Active inference as a general framework for modeling human driving behavior

Johan Engström, Matthew O'Kelly, Leif Johnson, Azadeh Dinparastdjadid, Shu-Yuan Liu, Joao Messias, Waymo LLC, CA, USA

Motivation

- Human driver behavior models play a key role in autonomous vehicle (AV) evaluation and in traffic safety research more broadly
- Active inference offers a common framework for understanding and modeling human driver behavior, suggesting a reconceptualization of traditional notions in driving behavior research [1]
- This poster summarizes recent and ongoing active inference-based computational driver behavior models, addressing three main aspects of driving behavior:
  1. Managing uncertainty
  2. Responding to urgent conflicts
  3. Social interaction

Starting point

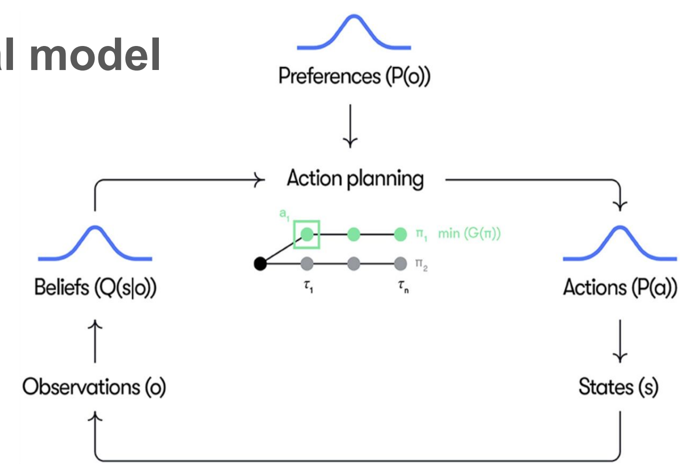
- *Driving behavior can be understood and modeled based on the single principle of minimizing expected free energy (EFE)*
- *EFE can be decomposed into* ***pragmatic*** *(goal related) and* ***epistemic*** *(information related) value*

$$EFE = G(\pi)$$
$$= -\underbrace{\mathbb{E}_{Q(\mathbf{o}|\pi)}[\log P(\mathbf{o})]}_{\text{Pragmatic value}} - \underbrace{\mathbb{E}_{Q(\mathbf{s},\mathbf{o}|\pi)} D_{KL}[Q(\mathbf{s}|\mathbf{o},\pi)||Q(\mathbf{s}|\pi)]}_{\text{Epistemic value}}$$

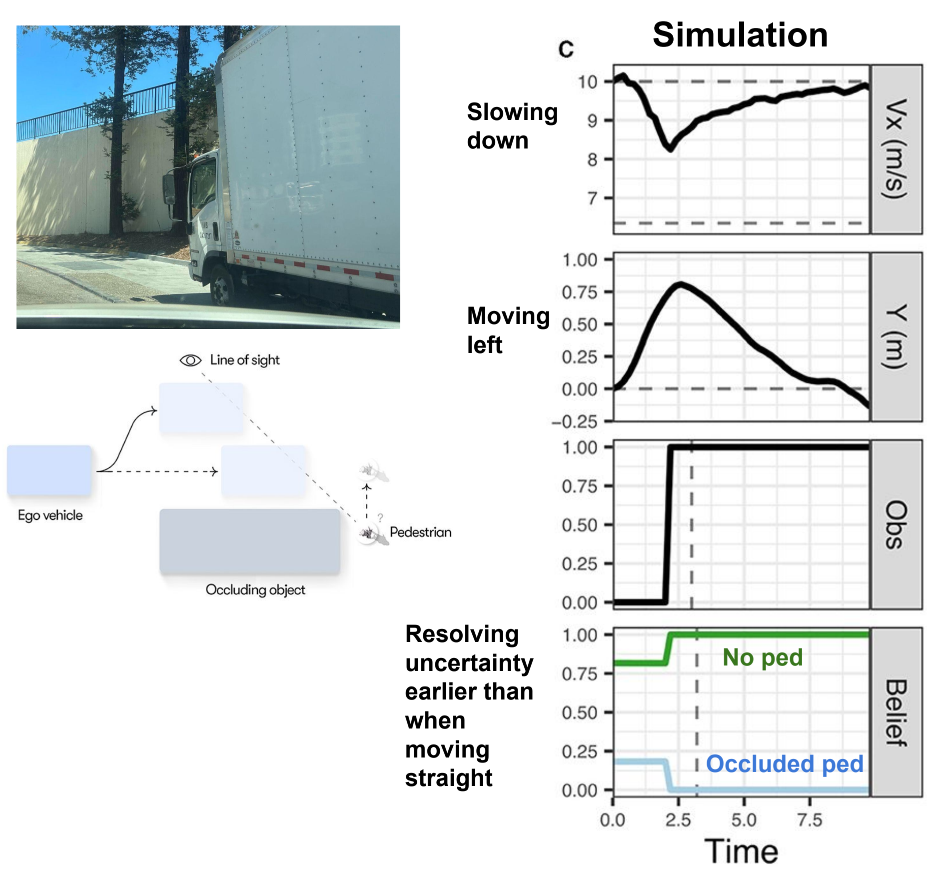
1. Managing uncertainty

Key idea: *Model adaptive driving behavior in terms of exploitation (goal achievement, pragmatic value) vs. exploration (uncertainty resolution, epistemic value) [2]*

General model

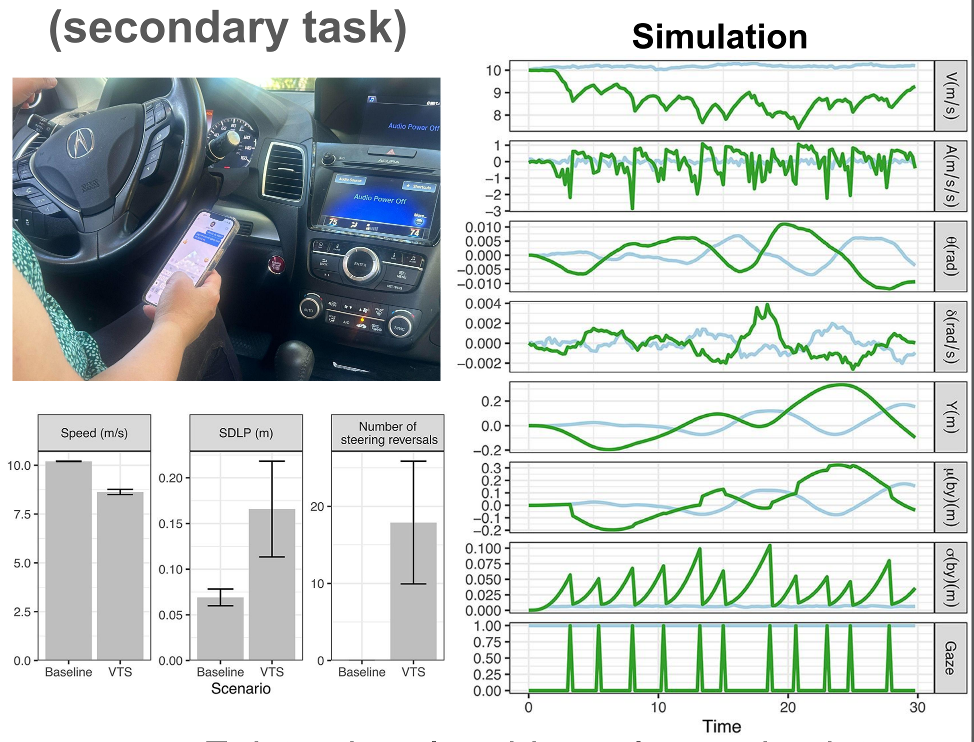


Simulation (a): Moving past an occlusion



- Epistemic value drives information seeking behavior ("looking around the corner") to resolve uncertainty
- Unlocks goal-directed pragmatic value (moving forward to make progress)
- Pragmatic and epistemic value seamlessly interact to optimize behavior (minimize EFE) -> "***resolves uncertainty on the fly"***

Simulation (b): Visual behavior (secondary task)

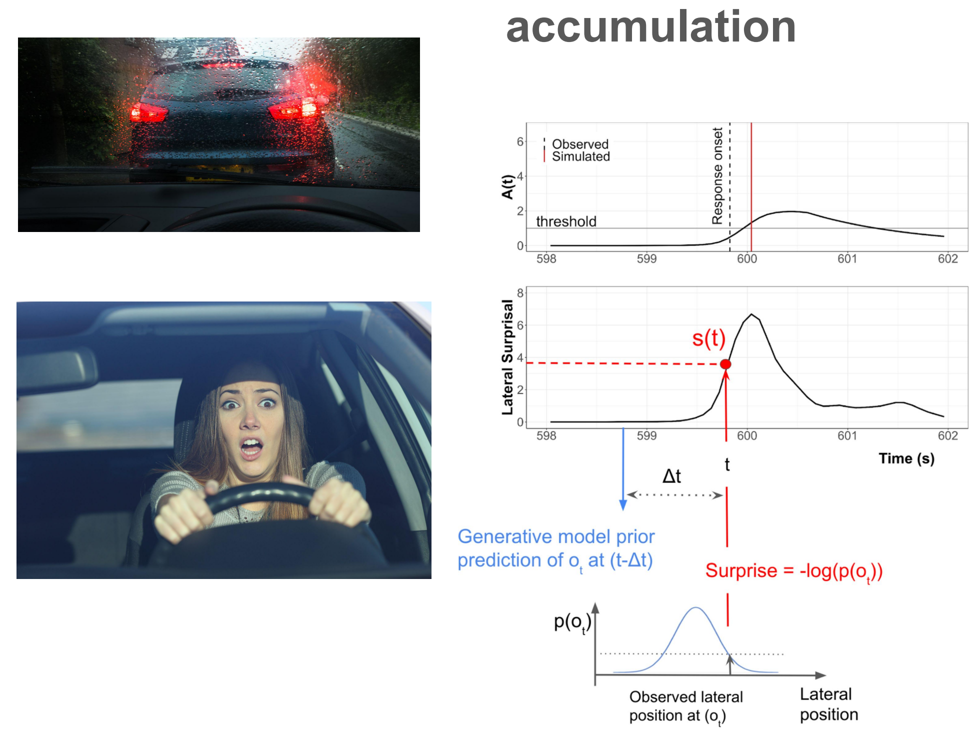


- Epistemic value drives glances back to the road to resolve uncertainty
- Traded against pragmatic value (preferred speed, lane keeping etc.) -> visual time sharing behavior
- Reproduces human data

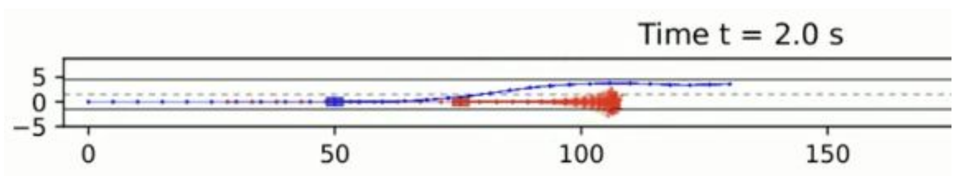
2. Responding to urgent conflicts

Key idea: *Model responses to urgent conflicts as driven by surprise (deviation from preferred / expected observations) [3] and evasive maneuver decisions as EFE minimization [4]*

Response timing as surprise accumulation



Full closed loop collision avoidance agent model [4] - see our companion poster (Schumann et al.)

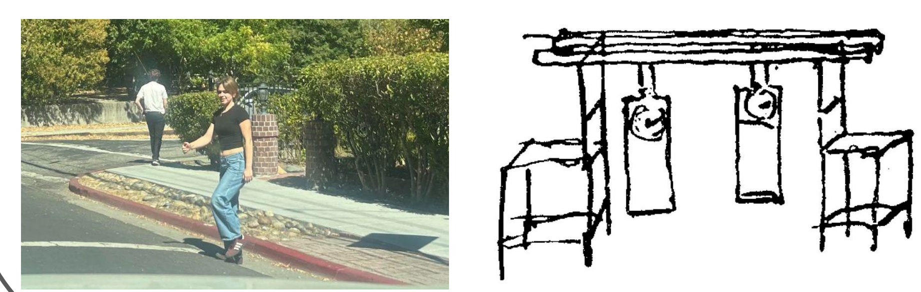


Surprise accumulation here represents the dynamics of the entire re-planning process along with other constraints such as perceptual limitations.

3. Social interaction

Key ideas (ongoing work):

- *Model communicative acts (gesturing, honking, intent yielding) as epistemic actions with the goal to reduce uncertainty*
- *Establishes a shared schema for how the situation will play out (e.g., who goes first).*
- *A shared schema may be obtained as the result of generalized synchrony between two agents with similar generative models [5]*



References

[1] Engström, J., Bärgman, J., Nilsson, D., Seppelt, B., Markkula, G., Piccinini, G. B., and Victor, T. Great expectations: a predictive processing account of automobile driving. Theoret. Iss. Ergon. Sci. 19, 156–194, (2018).

[2] Engström, J., Wei, R., McDonald, S.D., Garcia, A., O'Kelly, M. and Johnson, L. Resolving uncertainty on the fly: modeling adaptive driving behavior as active inference. Front. Neurorobot., 21, Volume 18, (2024).

[3] Engström, J., Liu, S.-Y., Dinparastdjadid, A., and Simoiu, C., Modeling road user response timing in naturalistic traffic conflicts: a surprise-based framework. Accid. Anal. Prevent. 198, (2024).

[4] Schumann, J., Engström, J., O'Kelly, M., Kober, J. and Zgonnikov, A. Active inference-based modeling of human driver collision avoidance behavior, Extended abstract Submitted to IWAI 2024.

[5] Friston, K., and Frith, C. A duet for one. Conscious. Cogn. 36, 390–405. (2015)