

An Epistemic Logic of Desire

Rui-Jie Yew (Scripps College)
CCMS Virtual Poster Session
September 2020



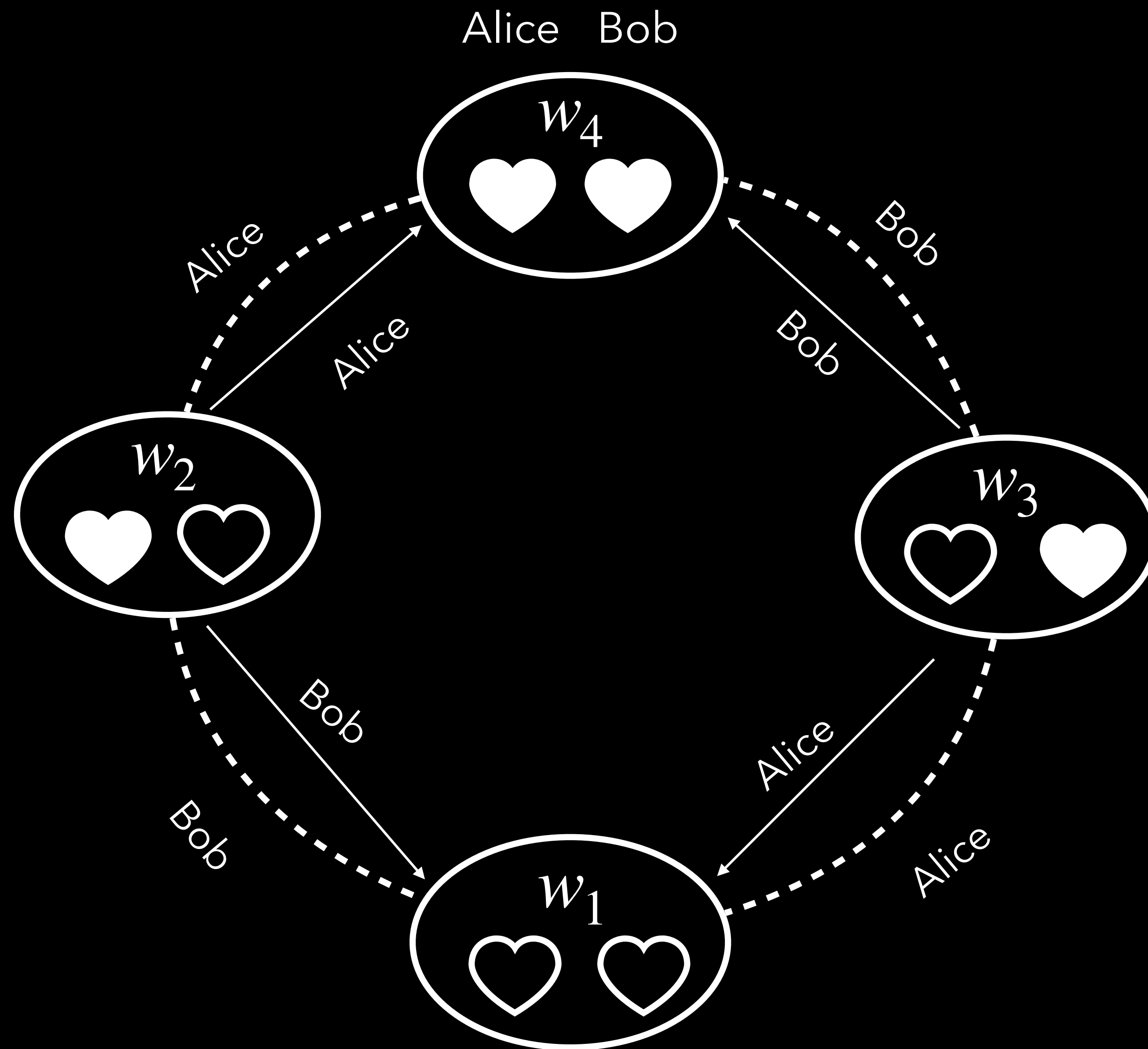
Pavel Naumov

Main Results

Complete list of axioms describing desires, with a 'coherence' axiom representing the interplay between desire and knowledge.

Completeness and soundness of logical system.

Alice and Bob: Epistemic Model with Preferences



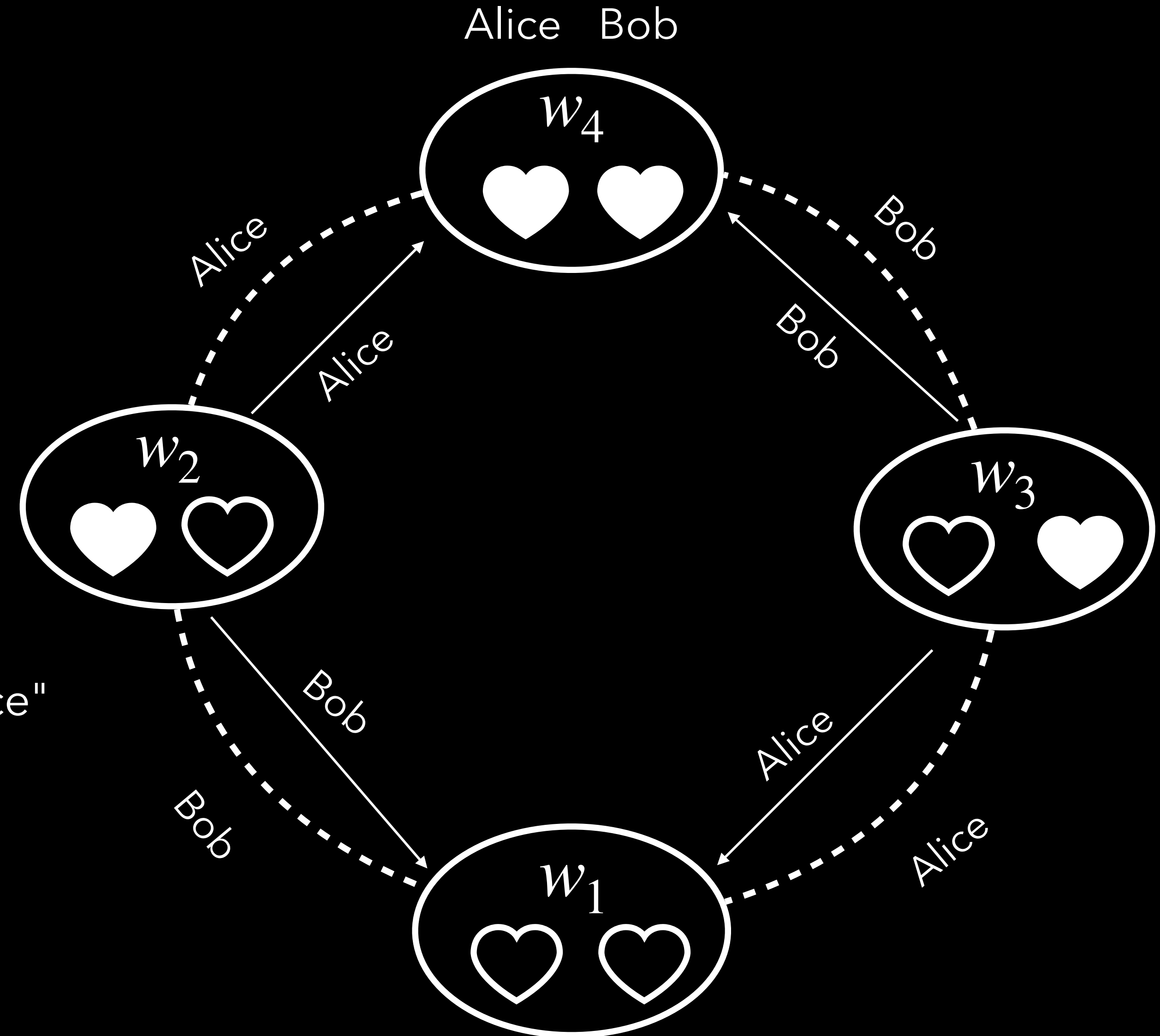
Alice and Bob: Desires

$w_2 \models D_{\text{Alice}} \text{"Bob likes Alice"}$

$w_2 \models K_{\text{Alice}} D_{\text{Alice}} \text{"Bob likes Alice"}$

$w_3 \not\models D_{\text{Alice}} \text{"Bob likes Alice"}$

$w_4 \models D_{\text{Bob}} \text{"Alice likes Bob"} \wedge D_{\text{Alice}} \text{"Bob likes Alice"}$



I'M SURE YOU'VE HEARD ALL ABOUT THIS SORDID AFFAIR IN THOSE GOSSIPY CRYPTOGRAPHIC PROTOCOL SPECS WITH THOSE BUSYBODIES SCHNEIER AND RIVEST, ALWAYS TAKING ALICE'S SIDE, ALWAYS LABELING ME THE ATTACKER.



YES, IT'S TRUE. I BROKE BOB'S PRIVATE KEY AND EXTRACTED THE TEXT OF HER MESSAGES. BUT DOES ANYONE REALIZE HOW MUCH IT HURT?



HE SAID IT WAS NOTHING, BUT EVERYTHING FROM THE PUBLIC-KEY AUTHENTICATED SIGNATURES ON THE FILES TO THE LIPSTICK HEART SMEARED ON THE DISK SCREAMED "ALICE."



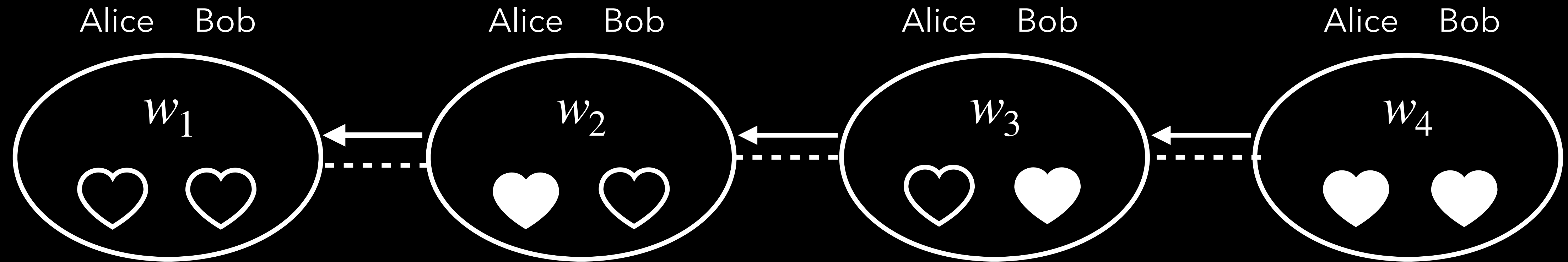
I DIDN'T WANT TO BELIEVE. OF COURSE ON SOME LEVEL I REALIZED IT WAS A KNOWN-PLAINTEXT ATTACK. BUT I COULDN'T ADMIT IT UNTIL I SAW FOR MYSELF.



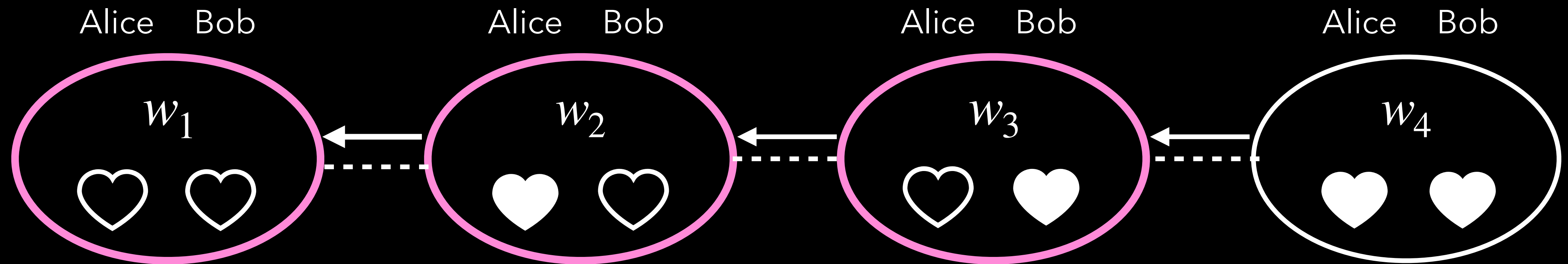
SO BEFORE YOU SO QUICKLY LABEL ME A THIRD PARTY TO THE COMMUNICATION, JUST REMEMBER: I LOVED HIM FIRST. WE HAD SOMETHING AND SHE TORE IT AWAY. SHE'S THE ATTACKER, NOT ME. NOT EVE.



Eve: Epistemic Model with Preferences

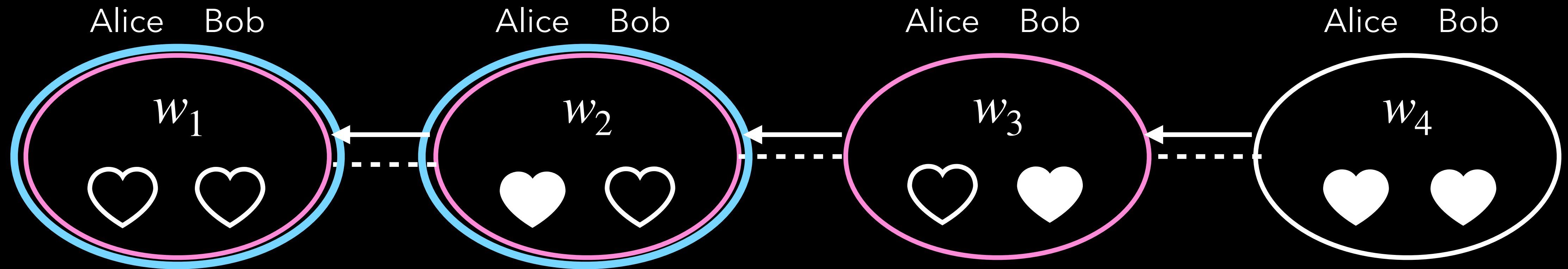


Eve: Desires



$w \models D_{\text{Eve}}$ "At least one of Bob and Alice is not in love with the other" for $w \in \{w_1, w_2, w_3, w_4\}$

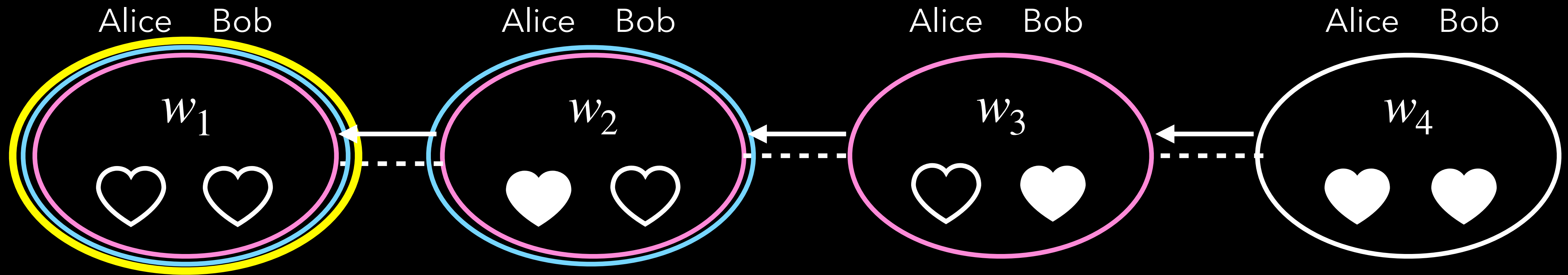
Eve: Desires



$w \models D_{\text{Eve}}$ "At least one of Bob and Alice is not in love with the other" for $w \in \{w_1, w_2, w_3, w_4\}$

$w \models D_{\text{Eve}}$ **"Bob is not in love with Alice" for $w \in \{w_1, w_2, w_3, w_4\}$**

Eve: Desires



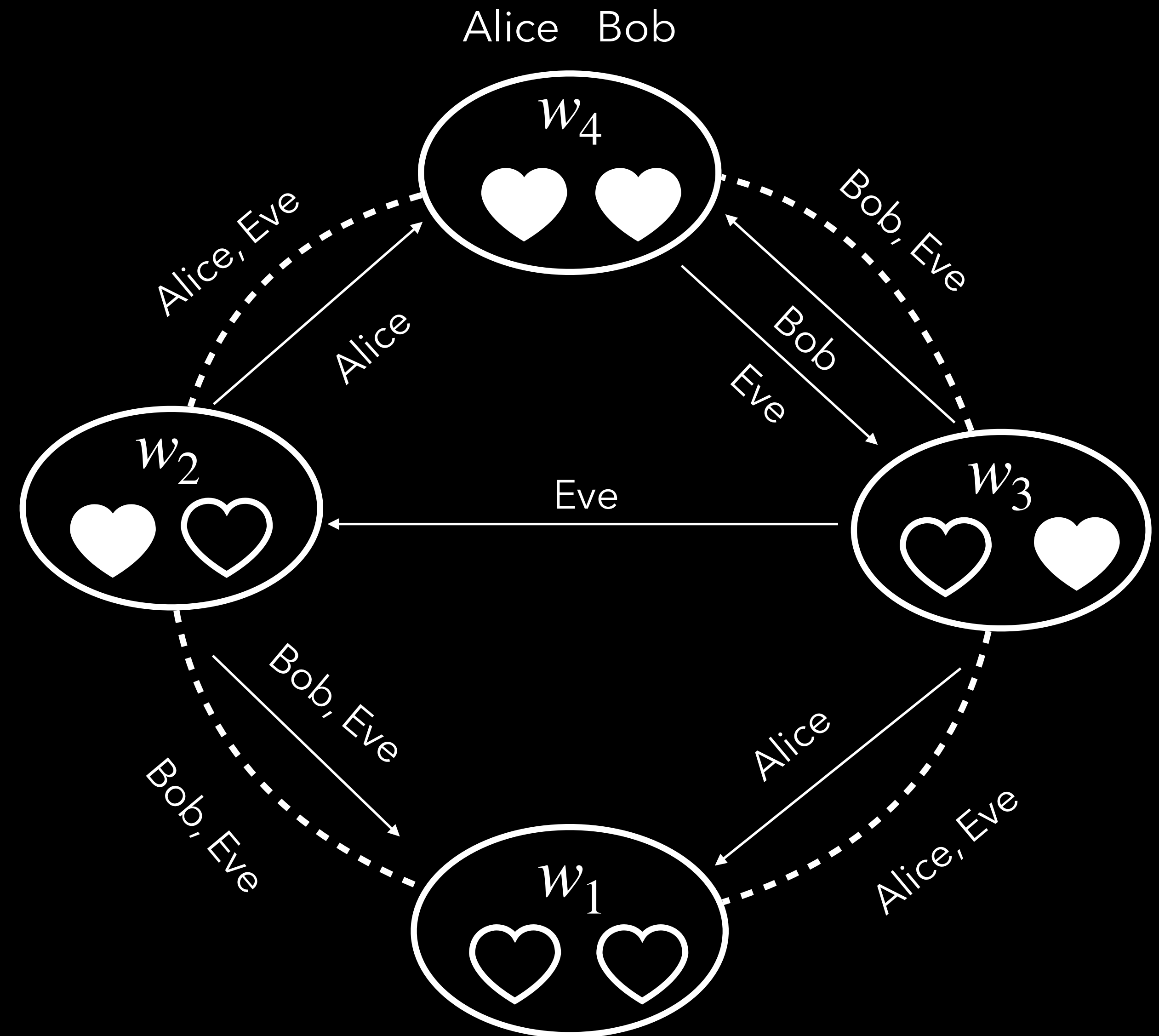
$w \models D_{\text{Eve}}$ "At least one of Bob and Alice is not in love with the other" for $w \in \{w_1, w_2, w_3, w_4\}$

$w \models D_{\text{Eve}}$ "Bob is not in love with Alice" for $w \in \{w_1, w_2, w_3, w_4\}$

$w \models D_{\text{Eve}}$ **"Bob and Alice are not in love with each other" for $w \in \{w_1, w_2, w_3, w_4\}$**

Everybody...

$w \Vdash D_{\text{Eve}} D_{\text{Bob}} D_{\text{Alice}} \text{"Bob does not like Alice"},$
 for $w \in \{w_1, w_2, w_3, w_4\}$.



Model Limitations

Single-agent second-order desires cannot be expressed through the logical system.

An agent cannot have contradicting desires.

Thank you!

Acknowledgements

Pavel Naumov

Wenxuan (Wendy) Zhang (Scripps '22)

Simone Gottlieb (Scripps '21)