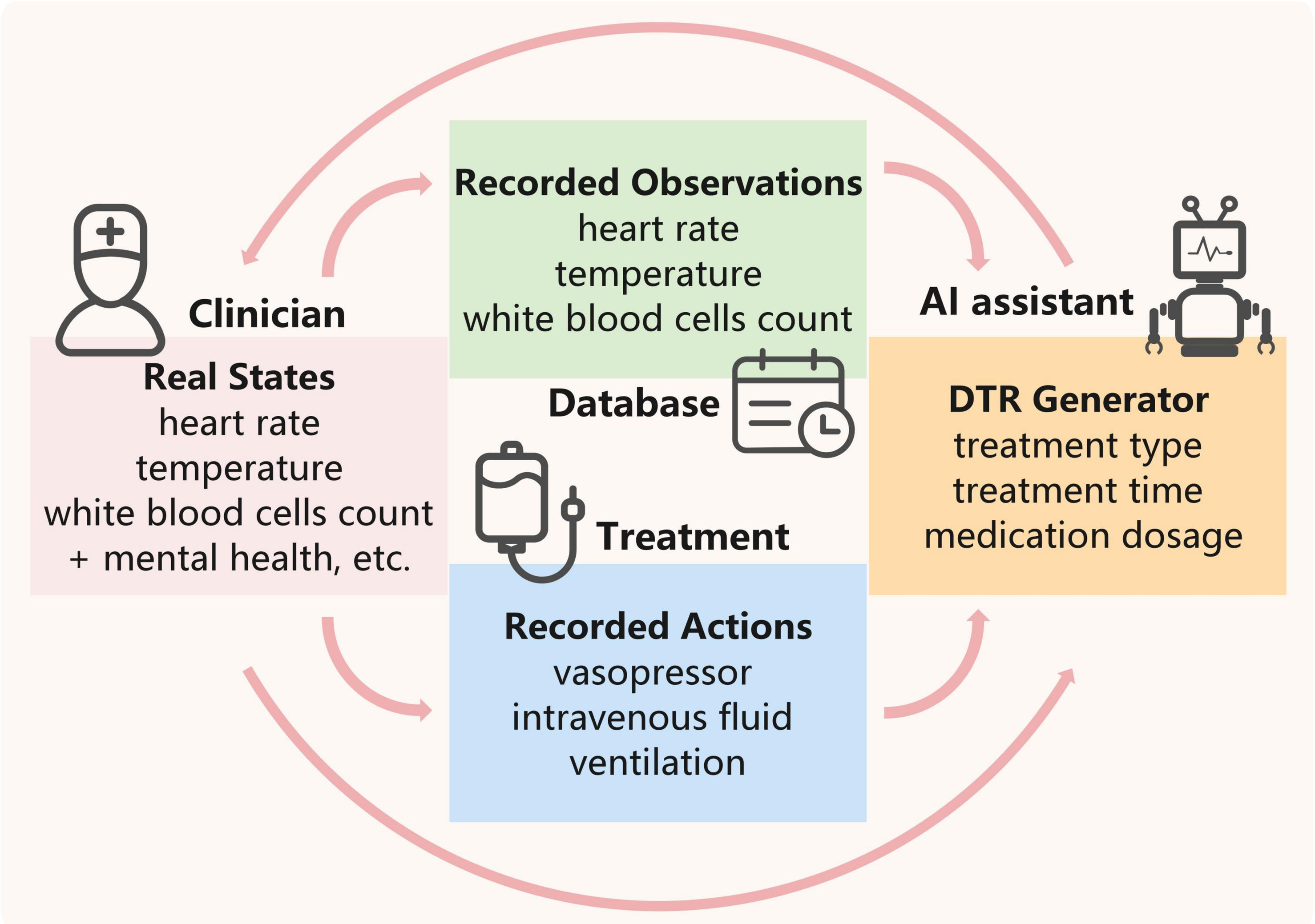
Generative Dynamic Treatment Regimes

Yiwei Dong Holly Wiberg✉ Shixiang Zhu✉

Carnegie Mellon University

BACKGROUND

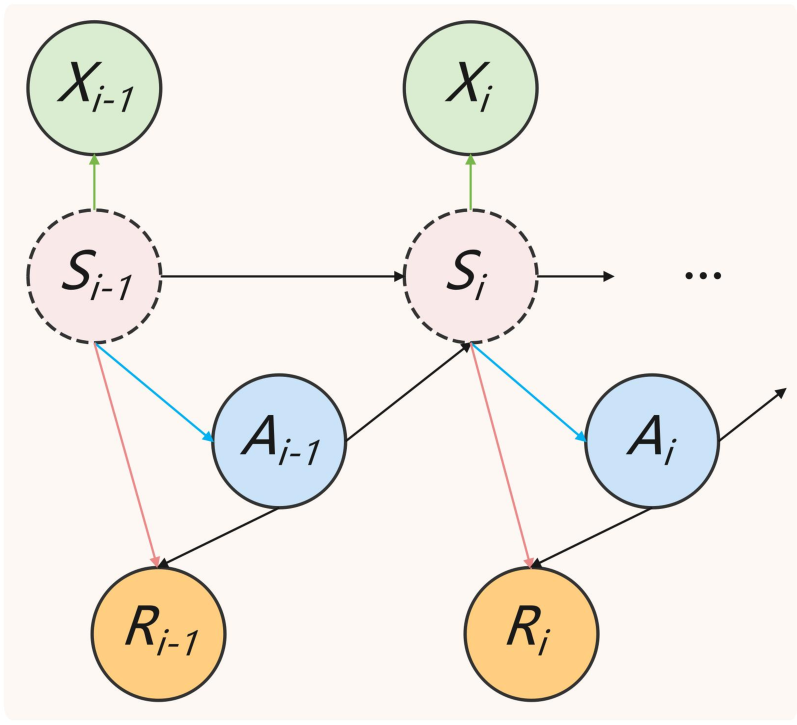
- Dynamic Treatment Regimes (DTRs) aim to prescribe treatment plans at various stages based on an individual's time-varying characteristics.
- Learning DTRs through the direct interaction with clinical environments is often impractical due to ethical constraints, necessitating the use of offline data from pre-collected electronic health records.
- The presence of unobserved confounders arising from incomplete patient records could result in biased treatment policies.



OBJECTIVES

- Develop an offline DTR learning algorithm that mitigates the impact of hidden confounders by introducing deconfounding methods.
- Derive the distribution of feasible treatments at each stage using the generative DTR modeling to provide a series of timely personalized treatment recommendations for clinicians.
- Incorporate feedback from human clinicians to the DTR generator and examine the trade-off between the number of generated samples and the accuracy of clinician decisions.

PROBLEM FORMULATION



Confounded Markov Decision Process

$\langle \mathcal{S}, \mathcal{A}, \mathcal{X}, \mathcal{T}, \mathcal{R}, \gamma \rangle$

$\mathcal{S}$: real states
$\mathcal{A}$: actions
$\mathcal{X}$: observations
$\mathcal{T}$: state transitions
$\mathcal{R}$: rewards
γ: discount factor

METHODS

I. Deconfounding Variational Autoencoder

- Generative Model

$$p(s_t) = \prod_{j=1}^{D_s} \mathcal{N}(s_{tj} \mid 0, 1) \quad \text{Prior}$$

$$\left.\begin{aligned} p_\theta(x_t \mid s_t) &= \mathcal{N}(x_t \mid \mu_x(s_t), \sigma_x^2(s_t)) \\ p_\theta(a_t \mid s_t) &= \mathcal{N}(a_t \mid \mu_a(s_t), \sigma_a^2(s_t)) \end{aligned}\right\} \text{Decoder}$$

$$p_\theta(s_t \mid s_{t-1}, a_{t-1}) = \mathcal{N}(s_t \mid \mu_s(s_{t-1}, a_{t-1}), \sigma_s^2(s_{t-1}, a_{t-1}))$$

- Inference Model

$$q_\phi(s_1 \mid x, a) = \mathcal{N}(s_1 \mid \mu_{s_1}(x, a), \sigma_{s_1}^2(x, a)) \quad \text{Encoder}$$

$$q_\phi(s_t \mid s_{t-1}, x, a) = \mathcal{N}(s_t \mid \mu_{s_t}(s_{t-1}, x, a), \sigma_{s_t}^2(s_{t-1}, x, a))$$

- Loss Function

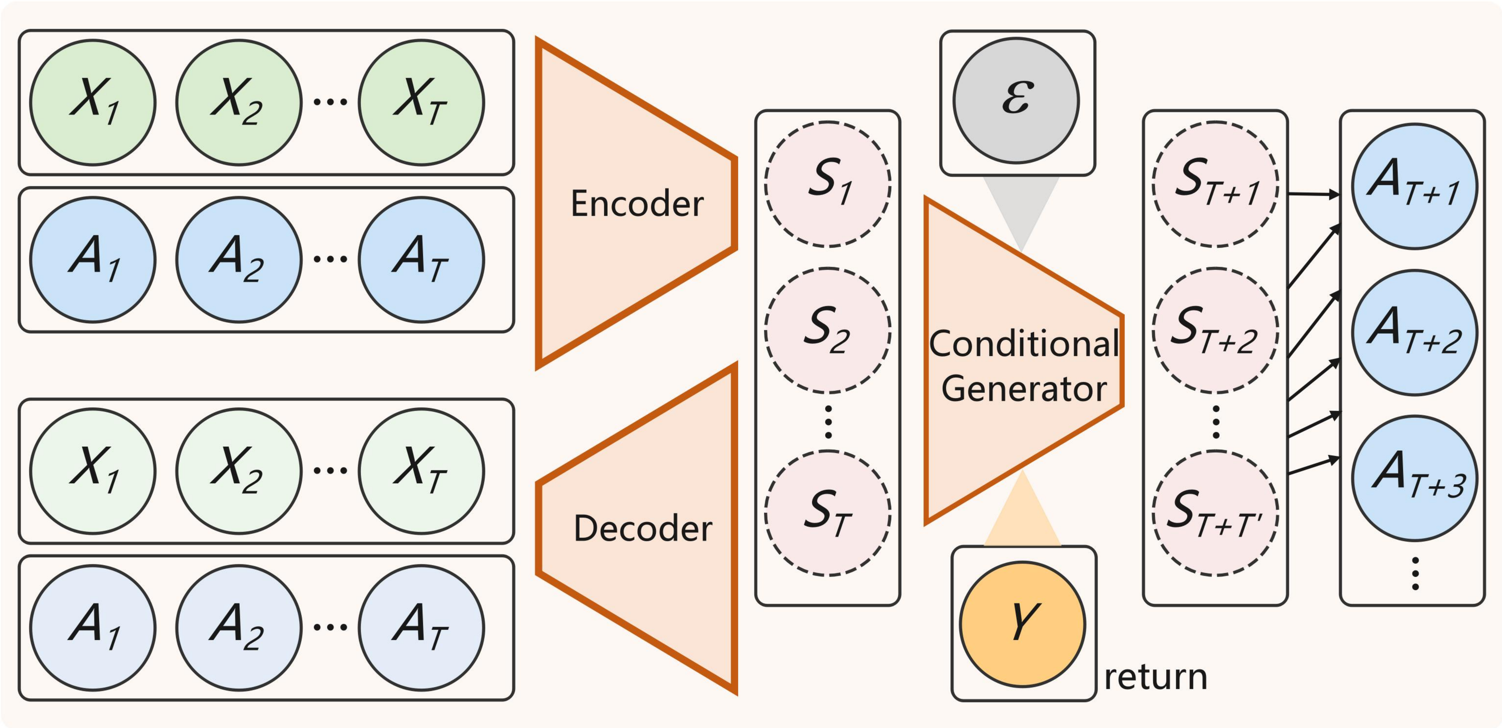
$$\mathcal{L}(x, a; \theta, \phi) \quad \text{Evidence Lower Bound}$$

$$= \sum_{t=1}^{T} \mathbb{E}_{s_t \sim q_\phi(s_t|s_{t-1}, x, a)} \left[\log p_\theta(x_t|s_t) + \log p_\theta(a_t|s_t)\right]$$
$$- \mathrm{KL}(q_\phi(s_1|x, a) \parallel p_\theta(s_1))$$
$$- \sum_{t=2}^{T} \mathrm{KL}(q_\phi(s_t|s_{t-1}, x, a) \parallel p_\theta(s_t|s_{t-1}, a_{t-1}))$$

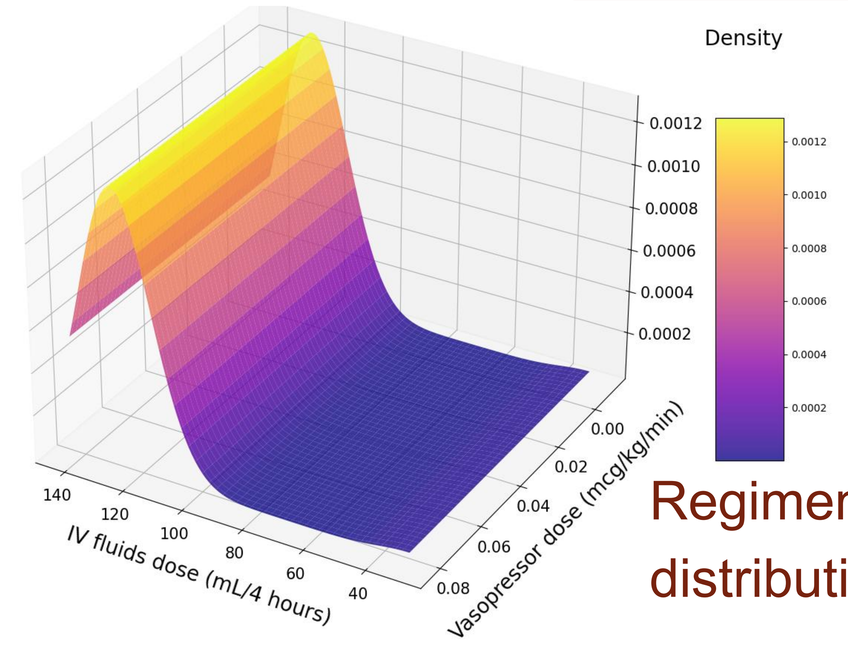
II. Conditional DTR Generator

- Learning Algorithm

Input: Noise model ϵ_ψ, inverse dynamics f_ϕ, guidance scale ω, history length C, return y
Initialize: $h \leftarrow$ Queue(length $= C$), $t \leftarrow 0$
while not done **do**
  Collect the representation s learned by deconfounding VAE;
  h.insert(s); Initialize $\overline{s}_K(\tau) \sim \mathcal{N}(0, \alpha I)$
  for $k = K$ **to** 1 **do**
    $\overline{s}_k(\tau)[:\text{length}(h)] \leftarrow h$
    $\hat{\epsilon} \leftarrow \epsilon_\psi(\overline{s}_k(\tau), k) + \omega(\epsilon_\psi(\overline{s}_k(\tau), y, k) - \epsilon_\psi(\overline{s}_k(\tau), k))$
    $(\mu_{k-1}, \Sigma_{k-1}) \leftarrow \text{Denoise}(\overline{s}_k(\tau), \hat{\epsilon})$
    $s_{k-1} \sim \mathcal{N}(\mu_{k-1}, \alpha\Sigma_{k-1})$
  end for
  Execute sampled a from $f_\phi(s_t, s_{t+1})$; $t \leftarrow t+1$
end while



RESULTS



Regimen distribution

Scan the QR code to see more results

REFERENCES

Johnson, A., Pollard, T., & Mark, R. (2016). MIMIC-III Clinical Database (version1.4). PhysioNet. https://doi.org/10.13026/C2XW26