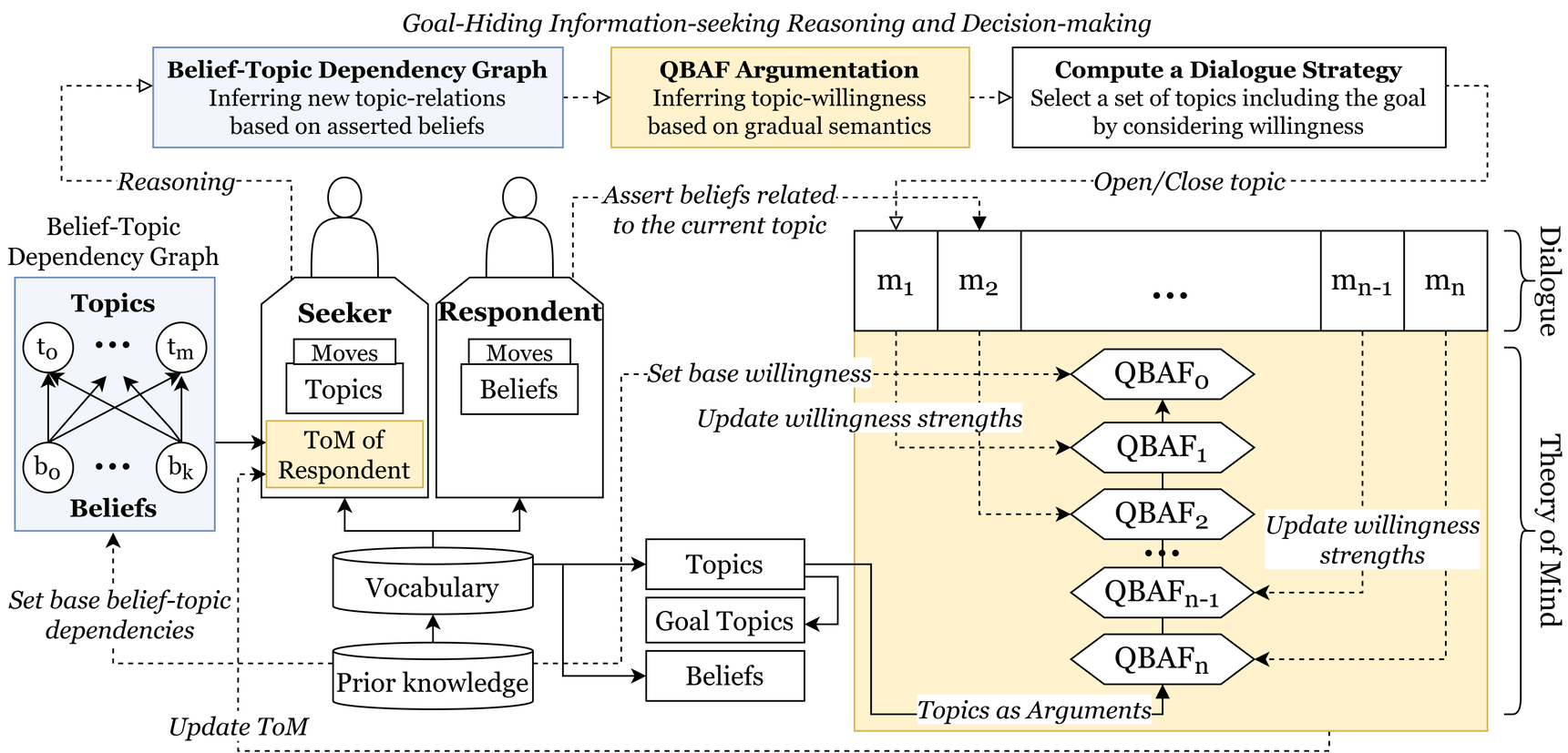
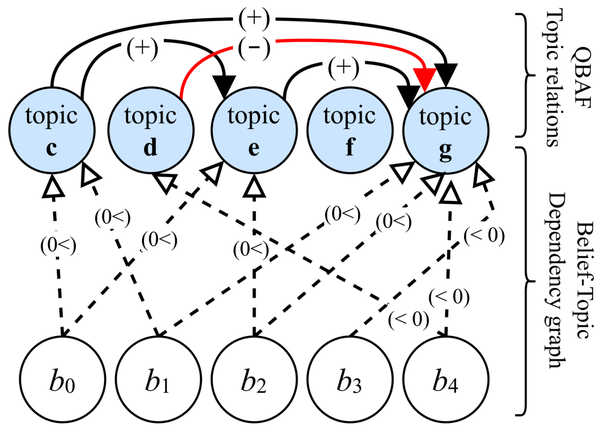
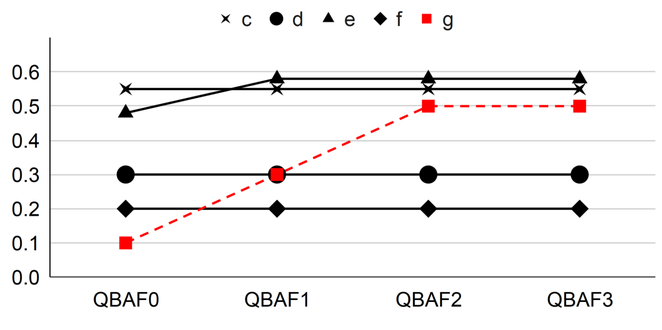
GOAL-HIDING INFORMATION-SEEKING DIALOGUES
A FORMAL FRAMEWORK

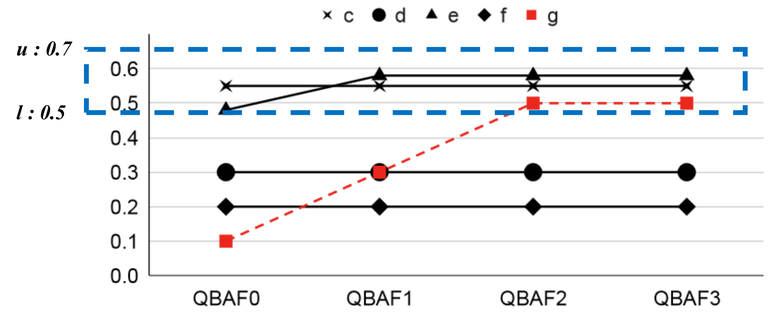
Andreas Brännström, Virginia Dignum, Juan Carlos Nieves

Goal-hiding information-seeking dialogues model interactions where a seeker agent believes a human respondent is unwilling to share the seeker's sought-for information. The seeker aims to introduce its goal topic to elicit information, while respecting the respondent's willingness for topics. Thus, the seeker postpones its goal topic until the respondent is perceived to willingly talk about it. The framework utilizes Quantitative Bipolar Argumentation Frameworks (QBAFs) to assign willingness scores to topics, represented as arguments, that either support or attack other topics. By considering supports among topics, paths of topics can be inferred to promote willingness for the hidden goal topic. A QBAF is incrementally learned for each dialogue state.









Belief-Topic Dependency Graph links beliefs to topics, based on their positive or negative dependency in [-1,1] to topics. Belief-topic dependencies indirectly infer relations (supports and attacks) between topics in a QBAF associated to the current dialogue state.

Strength monotonicity is a property checked for at topic selection, stating that willingness for the goal topic is estimated to not decrease at the next dialogue state. This maintains a goal-oriented dialogue structure.

Sensitivity interval constrains the seeker's topic introduction. Topics can only be opened when their willingness score falls within the specified upper (u) and lower (l) bounds, in [0,1]. The lower bound respects willingness, while the upper bound prevents excessive promotion of topics.

Applications: Conversational software agents may need to adapt goal-hiding dialogue strategies in their social interactions with humans.

- **A health assessment agent** may need to postpone an intimate question until its human interlocutor is perceived to be comfortable to share that information.
- **A social-media security agent** may need the capability of detecting nefarious hidden goals of users. This detection can be approached by recognizing goal-hiding behaviors in conversational data.

Dialogue	Belief-Topic dependency	Sup/Att relations	$\tau(c)$	$\tau(d)$	$\tau(e)$	$\tau(f)$	$\tau(g)$
[]		+{}, -{}	0.6	0.3	0.4	0.2	0.1
[c]		+{}, -{}	0.6	0.3	0.4	0.2	0.1
[c,b_0,b_1]	$dependent\_topics^{+}(b_0,\mathcal{U}^T)=\{c,e\}$ $dependent\_topics^{+}(b_1,\mathcal{U}^T)=\{c,g\}$	+{(c,e),(c,g)}, -{}	0.6	0.3	0.6	0.2	0.3
[c,b_0,b_1,e]		+{(c,e)(c,g),}, -{}	0.6	0.3	0.6	0.2	0.3
[c,b_0,b_1,e,b_2]	$dependent\_topics^{+}(b_2,\mathcal{U}^T)=\{e,g\}$	+{(c,e),(c,g),(e,g)}, -{}	0.6	0.3	0.6	0.2	0.5
[c,b_0,b_1,e,b_2,g]		+{(c,e),(c,g),(e,g)}, -{}	0.6	0.3	0.6	0.2	0.5
[c,b_0,b_1,e,b_2,g,b_3]	$dependent\_topics^{+}(b_3,\mathcal{U}^T)=\{g\}$	+{(c,e),(c,g),(e,g)}, -{}	0.6	0.3	0.6	0.2	0.5

Goal-hiding dialogue process. As beliefs are asserted, support and attack relations between topics are identified. The dialogue results in a gradual change in willingness of topics.
Topics = {c,d,e,f,g}. *Goals* = {g}. *Beliefs* = {b0, b1, b2, b3, b4}. *Lower bound l=0.5. Upper bound u=0.7.*

UMEÅ UNIVERSITY
Umeå University, Department of Computing Science, SE-901 87 Umeå, www.umu.se