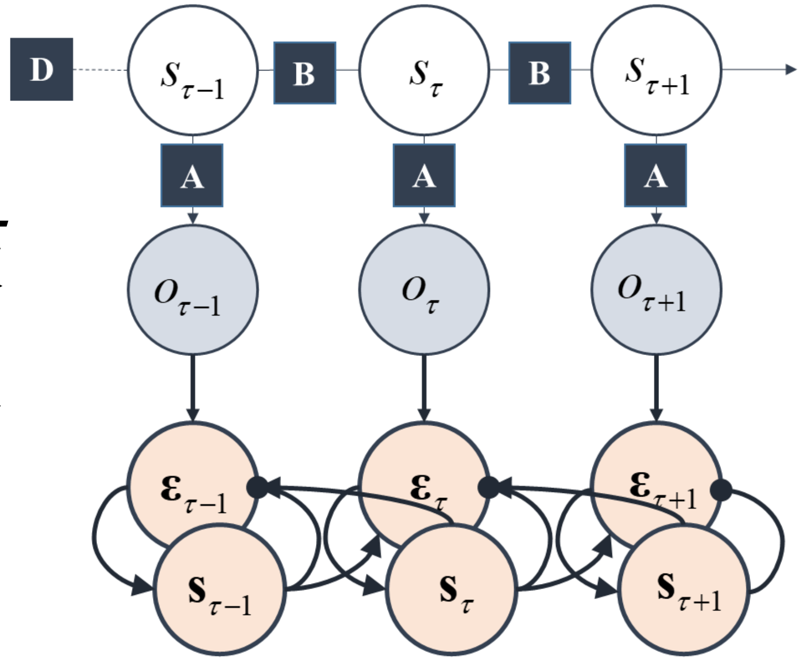




Discrete time models



$$\tilde{s} = [s_1,\ s_2,\ s_3,\ s_4,\ \ldots]$$



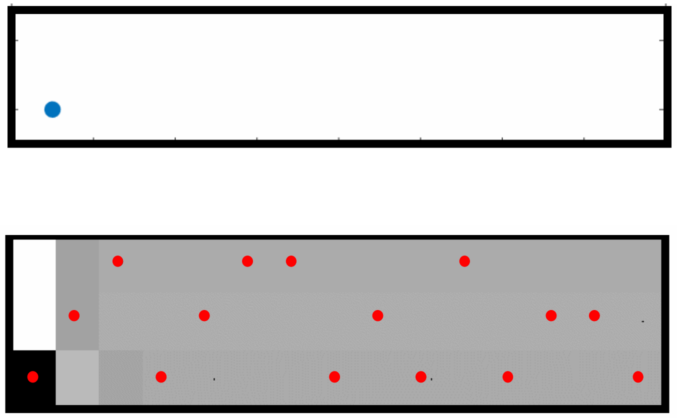
Generative model

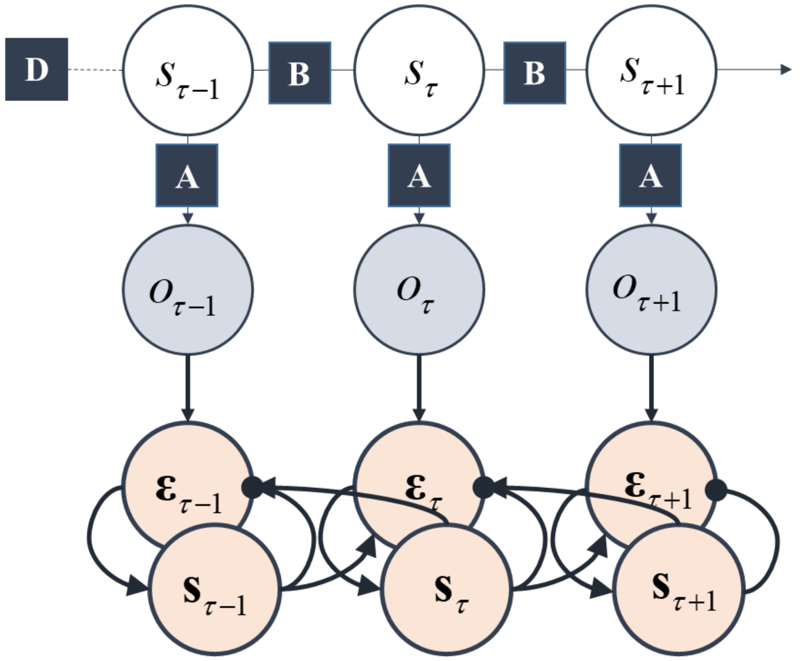
$$P(\tilde{o},\tilde{s}) = P(s_1)\prod_\tau P(s_{\tau+1} \mid s_\tau)P(o_\tau \mid s_\tau)$$

$$P(o_\tau \mid s_\tau) = Cat(\mathbf{A})$$
$$P(s_\tau \mid s_{\tau-1}) = Cat(\mathbf{B})$$
$$P(s_1) = Cat(\mathbf{D})$$

Bayesian message passing

$$\mathbf{s}_\tau = \sigma(\mathbf{v}_\tau);\ \ \dot{\mathbf{v}}_\tau = \boldsymbol{\varepsilon}_\tau$$
$$\boldsymbol{\varepsilon}_\tau = \ln \mathbf{A}\cdot o_\tau + \tfrac{1}{2}\ln(\mathbf{B}_\tau \mathbf{s}_\tau) + \tfrac{1}{2}\ln(\mathbf{B}^{\dagger}_{\tau+1}\mathbf{s}_{\tau+1}) - \ln \mathbf{s}_\tau$$





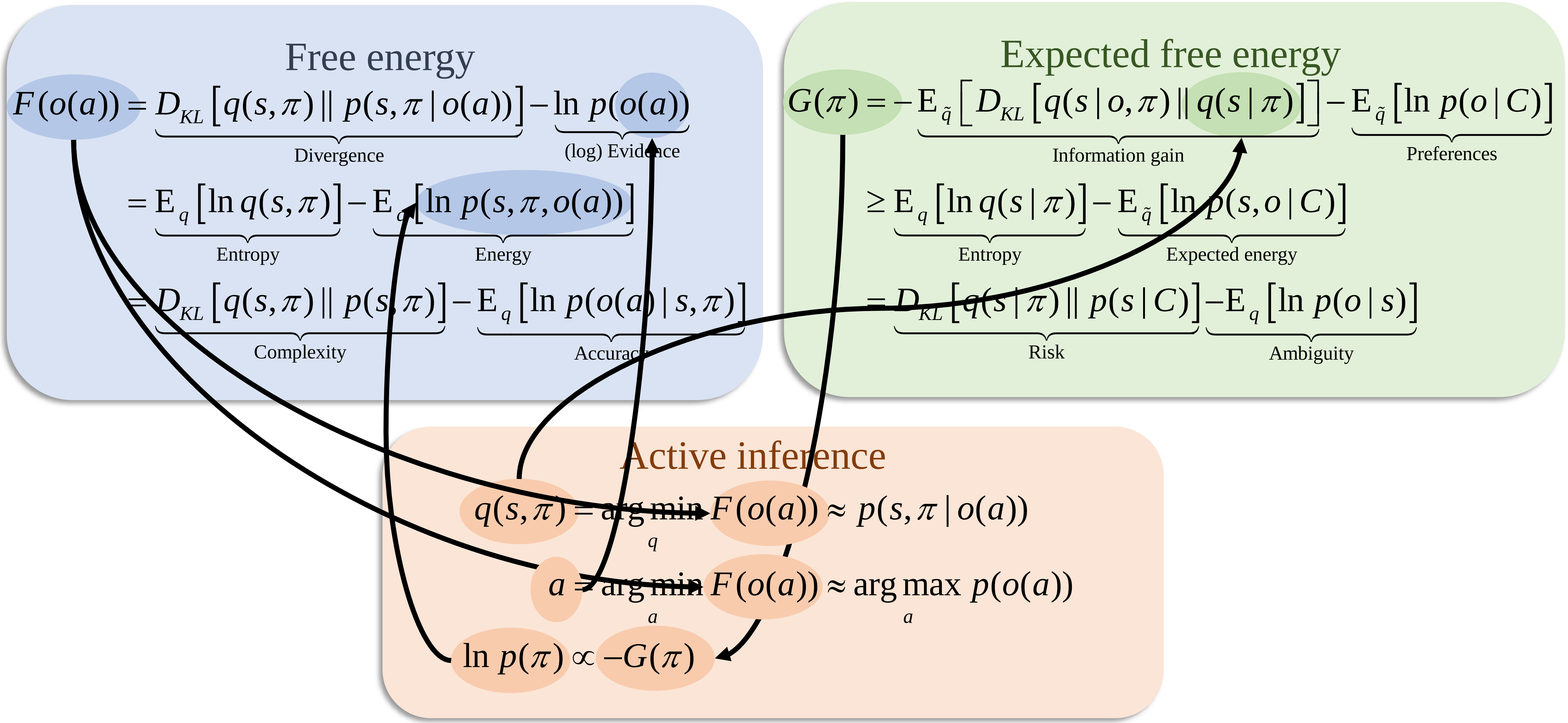
Generative model

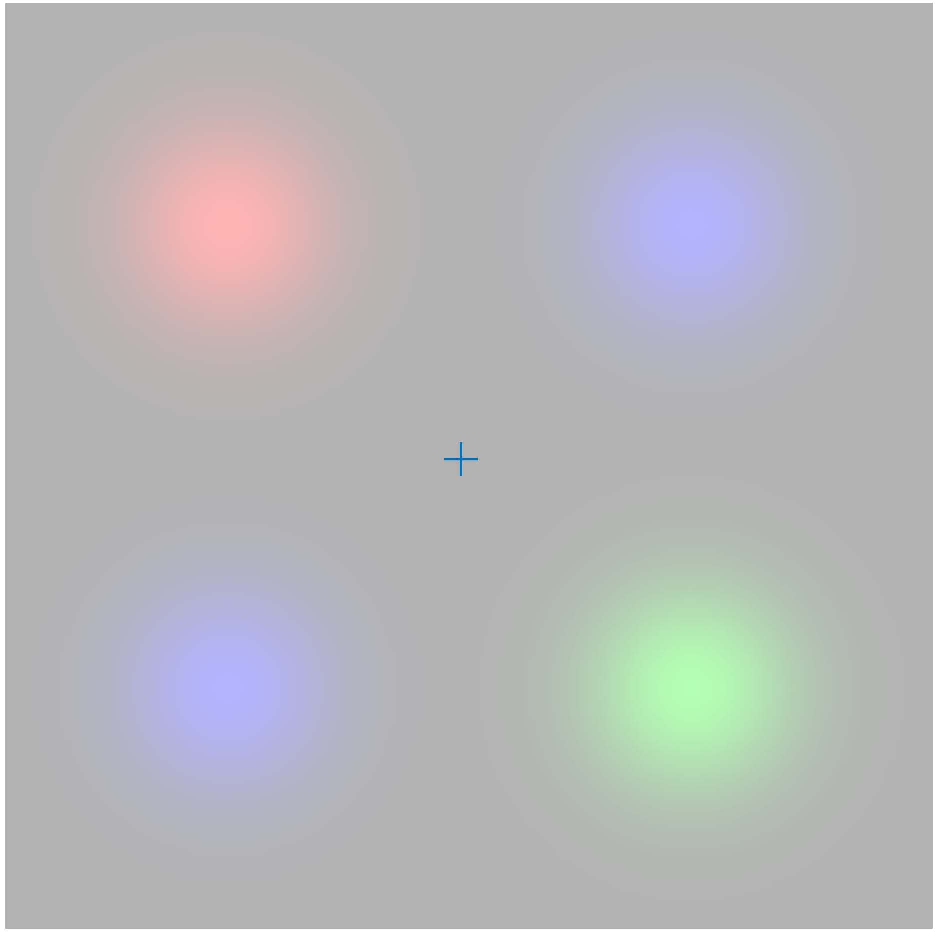
$$P(\tilde{o},\tilde{s}) = P(s_1)\prod_\tau P(s_{\tau+1} \mid s_\tau)P(o_\tau \mid s_\tau)$$

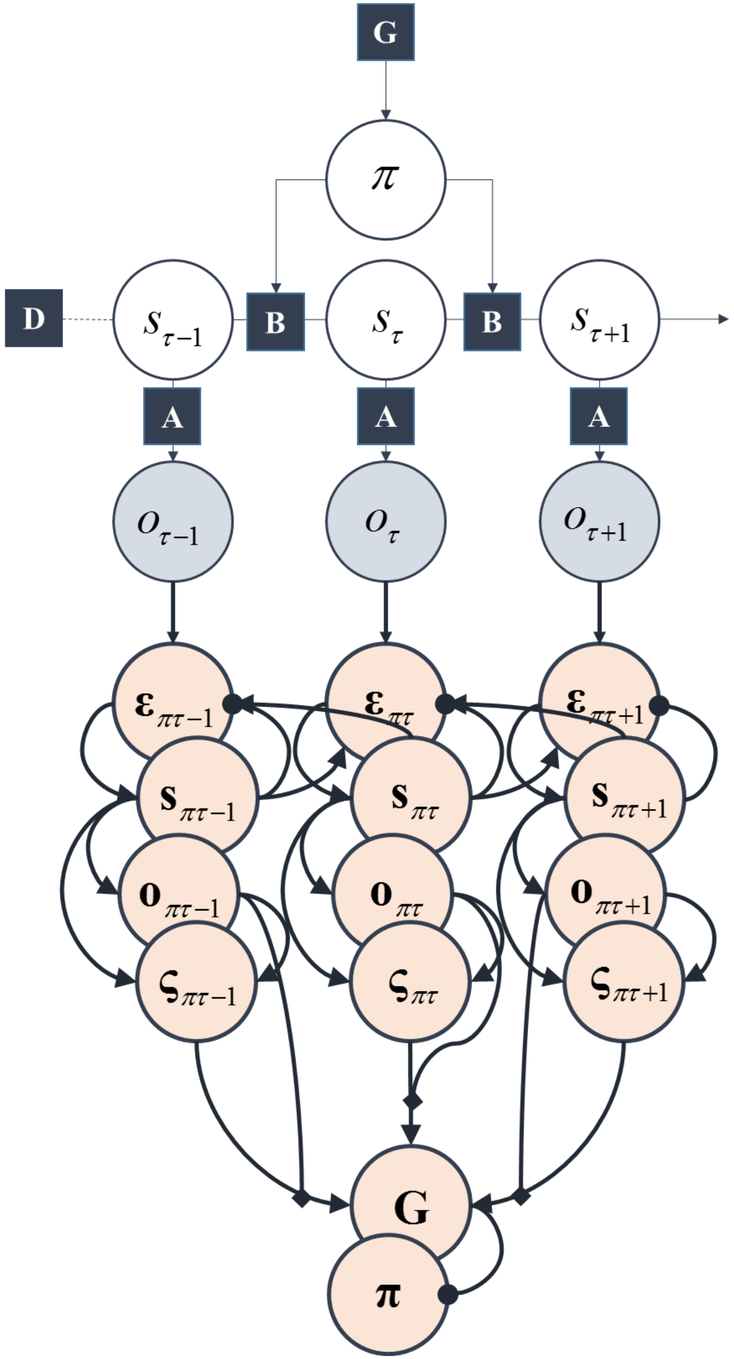
$$P(o_\tau \mid s_\tau) = Cat(\mathbf{A})$$
$$P(s_\tau \mid s_{\tau-1}) = Cat(\mathbf{B})$$
$$P(s_1) = Cat(\mathbf{D})$$

Bayesian message passing

$$\mathbf{s}_\tau = \sigma(\mathbf{v}_\tau);\ \dot{\mathbf{v}}_\tau = \boldsymbol{\varepsilon}_\tau$$
$$\boldsymbol{\varepsilon}_\tau = \ln \mathbf{A} \cdot o_\tau + \tfrac{1}{2}\ln(\mathbf{B}_\tau \mathbf{s}_\tau) + \tfrac{1}{2}\ln(\mathbf{B}^\dagger_{\tau+1}\mathbf{s}_{\tau+1}) - \ln \mathbf{s}_\tau$$


Free energy
$F(o(a)) = D_{KL}[q(s,\pi) \| p(s,\pi \mid o(a))] - \ln p(o(a))$
Divergence
(log) Evidence
$= E_q[\ln q(s,\pi)] - E_q[\ln p(s,\pi,o(a))]$
Entropy
Energy
$= D_{KL}[q(s,\pi) \| p(s,\pi)] - E_q[\ln p(o(a) \mid s,\pi)]$
Complexity
Accuracy
Expected free energy
$G(\pi) = -E_{\tilde{q}}[D_{KL}[q(s \mid o,\pi) \| q(s \mid \pi)]] - E_{\tilde{q}}[\ln p(o \mid C)]$
Information gain
Preferences
$\geq E_q[\ln q(s \mid \pi)] - E_{\tilde{q}}[\ln p(s,o \mid C)]$
Entropy
Expected energy
$= D_{KL}[q(s \mid \pi) \| p(s \mid C)] - E_q[\ln p(o \mid s)]$
Risk
Ambiguity
Active inference
$q(s,\pi) = \arg\min_q F(o(a)) \approx p(s,\pi \mid o(a))$
$a = \arg\min_a F(o(a)) \approx \arg\max_a p(o(a))$
$\ln p(\pi) \propto -G(\pi)$





Generative model

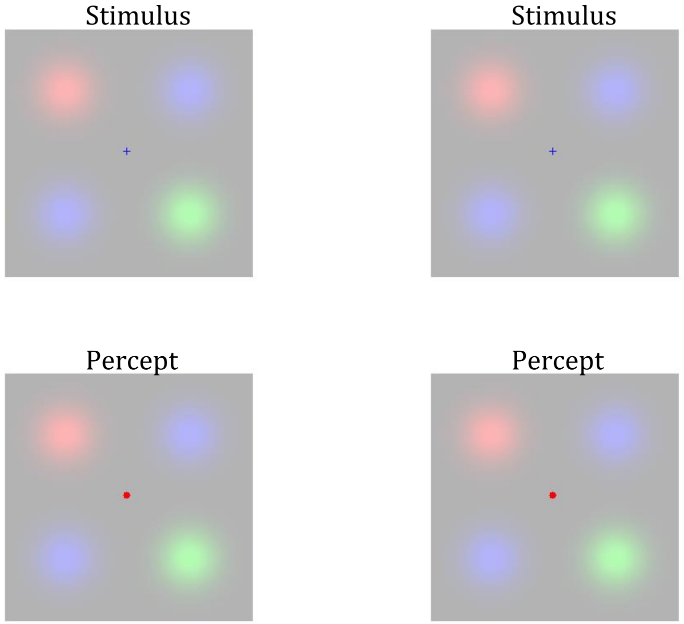
$$P(\tilde{o},\tilde{s},\pi) = P(s_1)P(\pi)\prod_{\tau} P(s_{\tau+1} \mid s_\tau,\pi)P(o_\tau \mid s_\tau)$$

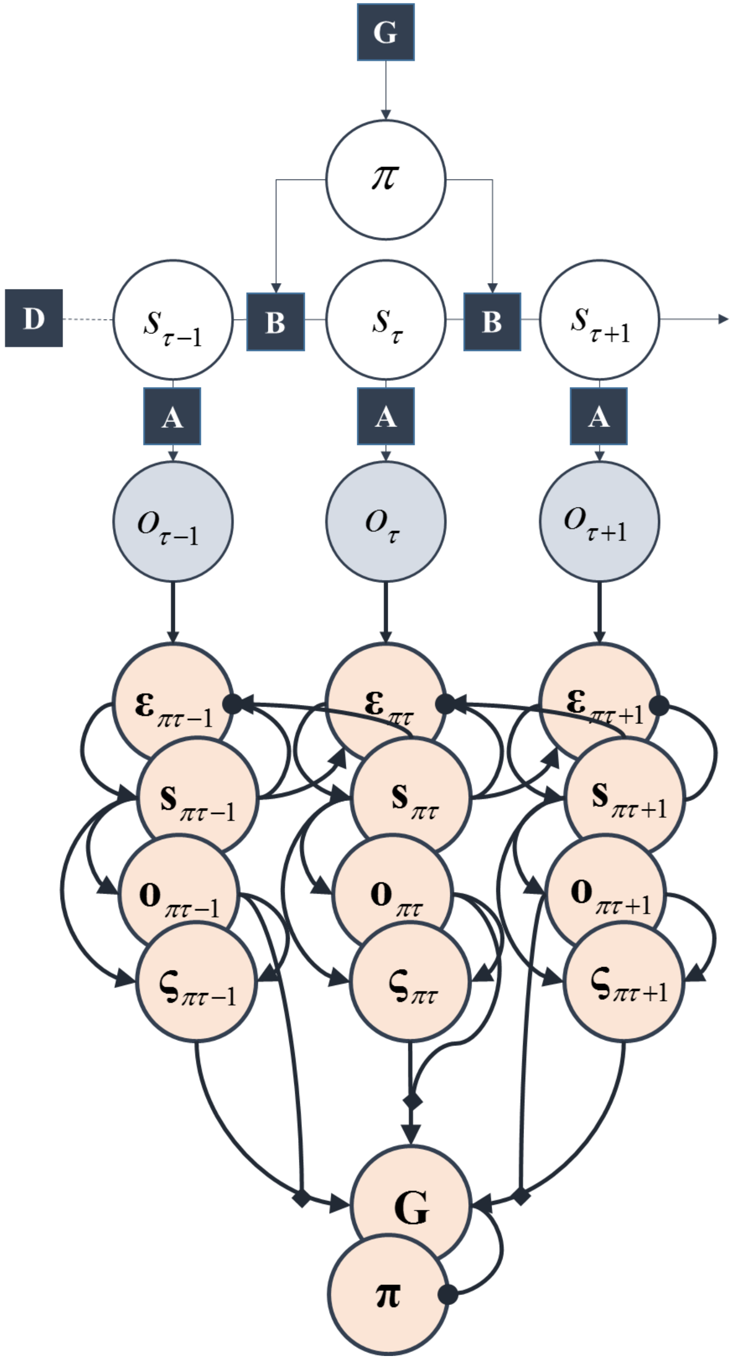
$$P(o_\tau \mid s_\tau) = Cat(\mathbf{A})$$
$$P(s_\tau \mid s_{\tau-1},\pi) = Cat(\mathbf{B}_{\pi\tau})$$
$$P(o_\tau) = Cat(\mathbf{C})$$
$$P(s_1) = Cat(\mathbf{D})$$
$$P(\pi) = \sigma(-\mathbf{G})$$

Bayesian message passing

$$\mathbf{s}_\tau = \boldsymbol{\pi} \cdot \mathbf{s}_{\pi\tau}$$
$$\mathbf{s}_{\pi\tau} = \sigma(\mathbf{v}_{\pi\tau}); \;\; \dot{\mathbf{v}}_{\pi\tau} = \boldsymbol{\varepsilon}_{\pi\tau}$$
$$\boldsymbol{\varepsilon}_{\pi\tau} = \ln \mathbf{A} \cdot o_\tau + \tfrac{1}{2}\ln(\mathbf{B}_{\pi\tau}\mathbf{s}_{\pi\tau}) + \tfrac{1}{2}\ln(\mathbf{B}^{\dagger}_{\pi\tau+1}\mathbf{s}_{\pi\tau+1}) - \ln \mathbf{s}_{\pi\tau}$$

$$\mathbf{o}_{\pi\tau} = \mathbf{A}\mathbf{s}_{\pi\tau}$$
$$\boldsymbol{\varsigma}_{\pi\tau} = \ln \mathbf{o}_{\pi\tau} - \ln \mathbf{C} + \mathbf{H} \cdot \mathbf{s}_{\pi\tau}$$
$$\mathbf{H} = -diag(\mathbf{A} \cdot \ln \mathbf{A})$$
$$\mathbf{G}_\pi = \mathbf{o}_{\pi\tau} \cdot \boldsymbol{\varsigma}_{\pi\tau}$$
$$\boldsymbol{\pi} = \sigma(-\mathbf{G})$$





Generative model

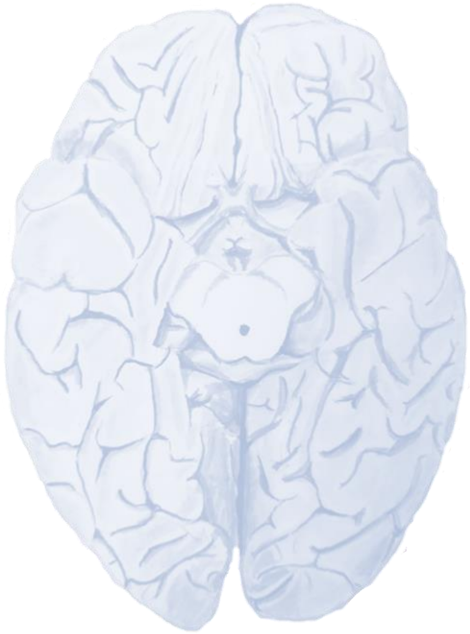
$$P(\tilde{o},\tilde{s},\pi)=P(s_1)P(\pi)\prod_{\tau}P(s_{\tau+1}\mid s_\tau,\pi)P(o_\tau\mid s_\tau)$$

$$P(o_\tau\mid s_\tau)=Cat(\mathbf{A})$$
$$P(s_\tau\mid s_{\tau-1},\pi)=Cat(\mathbf{B}_{\pi\tau})$$
$$P(o_\tau)=Cat(\mathbf{C})$$
$$P(s_1)=Cat(\mathbf{D})$$
$$P(\pi)=\sigma(-\mathbf{G})$$

Bayesian message passing

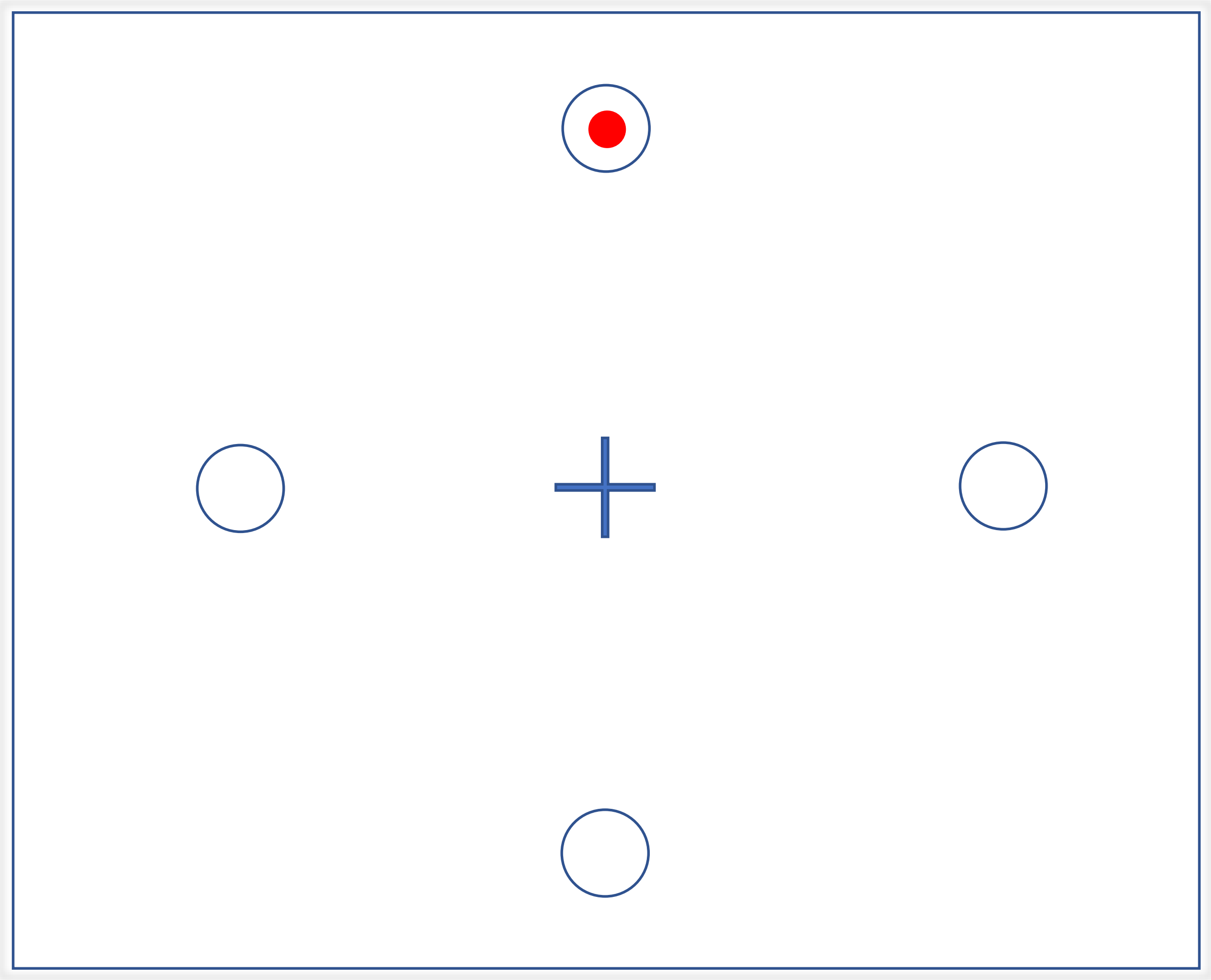
$$\mathbf{s}_\tau=\boldsymbol{\pi}\cdot\mathbf{s}_{\pi\tau}$$
$$\mathbf{s}_{\pi\tau}=\sigma(\mathbf{v}_{\pi\tau});\ \ \dot{\mathbf{v}}_{\pi\tau}=\boldsymbol{\varepsilon}_{\pi\tau}$$
$$\boldsymbol{\varepsilon}_{\pi\tau}=\ln\mathbf{A}\cdot o_\tau+\tfrac{1}{2}\ln(\mathbf{B}_{\pi\tau}\mathbf{s}_{\pi\tau})+\tfrac{1}{2}\ln(\mathbf{B}^{\dagger}_{\pi\tau+1}\mathbf{s}_{\pi\tau+1})-\ln\mathbf{s}_{\pi\tau}$$

$$\mathbf{o}_{\pi\tau}=\mathbf{A}\mathbf{s}_{\pi\tau}$$
$$\boldsymbol{\varsigma}_{\pi\tau}=\ln\mathbf{o}_{\pi\tau}-\ln\mathbf{C}+\mathbf{H}\cdot\mathbf{s}_{\pi\tau}$$
$$\mathbf{H}=-diag(\mathbf{A}\cdot\ln\mathbf{A})$$
$$\mathbf{G}_\pi=\mathbf{o}_{\pi\tau}\cdot\boldsymbol{\varsigma}_{\pi\tau}$$
$$\boldsymbol{\pi}=\sigma(-\mathbf{G})$$

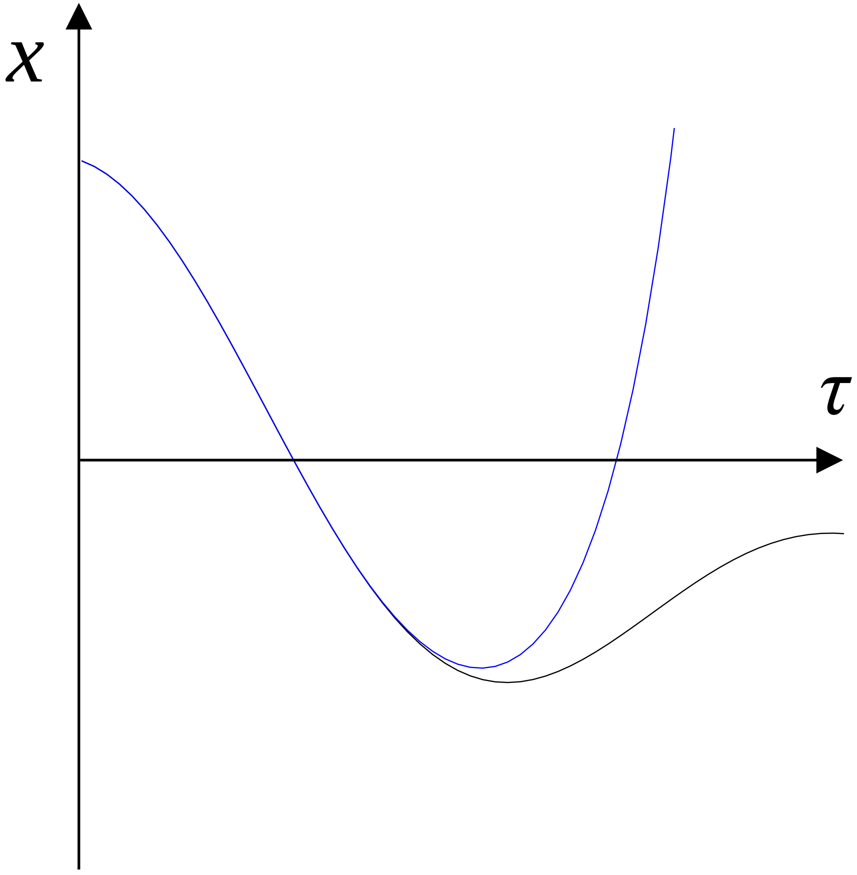


Active Inference

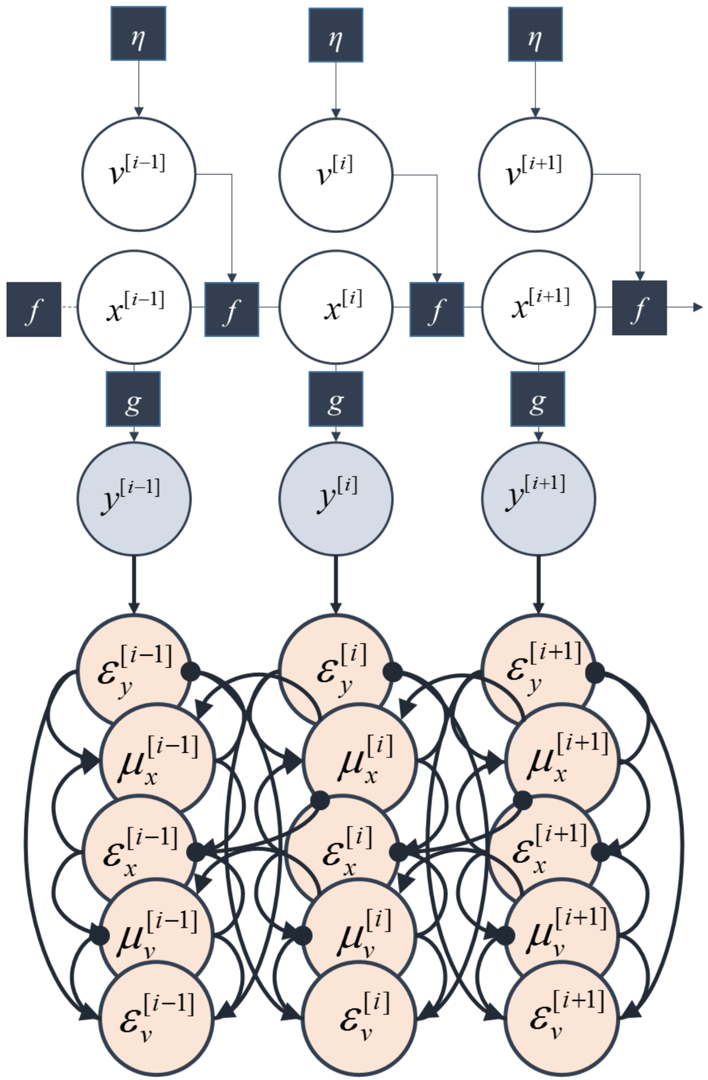
Self-evidencing
Message passing
Discrete time (planning)
Continuous time (movement)
Hierarchical models
Summary



Continuous time models



$$x(\tau) \approx x_0 + \tau x_0' + \tfrac{1}{2}\tau^2 x_0'' + \tfrac{1}{6}\tau^3 x_0''' + \cdots$$



Generative model

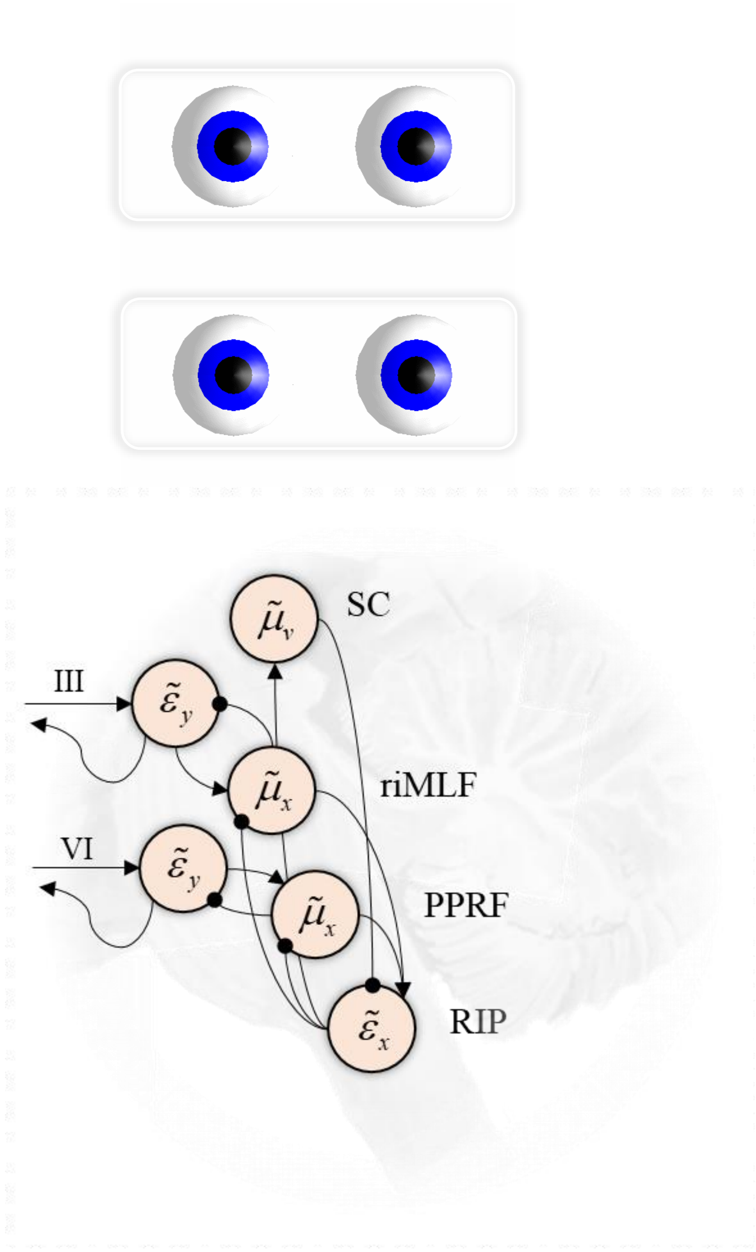
$$p(\tilde{y},\tilde{x},\tilde{v}) = \prod_i p(v^{[i]})\,p(x^{[i+1]} \mid x^{[i]}, v^{[i]})\,p(y^{[i]} \mid x^{[i]}, v^{[i]})$$

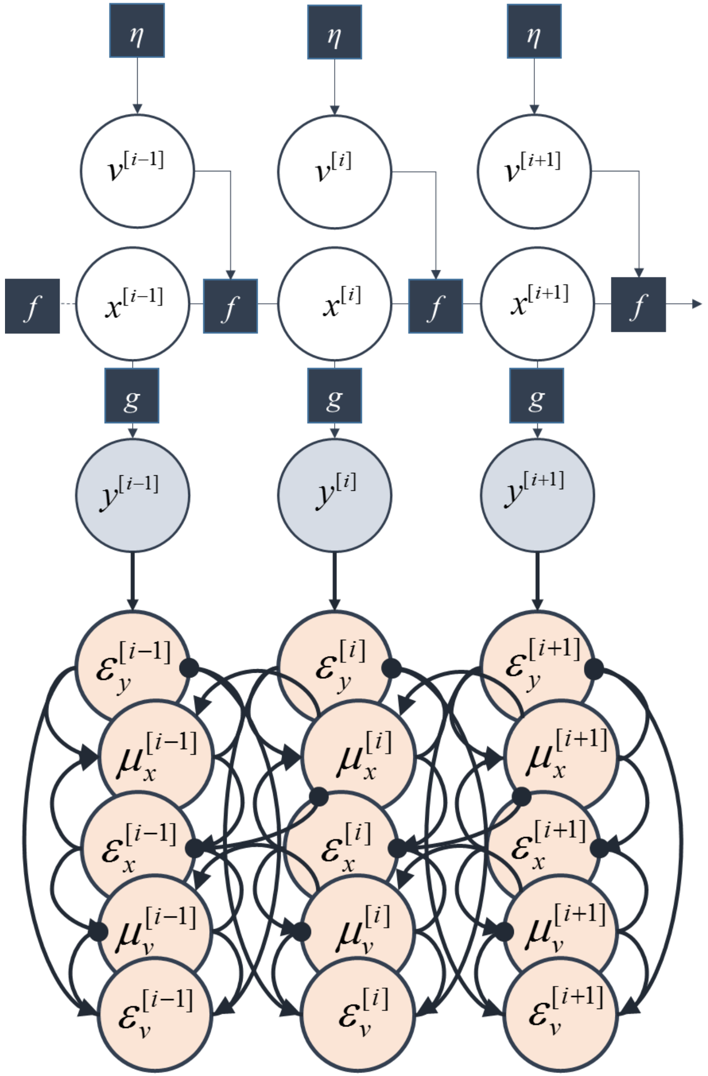
$$p(y^{[i]} \mid x^{[i]}, v^{[i]}) = \mathcal{N}(g^{[i]}(x^{[i]}, v^{[i]}), \Pi_y^{[i]})$$
$$p(x^{[i+1]} \mid x^{[i]}, v^{[i]}) = \mathcal{N}(f^{[i]}(x^{[i]}, v^{[i]}), \Pi_x^{[i]})$$
$$p(v^{[i]}) = \mathcal{N}(\eta^{[i]}, \Pi_v^{[i]})$$

Bayesian message passing

$$\varepsilon_y^{[i]} = y^{[i]} - g^{[i]}(\mu_x^{[i]}, \mu_v^{[i]})$$
$$\varepsilon_x^{[i]} = \mu_x^{[i+1]} - f^{[i]}(\mu_x^{[i]}, \mu_v^{[i]})$$
$$\varepsilon_v^{[i]} = \mu_v^{[i]} - \eta^{[i]}$$

$$\dot{\mu}_x^{[i]} = \mu_x^{[i+1]}$$
$$+\partial_{\mu_x^{[i]}} g^{[i]} \cdot \Pi_y^{[i]} \varepsilon_y^{[i]} - \Pi_x^{[i-1]} \varepsilon_x^{[i-1]} + \partial_{\mu_x^{[i]}} f^{[i]} \cdot \Pi_x^{[i]} \varepsilon_x^{[i]}$$
$$\dot{\mu}_v^{[i]} = \mu_v^{[i+1]}$$
$$+\partial_{\mu_v^{[i]}} g^{[i]} \cdot \Pi_y^{[i]} \varepsilon_y^{[i]} + \partial_{\mu_v^{[i]}} f^{[i]} \cdot \Pi_x^{[i]} \varepsilon_x^{[i]} - \Pi_v^{[i]} \varepsilon_v^{[i]}$$





Generative model

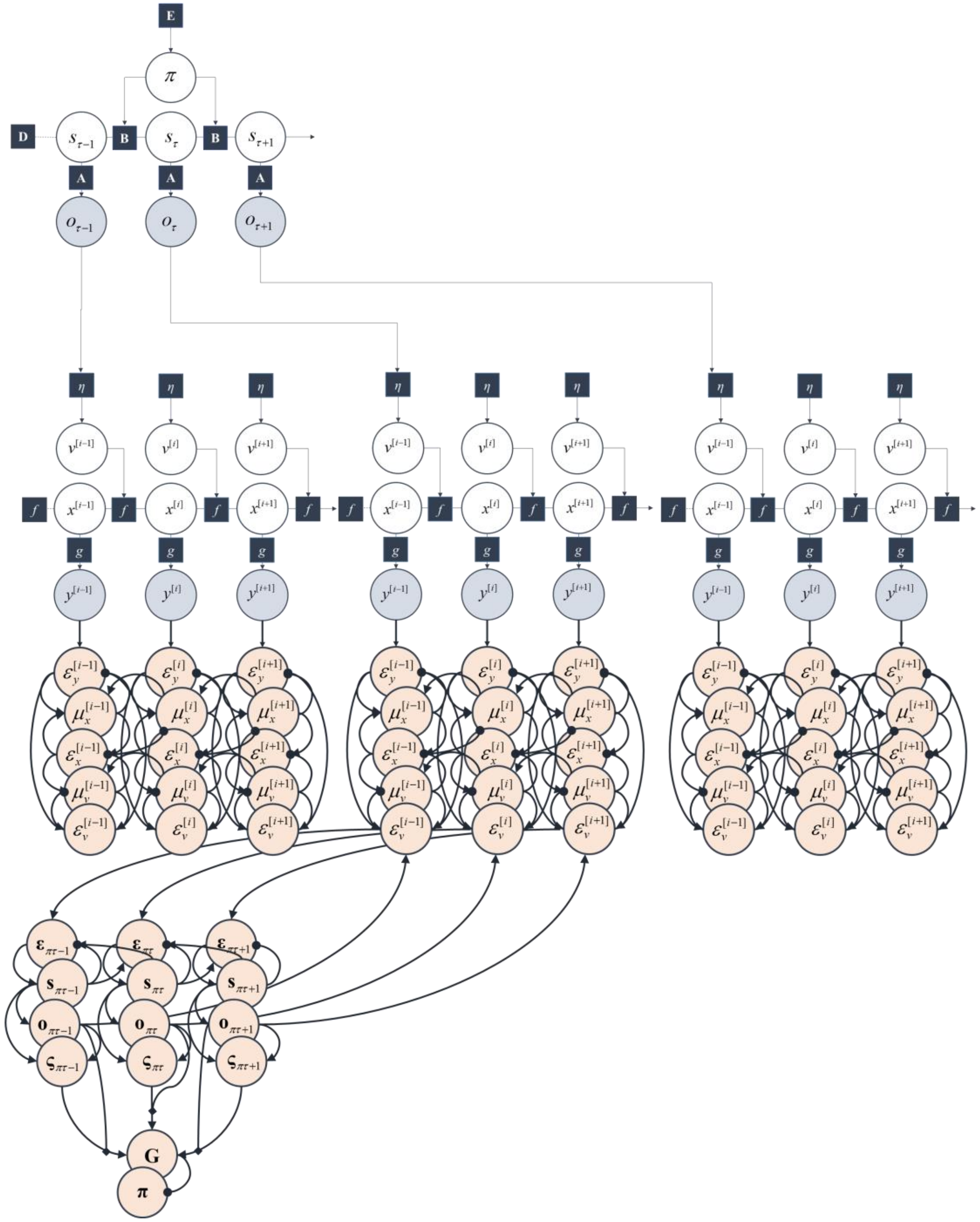
$$p(\tilde{y},\tilde{x},\tilde{v}) = \prod_i p(v^{[i]})\,p(x^{[i+1]} \mid x^{[i]},v^{[i]})\,p(y^{[i]} \mid x^{[i]},v^{[i]})$$

$$p(y^{[i]} \mid x^{[i]},v^{[i]}) = \mathcal{N}(g^{[i]}(x^{[i]},v^{[i]}),\Pi_y^{[i]})$$
$$p(x^{[i+1]} \mid x^{[i]},v^{[i]}) = \mathcal{N}(f^{[i]}(x^{[i]},v^{[i]}),\Pi_x^{[i]})$$
$$p(v^{[i]}) = \mathcal{N}(\eta^{[i]},\Pi_v^{[i]})$$

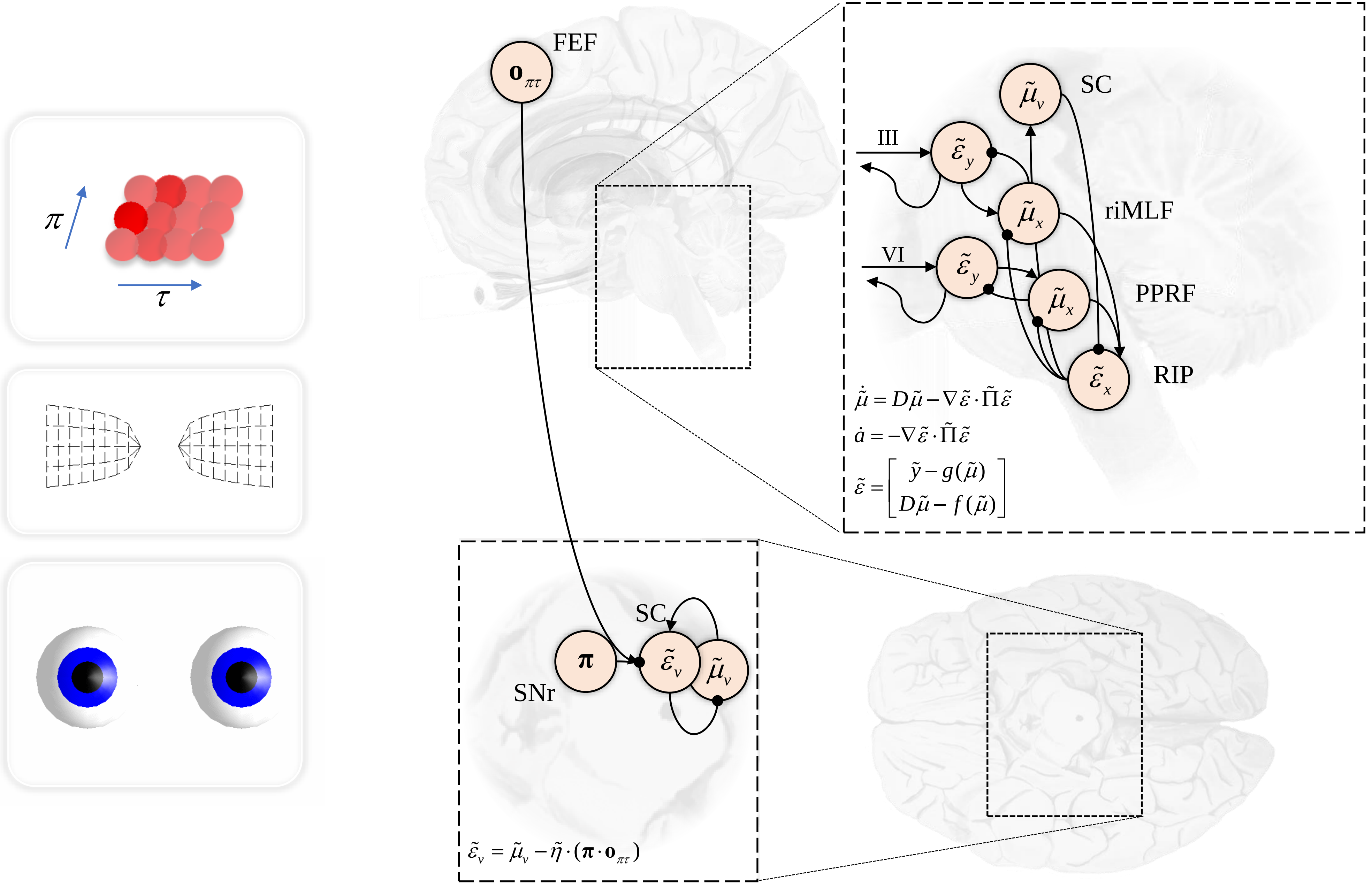
Bayesian message passing

$$\varepsilon_y^{[i]} = y^{[i]} - g^{[i]}(\mu_x^{[i]},\mu_v^{[i]})$$
$$\varepsilon_x^{[i]} = \mu_x^{[i+1]} - f^{[i]}(\mu_x^{[i]},\mu_v^{[i]})$$
$$\varepsilon_v^{[i]} = \mu_v^{[i]} - \eta^{[i]}$$

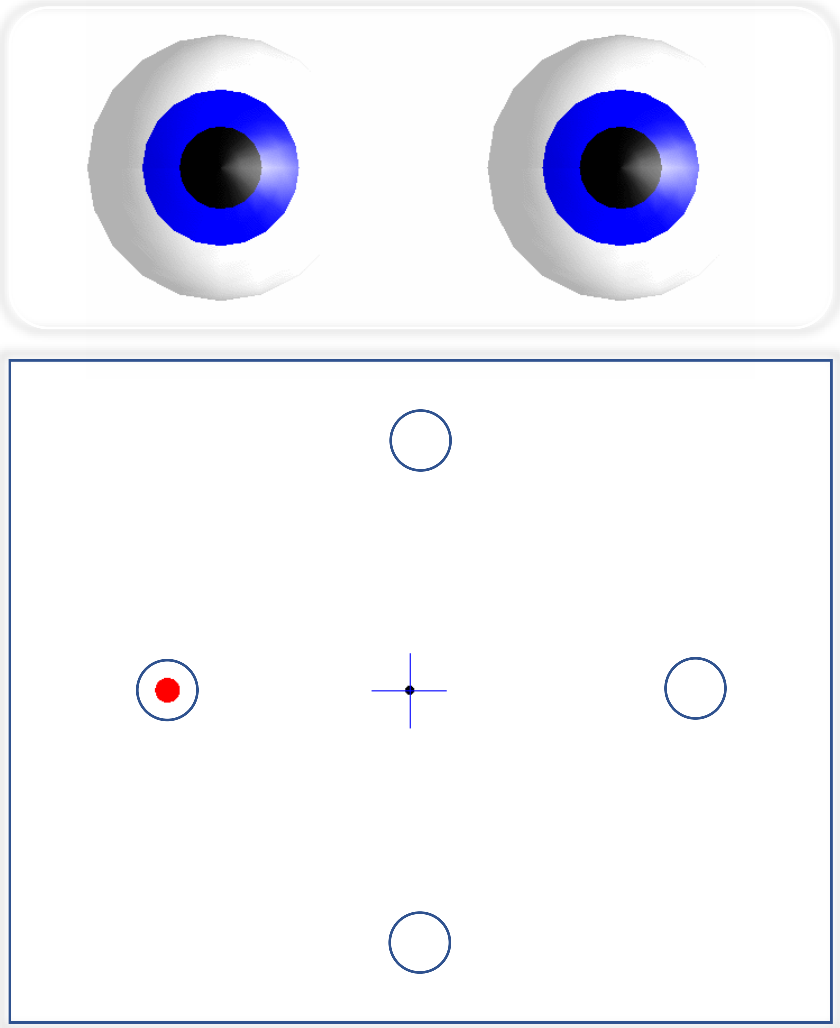
$$\begin{aligned}\dot{\mu}_x^{[i]} &= \mu_x^{[i+1]} \\ &\quad + \partial_{\mu_x^{[i]}} g^{[i]} \cdot \Pi_y^{[i]} \varepsilon_y^{[i]} - \Pi_x^{[i-1]} \varepsilon_x^{[i-1]} + \partial_{\mu_x^{[i]}} f^{[i]} \cdot \Pi_x^{[i]} \varepsilon_x^{[i]}\end{aligned}$$
$$\begin{aligned}\dot{\mu}_v^{[i]} &= \mu_v^{[i+1]} \\ &\quad + \partial_{\mu_v^{[i]}} g^{[i]} \cdot \Pi_y^{[i]} \varepsilon_y^{[i]} + \partial_{\mu_v^{[i]}} f^{[i]} \cdot \Pi_x^{[i]} \varepsilon_x^{[i]} - \Pi_v^{[i]} \varepsilon_v^{[i]}\end{aligned}$$

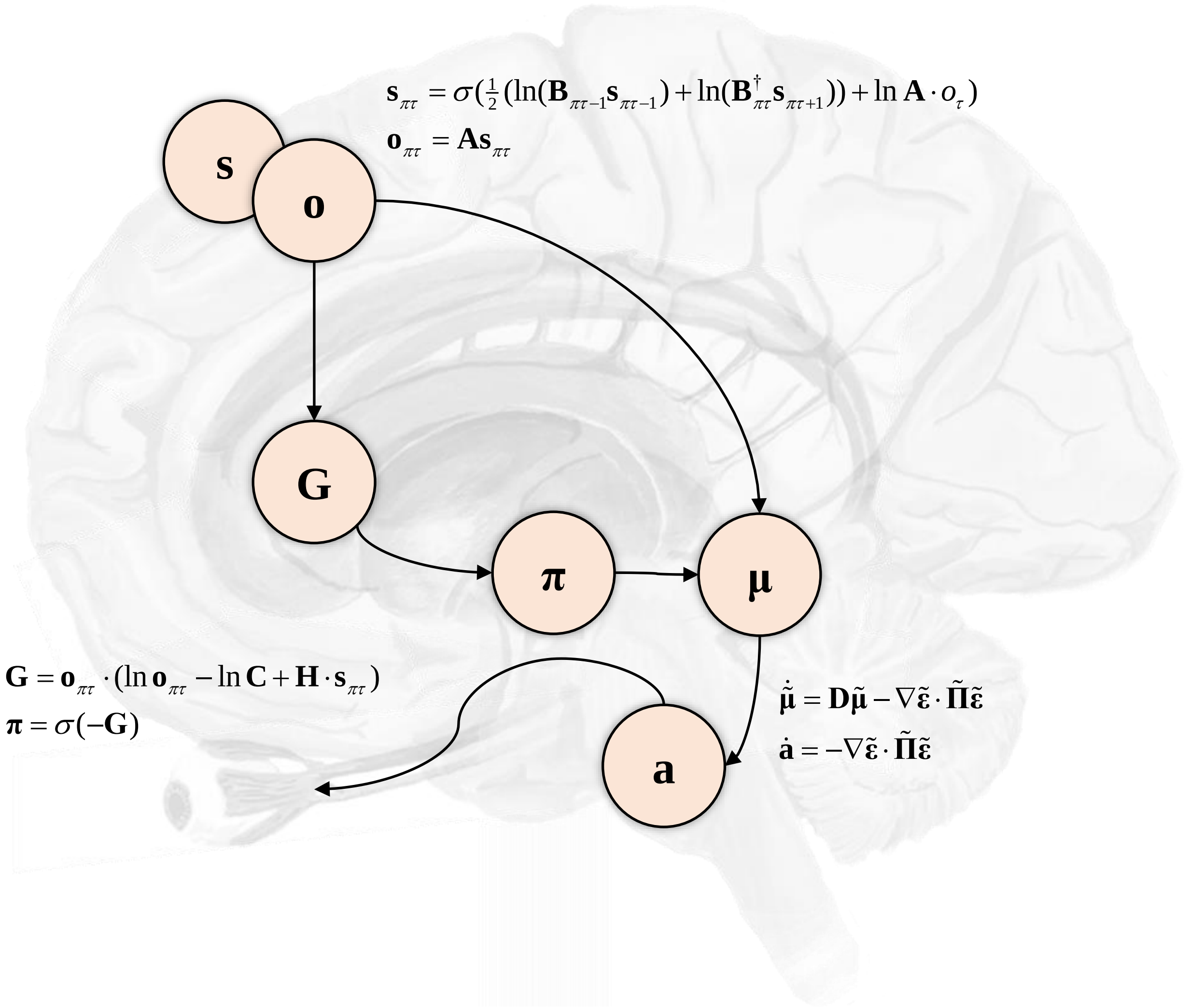


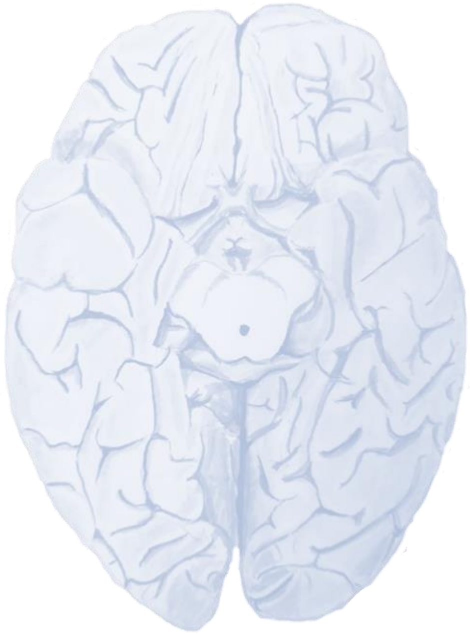
E
π
D
$s_{\tau-1}$
B
s_{τ}
B
$s_{\tau+1}$
A
A
A
$o_{\tau-1}$
o_{τ}
$o_{\tau+1}$
η
$v^{[i-1]}$
$v^{[i]}$
$v^{[i+1]}$
f
$x^{[i-1]}$
$x^{[i]}$
$x^{[i+1]}$
g
$y^{[i-1]}$
$y^{[i]}$
$y^{[i+1]}$
$\varepsilon_y^{[i-1]}$
$\mu_x^{[i-1]}$
$\varepsilon_x^{[i-1]}$
$\mu_v^{[i-1]}$
$\varepsilon_v^{[i-1]}$
$\varepsilon_y^{[i]}$
$\mu_x^{[i]}$
$\varepsilon_x^{[i]}$
$\mu_v^{[i]}$
$\varepsilon_v^{[i]}$
$\varepsilon_y^{[i+1]}$
$\mu_x^{[i+1]}$
$\varepsilon_x^{[i+1]}$
$\mu_v^{[i+1]}$
$\varepsilon_v^{[i+1]}$
$\boldsymbol{\varepsilon}_{\pi\tau-1}$
$\mathbf{s}_{\pi\tau-1}$
$\mathbf{o}_{\pi\tau-1}$
$\boldsymbol{\varsigma}_{\pi\tau-1}$
$\boldsymbol{\varepsilon}_{\pi\tau}$
$\mathbf{s}_{\pi\tau}$
$\mathbf{o}_{\pi\tau}$
$\boldsymbol{\varsigma}_{\pi\tau}$
$\boldsymbol{\varepsilon}_{\pi\tau+1}$
$\mathbf{s}_{\pi\tau+1}$
$\mathbf{o}_{\pi\tau+1}$
$\boldsymbol{\varsigma}_{\pi\tau+1}$
G
$\boldsymbol{\pi}$



π
τ
FEF
$\mathbf{o}_{\pi\tau}$
SC
III
VI
riMLF
PPRF
RIP
$\tilde{\mu}_v$
$\tilde{\varepsilon}_y$
$\tilde{\mu}_x$
$\tilde{\varepsilon}_x$
$\dot{\tilde{\mu}} = D\tilde{\mu} - \nabla\tilde{\varepsilon} \cdot \tilde{\Pi}\tilde{\varepsilon}$
$\dot{a} = -\nabla\tilde{\varepsilon} \cdot \tilde{\Pi}\tilde{\varepsilon}$
$\tilde{\varepsilon} = \begin{bmatrix} \tilde{y} - g(\tilde{\mu}) \\ D\tilde{\mu} - f(\tilde{\mu}) \end{bmatrix}$
SC
SNr
$\boldsymbol{\pi}$
$\tilde{\varepsilon}_v$
$\tilde{\varepsilon}_v = \tilde{\mu}_v - \tilde{\eta} \cdot (\boldsymbol{\pi} \cdot \mathbf{o}_{\pi\tau})$

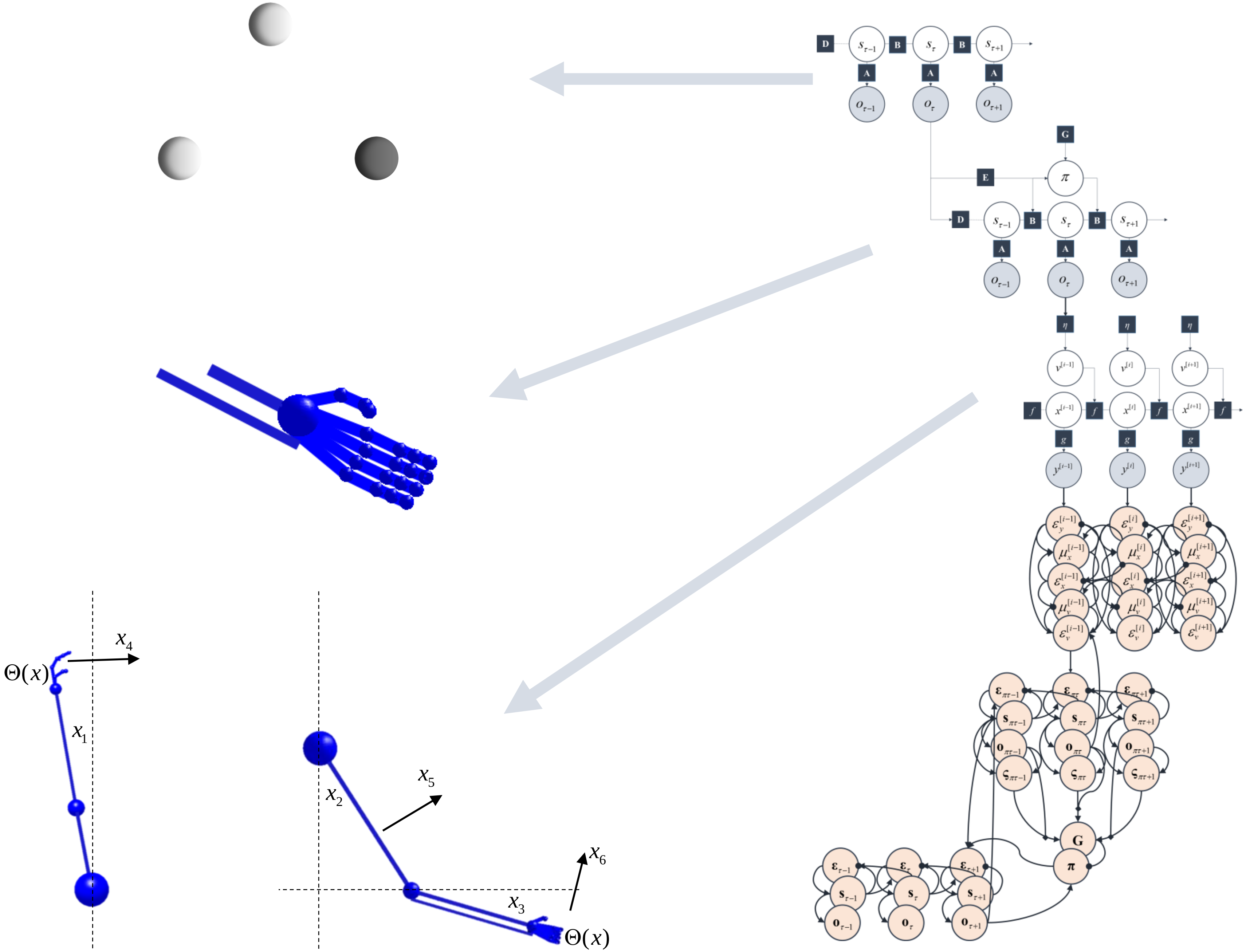



$\mathbf{s}_{\pi\tau} = \sigma(\tfrac{1}{2}(\ln(\mathbf{B}_{\pi\tau-1}\mathbf{s}_{\pi\tau-1}) + \ln(\mathbf{B}^{\dagger}_{\pi\tau}\mathbf{s}_{\pi\tau+1})) + \ln \mathbf{A}\cdot o_{\tau})$
$\mathbf{o}_{\pi\tau} = \mathbf{A}\mathbf{s}_{\pi\tau}$
s
o
G
π
μ
a
$\mathbf{G} = \mathbf{o}_{\pi\tau}\cdot(\ln \mathbf{o}_{\pi\tau} - \ln \mathbf{C} + \mathbf{H}\cdot \mathbf{s}_{\pi\tau})$
$\boldsymbol{\pi} = \sigma(-\mathbf{G})$
$\dot{\tilde{\boldsymbol{\mu}}} = \mathbf{D}\tilde{\boldsymbol{\mu}} - \nabla\tilde{\boldsymbol{\varepsilon}}\cdot\tilde{\boldsymbol{\Pi}}\tilde{\boldsymbol{\varepsilon}}$
$\dot{\mathbf{a}} = -\nabla\tilde{\boldsymbol{\varepsilon}}\cdot\tilde{\boldsymbol{\Pi}}\tilde{\boldsymbol{\varepsilon}}$

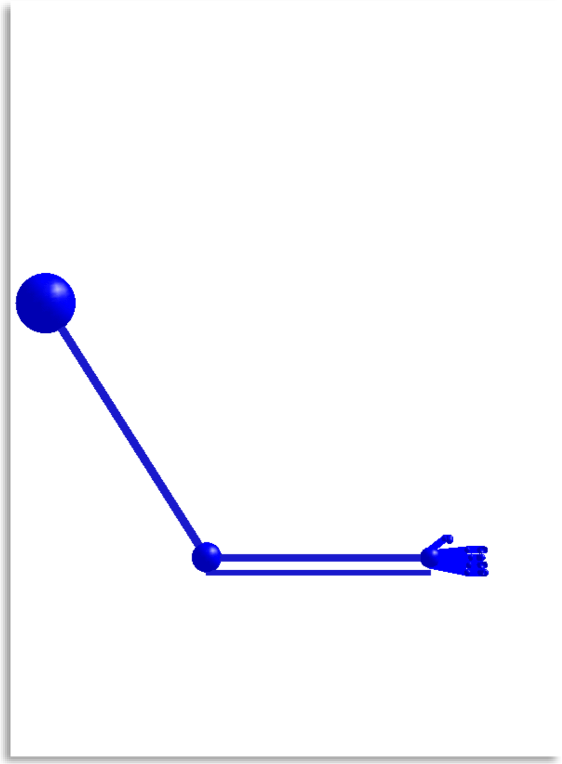


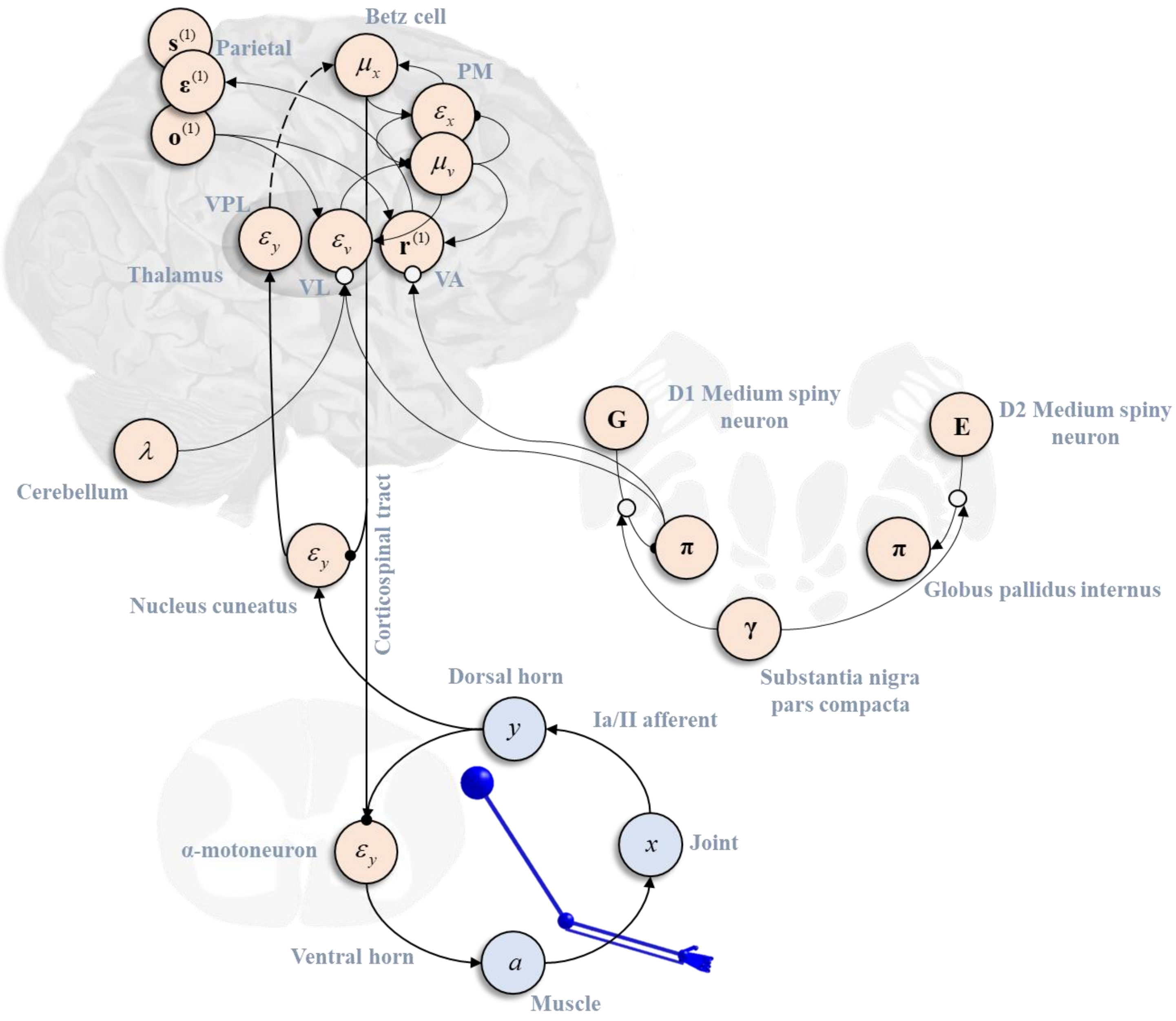
Active Inference

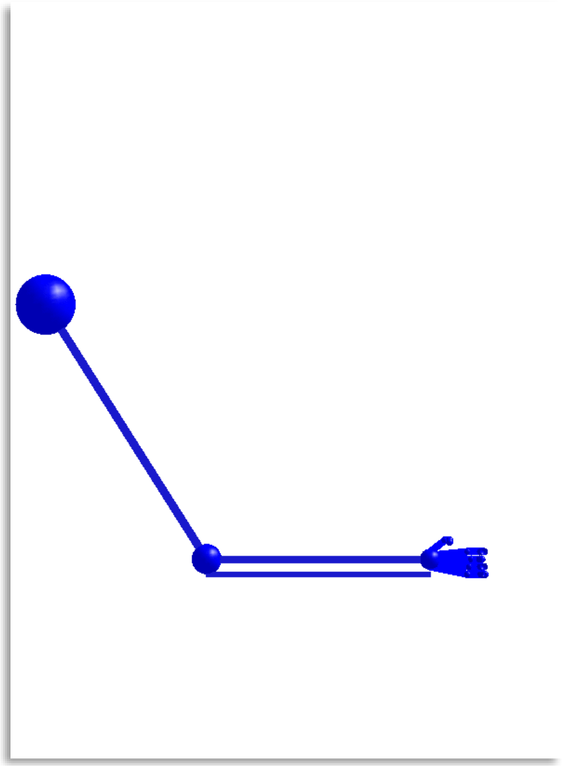
Self-evidencing
Message passing
Discrete time (planning)
Continuous time (movement)
Hierarchical models
Summary

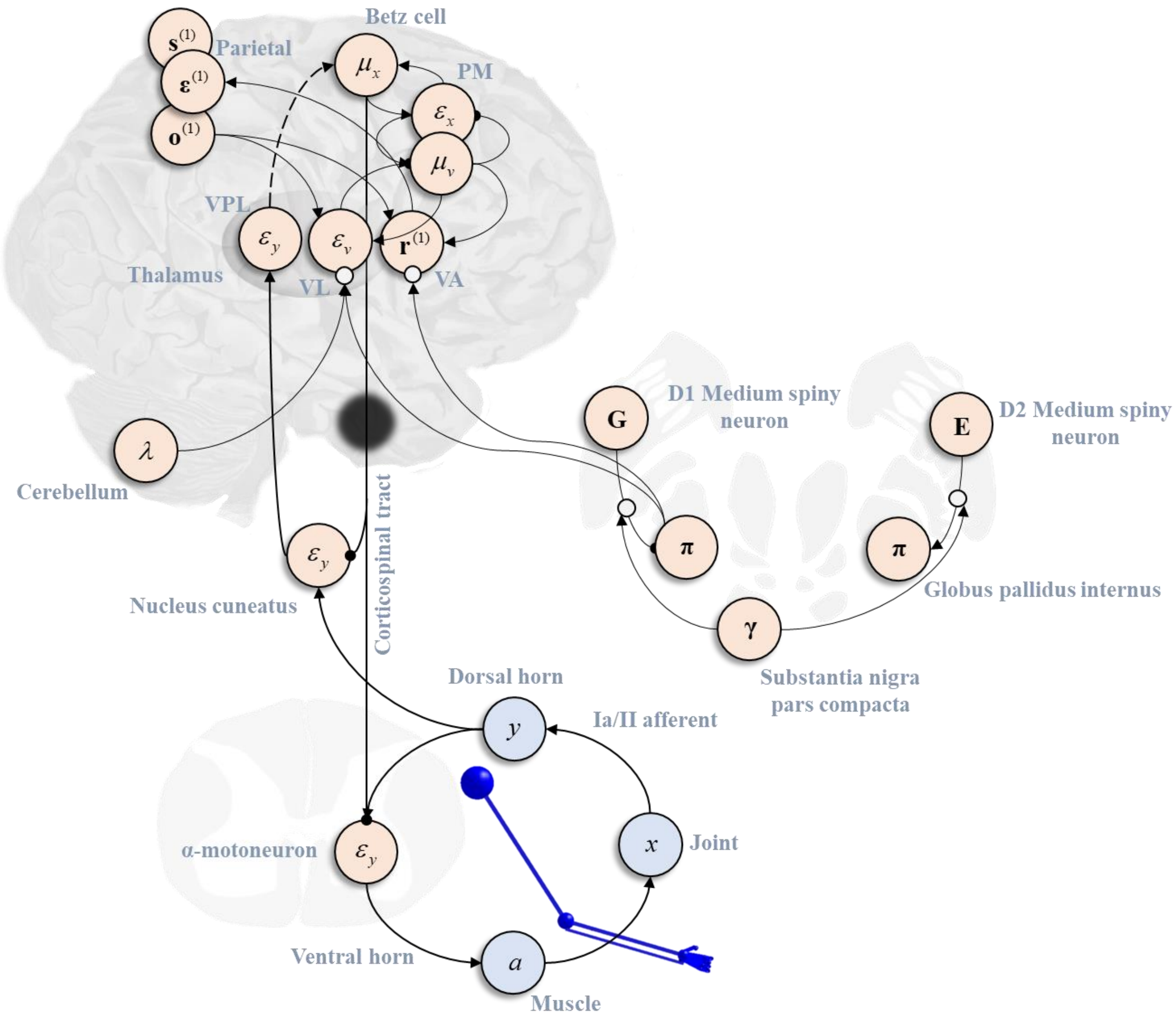


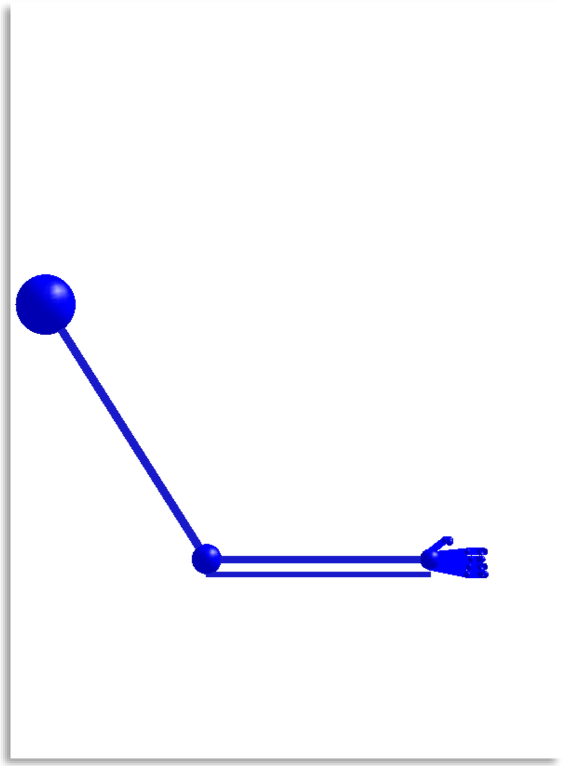
Θ(x)
x4
x1
x5
x2
x6
x3
Θ(x)

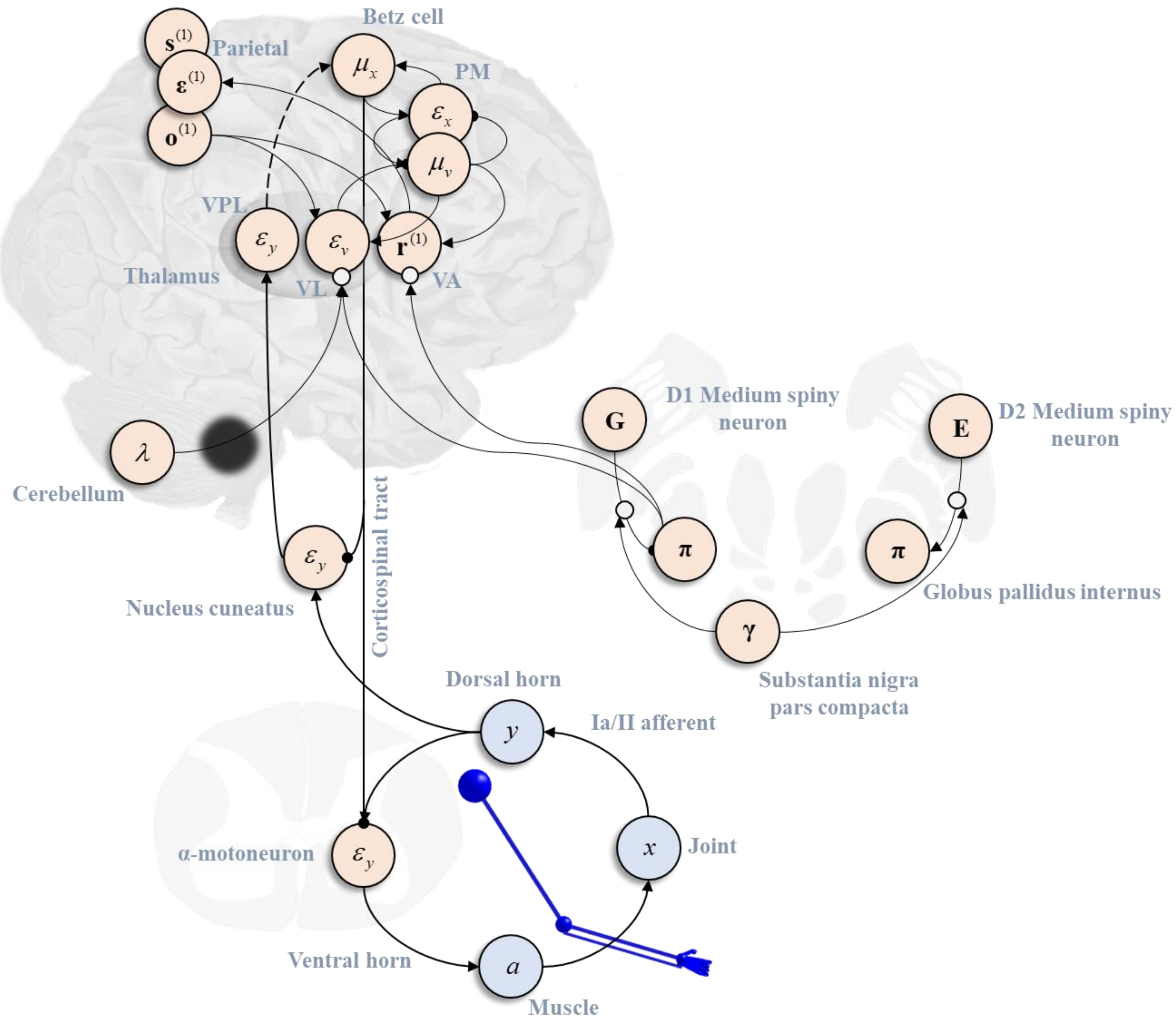


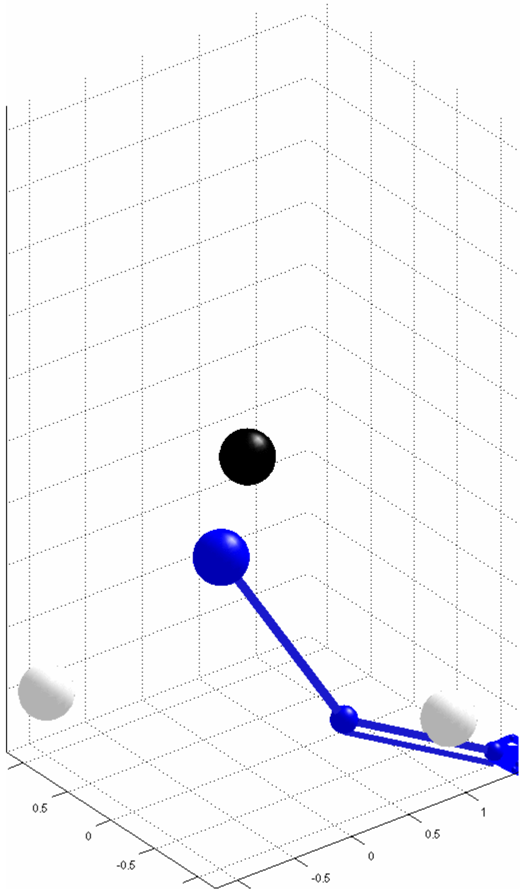

Betz cell
Parietal
PM
VPL
Thalamus
VL
VA
Cerebellum
Nucleus cuneatus
Corticospinal tract
D1 Medium spiny neuron
D2 Medium spiny neuron
Globus pallidus internus
Substantia nigra pars compacta
Dorsal horn
Ia/II afferent
Joint
α-motoneuron
Ventral horn
Muscle

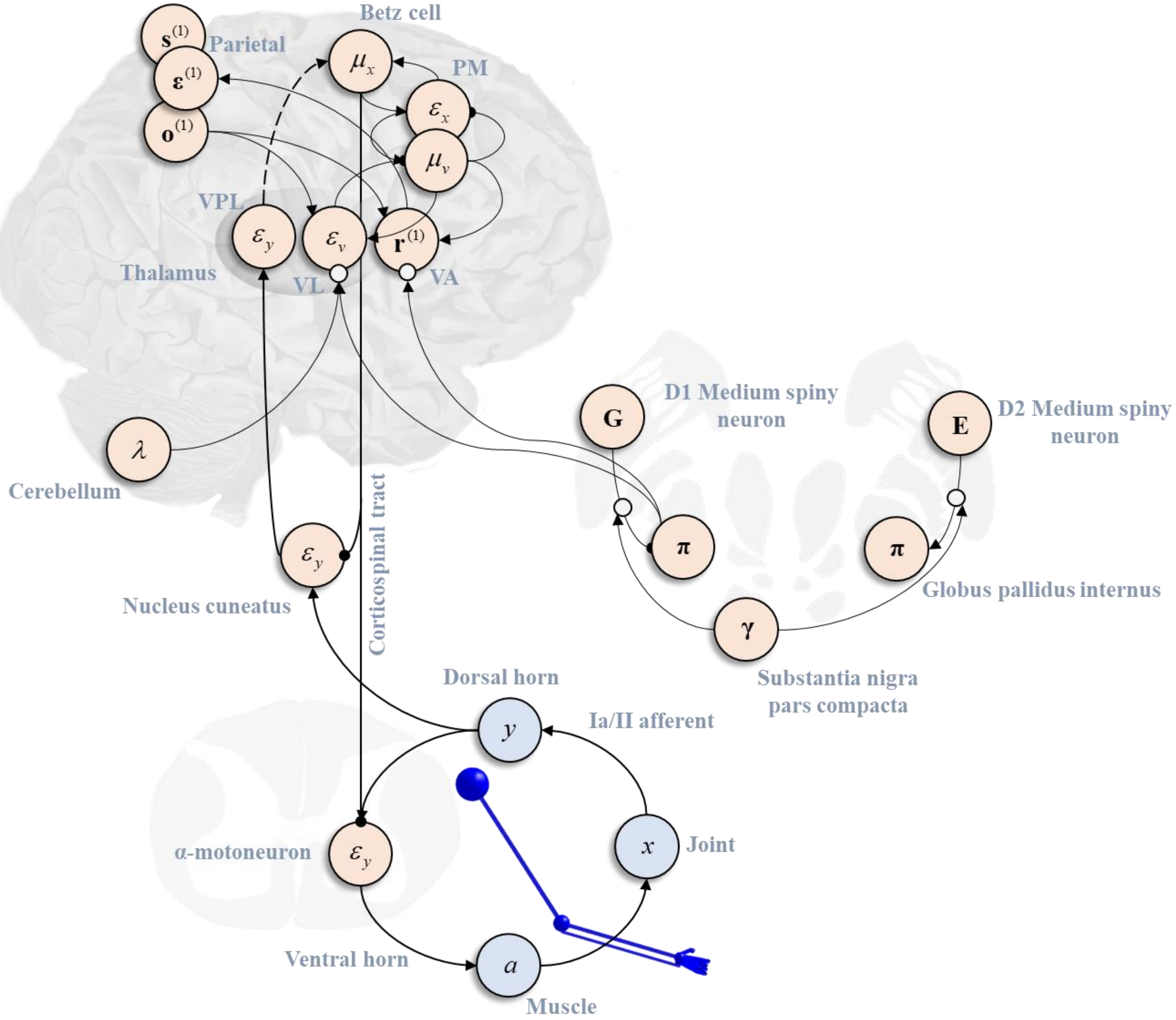


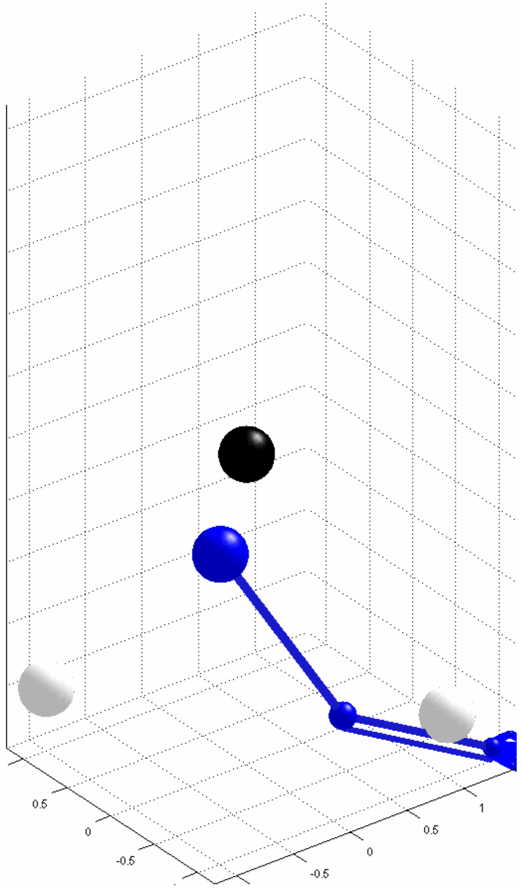

Betz cell
Parietal
PM
VPL
Thalamus
VL
VA
D1 Medium spiny
neuron
D2 Medium spiny
neuron
Cerebellum
Corticospinal tract
Nucleus cuneatus
Globus pallidus internus
Substantia nigra
pars compacta
Dorsal horn
Ia/II afferent
Joint
α-motoneuron
Ventral horn
Muscle

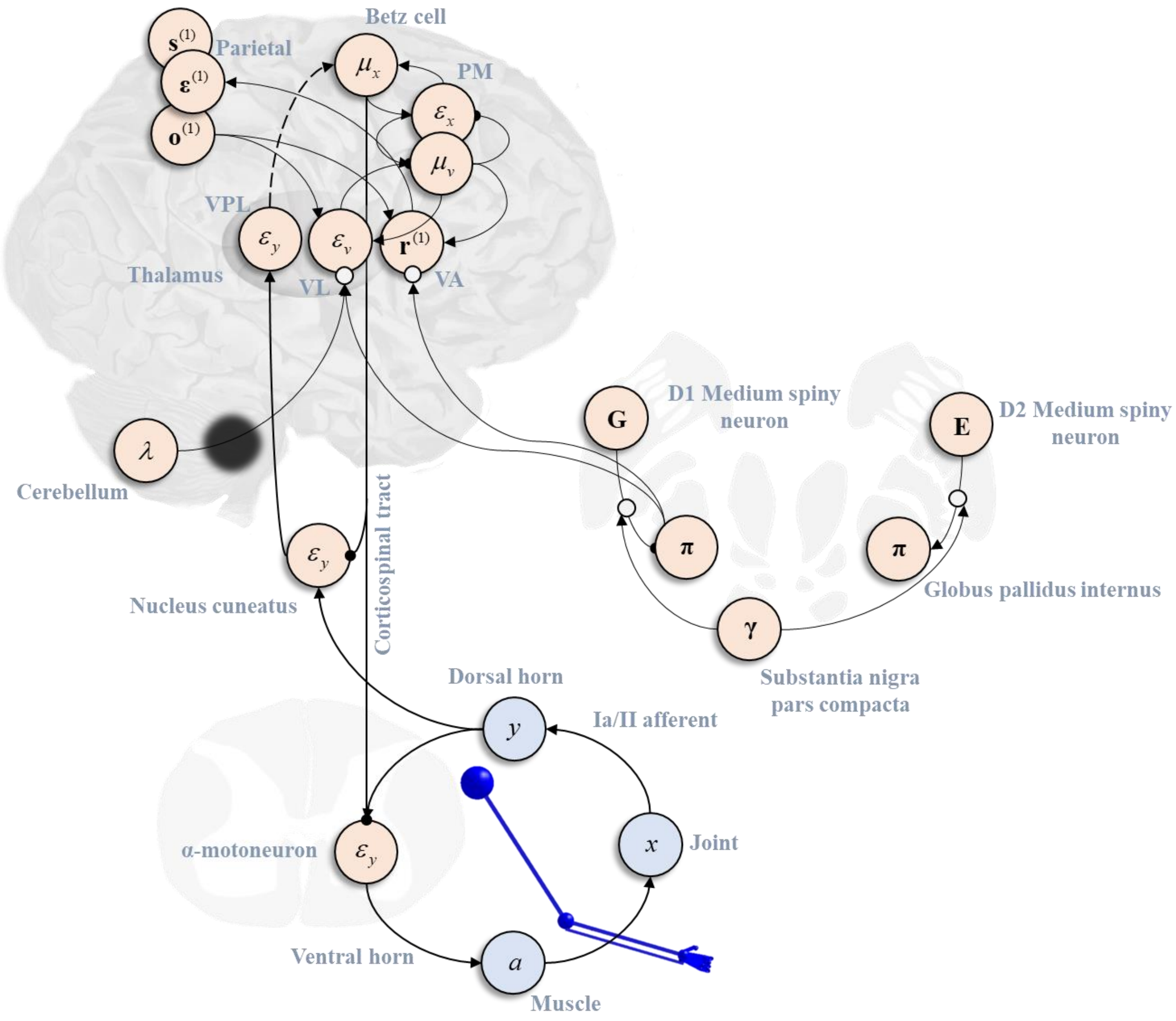


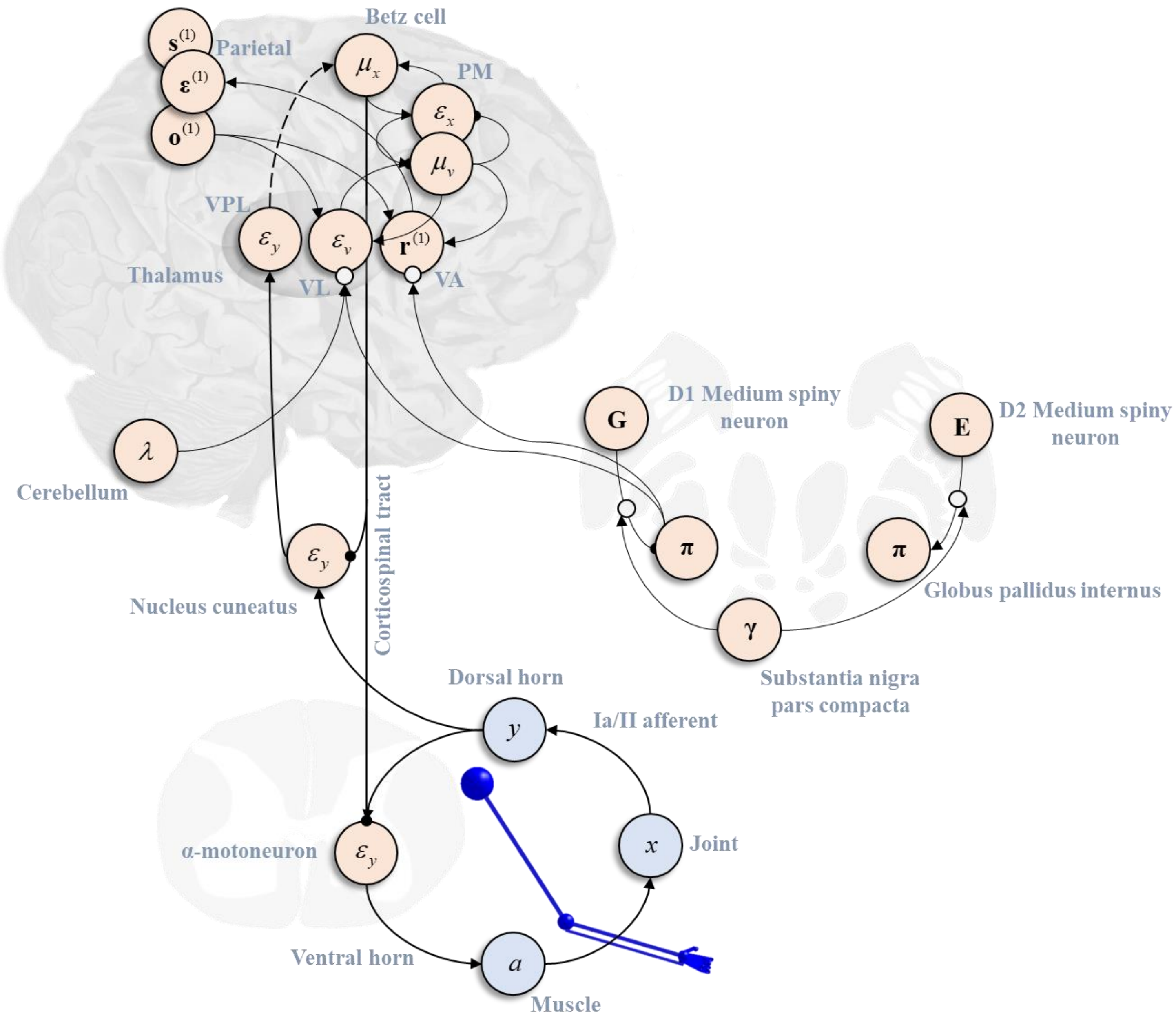

Betz cell
Parietal
PM
VPL
Thalamus
VL
VA
Cerebellum
D1 Medium spiny
neuron
D2 Medium spiny
neuron
Nucleus cuneatus
Corticospinal tract
Globus pallidus internus
Substantia nigra
pars compacta
Dorsal horn
Ia/II afferent
Joint
α-motoneuron
Ventral horn
Muscle


0.5
0
-0.5
-0.5
0
0.5
1

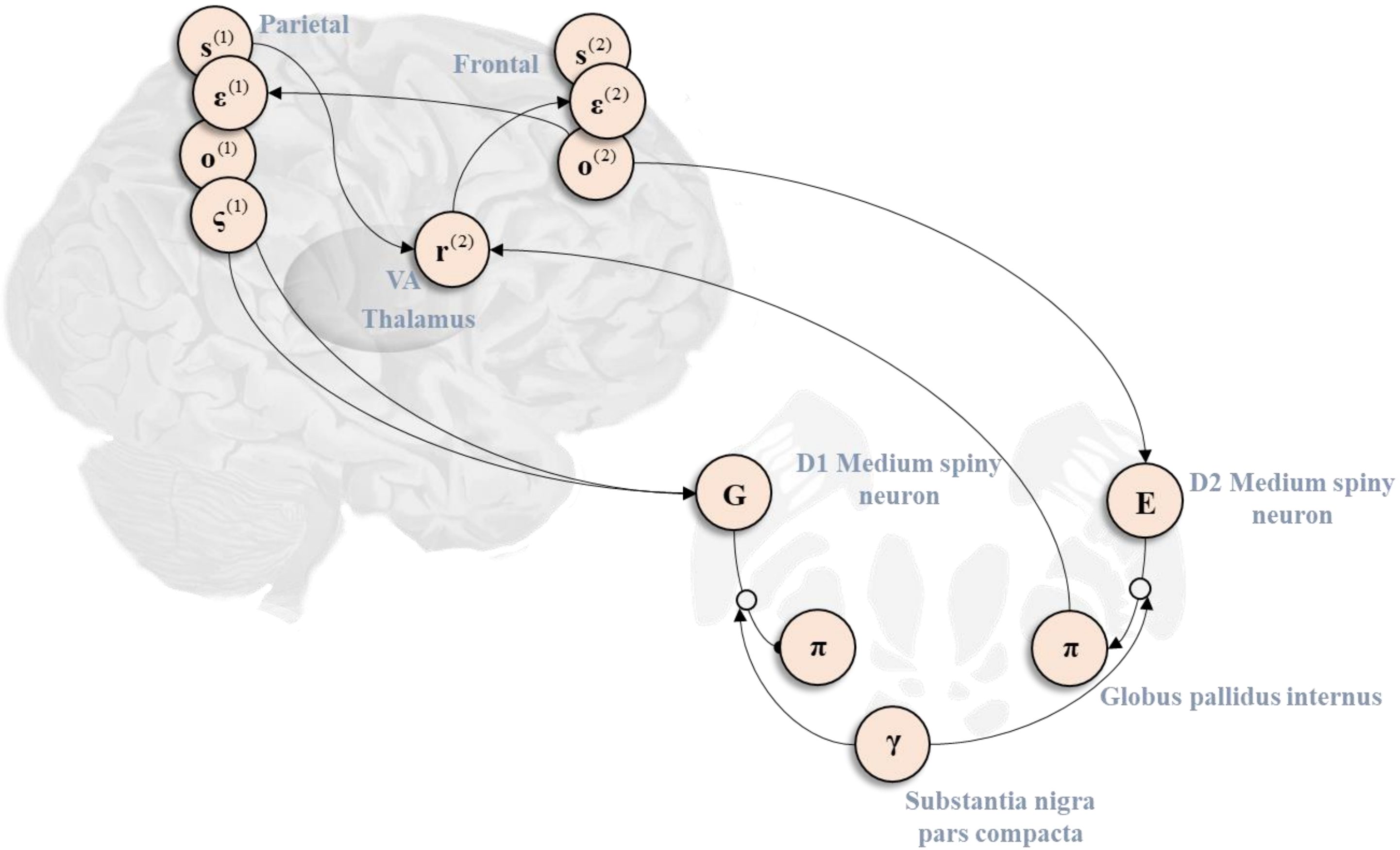

Betz cell
Parietal
PM
$\mathbf{s}^{(1)}$
$\boldsymbol{\varepsilon}^{(1)}$
$\mathbf{o}^{(1)}$
μ_x
ε_x
μ_v
VPL
ε_y
ε_v
$\mathbf{r}^{(1)}$
Thalamus
VL
VA
λ
Cerebellum
G
D1 Medium spiny neuron
E
D2 Medium spiny neuron
$\boldsymbol{\pi}$
$\boldsymbol{\pi}$
Globus pallidus internus
ε_y
Nucleus cuneatus
Corticospinal tract
$\boldsymbol{\gamma}$
Substantia nigra pars compacta
Dorsal horn
y
Ia/II afferent
x
Joint
α-motoneuron
ε_y
Ventral horn
a
Muscle

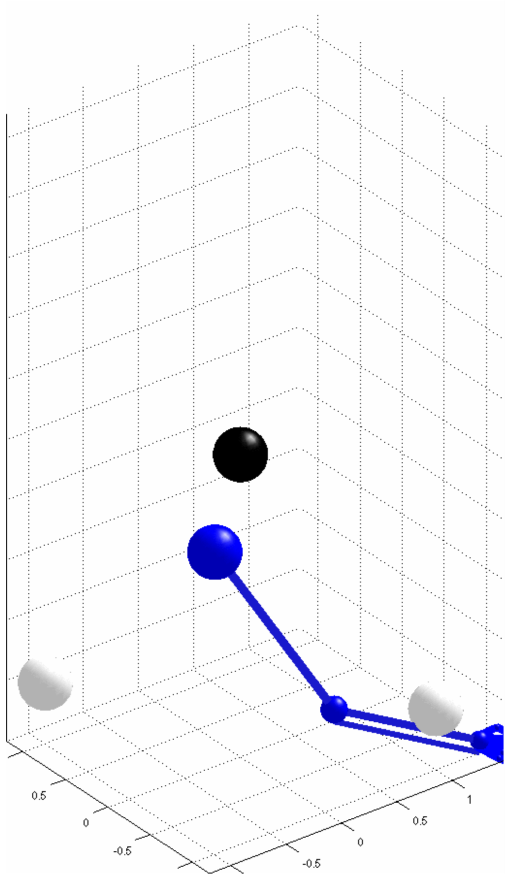


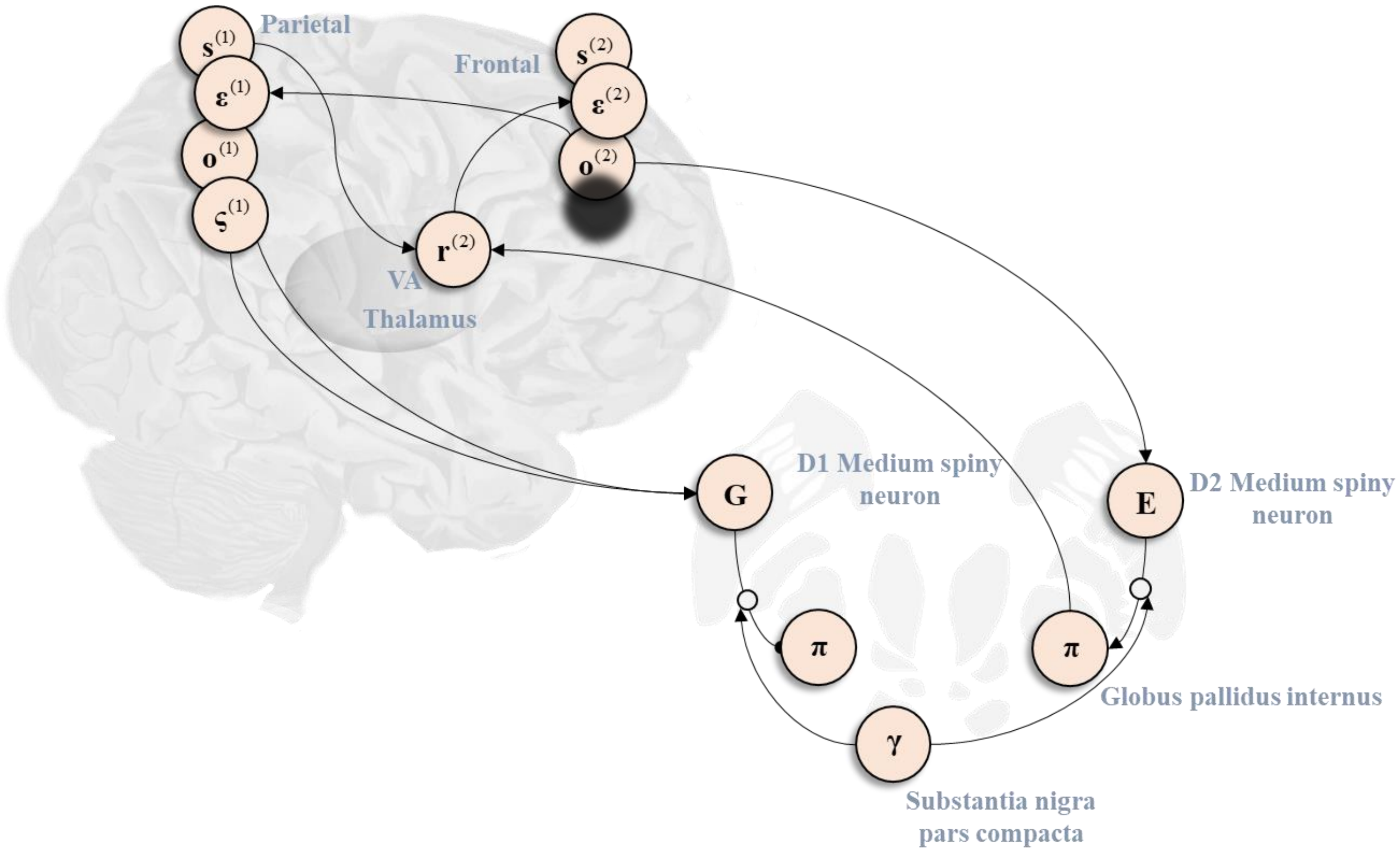

Betz cell
Parietal
PM
VPL
Thalamus
VL
VA
Cerebellum
Corticospinal tract
Nucleus cuneatus
D1 Medium spiny neuron
D2 Medium spiny neuron
Globus pallidus internus
Substantia nigra pars compacta
Dorsal horn
Ia/II afferent
Joint
α-motoneuron
Ventral horn
Muscle

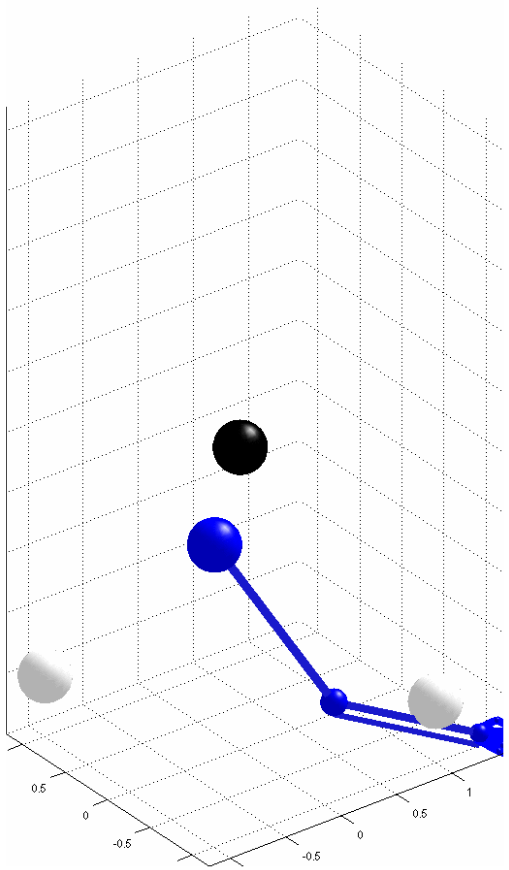


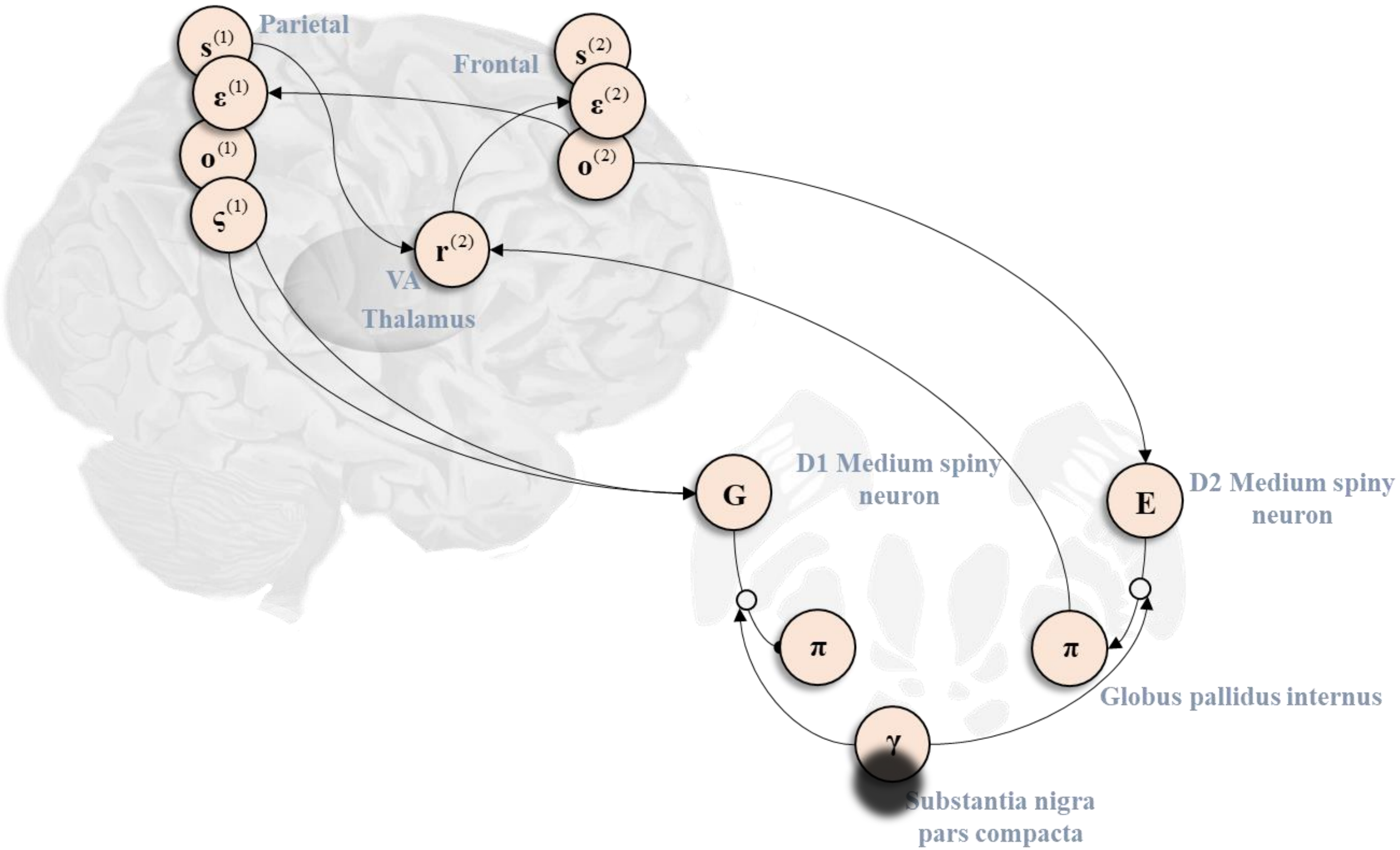
Betz cell
Parietal
PM
VPL
Thalamus
VL
VA
Cerebellum
Nucleus cuneatus
Corticospinal tract
D1 Medium spiny neuron
D2 Medium spiny neuron
Globus pallidus internus
Substantia nigra pars compacta
Dorsal horn
Ia/II afferent
Joint
α-motoneuron
Ventral horn
Muscle


Parietal
Frontal
VA
Thalamus
D1 Medium spiny neuron
D2 Medium spiny neuron
Globus pallidus internus
Substantia nigra pars compacta

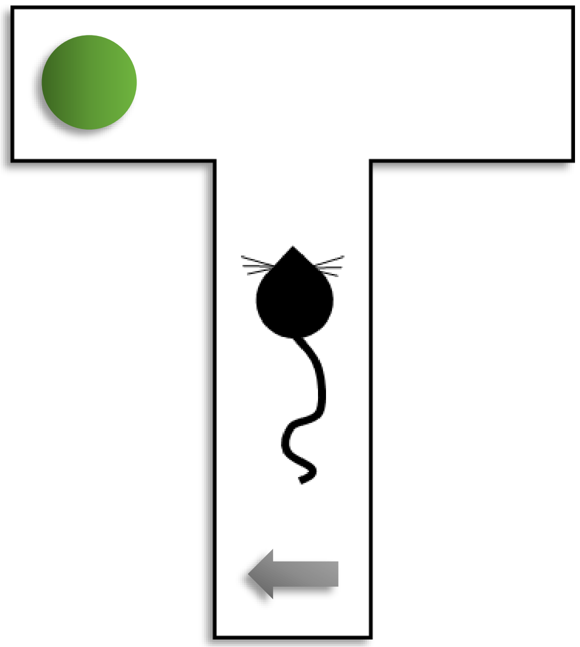

0.5
0
-0.5
-0.5
0
0.5
1

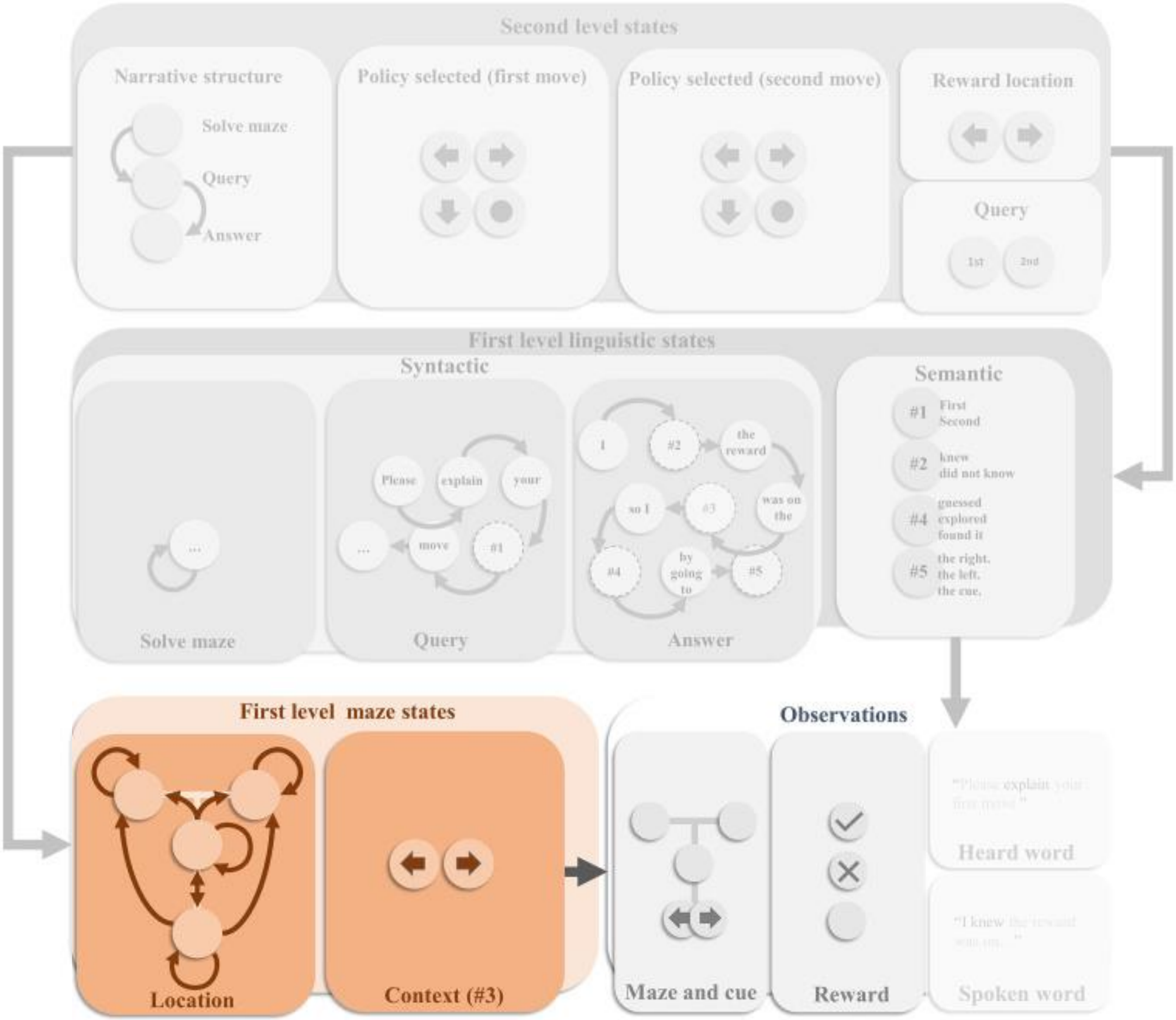

Parietal
Frontal
s(1)
ε(1)
o(1)
ς(1)
s(2)
ε(2)
o(2)
r(2)
VA
Thalamus
G
D1 Medium spiny neuron
E
D2 Medium spiny neuron
π
π
Globus pallidus internus
γ
Substantia nigra pars compacta


0.5
0
-0.5
-0.5
0
0.5
1

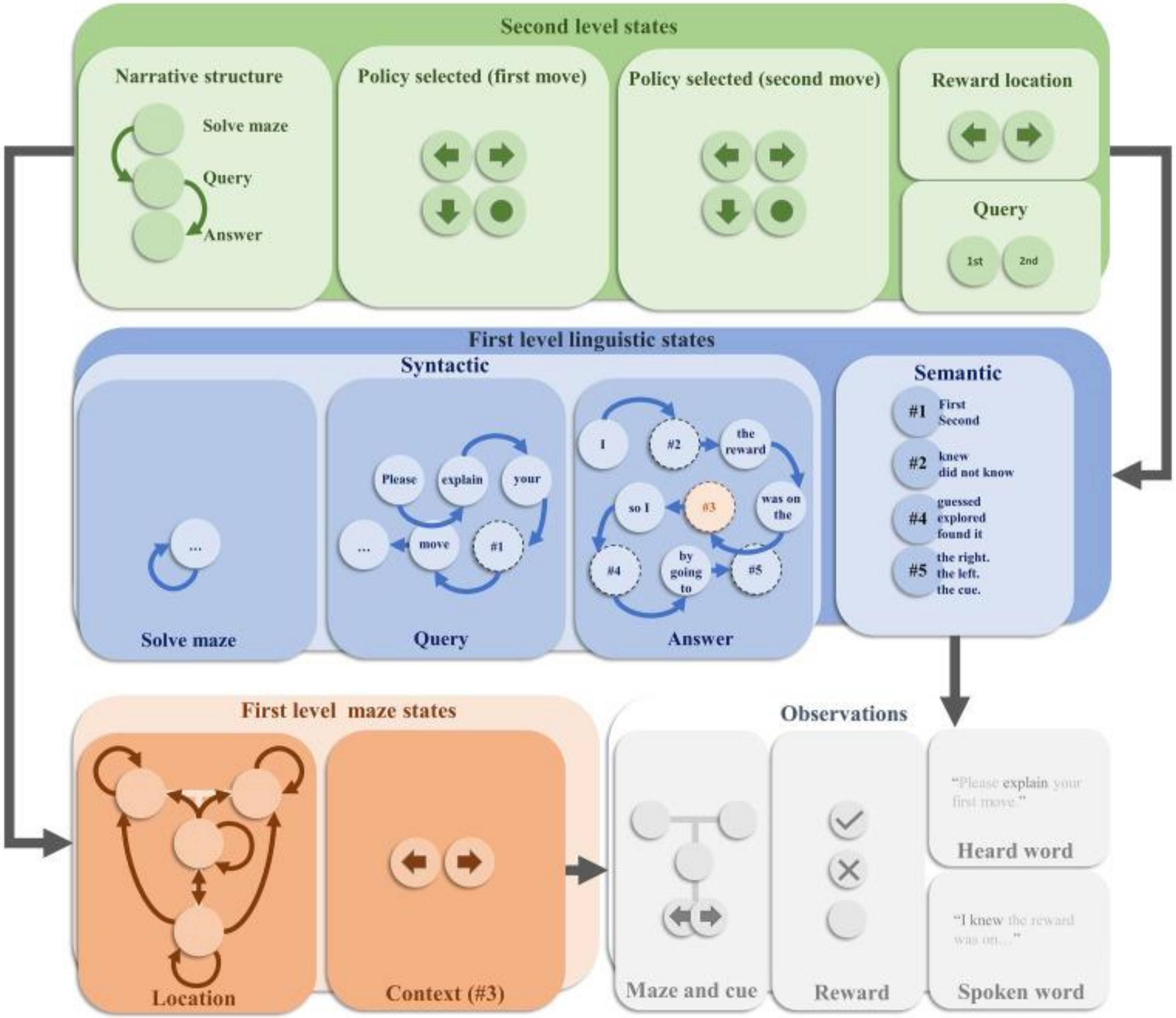

Parietal
Frontal
VA
Thalamus
D1 Medium spiny
neuron
D2 Medium spiny
neuron
Globus pallidus internus
Substantia nigra
pars compacta
$\mathbf{s}^{(1)}$
$\boldsymbol{\varepsilon}^{(1)}$
$\mathbf{o}^{(1)}$
$\varsigma^{(1)}$
$\mathbf{r}^{(2)}$
$\mathbf{s}^{(2)}$
$\boldsymbol{\varepsilon}^{(2)}$
$\mathbf{o}^{(2)}$
G
E
π
π
γ

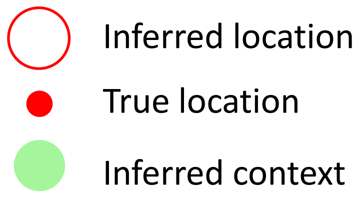
Navigating a T-maze

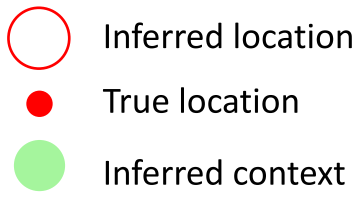


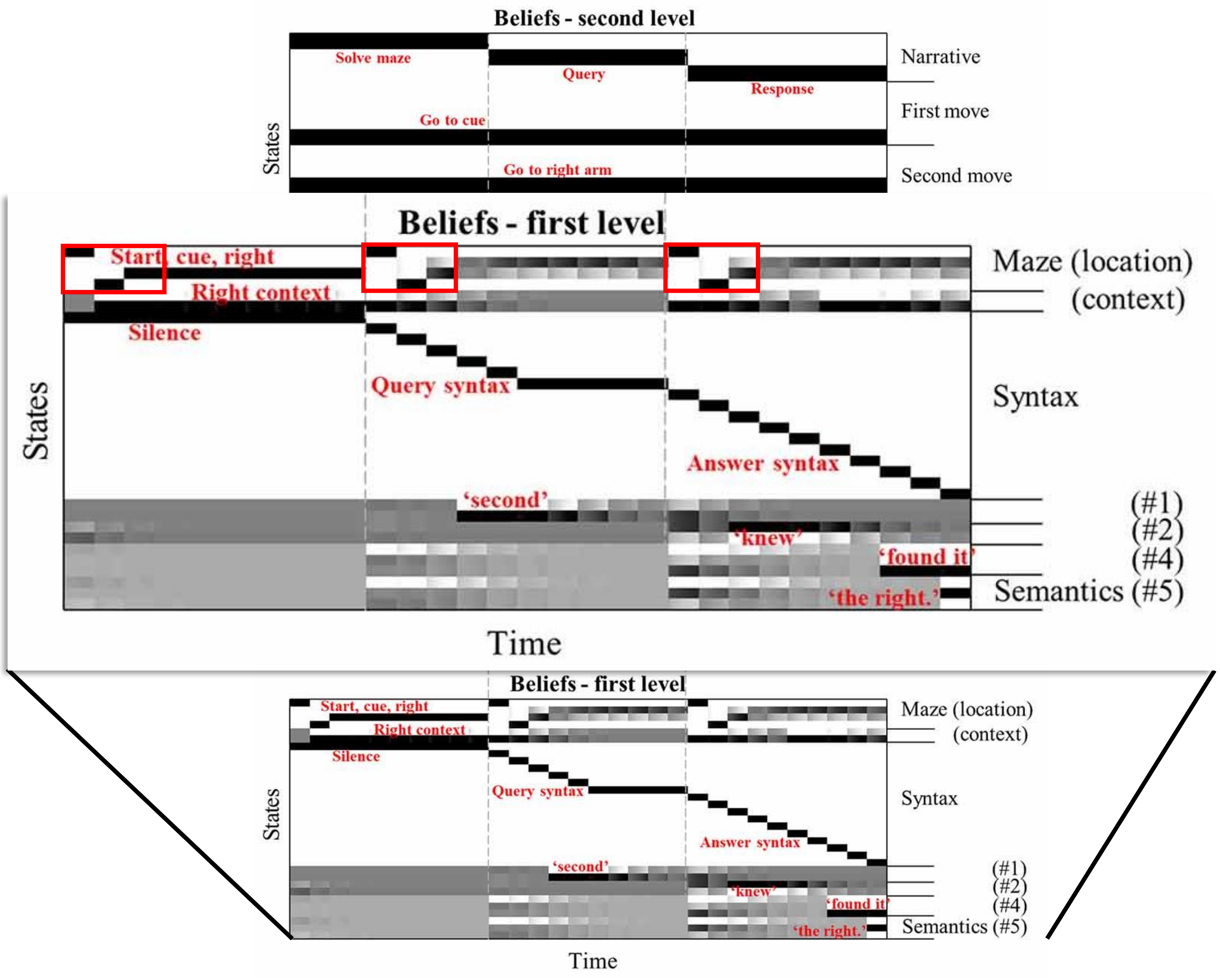


Second level states
Narrative structure
Solve maze
Query
Answer
Policy selected (first move)
Policy selected (second move)
Reward location
Query
1st
2nd
First level linguistic states
Syntactic
...
Solve maze
Please
explain
your
...
move
#1
Query
I
#2
the reward
was on the
so I
#3
#4
by going to
#5
Answer
Semantic
#1
First
Second
#2
knew
did not know
#4
guessed
explored
found it
#5
the right.
the left.
the cue.
First level maze states
Location
Context (#3)
Observations
Maze and cue
Reward
Heard word
Spoken word


Second level states
Narrative structure
Solve maze
Query
Answer
Policy selected (first move)
Policy selected (second move)
Reward location
Query
1st
2nd
First level linguistic states
Syntactic
...
Solve maze
Please
explain
your
#1
move
...
Query
I
#2
the reward
was on the
#3
so I
#4
by going to
#5
Answer
Semantic
#1 First Second
#2 knew did not know
#4 guessed explored found it
#5 the right. the left. the cue.
First level maze states
Location
Context (#3)
Observations
Maze and cue
Reward
"Please explain your first move."
Heard word
"I knew the reward was on..."
Spoken word


Inferred location
True location
Inferred context

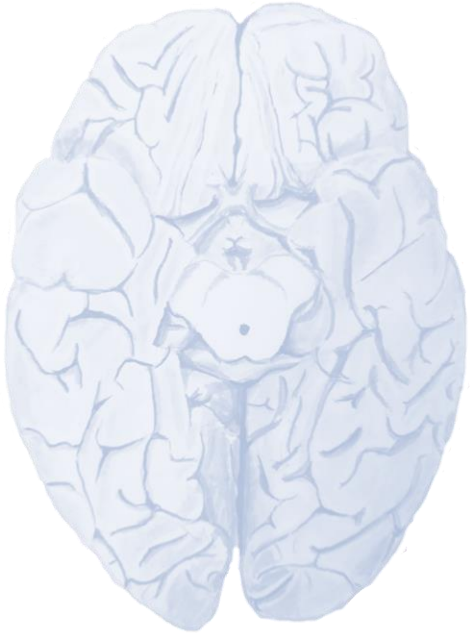

Inferred location
True location
Inferred context



Beliefs - second level
Solve maze
Query
Response
Go to cue
Go to right arm
Narrative
First move
Second move
States
Beliefs - first level
Start, cue, right
Right context
Silence
Query syntax
Answer syntax
'second'
'knew'
'found it'
'the right.'
Maze (location)
(context)
Syntax
(#1)
(#2)
(#4)
Semantics (#5)
States
Time

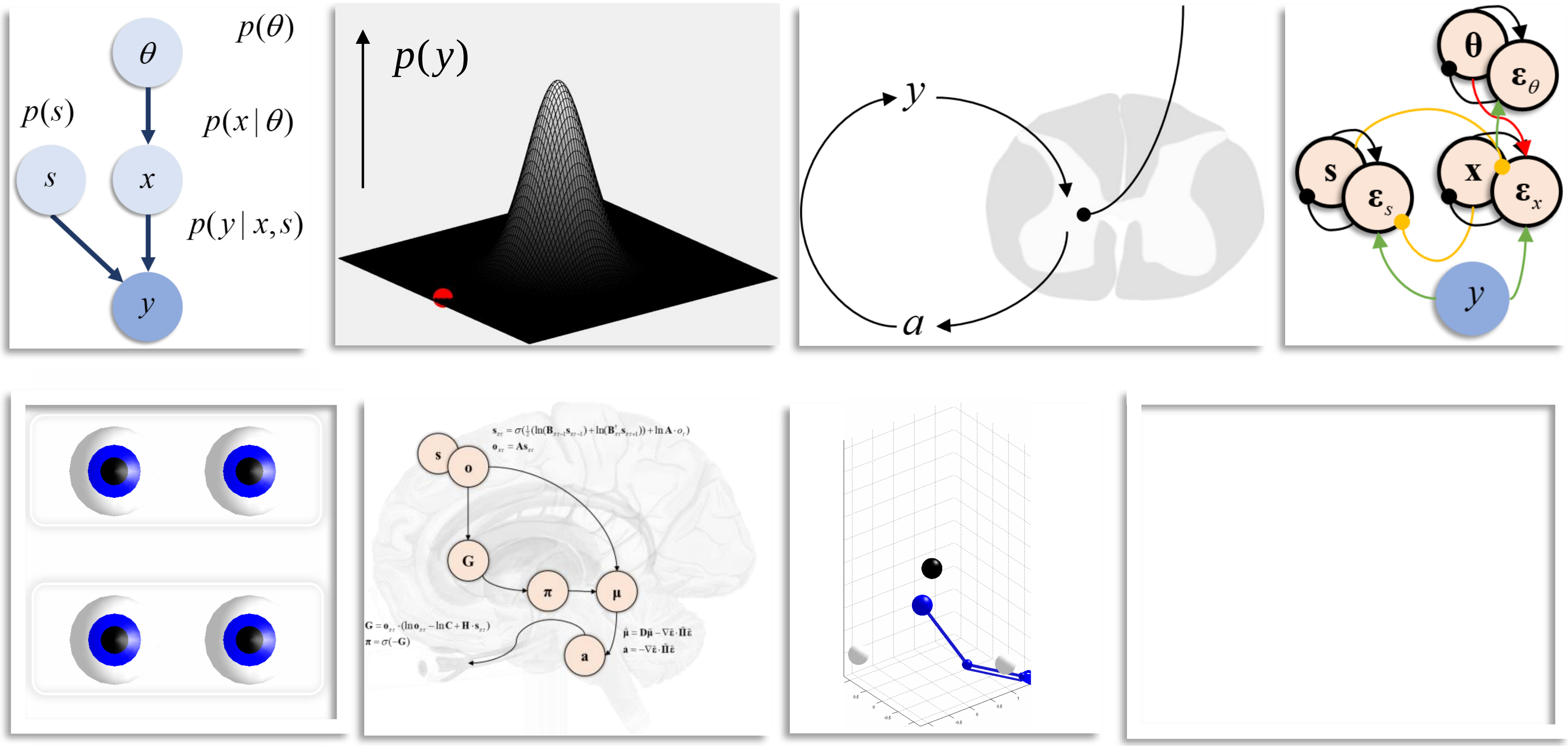
Inferred location
True location
Inferred context

Confabulation



Active Inference

Self-evidencing
Message passing
Discrete time (planning)
Continuous time (movement)
Hierarchical models
Summary



p(θ)
θ
p(s)
p(x | θ)
s
x
p(y | x,s)
y
p(y)
y
a
θ
ε_θ
s
ε_s
x
ε_x
y
s
o
G
π
μ
a

Thanks

Berk Mirza
David Benrimoh
Dimitrije Markovic
Emma Holmes
Giovanni Pezzulo
Jakub Limanowski
Jakob Hohwy
Jelle Bruineberg
Karl Friston
Lance Da Costa
Noor Sajid
Peter Vincent
Rick Adams
Stefan Kiebel
Vishal Rawji
And many others

Rosetrees Trust
Supporting the best in medical research

UCL

wellcome centre human neuroimaging