Sequential Neural Processes


Gautam Singh*[1]


Jaesik Yoon*[2]

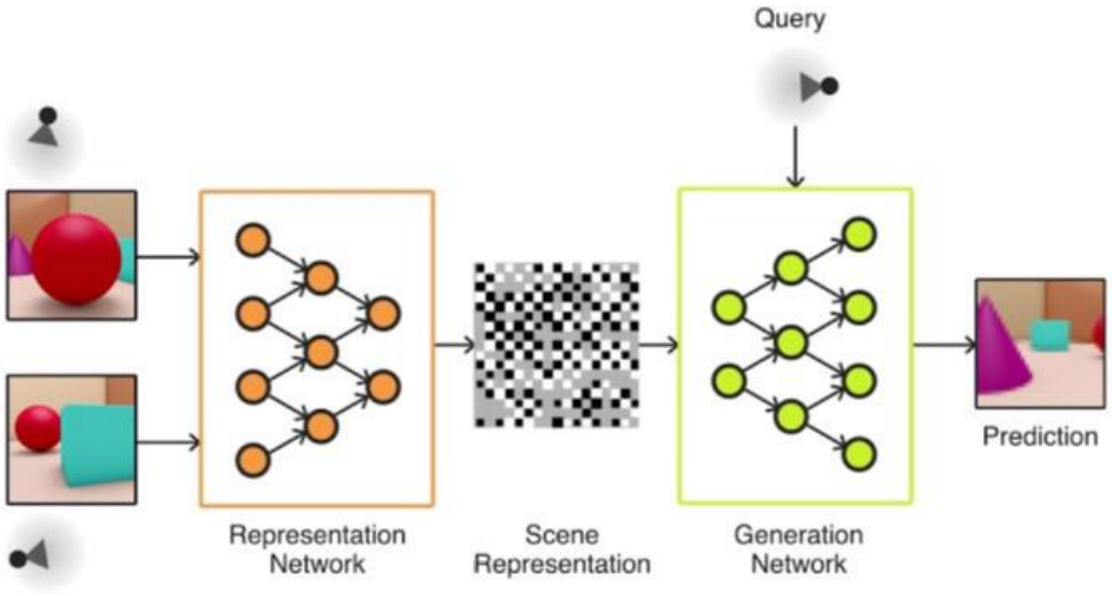

Youngsung Son[3]

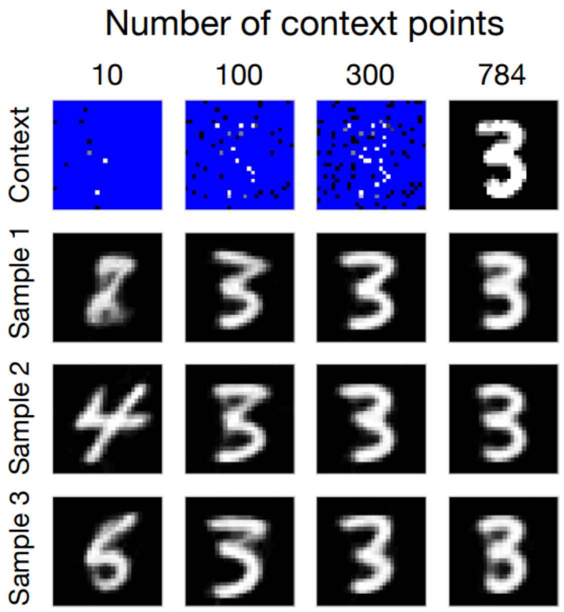

Sungjin Ahn[1]

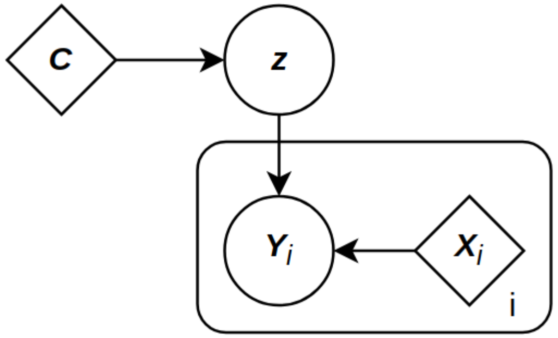
[1]Rutgers University
[2]SAP
[3]ETRI

*Equal Contribution

Background: GQN, NP and Meta-Learning





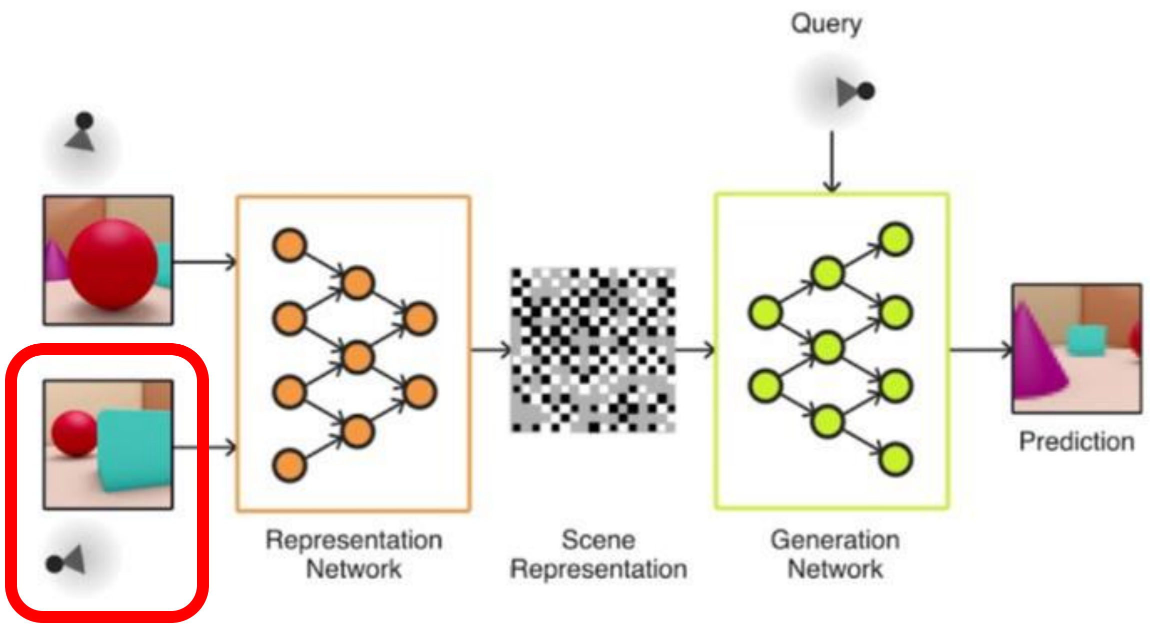


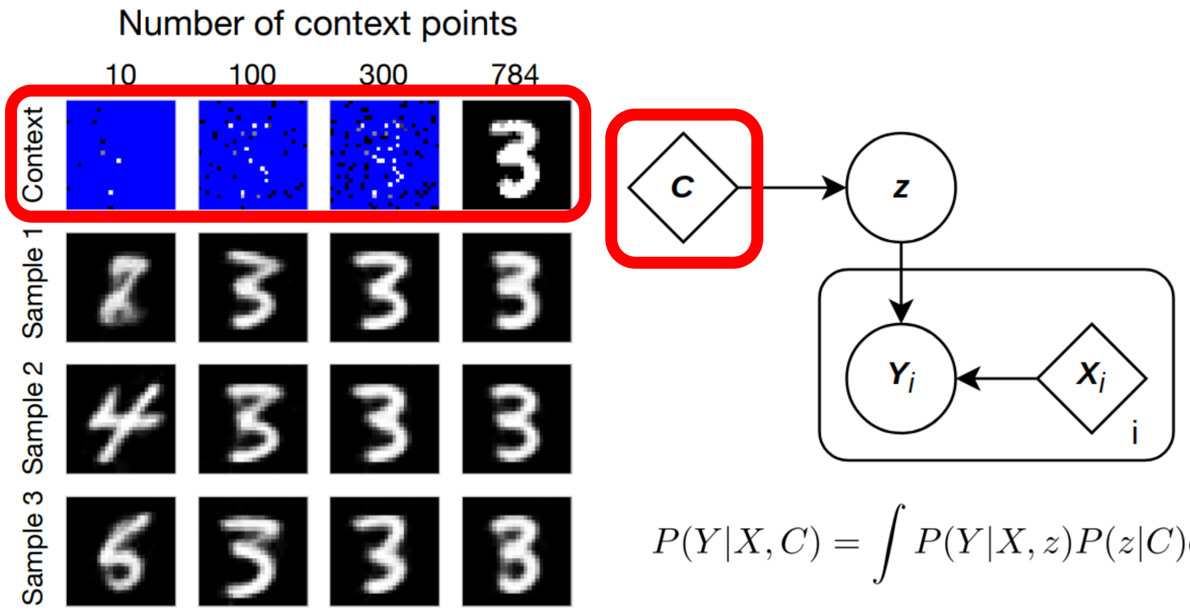
$$P(Y|X,C) = \int P(Y|X,z)P(z|C)\mathrm{d}z$$

"What if the stochastic process also had some underlying temporal dynamics?"

[1] Eslami, SM Ali, et al. "Neural scene representation and rendering." *Science* 360.6394 (2018): 1204-1210.
[2] Garnelo, Marta, et al. "Neural processes." *arXiv preprint arXiv:1807.01622* (2018).

Background: GQN, NP and Meta-Learning



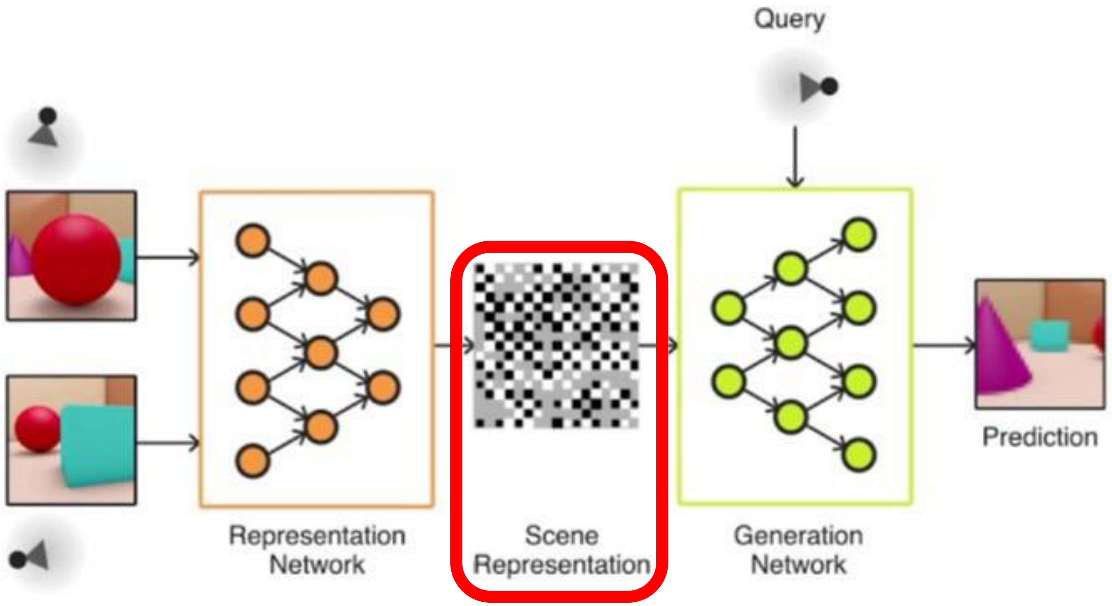


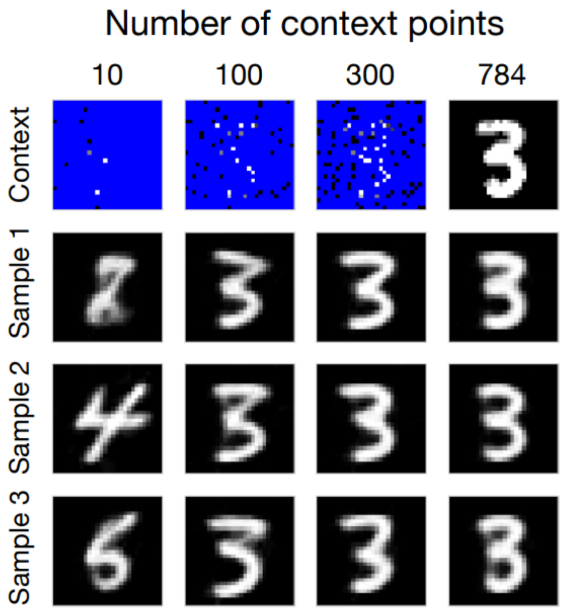
$$P(Y|X,C) = \int P(Y|X,z)P(z|C)\mathrm{d}z$$

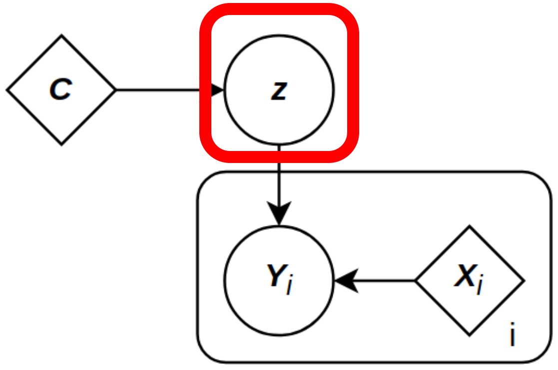
"What if the stochastic process also had some underlying temporal dynamics?"

[1] Eslami, SM Ali, et al. "Neural scene representation and rendering." *Science* 360.6394 (2018): 1204-1210.
[2] Garnelo, Marta, et al. "Neural processes." *arXiv preprint arXiv:1807.01622* (2018).

Background: GQN, NP and Meta-Learning





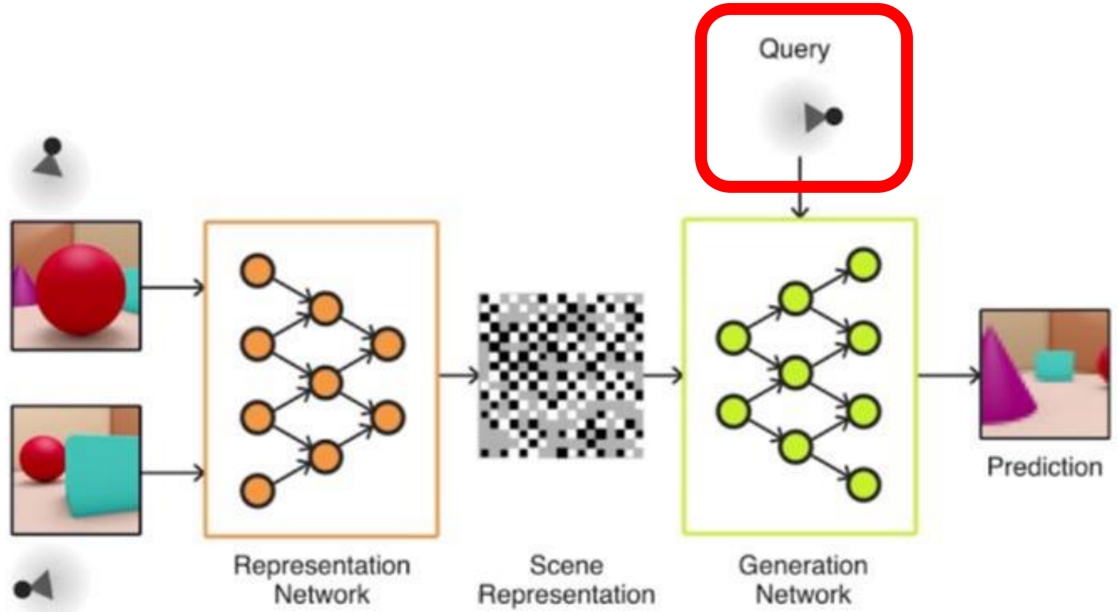


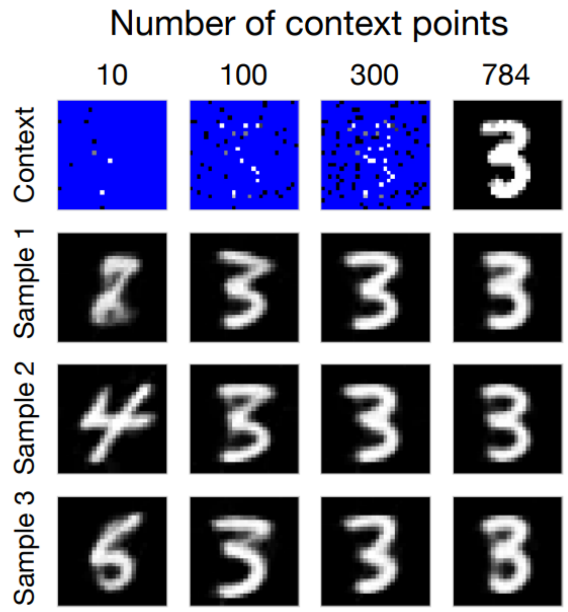
$$P(Y|X,C) = \int P(Y|X,z)P(z|C)\mathrm{d}z$$

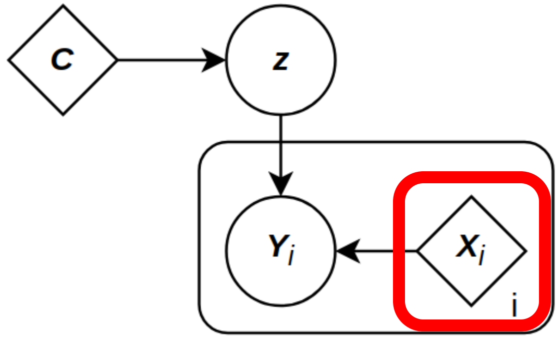
"What if the stochastic process also had some underlying temporal dynamics?"

[1] Eslami, SM Ali, et al. "Neural scene representation and rendering." *Science* 360.6394 (2018): 1204-1210.
[2] Garnelo, Marta, et al. "Neural processes." *arXiv preprint arXiv:1807.01622* (2018).

Background: GQN, NP and Meta-Learning





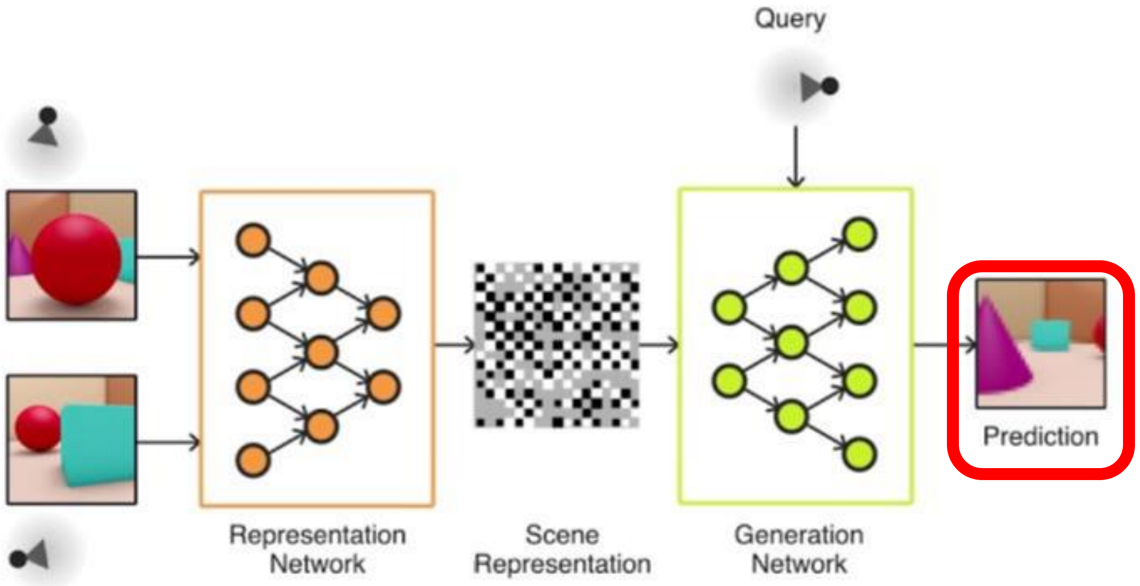


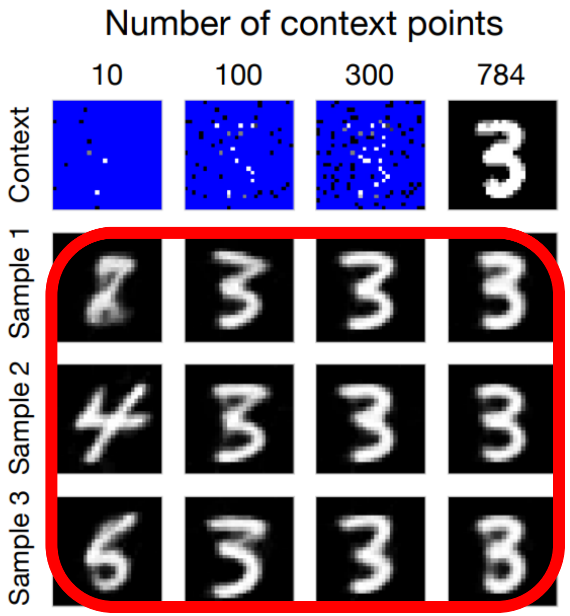
$$P(Y|X,C) = \int P(Y|X,z)P(z|C)\mathrm{d}z$$

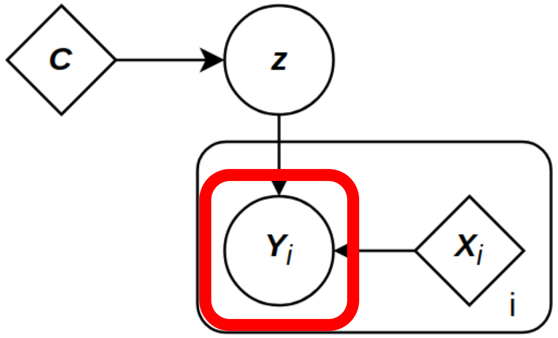
"What if the stochastic process also had some underlying temporal dynamics?"

[1] Eslami, SM Ali, et al. "Neural scene representation and rendering." *Science* 360.6394 (2018): 1204-1210.
[2] Garnelo, Marta, et al. "Neural processes." *arXiv preprint arXiv:1807.01622* (2018).

Background: GQN, NP and Meta-Learning





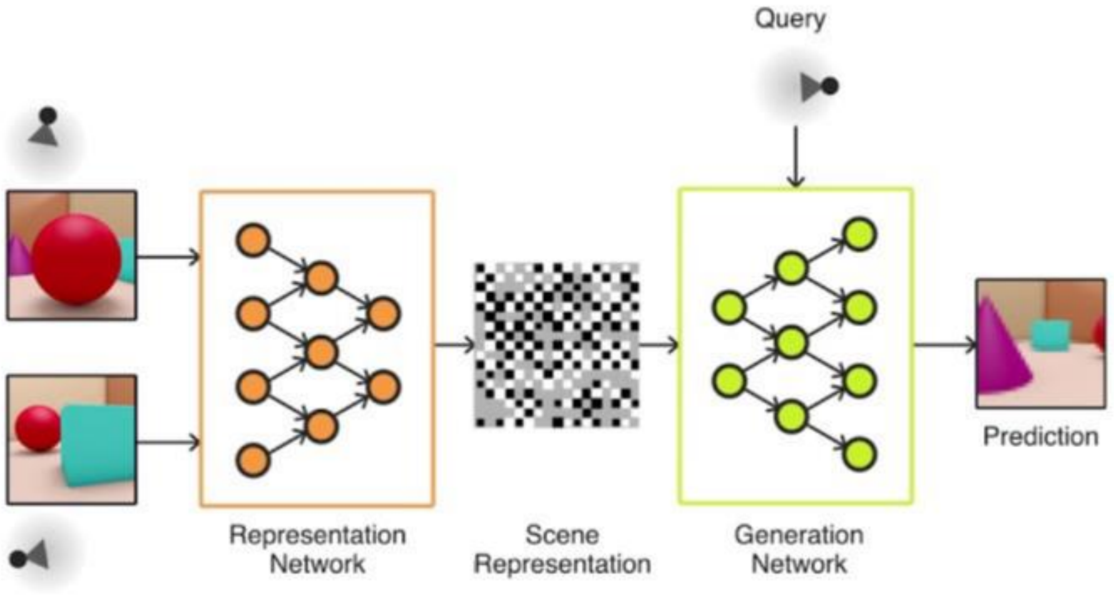


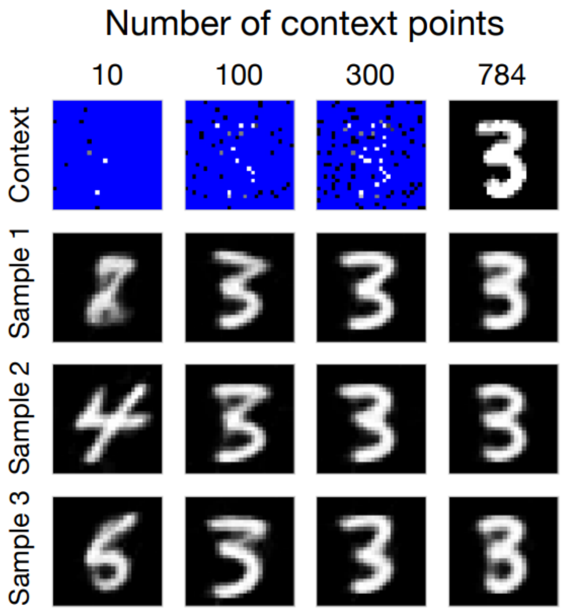
$$P(Y|X,C) = \int P(Y|X,z)P(z|C)\mathrm{d}z$$

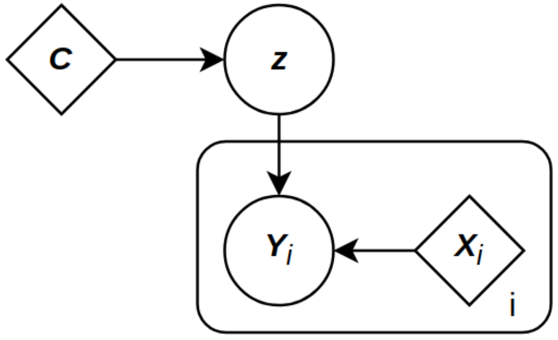
"What if the stochastic process also had some underlying temporal dynamics?"

[1] Eslami, SM Ali, et al. "Neural scene representation and rendering." *Science* 360.6394 (2018): 1204-1210.
[2] Garnelo, Marta, et al. "Neural processes." *arXiv preprint arXiv:1807.01622* (2018).

Background: GQN, NP and Meta-Learning





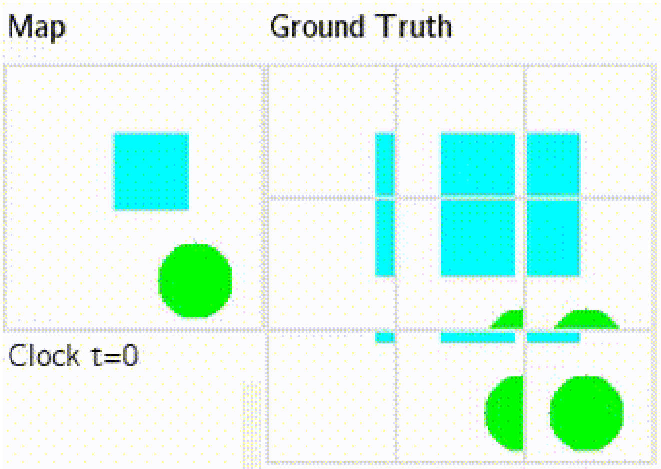


$$P(Y|X,C) = \int P(Y|X,z)P(z|C)\mathrm{d}z$$

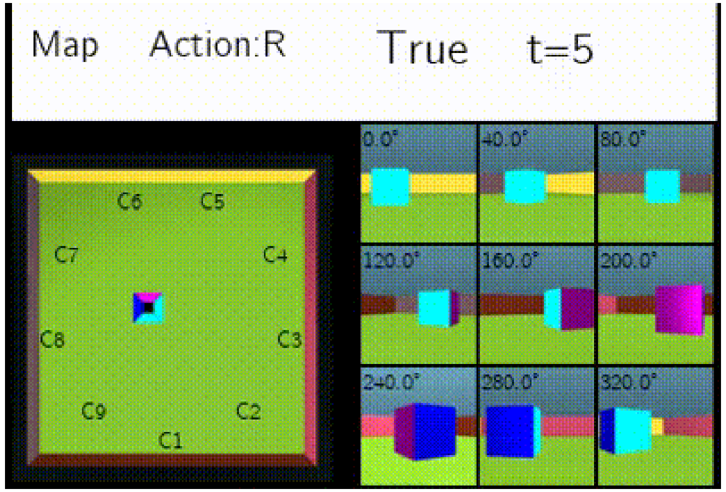
"What if the stochastic process also had some underlying temporal dynamics?"

[1] Eslami, SM Ali, et al. "Neural scene representation and rendering." *Science* 360.6394 (2018): 1204-1210.
[2] Garnelo, Marta, et al. "Neural processes." *arXiv preprint arXiv:1807.01622* (2018).

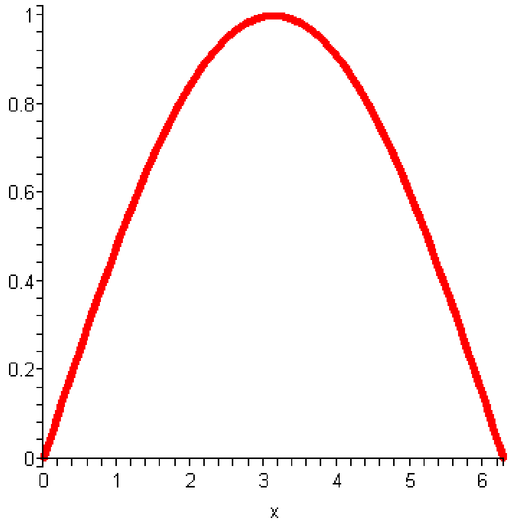
Stochastic Processes with Time Structure



Bouncing 2D Shapes

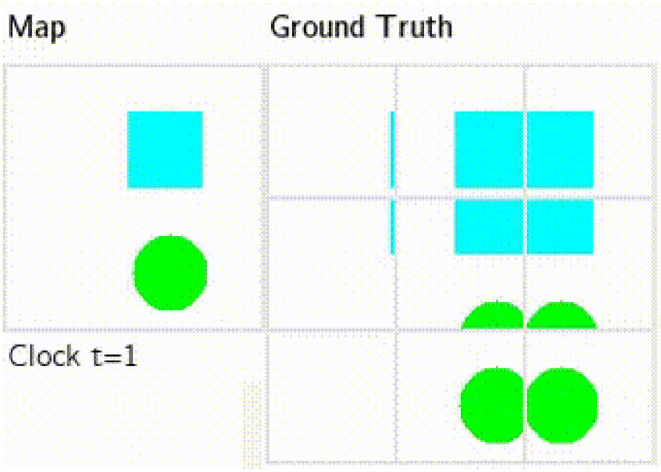


Moving 3D Object

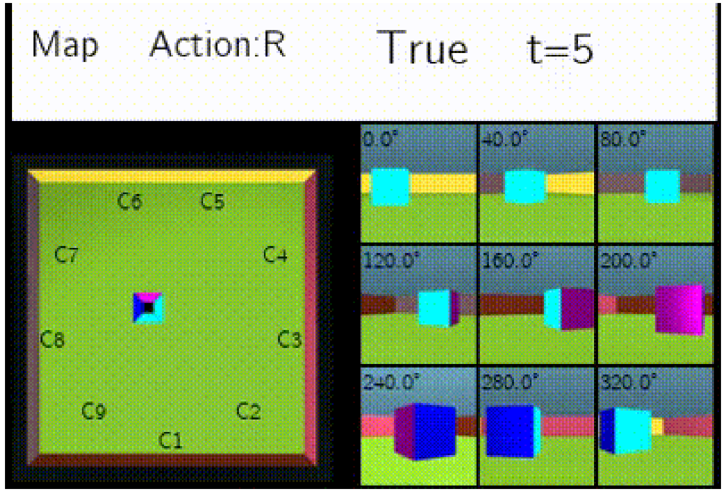


Temperature of a rod changing with time

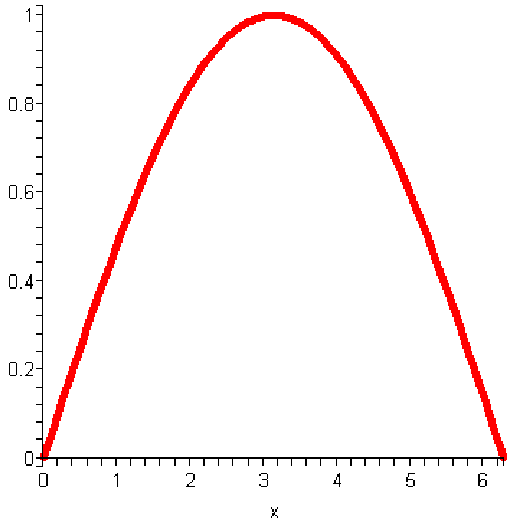
Stochastic Processes with Time Structure



Bouncing 2D Shapes

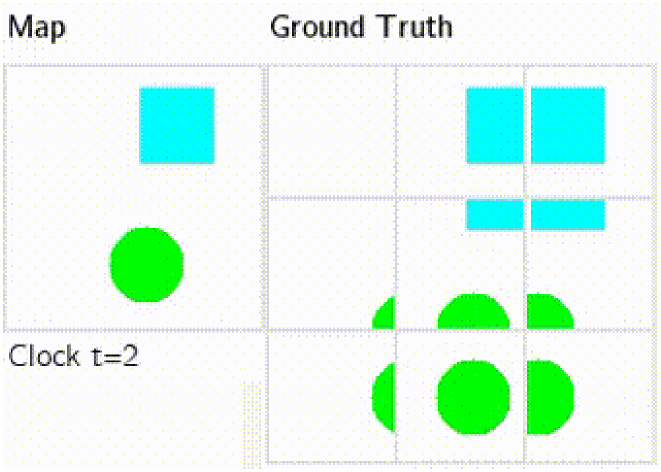


Moving 3D Object

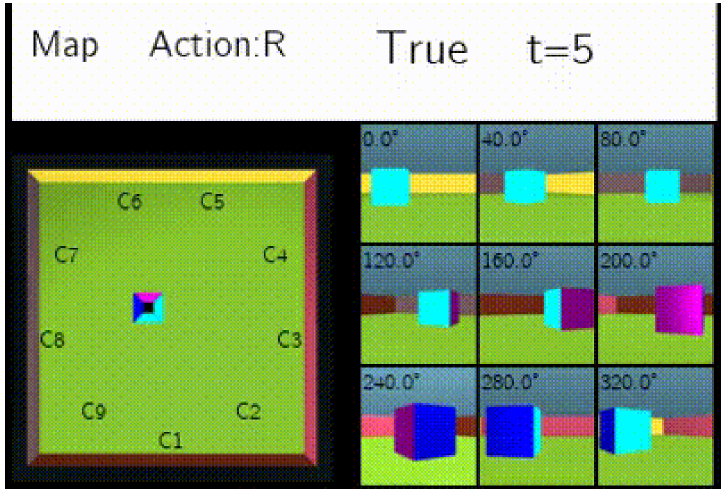


Temperature of a rod changing with time

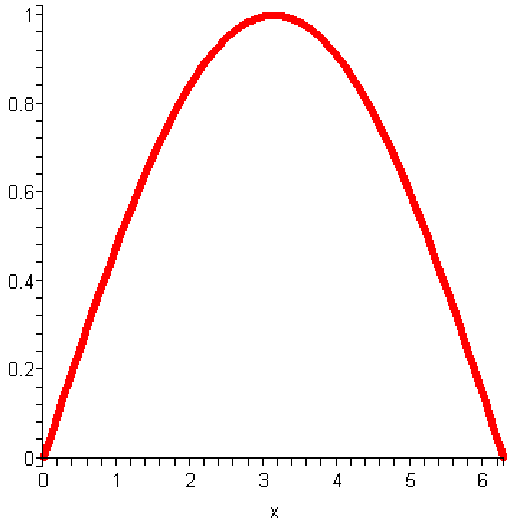
Stochastic Processes with Time Structure



Bouncing 2D Shapes

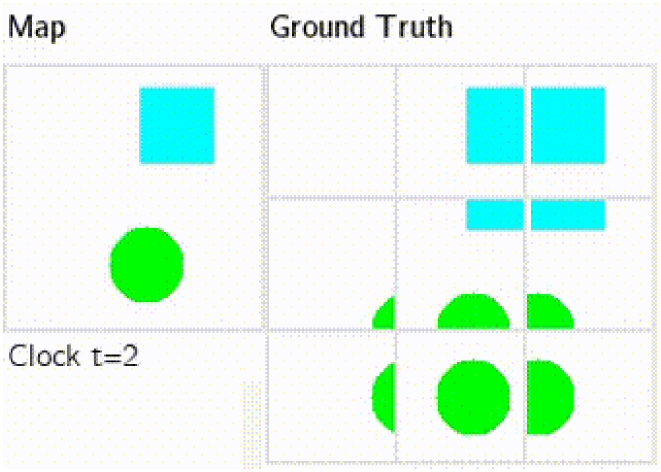


Moving 3D Object

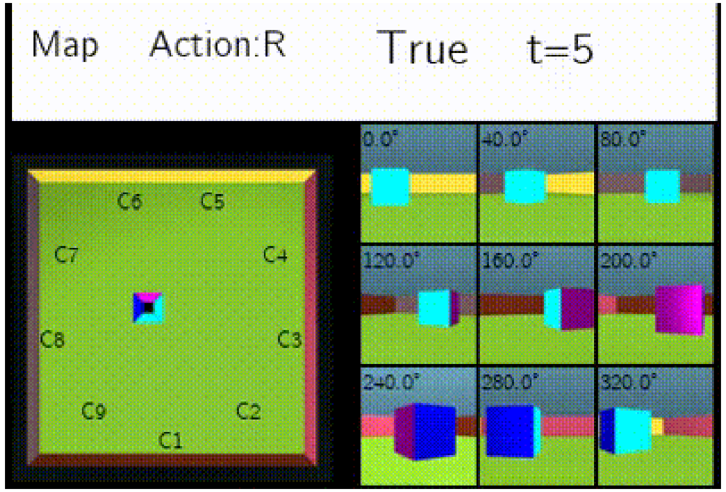


Temperature of a rod changing with time

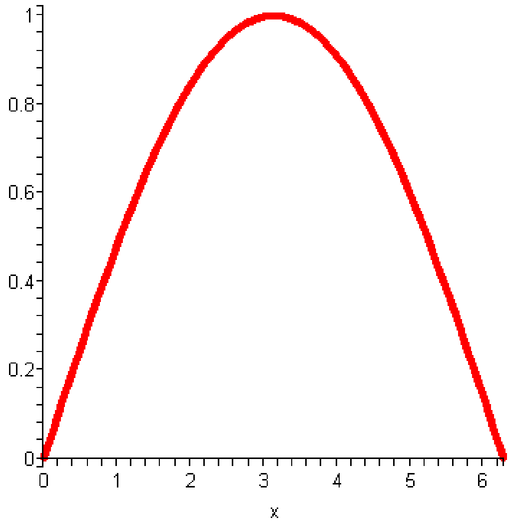
Stochastic Processes with Time Structure



Bouncing 2D Shapes

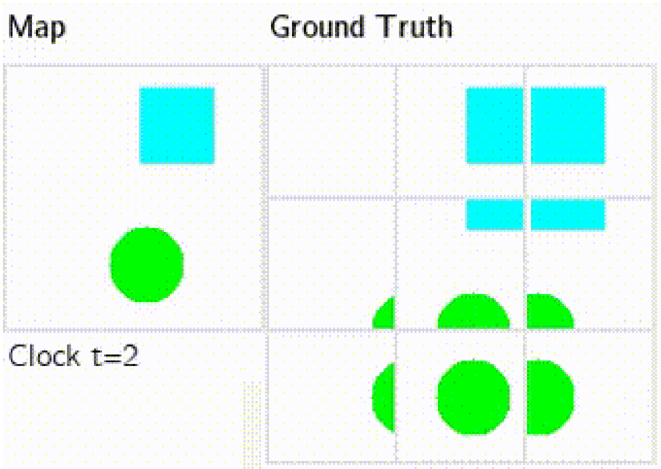


Moving 3D Object

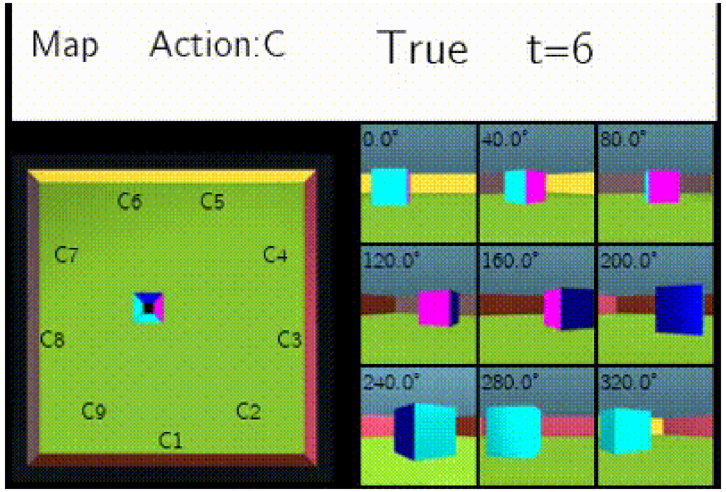


Temperature of a rod changing with time

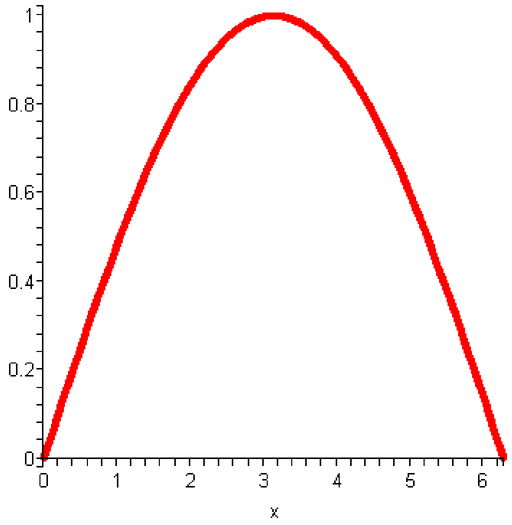
Stochastic Processes with Time Structure



Bouncing 2D Shapes

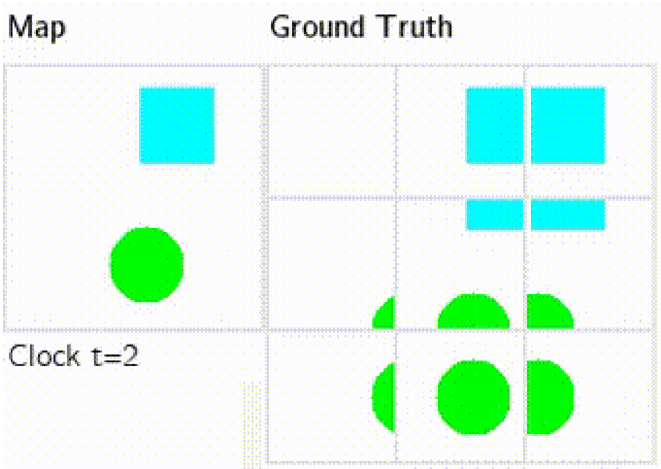


Moving 3D Object

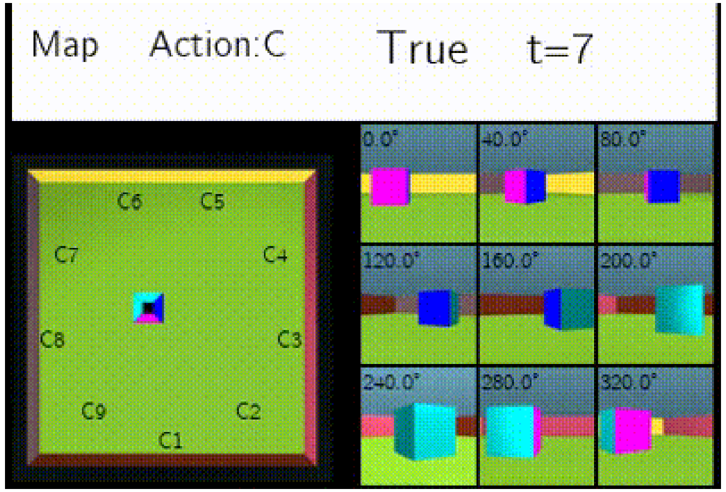


Temperature of a rod changing with time

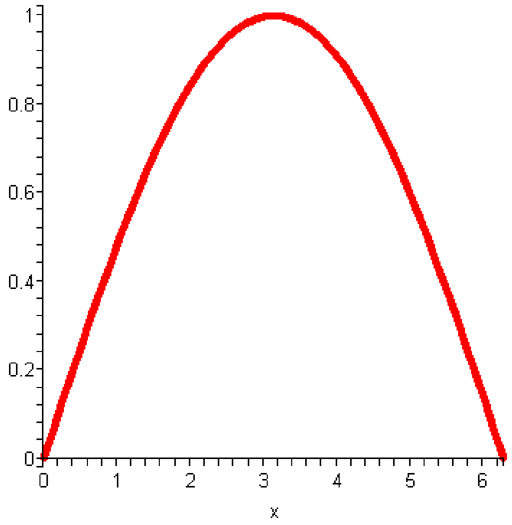
Stochastic Processes with Time Structure



Bouncing 2D Shapes

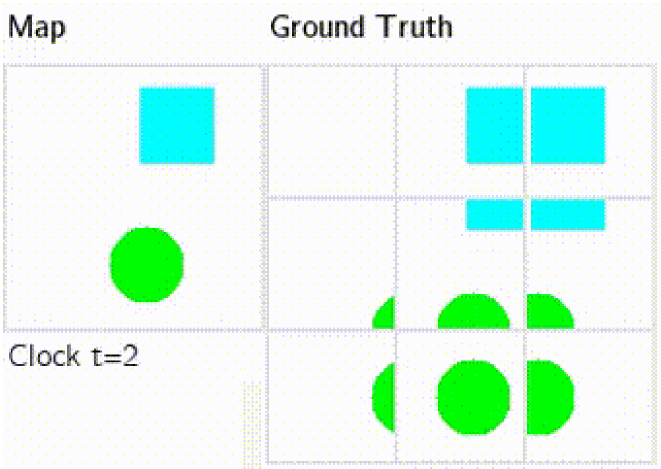


Moving 3D Object

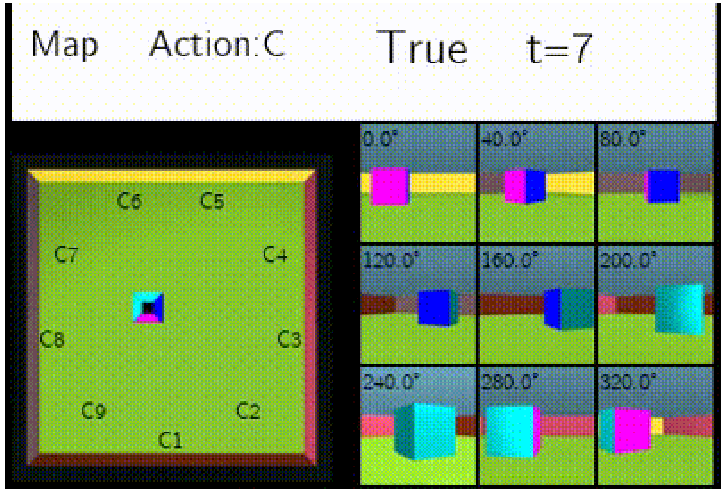


Temperature of a rod changing with time

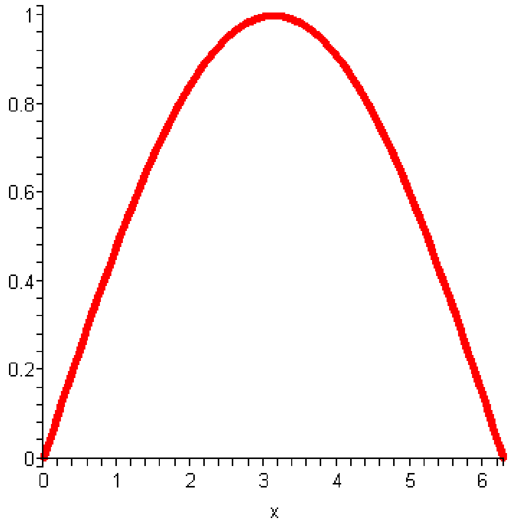
Stochastic Processes with Time Structure



Bouncing 2D Shapes

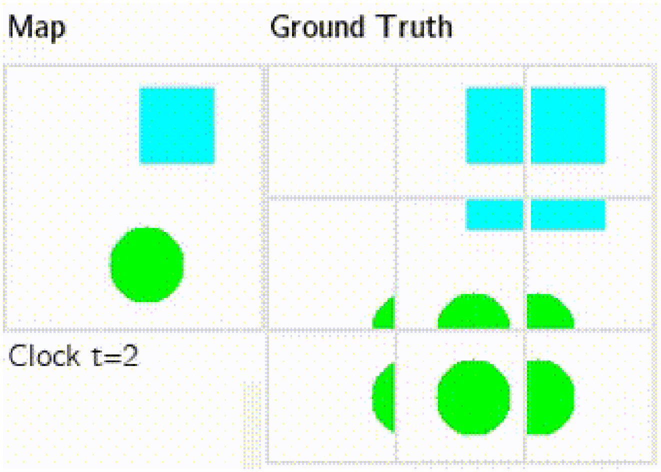


Moving 3D Object

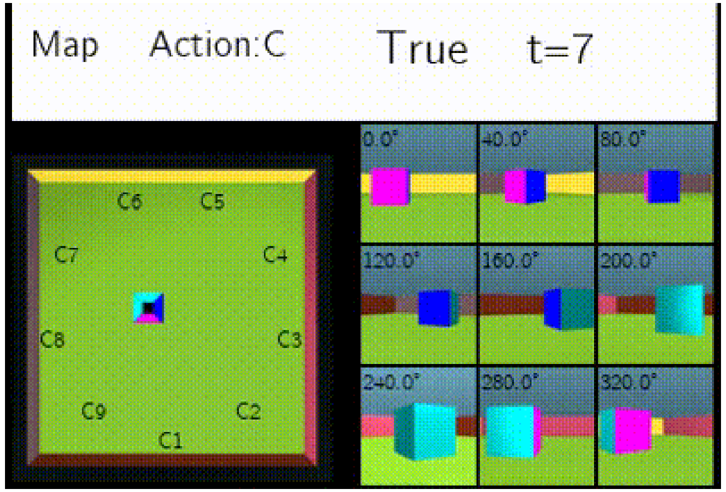


Temperature of a rod changing with time

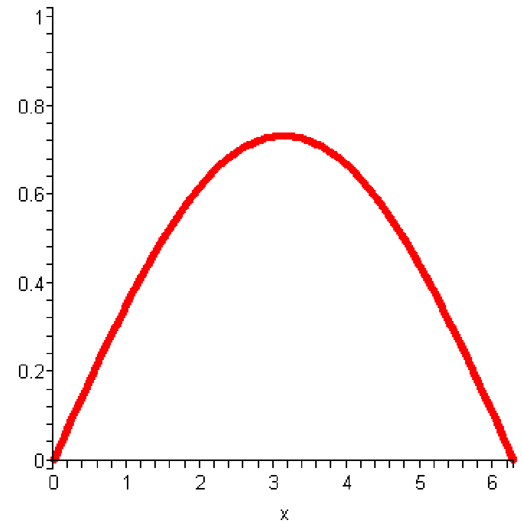
Stochastic Processes with Time Structure



Bouncing 2D Shapes

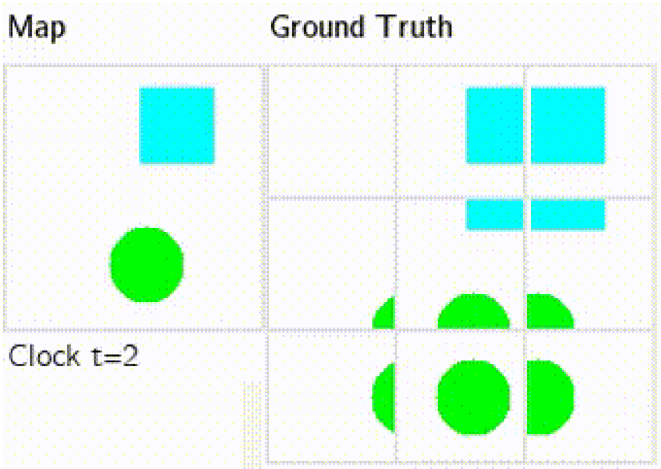


Moving 3D Object

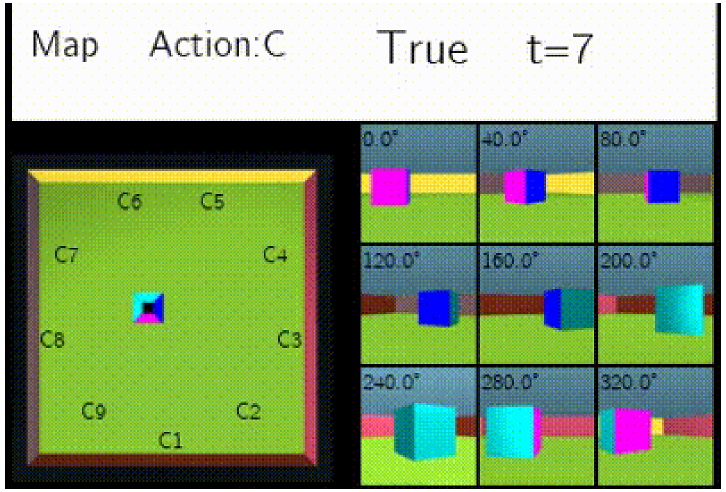


Temperature of a rod
changing with time

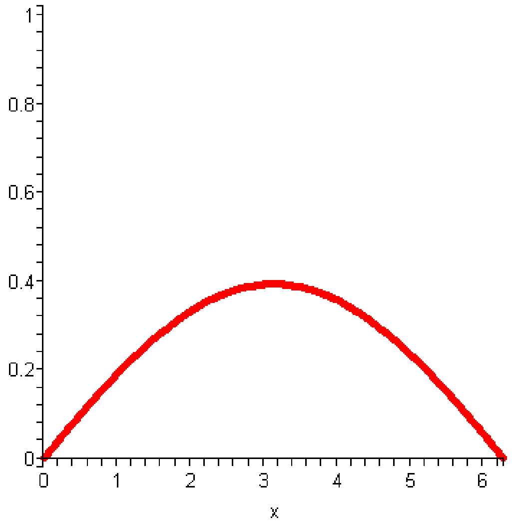
Stochastic Processes with Time Structure



Bouncing 2D Shapes

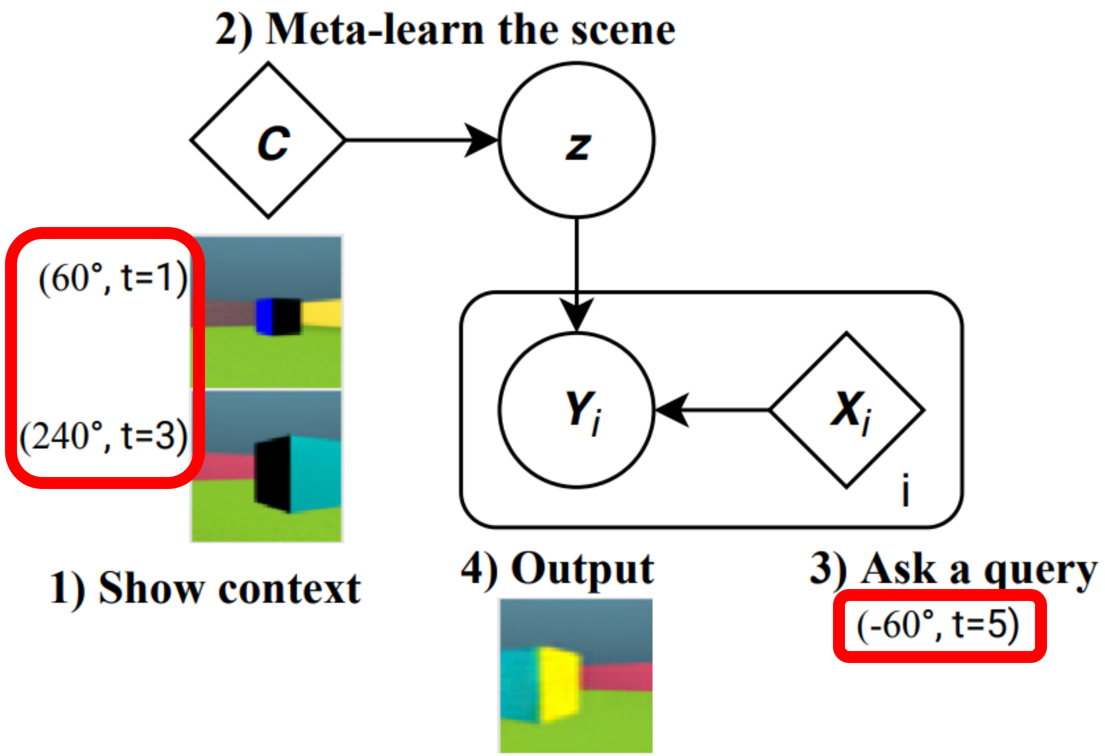


Moving 3D Object



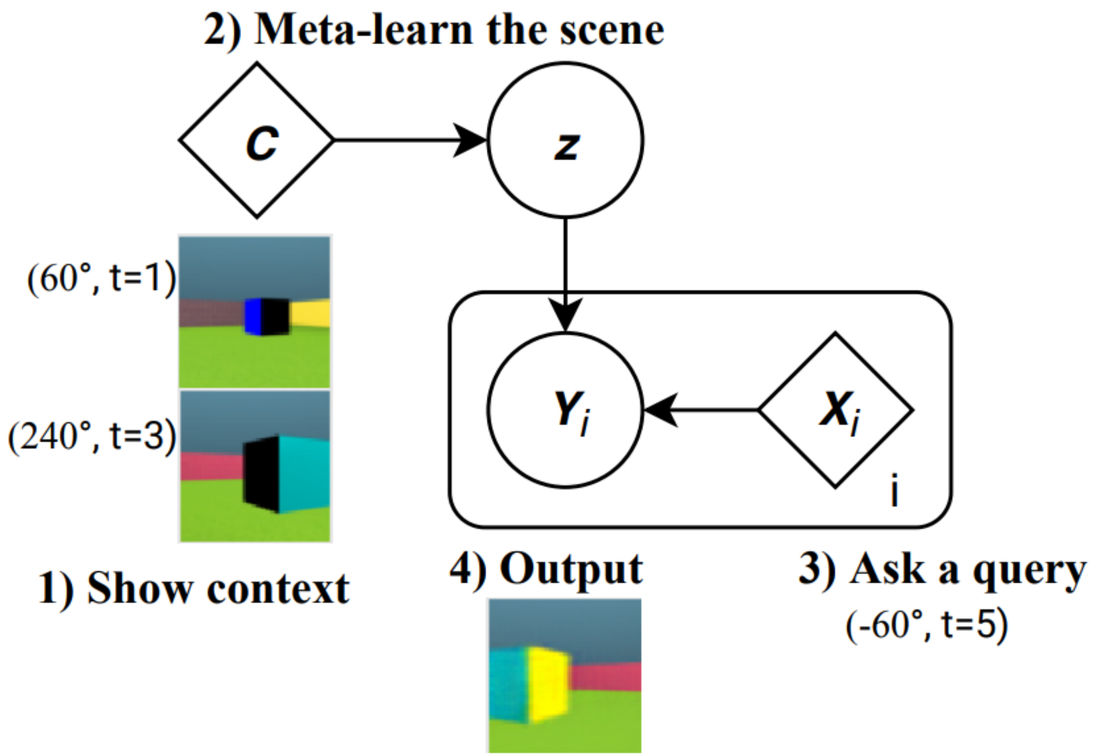
Temperature of a rod changing with time

Simple Extension of the Baselines



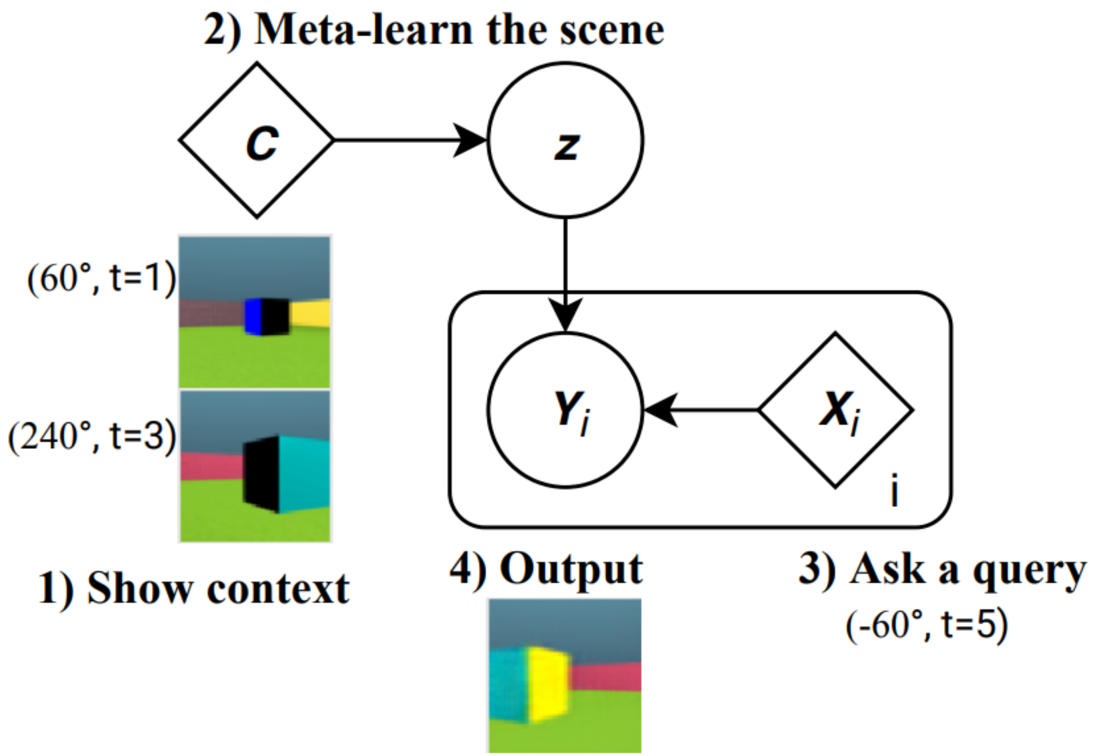
- Append time *t* to the query in Neural Processes or GQN.
- Our findings show that this does not work well *since it does not model time explicitly.*
  - Poor generation quality
  - Cannot generalize to long time-horizons

Simple Extension of the Baselines



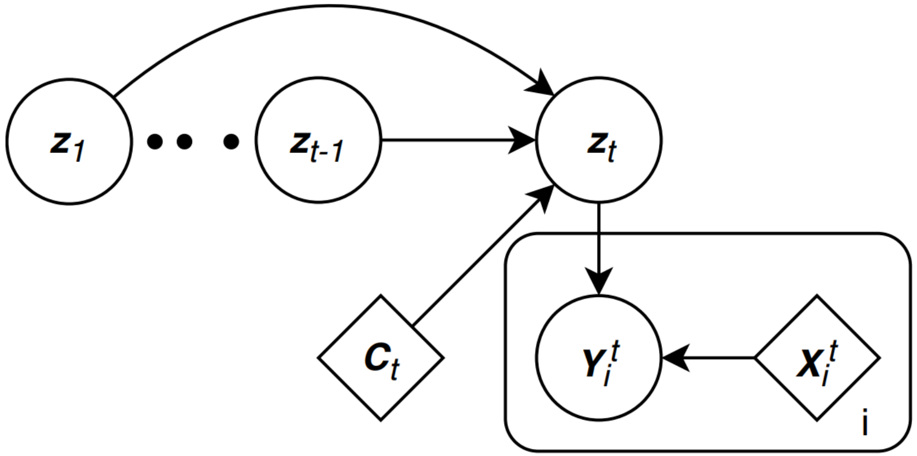
- Append time *t* to the query in Neural Processes or GQN.
- Our findings show that this does not work well *since it does not model time explicitly.*
  - Poor generation quality
  - Cannot generalize to long time-horizons

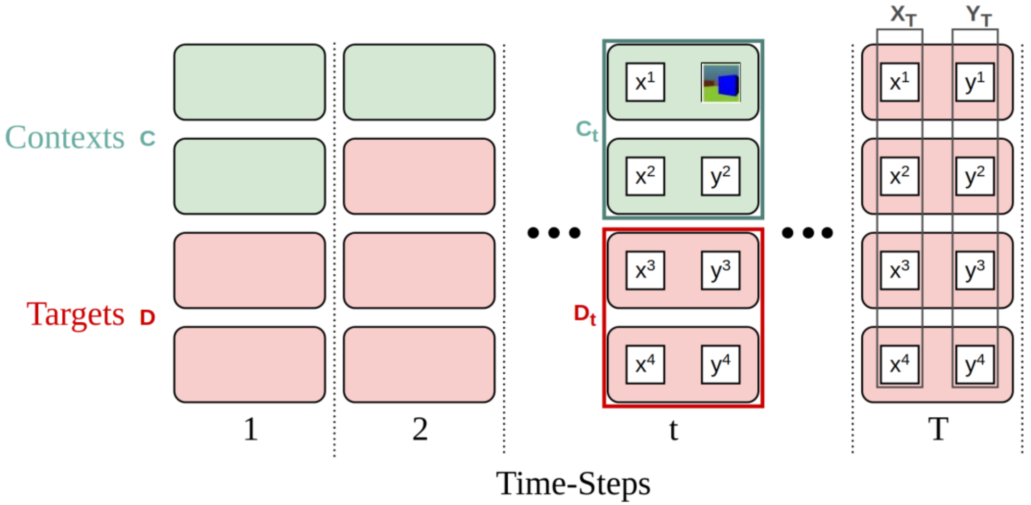
Simple Extension of the Baselines



- Append time *t* to the query in Neural Processes or GQN.
- Our findings show that this does not work well *since it does not model time explicitly.*
  - Poor generation quality
  - Cannot generalize to long time-horizons

Sequential Neural Processes

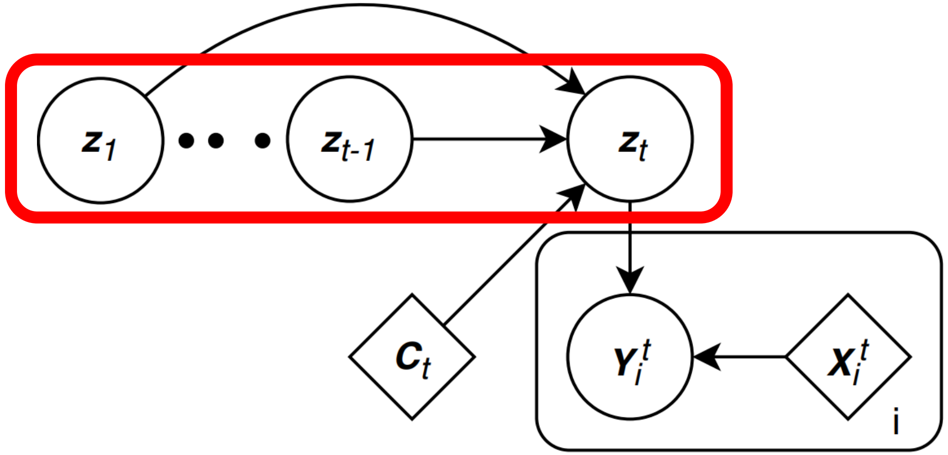


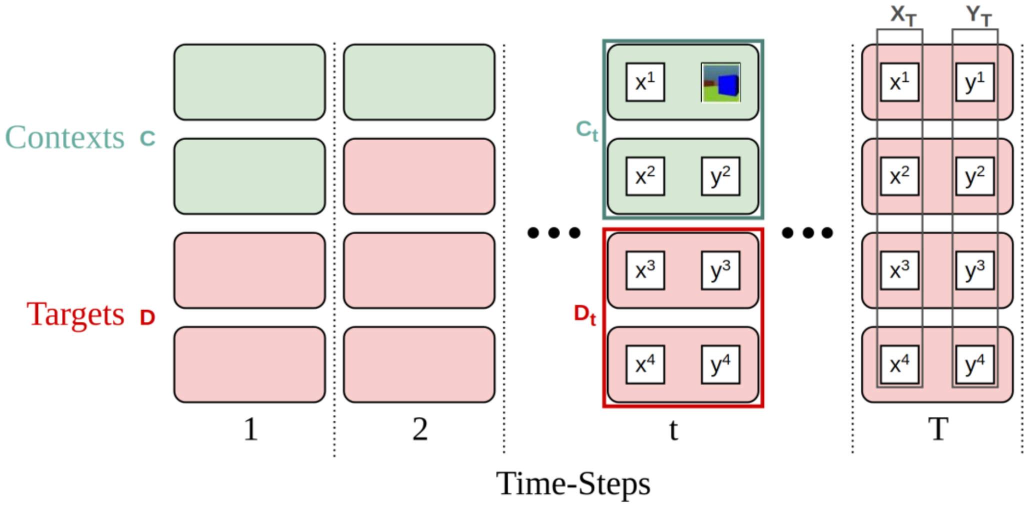


$$P(Y, Z|X, C) = \prod_{t=1}^{T} P\left(Y_t|X_t, z_t\right) P\left(z_t|z_{<t}, C_t\right)$$

Meta-Transfer Learning.
"We not only learn from the current context but also utilize our knowledge of the past stochastic processes"

Sequential Neural Processes



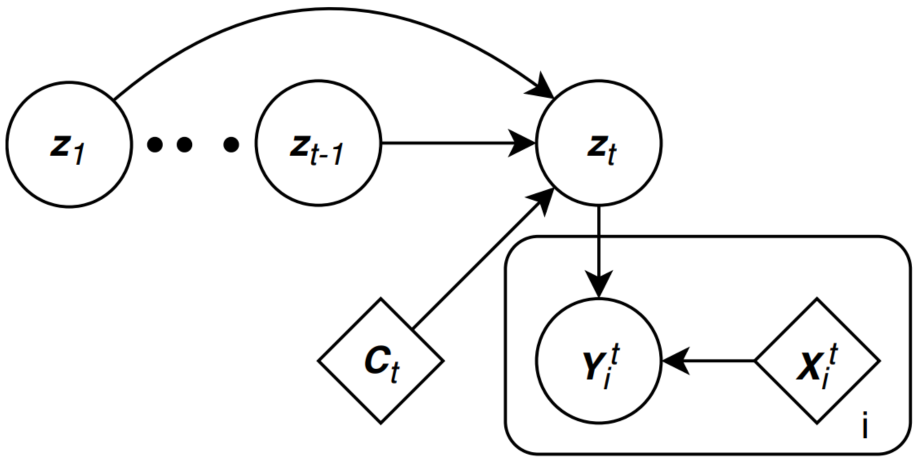


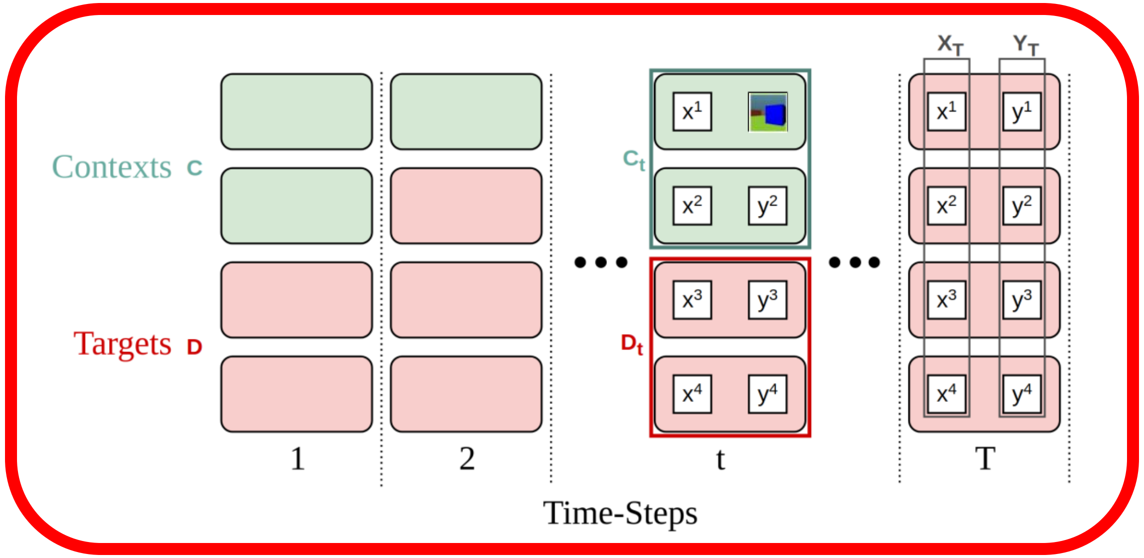
Transition Model

$$P(Y, Z|X, C) = \prod_{t=1}^{T} P(Y_t|X_t, z_t) \boxed{P(z_t|z_{<t}, C_t)}$$

Meta-Transfer Learning.
"We not only learn from the current context but also utilize our knowledge of the past stochastic processes"

Sequential Neural Processes

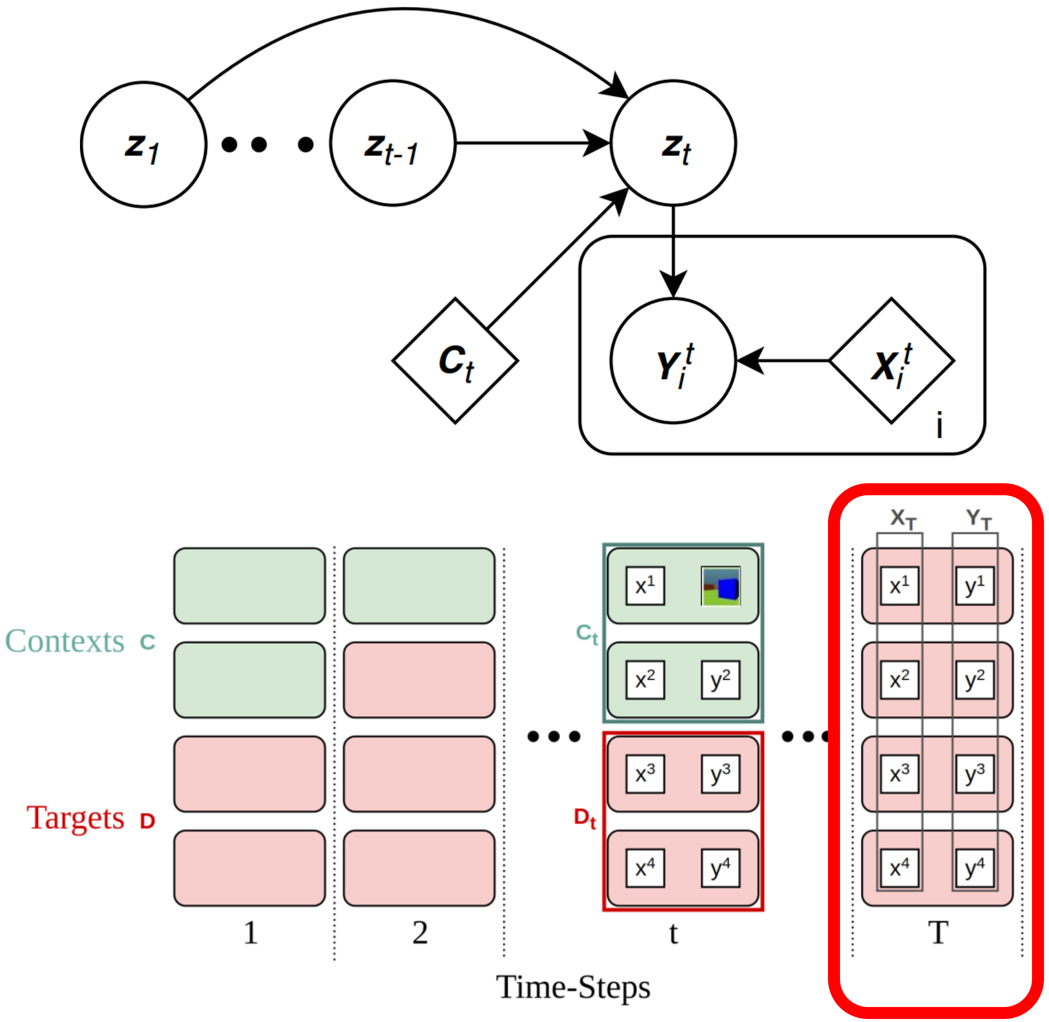




$$P(Y, Z|X, C) = \prod_{t=1}^{T} P\left(Y_t|X_t, z_t\right) P\left(z_t|z_{<t}, C_t\right)$$

Meta-Transfer Learning.
"We not only learn from the current context but also utilize our knowledge of the past stochastic processes"

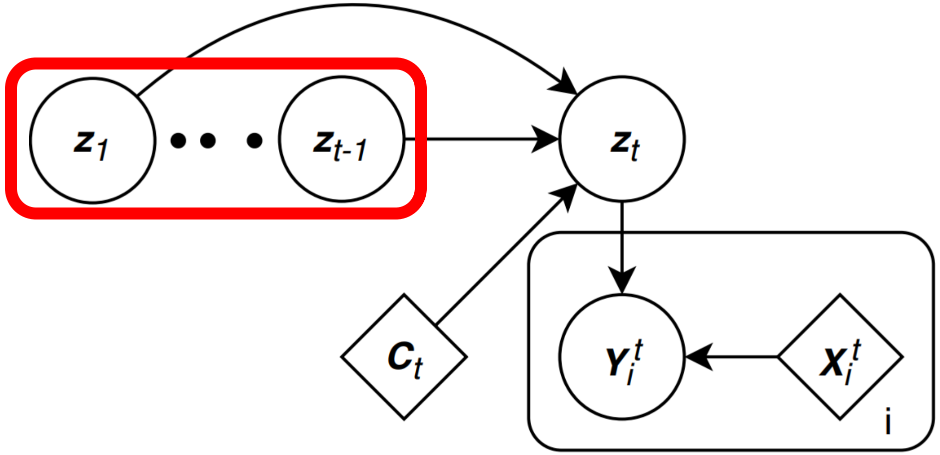
Sequential Neural Processes

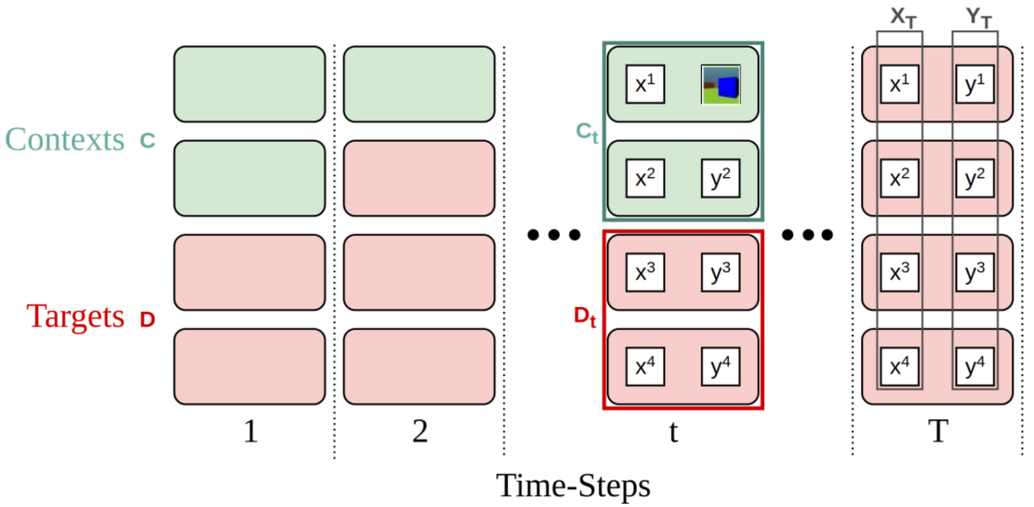


$$P(Y, Z|X, C) = \prod_{t=1}^{T} P(Y_t|X_t, z_t) P(z_t|z_{<t}, C_t)$$

Meta-Transfer Learning.
"We not only learn from the current context but also utilize our knowledge of the past stochastic processes"

Sequential Neural Processes





$$P(Y, Z|X, C) = \prod_{t=1}^{T} P(Y_t|X_t, z_t)\, P(z_t|\boxed{z_{<t}}, C_t)$$

Meta-Transfer Learning.
"We need not learn everything from current context but only use it to update our prior hypothesis."

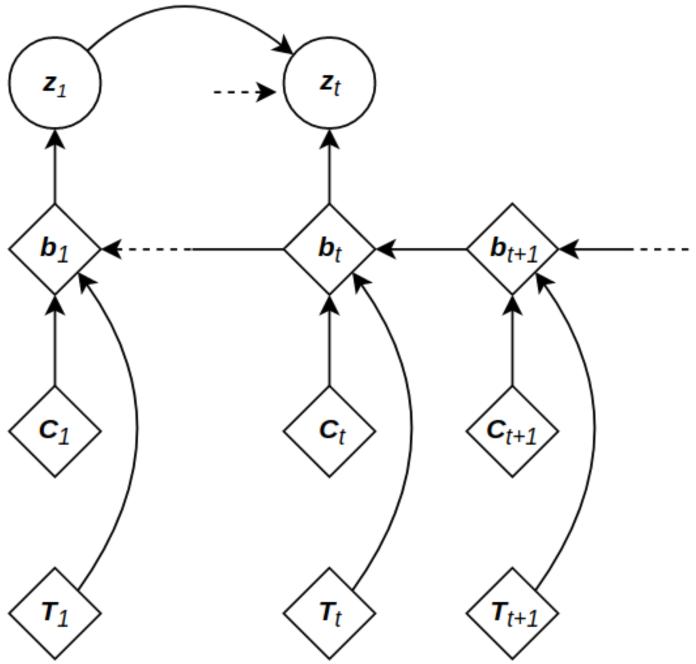
Inference and Learning

- We train the model via a variational approximation.

$$P(Z|C,D) \approx \prod_{t=1}^{T} Q_{\phi}(z_t|z_{<t}, C, D)$$

- This leads to the following ELBO training objective.

$$\begin{aligned}\log P(Y|X,C) &\geq \mathcal{L}_{\text{SNP}}(\theta,\phi) \\ &= \sum_{t=1}^{T} \mathbb{E}_{Q_{\phi}(z_t|\mathcal{V})}\left[\log P_{\theta}(Y_t|X_t, z_t)\right] \\ &\quad - \mathbb{E}_{Q_{\phi}(z_{<t}|\mathcal{V})}\left[\mathbb{KL}(Q_{\phi}(z_t|z_{<t},\mathcal{V}) \parallel P_{\theta}(z_t|z_{<t}, C_t))\right]\end{aligned}$$



A realization of the inference model using a backward RNN.

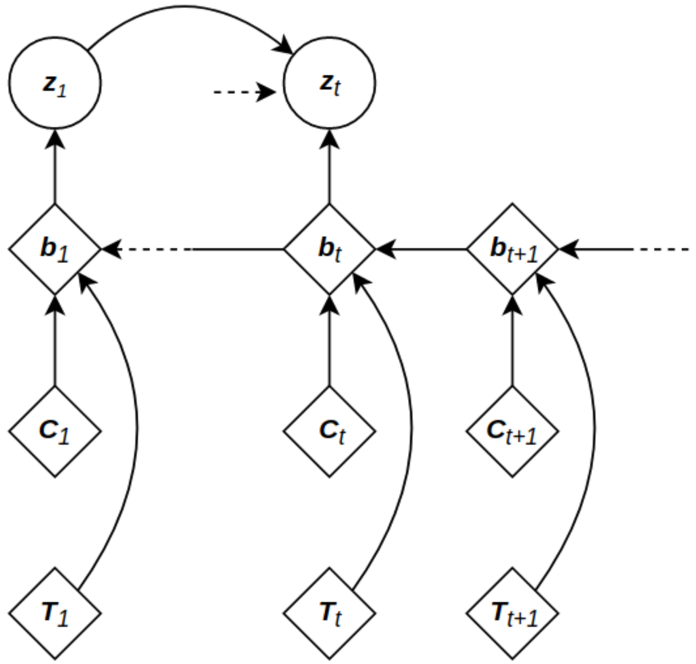
Inference and Learning

- We train the model via a variational approximation.

$$P(Z|C,D) \approx \prod_{t=1}^{T} \boxed{Q_\phi(z_t|z_{<t}, C, D)}$$

- This leads to the following ELBO training objective.

$$\begin{aligned}\log P(Y|X,C) &\geq \mathcal{L}_{\text{SNP}}(\theta,\phi) \\ &= \sum_{t=1}^{T} \mathbb{E}_{Q_\phi(z_t|\mathcal{V})}\left[\log P_\theta(Y_t|X_t, z_t)\right] \\ &\quad - \mathbb{E}_{Q_\phi(z_{<t}|\mathcal{V})}\left[\mathbb{KL}(Q_\phi(z_t|z_{<t},\mathcal{V}) \parallel P_\theta(z_t|z_{<t}, C_t))\right]\end{aligned}$$



A realization of the inference model using a backward RNN.

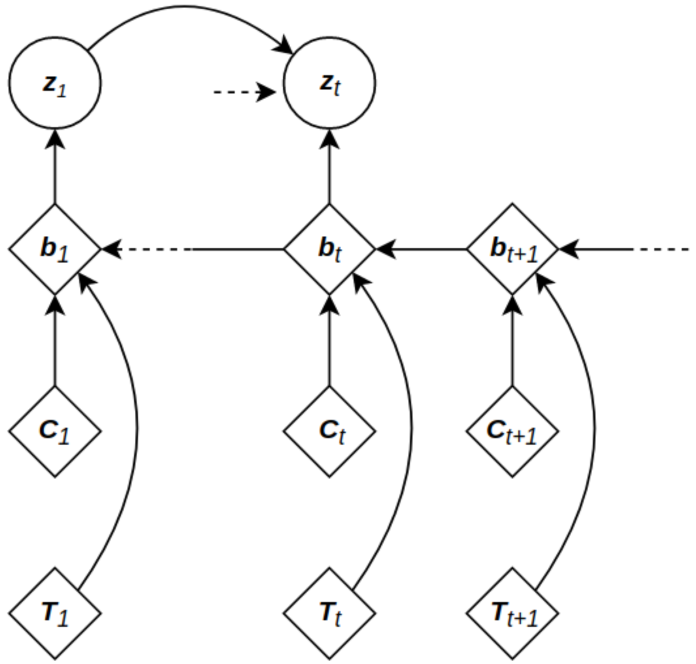
Inference and Learning

- We train the model via a variational approximation.

$$P(Z|C,D) \approx \prod_{t=1}^{T} Q_\phi(z_t|z_{<t}, C, D)$$

- This leads to the following ELBO training objective.

$$\begin{aligned}\log P(Y|X,C) &\geq \mathcal{L}_{\text{SNP}}(\theta,\phi) \\ &= \sum_{t=1}^{T} \mathbb{E}_{Q_\phi(z_t|\mathcal{V})}[\log P_\theta(Y_t|X_t, z_t)] \\ &\quad - \mathbb{E}_{Q_\phi(z_{<t}|\mathcal{V})}[\mathbb{KL}(Q_\phi(z_t|z_{<t},\mathcal{V}) \parallel P_\theta(z_t|z_{<t}, C_t))]\end{aligned}$$



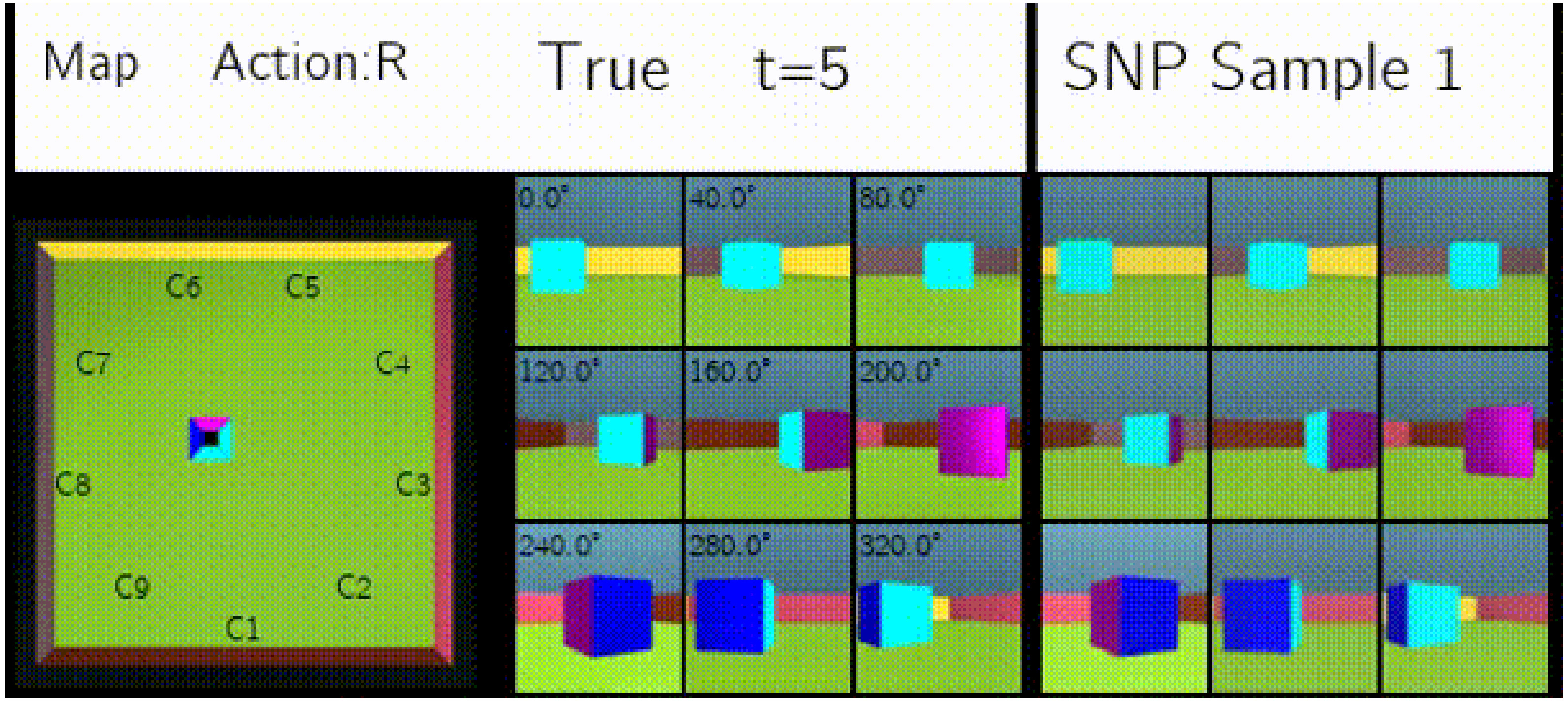
A realization of the inference model using a backward RNN.

Demonstrations

Color Cube

Context is shown in the first 5 time-steps and the remaining are predicted purely on the command of the actions provided to the object. The actions can be translation (L, R, U, D) or rotations (Clockwise, A-Clockwise)

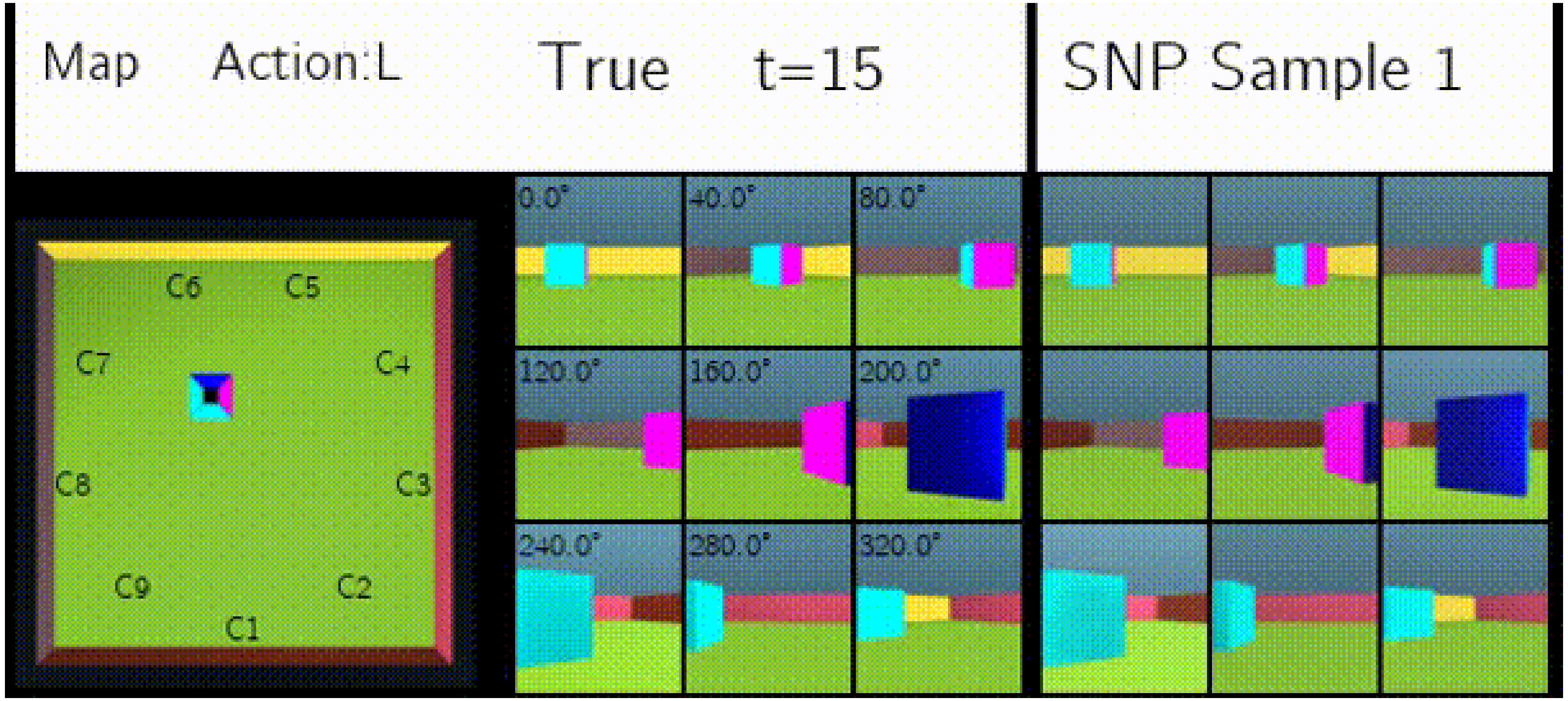
1st time-step without context



Color Cube

Context is shown in the first 5 time-steps and the remaining are predicted purely on the command of the actions provided to the object. The actions can be translation (L, R, U, D) or rotations (Clockwise, A-Clockwise)

10th time-step without context

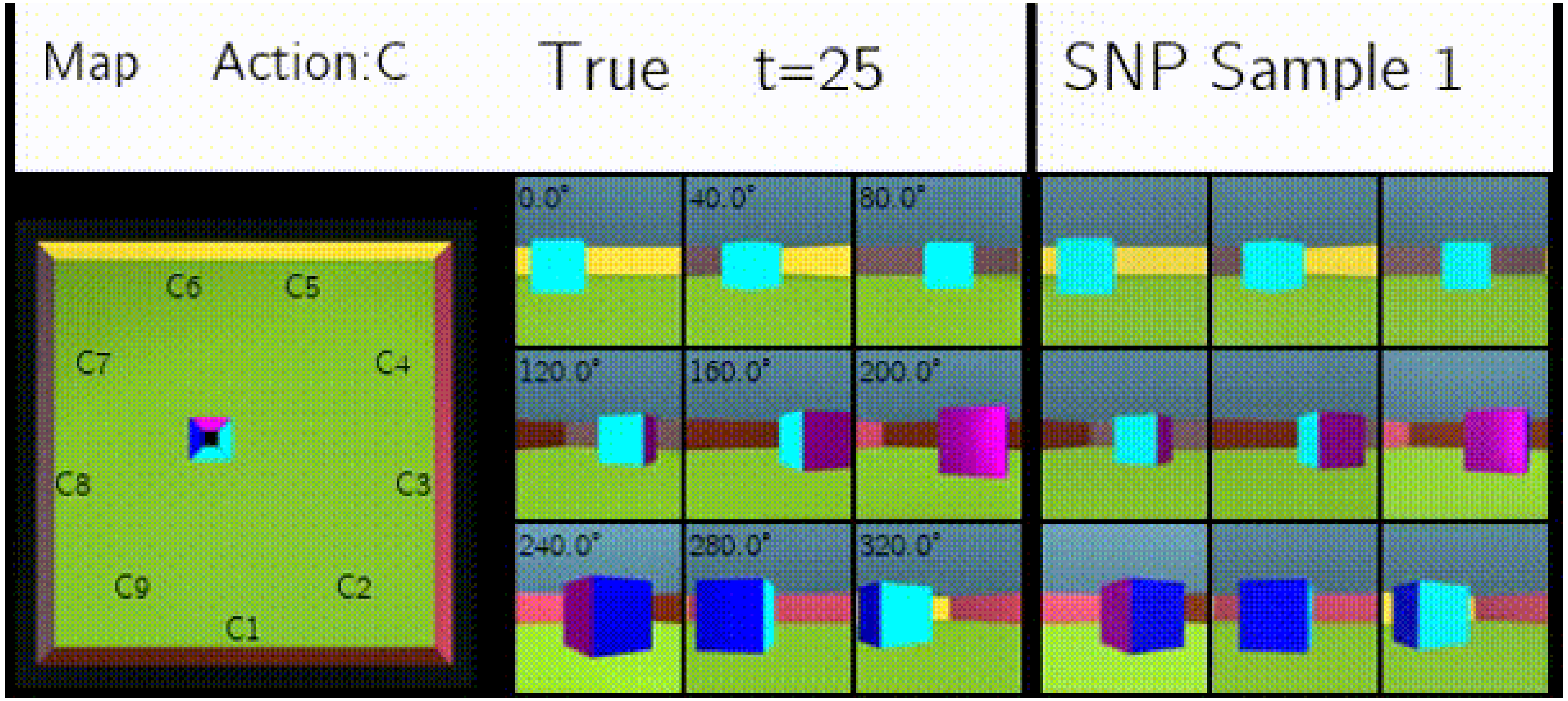


Color Cube

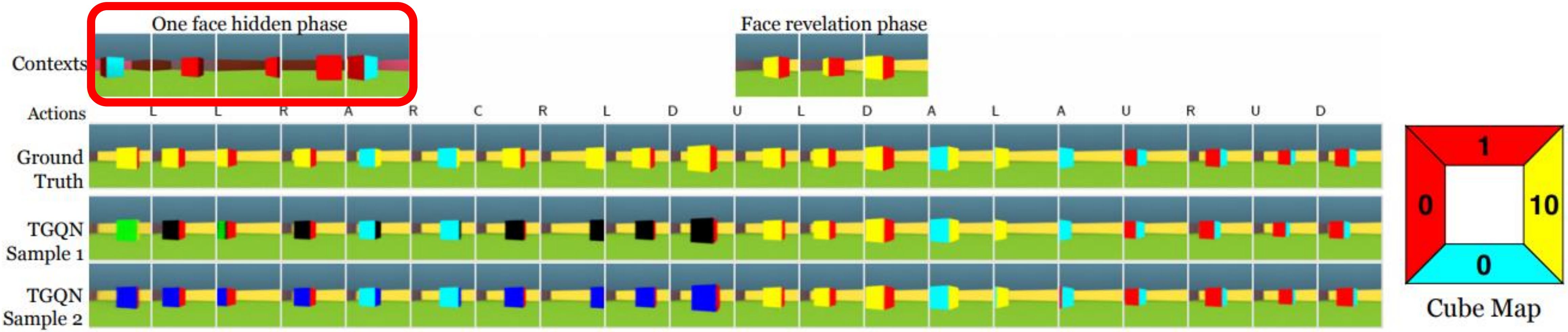
Context is shown in the first 5 time-steps and the remaining are predicted purely on the command of the actions provided to the object. The actions can be translation (L, R, U, D) or rotations (Clockwise, A-Clockwise)

20th time-step without context.

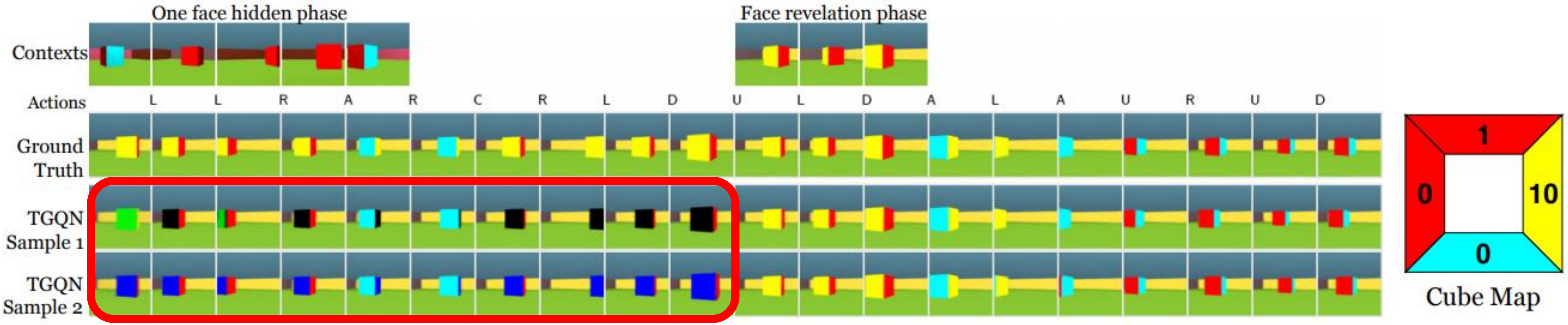
Beyond training time horizon



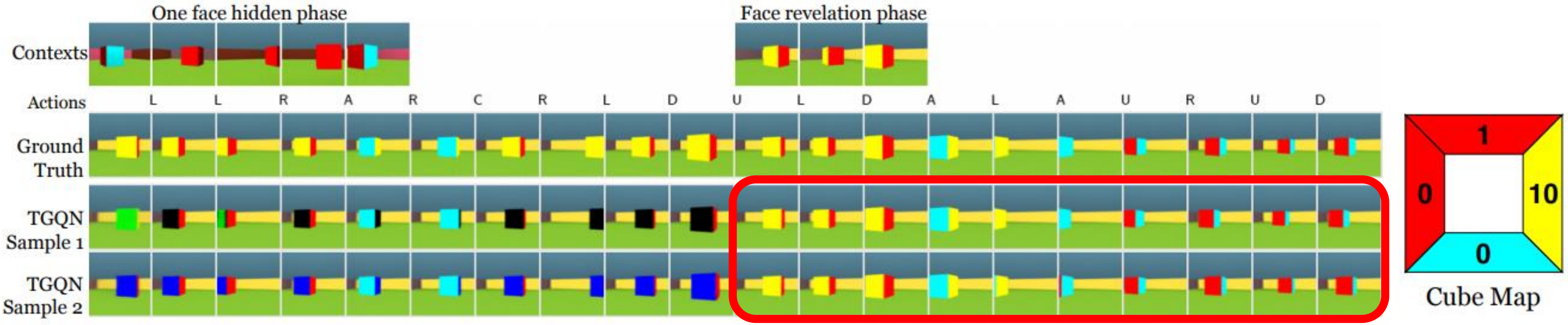
Meta-Transfer Learning



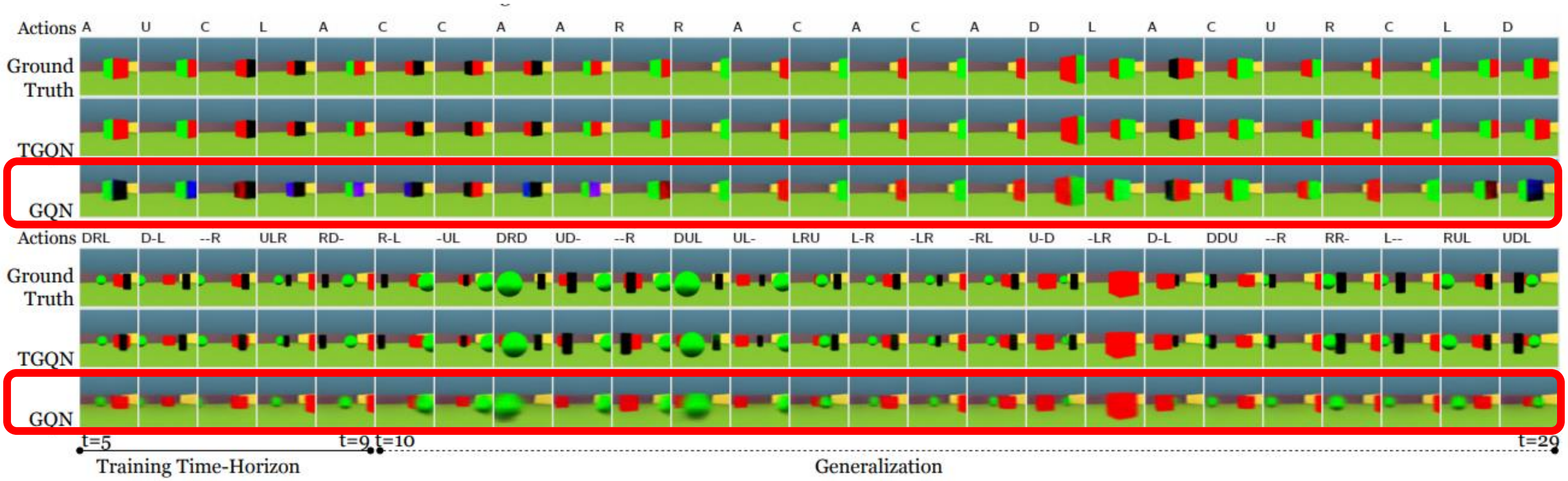
Meta-Transfer Learning



Meta-Transfer Learning

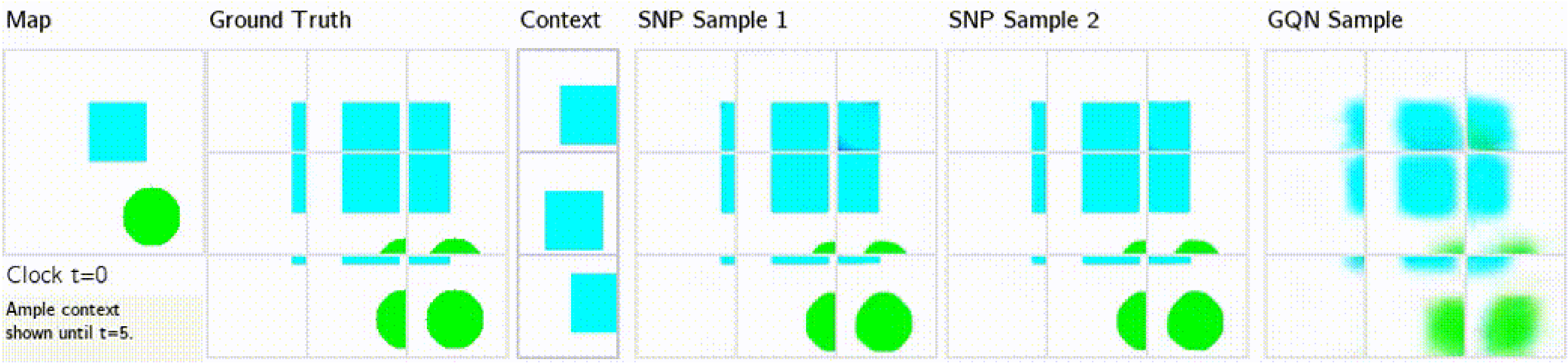


Comparing against GQN



Color Shapes : Tracking and Updating

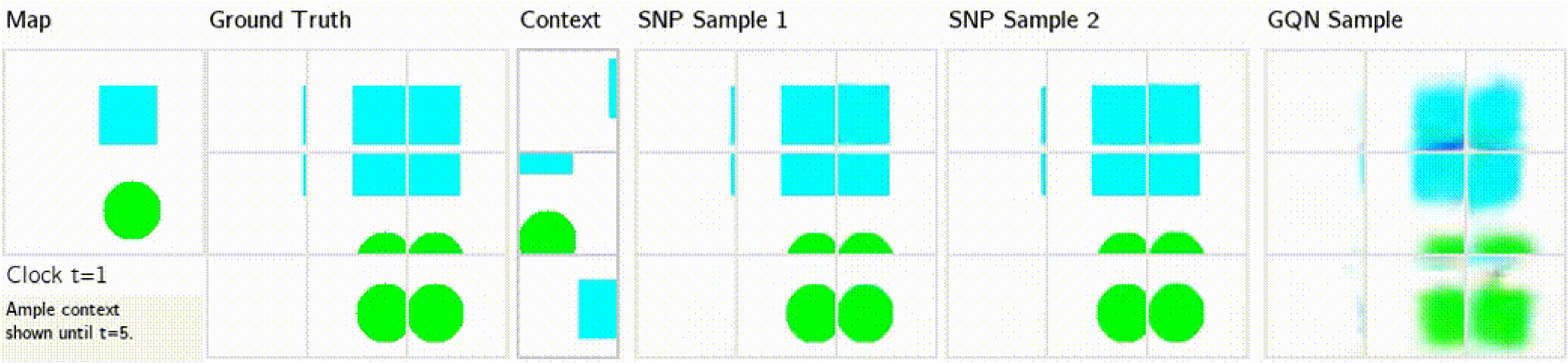
Context is shown intermittently and we allow the predictions to diverge from the true. On seeing the context, we observe that the belief about the object is updated.



Context is being shown.

Color Shapes : Tracking and Updating

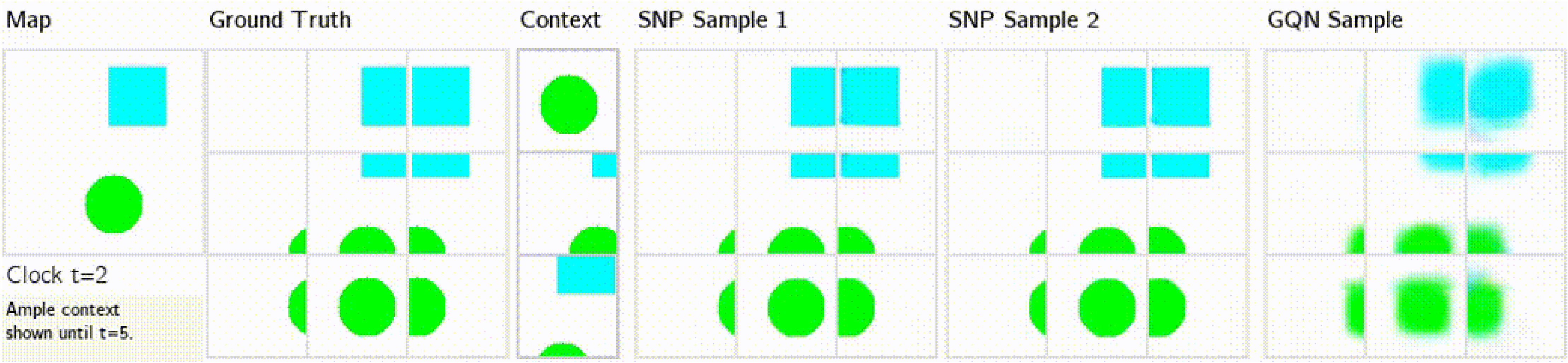
Context is shown intermittently and we allow the predictions to diverge from the true. On seeing the context, we observe that the belief about the object is updated.



Context is being shown.

Color Shapes : Tracking and Updating

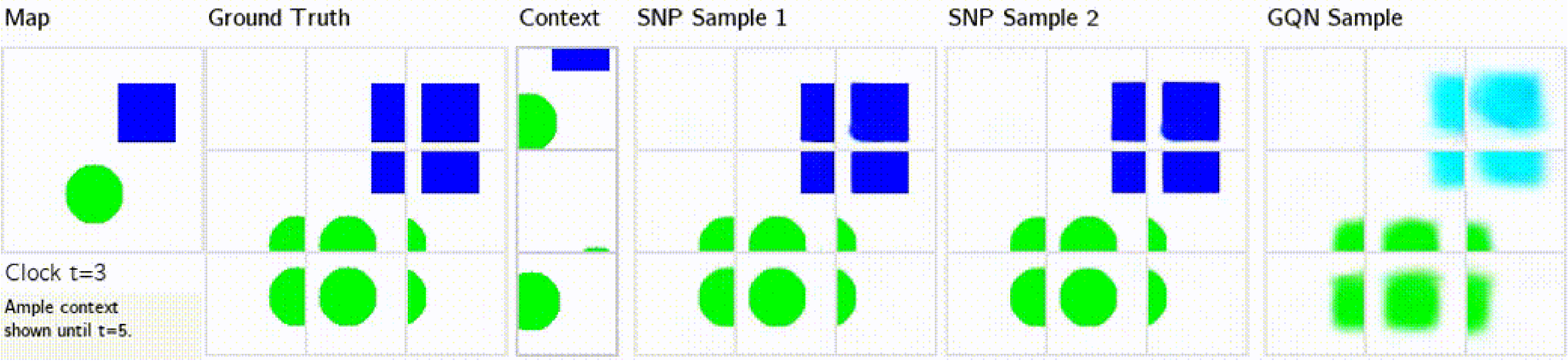
Context is shown intermittently and we allow the predictions to diverge from the true. On seeing the context, we observe that the belief about the object is updated.



Context is being shown.

Color Shapes : Tracking and Updating

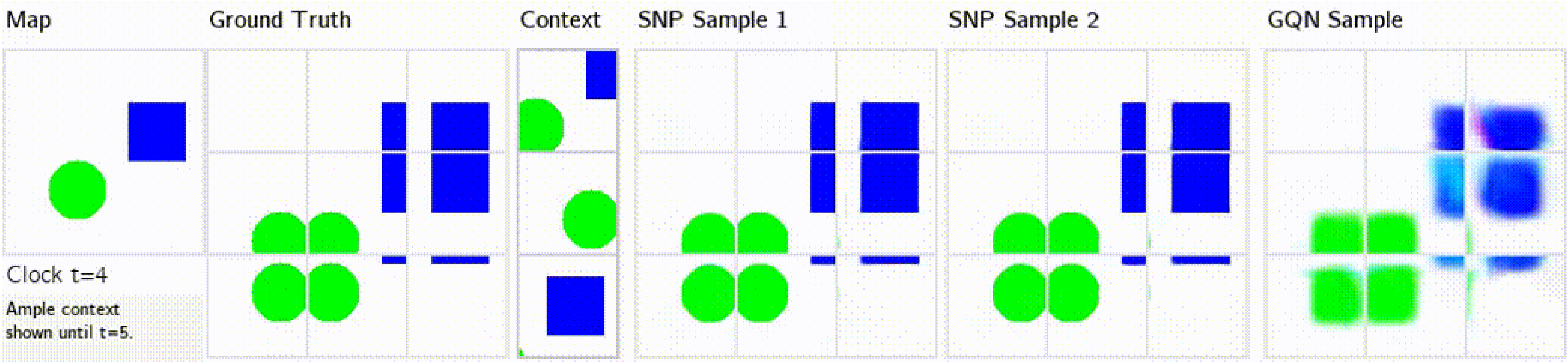
Context is shown intermittently and we allow the predictions to diverge from the true. On seeing the context, we observe that the belief about the object is updated.



Context is being shown.

Color Shapes : Tracking and Updating

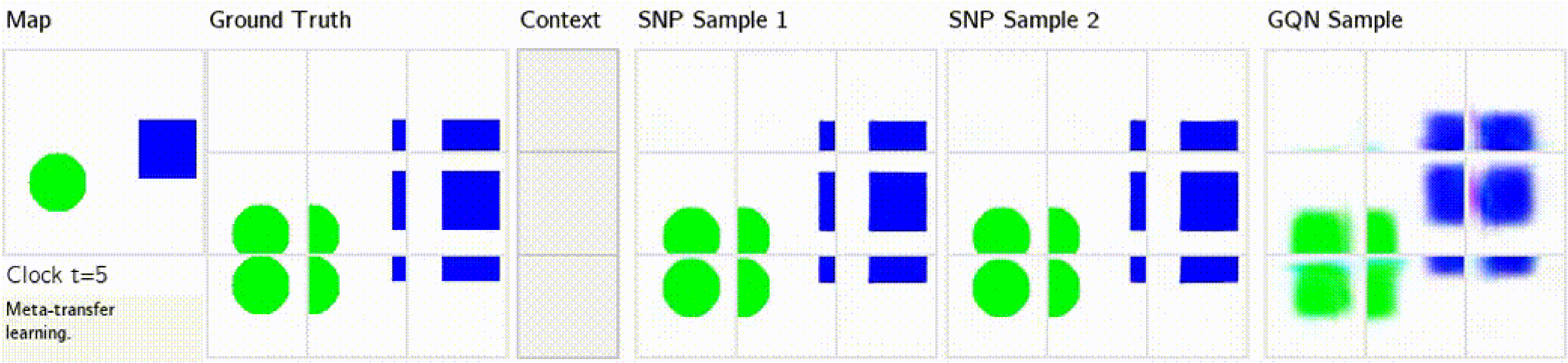
Context is shown intermittently and we allow the predictions to diverge from the true. On seeing the context, we observe that the belief about the object is updated.



Context is being shown.

Color Shapes : Tracking and Updating

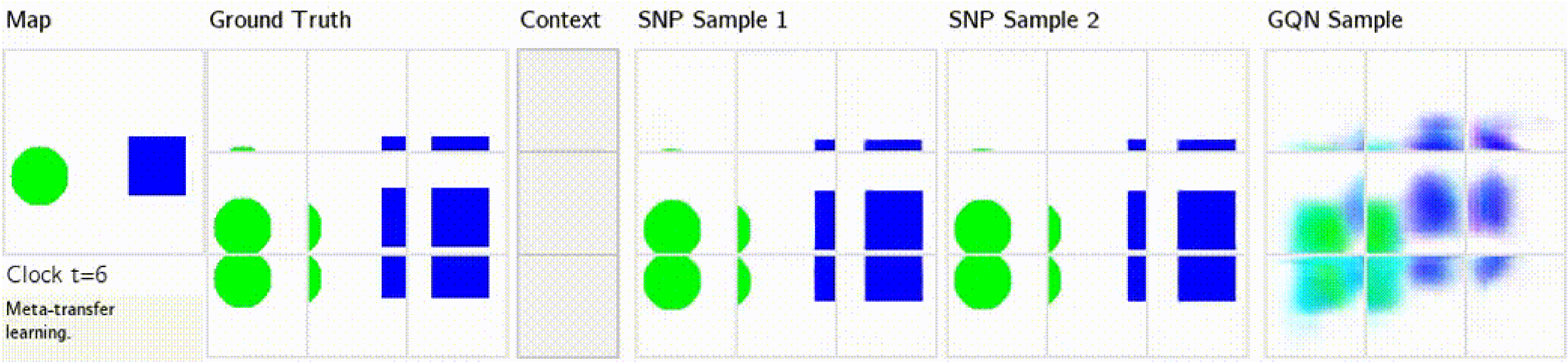
Context is shown intermittently and we allow the predictions to diverge from the true. On seeing the context, we observe that the belief about the object is updated.



Context is removed.

Color Shapes : Tracking and Updating

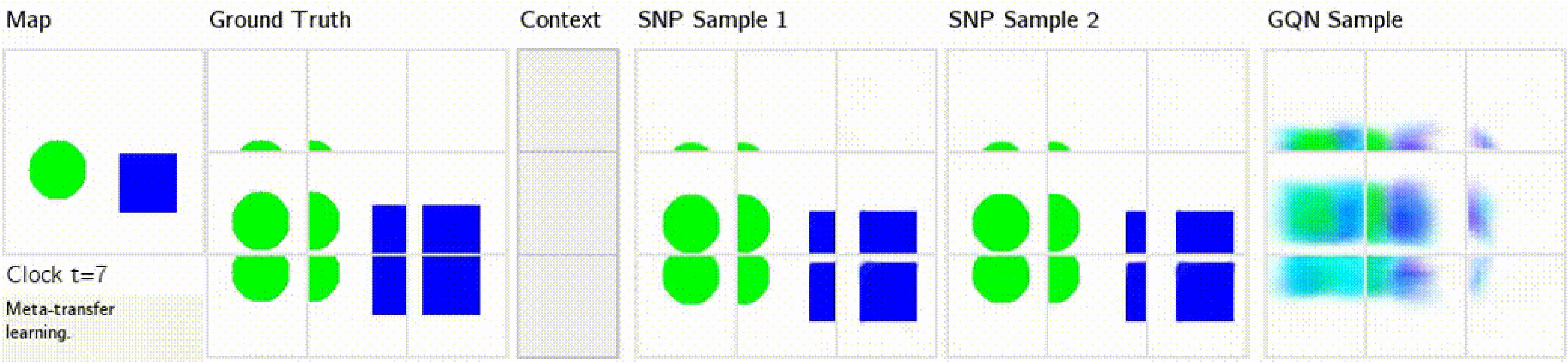
Context is shown intermittently and we allow the predictions to diverge from the true. On seeing the context, we observe that the belief about the object is updated.



Context is removed.

Color Shapes : Tracking and Updating

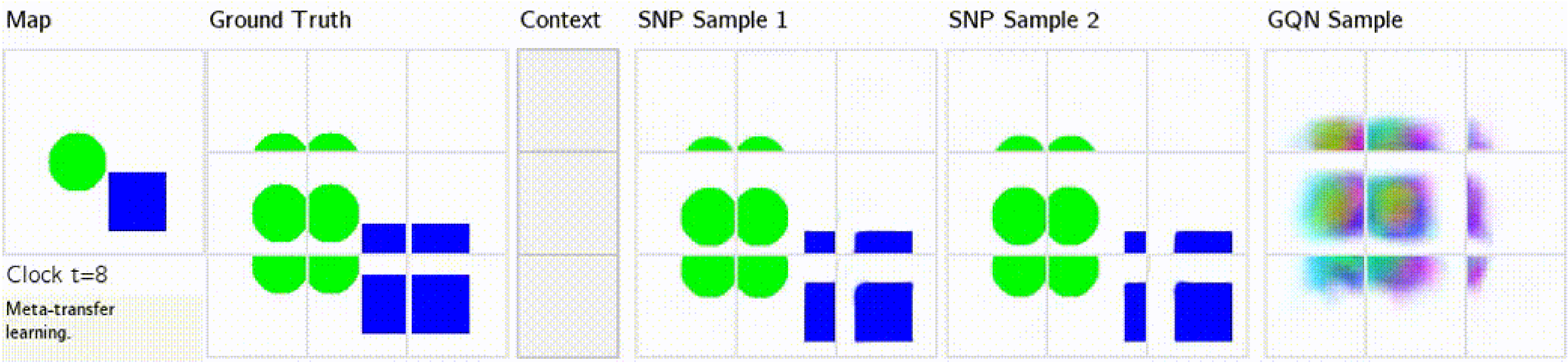
Context is shown intermittently and we allow the predictions to diverge from the true. On seeing the context, we observe that the belief about the object is updated.



Context is removed.

Color Shapes : Tracking and Updating

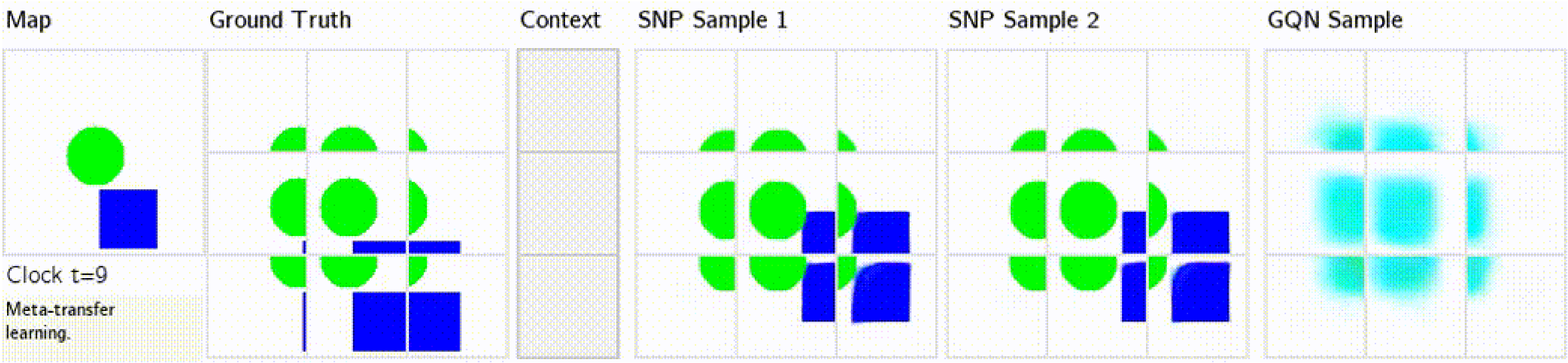
Context is shown intermittently and we allow the predictions to diverge from the true. On seeing the context, we observe that the belief about the object is updated.



Context is removed.

Color Shapes : Tracking and Updating

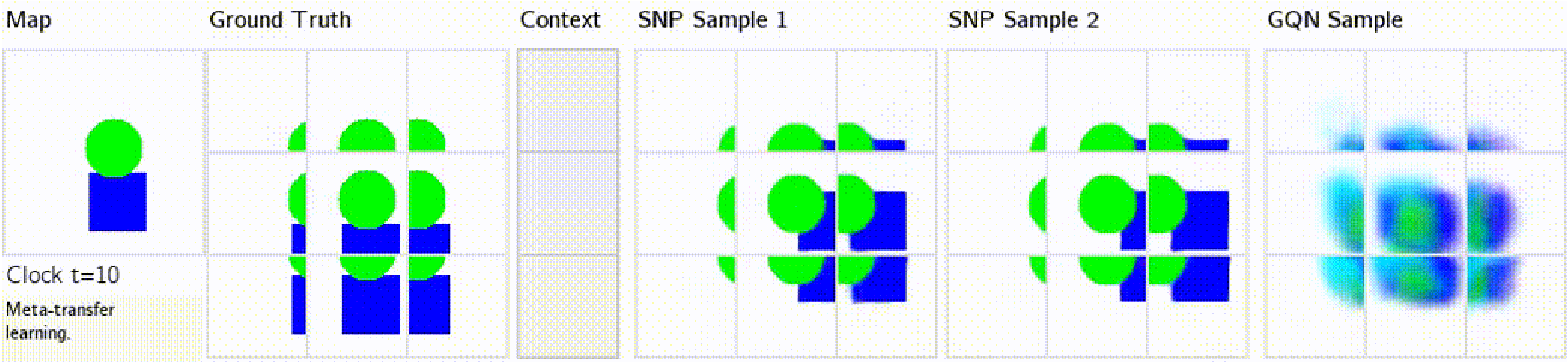
Context is shown intermittently and we allow the predictions to diverge from the true. On seeing the context, we observe that the belief about the object is updated.



Context is removed.

Color Shapes : Tracking and Updating

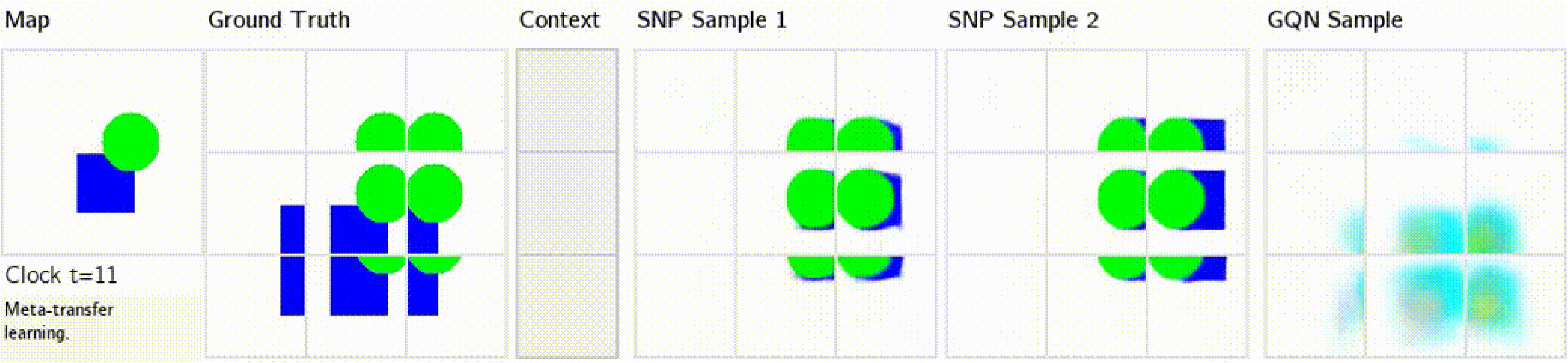
Context is shown intermittently and we allow the predictions to diverge from the true. On seeing the context, we observe that the belief about the object is updated.



Context is removed.

Color Shapes : Tracking and Updating

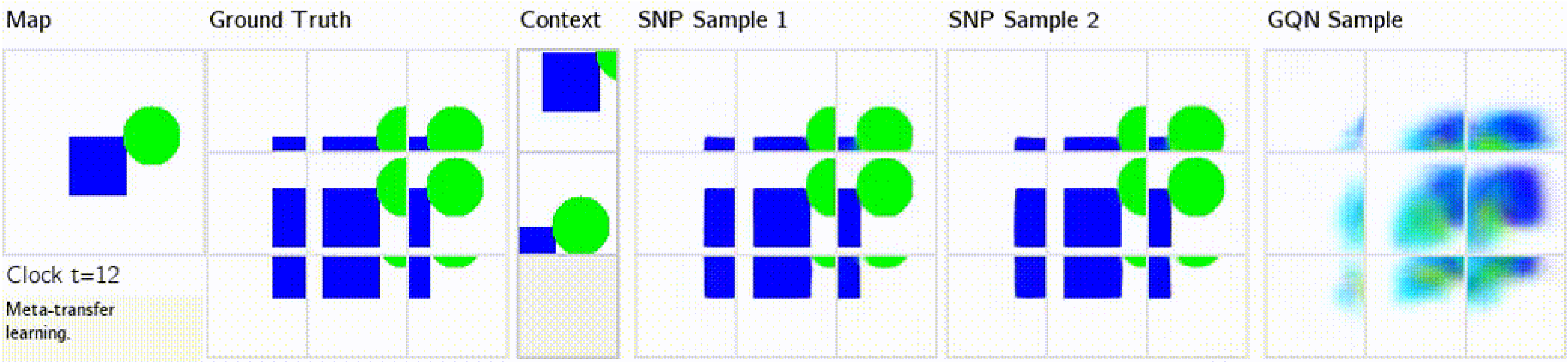
Context is shown intermittently and we allow the predictions to diverge from the true. On seeing the context, we observe that the belief about the object is updated.



Context is removed.

Color Shapes : Tracking and Updating

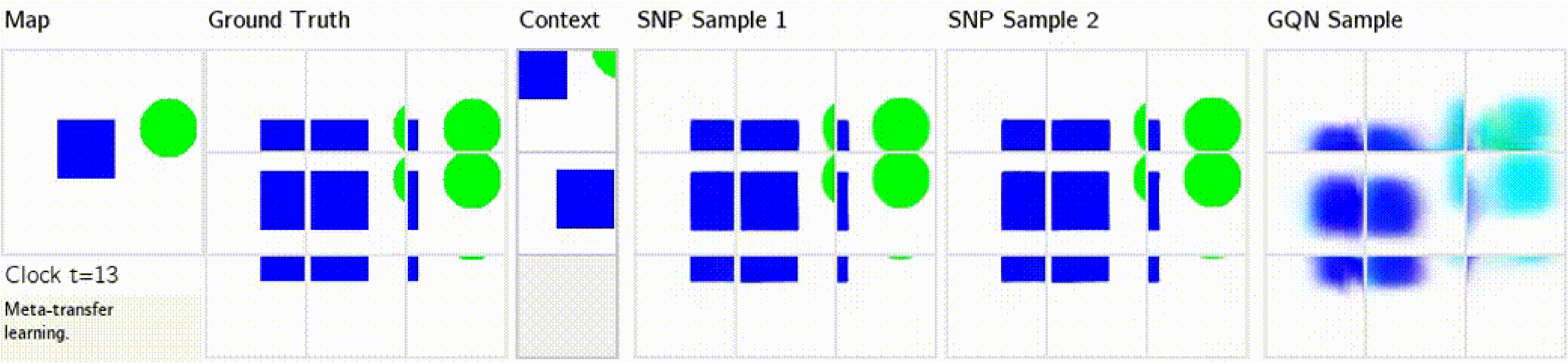
Context is shown intermittently and we allow the predictions to diverge from the true. On seeing the context, we observe that the belief about the object is updated.



Context is shown again.

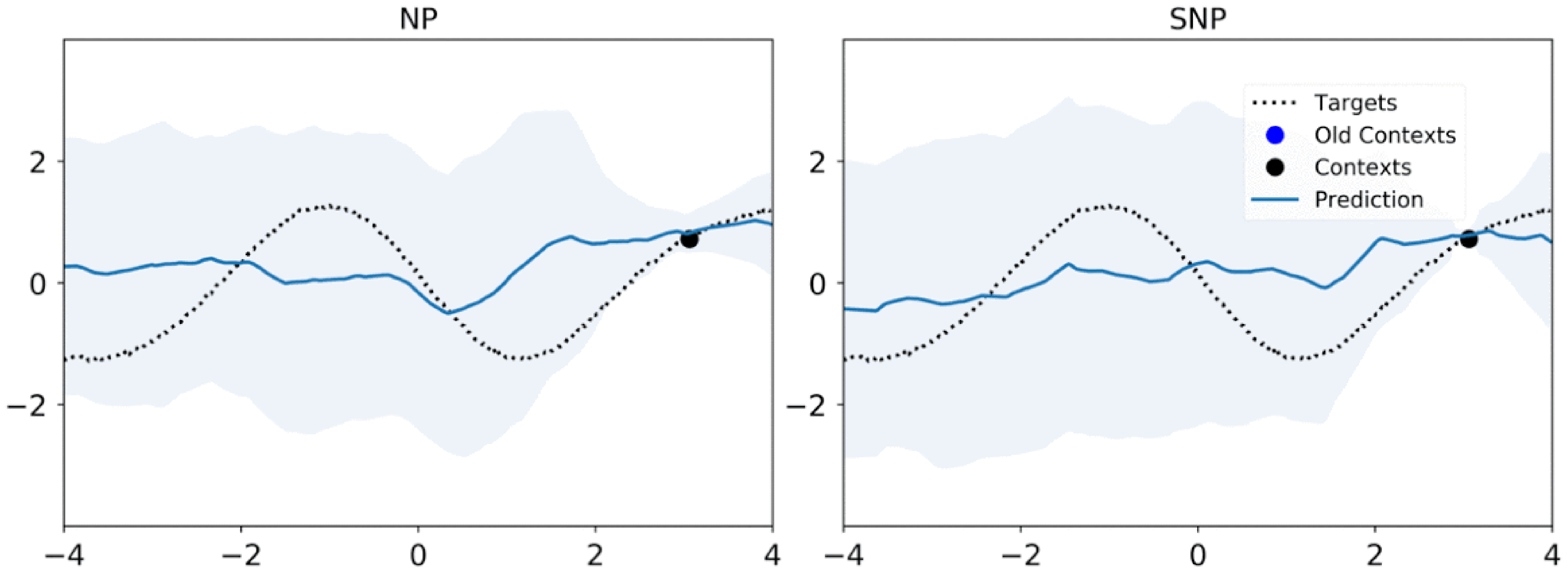
Color Shapes : Tracking and Updating

Context is shown intermittently and we allow the predictions to diverge from the true. On seeing the context, we observe that the belief about the object is updated.

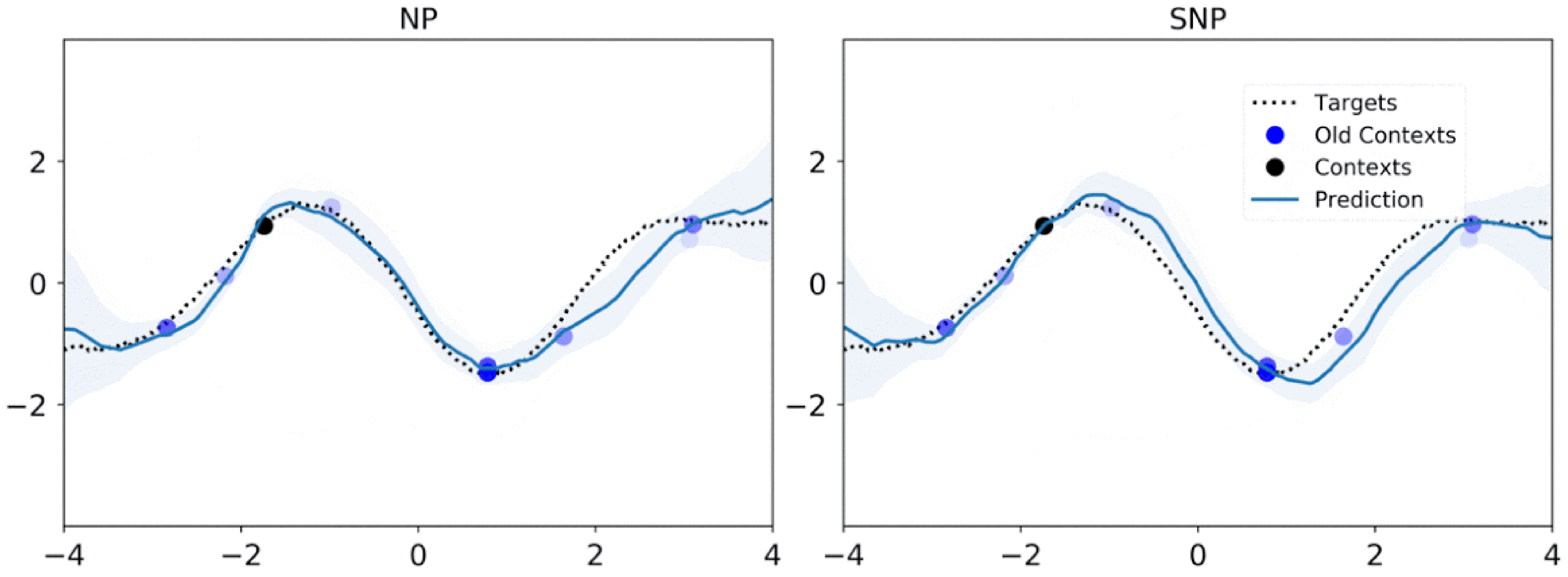


Context is shown again.

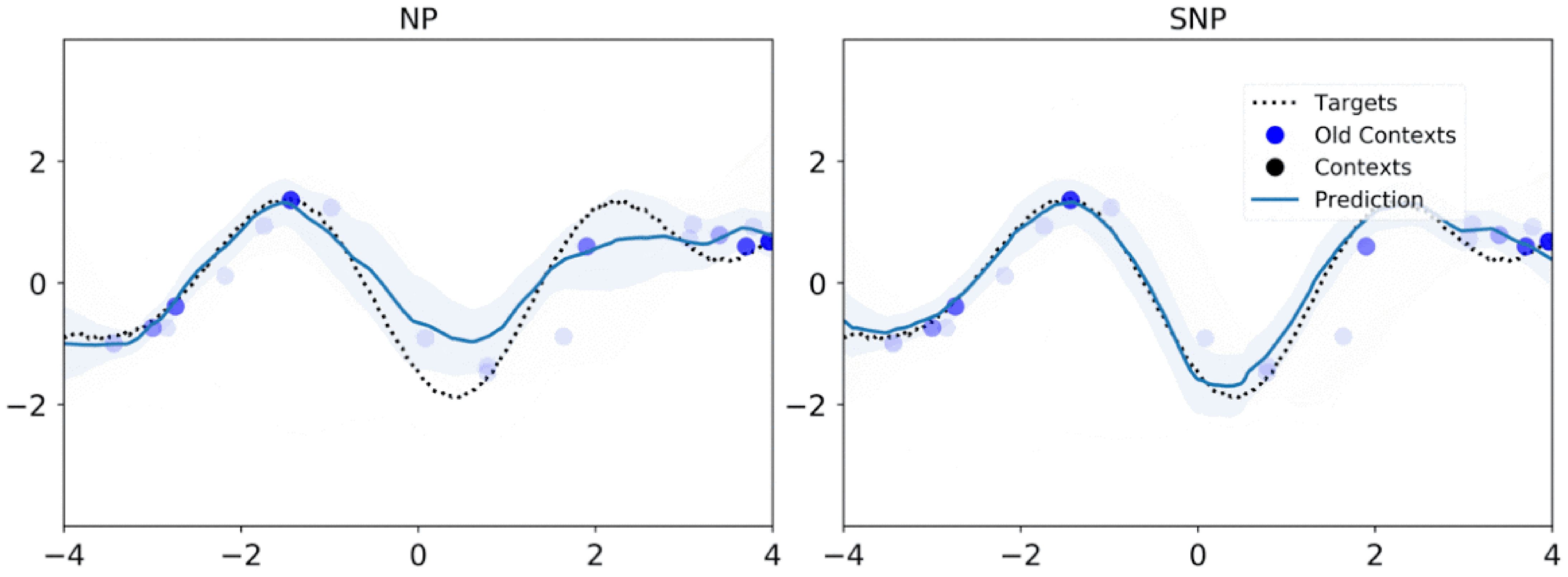
1D Regression



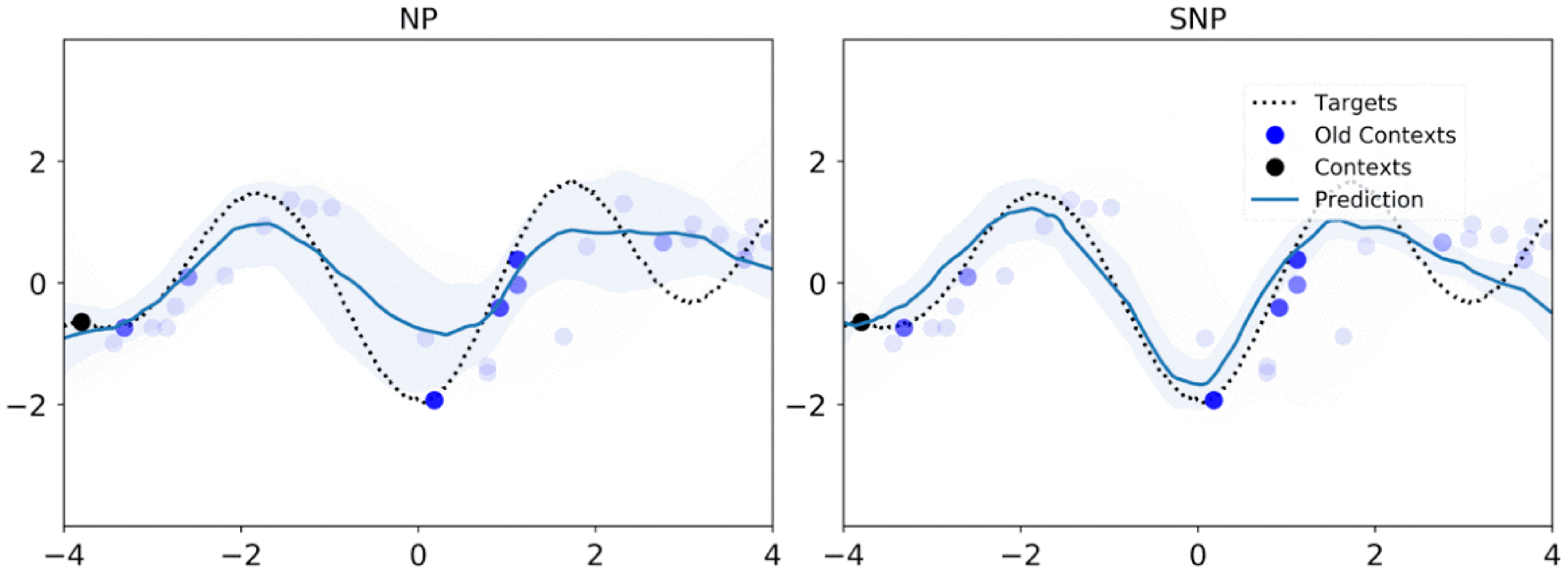
1D Regression



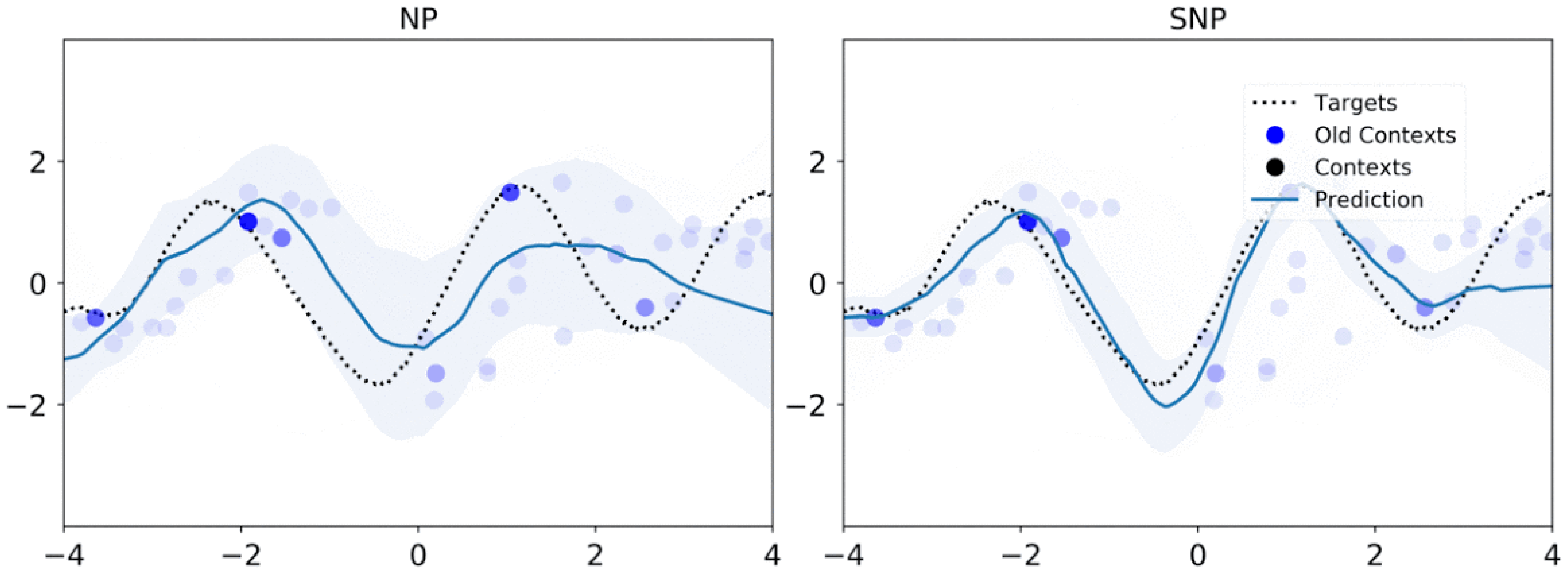
1D Regression



1D Regression



1D Regression



Negative Log-Likelihood

Dataset	Regime	T	GQN	TGQN	
				no PD	PD
Color Shapes	Predict	20	5348	489	564
Color Cube (*Det.*)	Predict	10	380	221	226
Multi-Object (*Det.*)	Predict	10	844	346	357
Color Shapes	Track	20	5285	482	513
Color Cube (*Jit.*)	Track	20	783	153	156
Multi-Object (*Jit.*)	Track	20	1777	450	475

Table 1: Negative $\log p(Y|X, C)$ estimated using importance-sampling from posterior with $K = 40$.

Do visit our poster

Today
10:45 AM - 12:45 PM

East Exhibition Hall B + C
Poster 132

Thank You

RUTGERS

SAP

ETRI