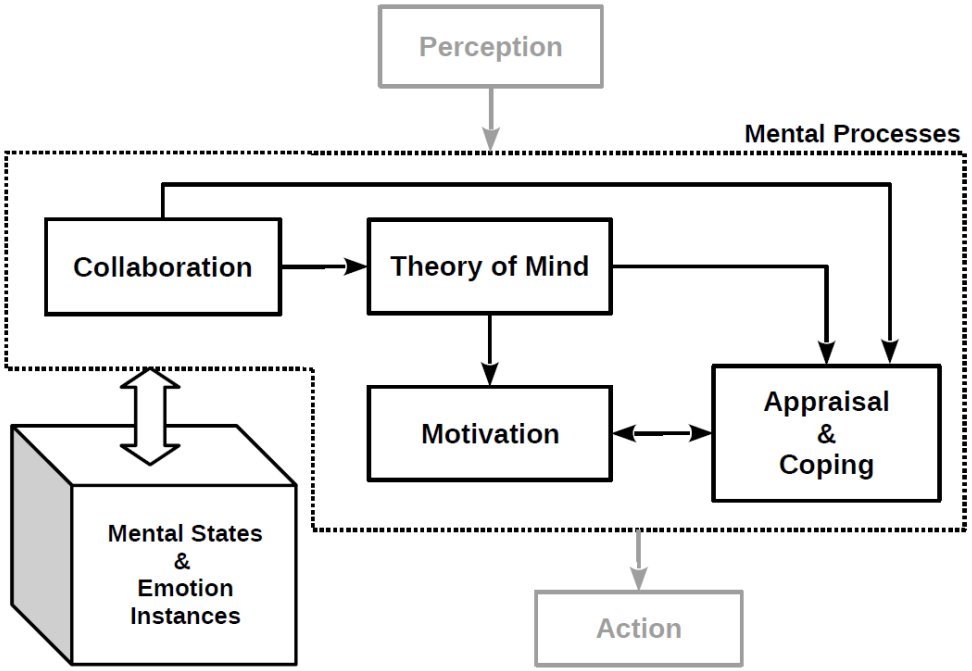
Example 1 - Agreeing on a Shared Goal:

Emotion-awareness	Emotion-Ignorance (neutral)
A1. Astronaut: *Oh no! Finishing the quality check of our installation with this measurement problem is so frustrating. I think we should stop now!* *[Astronaut is frustrated.]* ***A2. Robot:*** *I see. This is frustrating. But, I can help you with the measurement tool and we can finish the task as originally planned.* *[Robot perceives Astronaut's frustration and acknowledges that.]* ***A3. Astronaut:*** *Can you fix the measurement tool?* *[Astronaut's emotion is neutral.]* ***A4. Robot:*** *The next task is fixing the panel and it needs you to prepare and attach the welding rod to your welding tool. To save our time, I will fetch another measurement tool while you are preparing your welding tool.* *[Robot perceives Astronaut's neutral emotion, and tries to negotiate and provide a fair offer.]* ***A5. Astronaut:*** *That would be great!* *[Astronaut is content.]* *[Robot perceives Astronaut's contentment, and interprets it as his agreement. Mission continues.]*	***B1. Astronaut:*** *Oh no! Finishing the quality check of our installation with this measurement problem is so frustrating. I think we should stop now!* *[Astronaut is frustrated.]* ***B2. Robot:*** *I see. I can help you with the measurement tool, or we can terminate this task. What do you want me to do?* *[Robot does not perceive Astronaut's frustration.]* ***B3. Astronaut:*** *As I said the measurement tool does not work properly. We can not continue.* *[Astronaut is frustrated.]* ***B4. Robot:*** *Okay. Do you want me to fix this problem or terminate the task?* *[Robot does not perceive Astronaut's frustration.]* ***B5. Astronaut:*** *Can you fix my measurement tool?* *[Astronaut is frustrated, even more.]* ***B6. Robot:*** *I cannot fix your measurement tool, but I can fetch another one for you if you want?* *[Despite Astronaut's strong frustration, Robot tries to negotiate.]* ***B7. Astronaut:*** *No, I don't want another measurement tool! We don't have time for that!* *[Astronaut is angry.]* ***B8. Robot:*** *Okay. You want me to terminate this task. Terminating this task can influence the quality of installing this solar panel which can cause the mission to fail. Or, do you want us to work on another task? This can help us to install the panel using your welding tool, but I do not know whether the quality of our installation will be acceptable.* *[Not only the Robot does not perceive Astronaut's anger, but also contniues to negotiate the next step based on the shared plan to select proper action.]* ***B9. Astronaut:*** *I told you we have this problem and we should terminate the mission! We cannot continue without the measurement tool!* *[Astronaut is angry.]* *[Mission is unsuccessfully terminated.]*

NOTE: This example shows that agreeing on a shared goal requires the robot to be aware of its collaborator's frustration. Otherwise a collaborative robot will try to maintain the status of the shared goal and prevent it from failure without considering its collaborator's negative emotion. First, the emotion-ignorant robot does not acknowledge the Astronaut's frustration (i.e., B2 in compare to A2), since it does not perceive that. Then, while negotiating the shared goal the robot fails to offer a potential solution based on the Astronaut's emotional state, resulting in the failure of the negotiation during collaboration.

Emotions [A, V, S]:
Frustration = [-0.5, -0.5, 0.1]
Neutral = [0, 0, 0]

NOTE: I use a convention to refer to the utterances of the left column (in the first page) by A# and the right one by B# in my walk-through explanations.



Awareness: Ability of the robot to a) manage and understand its own emotions and those of the other, b) express emotions accordingly.

If Robot perceives Astronaut's **frustration** correctly: (Frustration = [-0.5, -0.5, 0.1])

1. **(Perception)**: Robot perceives human's utterances and emotion. **(A1)**

2. **(Collaboration , Theory of Mind → Reverse Appraisal):** Robot uses *reverse appraisal* to understand the meaning of *frustration* according to the *collaborative task status* (e.g., precondition status, postcondition status, required resources, shared goal). Robot updates human's *user model* respectively.

3. (**Appraisal**): Robot *appraises* Astronaut's utterances and emotion.

4. (**Collaboration , Motivation → Motive Formation → Belief Formation → Intention Formation**): Robot forms new motives according to the result of a) *appraisal* with respect to the *shared goal*, and b) *reverse appraisal* of the human emotion. Robot forms new beliefs about a) human's *inferred mental states* (belief, intention, motive, goal, emotion), and b) the task status. Robot forms new intention(s) with respect to the new beliefs.

5. **(Mental States, Coping)**: Based on the current mental states, Robot chooses an *emotion-focused* coping strategy and decides to acknowledge Astronaut's emotion, and provide alternative solution. **(A2)**

If Robot **does not** perceive Astronaut's emotions: (Neutral = [0, 0, 0])

1. **(Perception)**: Robot only **perceives** human's utterances. **(B1)**

2. **(Collaboration , Theory of Mind → Reverse Appraisal):** Robot uses *reverse appraisal* to understand the meaning of *neutral* according to the *collaborative task status* (e.g., precondition status, postcondition status, required resources, shared goal). Robot updates human's *user model* respectively.

[Robot will not be able to update human's user model correctly based on his reverse appraisal, since it is missing human's frustration (consequence on human's **user model**).]

3. (**Appraisal**): Robot *appraises* Astronaut's utterances.

[Robot will not be able to appraise the external event correctly, since it misses human's emotion (consequence on A**ppraisal mechanism** → As a side effect of wrong appraisal the **Coping mechanism** will act incorrectly → As another side effect, this leads to incorrect values of mental state attributes, e.g., **Accuracy** of a belief, **Certainty** or **Affective-Deliberative Consistency** of an intention, **Difficulty** of a goal, **Failure Disruptiveness** or **Importance** of a motive).]

4. (**Collaboration , Motivation → Motive Formation → Belief Formation → Intention Formation**): Robot forms new motives according to

the result of a) *appraisal* with respect to the *shared goal*, and b) *reverse appraisal* of the human's wrong emotion (i.e., neutral). Robot forms new beliefs about a) human's *inferred mental states* (belief, intention, motive, goal, emotion), and b) the task status. Robot forms new intention(s) with respect to the new beliefs.

- [Robot will not be able to form required/appropriate motive according to the human's real emotion (i.e., frustration). As a result, there will be no motive to remove Astronaut's frustration, and similarly no correct motive to agree on the shared goal (consequence on **forming motives**).]
- [Robot will not be able to form proper beliefs according to the task status, human's mental states and appraisal of the event (consequence on **forming beliefs**).]
- [Robot will not be able to form proper intentions according to the corresponding beliefs (consequence on **forming intentions**).]

5. **(Mental States, Coping)**: Based on the current mental states, Robot decides to use *problem-focused* coping strategy of seeking information to be able to choose between two available actions and reduce the current amount of uncertainty. **(B2)**

NOTE: As a result Robot will show emotionally detached behaviors. (Against what AI in general claims: proper behavior)

Mediation: Collaboration needs the collaborators to be seeking to discover a common ground and reach agreement to settle matter of mutual concern or resolve a conflict.

If Robot perceives Astronaut's **frustration** correctly: (Frustration = [-0.5, -0.5, 0.1])

* *The process in the awareness function will be run again (similar to what we had above) for the new utterance of the Astronaut* ***(A3)*** before the following processes.

6. **(Collaboration , Mental States → Attributes)**: Robot extracts the current *focus of attention* during collaboration and also computes the *distance* between its own and human's mental states (distance is a function of mental states' attributes and human's emotion).

7. **(Collaboration , Appraisal → Belief Formation → Intention Formation)**: Robot appraises the human's perceived emotion with respect to the *current focus of attention* (on a task) and *distance* between collaborators' mental states. Robot *forms/updates its mental states* (beliefs and intentions) with respect to the result of appraising human's *perceived emotion*.

8. **(Theory of Mind → User Model)**: Robot uses its new mental states to *update* human's *user model*.

9. (**Mental States , Collaboration , Theory of Mind , Motivation → Motive Formation → Belief Formation → Intention Formation**): Robot forms a new motive based on how the collaboration can proceed with respect to the Astronaut's inferred mental states.

10. **(Mental States, Coping)**: Based on the current mental states, Robot chooses to negotiate and offer an alternative action to the Astronaut. **(A4)**

If Robot **does not** perceive Astronaut's emotions: (Neutral = [0, 0, 0])

** *Astronaut's new utterance* ***(B3)*** *does not change Robot's mental state about the Astronaut. Thus the Robot repeats itself* ***(B4)***.

** *The process in the awareness function will be run again (similar to what we had above) for the new utterance of the Astronaut* ***(B5)*** *before the following processes.*

6. **(Collaboration , Mental States → Attributes)**: Robot extracts the current *focus of attention* during collaboration and also computes the *distance* between its own and human's mental states (distance is a function of mental states' attributes and human's emotion).

[Robot will not be able to compute the value of the distance function correctly since it is a function of mental states' attributes and human's emotions (consequence on understanding of the negotiation topic, i.e., shared goal).]

7. **(Collaboration , Appraisal → Belief Formation → Intention Formation)**: Robot appraises the human's perceived emotion with respect to the *current focus of attention* (on a task) and *distance* between collaborators' mental states. Robot *forms/updates its mental states* (beliefs and intentions) with respect to the result of appraising human's *perceived emotion*.

[Robot will not be able to appraise the external event correctly with respect to the distance between collaborators' mental states (consequence on A**ppraisal mechanism** → As a side effect of wrong appraisal the **Coping mechanism** will act incorrectly → As another side effect, this leads to incorrect values of mental state attrbutes, e.g., ***Accuracy*** of a belief, ***Certainty*** or ***Affective-Deliberative Consistency*** of an intention, ***Difficulty*** of a goal, ***Failure Disruptiveness*** or ***Importance*** of a motive).]

8. **(Theory of Mind → User Model): Robot uses its new mental states to** *update* **human's** *user model***.**

[Robot will not be able to update human's user model correctly based on its own appraisal (consequence on human's user model).]

9. (**Mental States , Collaboration , Theory of Mind , Motivation → Motive Formation → Belief Formation → Intention Formation**): Robot forms a new motive based on how the collaboration can proceed with respect to the Astronaut's inferred mental states.

[Robot will not be able to form a new motive correctly based on human's mental states and emotion (consequence on negotiation).]

10. **(Mental States, Coping)**: Based on the current mental states, Robot chooses to negotiate and offer an alternative action to the Astronaut. **(B6)**

*** The process in the awareness function will be run again (similar to what we had above) for the new utterance of the Astronaut **(B7)** before the following processes.*

** Robot repeats processes #6 to #10 to prevent failure of the collaboration after Astronaut's response. **(B8)**

NOTE: As a result the Robot will not be able to negotiate to converge to a fair agreement with the Astronaut.