A Formal Architecture of Shared Mental Models for Computational Improvisational Agents

Rania Hodhod[1,2], Andreya Piplica[1], Brian Magerko[1]

[1]School of Literature, Communicationa and Culture, Georgia Institute of Technology
[2]Faculty of Computer and Information Sciences, Ain Shams University
{rhodhod; piplica; magerko}@gatech.edu

Abstract. This paper proposes a formal approach of constructing shared mental models between computational improvisational agents (*improv agents*) and human interactors based on our socio-cognitive studies of human improvisers. Creating shared mental models helps improv agents co-create stories with each other and interactors in real-time interactive narrative experiences. The approach described here allows flexible modeling of non-Boolean (i.e. fuzzy) knowledge about scene and background concepts through the use of fuzzy rules and confidence factors in order to allow reasoning under uncertainty. It also allows improv agents to infer new knowledge about a scene from existing knowledge, recognize when new knowledge may be divergent from the other actor's mental model, and attempt to resolve this divergence to reach cognitive consensus despite the absence of explicit goals in the story environment.

Keywords: Improv agents, shared mental models, computational creativity

1 Introduction

While there have been systems that have explored improv theatre as a model for creating interactive narratives [1, 2, 3], to the extent of our knowledge none have focused on the co-construction of story between an AI and a human interactor (i.e. where neither agent has privileged or pre-authored story knowledge). Achieving this goal requires the agents to be able to construct *shared mental models* (i.e. shared understandings about scene content) while collaboratively performing an improvised story. Shared mental models (SMMs) are a cognitive construct that incorporates the development of mutual beliefs from individuals' mental models until a common mental model is reached by the group, either explicitly or implicitly [5, 6, 7]. Agents in a co-constructive interactive narrative also must be able to reason about ambiguous knowledge in an uncertain environment and to reach a shared understanding about scene elements with the other actor *without any collaboration outside of the performance*. We have developed an interactive narrative within the domain of the Old West based on the improv game *Three Line Scene*, which focuses on establishing the *platform* (i.e. the characters, setting, and joint activity of a scene) in only three lines of dialogue. *Three Line Scene* allows users to provide gestural input through Kinect to an AI-controlled avatar that is in a scene with another AI-controlled character. This paper describes a formal approach to shared mental models for interactive narrative agents that is flexible enough to allow human interactors to act as an equal co-creator in an improvised scene.

2 Shared Mental Models in Improvisation

The ambiguous actions in a scene (e.g. if one actor holds their fists one on top of the other and moves them from side to side, another actor could interpret this as either raking or sweeping among other possibilities) and the ease with which they can be misinterpreted can cause an improviser to develop a mental model that differs from the other improvisers' models. The state where improvisers' mental models differ is called *cognitive divergence* [4]. Improvisers repair their divergent mental models to reach a state of *cognitive consensus*, where everyone either implicitly or explicitly agrees on a shared mental model, through the process of *cognitive convergence*[1] [6]. Cognitive consensus can be thought of as the process of "getting on the same page."

Improvisers employ *repair strategies* to deal with divergences as they occur [4], which are either *other-oriented* or *self-oriented***.** Other-oriented repair strategies aim to affect another's mental model through presenting new concepts (*presentation*) or correcting divergences (*clarification*). Self-oriented repair strategies try to align one's own mental model with those of others. For example, an actor may state an unsure idea about what is going on in the scene so that others may confirm it (*verification*). Alternatively, the actor may introduce new, vague information to the scene to observe how the others react, hoping that this will reveal some enlightening information (*blind offer*). Repair strategies help improvisers update their mental models and approach a cognitive consensus that reflects their common understanding regarding how key issues are defined and conceptualized, which is essential in story co-creation.

3 Computational Shared Mental Models

Improvisers interact through the process of proposing and responding to *offers* (i.e. proposals made by improvisers in a performance to add something to a scene) [8]. While making or responding to offers, an improviser is able to model other actors' mental models in the scene, evaluate the outcome of actions, and update goals, which can be referred to as theory of mind [6].

Based on our understanding of how human improvisers construct shared mental models in an improvised scene, we developed a shared mental model for interactive narrative agents that preserves the above-mentioned cognitive behavior related to theory of mind. The model consists of three components, as described below: beliefs, commitments, and reasoning and decision making modules.

The *beliefs component* models the agent's beliefs about itself, about others, and about others' beliefs (see Section 4). Those beliefs are associated with confidence factors that show how much the agent thinks its belief is correct (see Section 4). Confidence factors (CFs) are fuzzy values (degrees of membership values on a scale from 0 to 1) that allow the agent to compute the strength of its beliefs in a specific world frame, such as the emerging platform (i.e. initial scene details). CFs guide the actions that the agent takes either towards advancing the scene (by adding new information to the scene) or correcting cognitive divergences (by taking steps to repair its mental model). The *commitments component* encapsulates any obligations (commitments) the agent might have towards others. Finally, the *reasoning and decision making component* provides the process for reasoning about fuzzy knowledge and dynamically updates beliefs, checks for any inconsistencies that exist (i.e. a divergence), and allows the agent to decide on its next action(s). Consequently, the mental model may turn to one of the previously described repair strategies to resolve an observed divergence in order to "get on the same page" (i.e. reach cognitive convergence).

We utilize a hybrid model to describe the components of a computational shared mental model which incorporates fuzzy logic that allows reasoning about degrees of truth rather than exact knowledge. This logical representation has mapped well to Lakoff's prototype theory [10], which describes a view on how humans represent concepts in the world, and the way we have seen improvisers describe their own characters in a scene [12]. We represent knowledge about the story domain (e.g. the association between characters and joint activities) as degrees of association (DoA), which are essentially bi-directional fuzzy memberships (to know more about how DoA is utilized in this work, please see [13]). By representing domain knowledge this way, agents can compute the *iconicity* of memberships. Agents do not always want to present an element (e.g. a character) that has the highest DoA to the other elements in its mental model as there might also be iconic elements that hold low DoA values (e.g. a *monk* has a higly iconic low DoA with *talking*). Rather, the agent should present something that is iconic that conveys a lot of information about the agent's mental model. When an agent makes an iconic presentation, another improviser is less likely to misinterpret the presentation, which reduces cognitive divergences and aids in repairing the shared mental model. For example, only a few characters in the Old West, like an outlaw, are highly associated with robbing a bank. Therefore, robbing a bank is an iconic activity for an outlaw.

Iconicity calculations are similar to those used in [9] for *Party Quirks*, but these are updated to account for the different element categories of the knowledge structure: actions, characters, and joint activities. An agent can use the iconicities of the elements in its mental model to determine the probability that a certain element will be selected from its knowledge structure category given its current mental model. The necessity of probabilistic selection is raised to model the unpredictable nature of human behavior in interpreting context in a scene. Moreover, probability helps the agent to estimate the level of confidence it might have in its mental model.

4 Application of Fuzzy Rules

Improv actors execute motions on stage that are ambiguous. Even a simple motion like waving a hand can be interpreted multiple ways (e.g. waving to someone, erasing a board, or cleaning a window). Improv agents need to entail knowledge and future actions after observing these ambiguous presentations from a human (via Kinect) or other AI agent. To handle this ambiguity, agents need to measure their confidence in their entailed knowledge because it may be based on assumptions that diverge from others' mental models.

Fuzzy rules (a procedural representation in the form of "*if… then…*" rules for handling the ambiguous kinds of knowledge we have seen in our observations of improvisers) are designed based on the way humans reason about the likeliness of a (random) event to occur that is: the higher the probability of an event, the more certain we are that the event will occur. However, even if this is the case other elements should be considered in order to come up with a confidence value that really reflects the whole situation, such as iconicity in this work. For example, the agent would use the following fuzzy rules to determine its confidence in the character it thinks another agent is portraying:

Fuzzy Rule One: *If* $agent_B$ thinks that $action_1$ done by $agent_A$ has medium probability to occur and has low iconicity to $character_X$ portrayed by $agent_B$, *then* $agent_B$'s confidence in $agent_A$ portraying $character_X$ is low.
Fuzzy Rule Two: *If* $agent_B$ thinks that $action_2$ done by $agent_A$ has high probability to occur and medium iconicity to $Character_Z$ portrayed by $agent_B$, *then* $agent_B$'s confidence in $agent_A$ portraying $character_Z$ is medium.

For this purpose, we use Trapezoidal and Triangular Membership Functions that use three fuzzy values (low, medium, high), which provides high quality results when compared to other membership functions, see Fig. 1. The x axis represents the inputs of the probabilistic or iconicity values φ. The y axis represents the degree of membership μ of element φ in the fuzzy sets *low*, *medium*, and *high*, where each term in μ(probability) is characterized by a fuzzy set in a universe of discourse U=[0, 1].

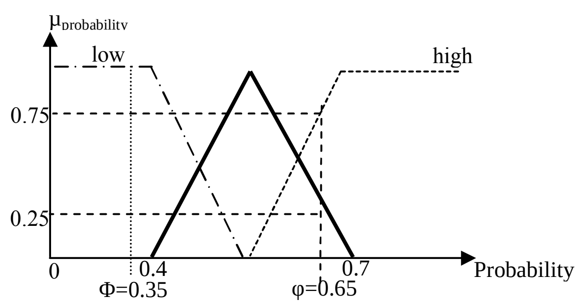

Fig. 1. Diagrammatic representation of fuzzy probabilities using Trapezoidal and Triangular Membership Functions. The low, medium, and high fuzzy values are used to determine the agent's confidence based on the probability and iconicity inputs..

In order to illustrate the computation of the confidence factor using fuzzy rules 1 and 2, consider a scene set in the Old West where $agent_B$ believes that $agent_A$ is presenting the joint activity *apprehending a criminal*. $Agent_B$ wants to extrapolate this knowledge to learn what character $agent_A$ might be portraying. $Agent_B$ considers that $agent_A$ might be portraying the character *sheriff* or the character *outlaw*. Assume <*apprehending a criminal, sheriff*> has *medium iconicity* = 0.4 *and* <*apprehending a criminal, outlaw*> has *low iconicity* = 0.22. Based on the iconicities of those characters with the joint activity *apprehending a criminal*, the probabilistic values for $agent_A$ portraying a *sheriff and an outlaw are* φ=0.65 and φ=0.35 consecutively. These values will act as the inputs to the Trapezoidal and Triangular Membership Functions to compute their membership to the fuzzy sets: low, medium, and high. It is worth noting that we are using the same membership function shown in Fig. 1 for both iconcity and probability fuzzy variables. In order to compute the certainty factors for the portrayed characters, the agent need to apply the following three steps for Fuzzy Rule One and Fuzzy Rule Two:

Step1: Fuzzify inputs: Resolve all fuzzy statements in the antecedent to a degree of membership between 0 and 1.

Fig. 1 shows that the probabilistic value 0.65 cuts the medium and high fuzzy sets (y axis) in 0.25 and 0.75 respectively. This means that the probability of being a *sheriff* character has the following degrees of membership: $\mu_{low}(0.65)=0$, $\mu_{medium}(0.65)=0.25$ and $\mu_{high}(0.65)=0.75$. Similarly, the *outlaw* character has the following degrees of membership: $\mu_{low}(0.35)=1$ and $\mu_{medium}(0.35)=0$, $\mu_{high}(0.35)=0$.

Repeat Step 1 for the iconicity values 0.4 and 0.22, which provides the following results: $\mu_{low}(0.4)=1$, $\mu_{medium}(0.4)=0$, $\mu_{high}(0.4)=0$ for the *sheriff* character, and $\mu_{low}(0.22)=1$, $\mu_{medium}(0.22)=0$ and $\mu_{high}(0.22)=0$ for the *outlaw* character.

Step 2: Apply fuzzy operators to multiple part antecedents: The "fuzzy and" operator is the minimum of the degree of memberships in the antecedents, while the "fuzzy or" operator is the maximum of the degree of memberships in the antecedents.

Applying this step on the antecedent part of Fuzzy Rule One, we will find that the *sheriff* character has $\mu_{low}(0.4)=1$ for iconicity and $\mu_{medium}(0.65)=0.25$ for probability. Next, apply the "fuzzy and" operator to these results; $CF_{rule1}= \min\{1, 0.25\} = 0.25$.

Repeat Step 2 for Fuzzy Rule Two. We will find that the *sheriff* character has $\mu_{medium}(0.4)=0$ for iconicity and $\mu_{high}(0.65)=0.75$ for probability. Again, apply the "fuzzy and" operator to these results; $CF_{rule2}= \min\{0, 0.75\} = 0$.

Step 3: Defuzzify outputs: For a group of rules, defuzzify the outputs by aggregating all the rules' outputs to produce one 'crisp' value using the Centroid Defuzzification Method. Final crisp value for a group of rules = $\sum_{i=1}^{n} m_i w_i \ / \ \sum_{i=1}^{n} m_i$, where m_i is the membership of the output of each rule, and w_i is the centre of gravity of each fuzzy value area. Applying the Centroid Defuzzification Method to the values obtained from steps 1 and 2, we obtain: $CF_{(rule1and\ rule2)}= (0.25*0.4+0*0.55) / (0.25+0) = 0.4$

Now the agent can use this confidence factor in the representation of knowledge in his mental model as shown in the following form of logical predicate:

```
bel(agentB, is_a(agentA, sheriff), medium, 0.4)
```

In plain English, this predicate can be read as: "$agent_B$ believes that $agent_A$ is a *sheriff* with medium confidence 0.4" It is worth of noting that in the transformation process of probability to confidence, iconicity acts as an adjustment factor. The same procedure would be followed to represent a shared belief about the outlaw character as shown below:

```
mutual_belief (agentB, bel(agentA, is_a(agentB, cowboy),
high, 0.7))
```

In plain English, this predicate can be read as: "$Agent_B$ has a shared belief that $agent_A$ believes that $agent_B$ is a *cowboy* with high confidence 0.7" Fuzzy rules are also needed to update the agent's confidence about his beliefs after each interaction with the other agents. An example of these fuzzy rules is shown below:

Fuzzy Rule Three: *If* $agent_B$ has <u>low</u> confidence about $agent_A$'s motion *and* has <u>medium</u> confidence about $agent_A$ as $character_Y$, *Then* <u>decrease</u> $agent_B$'s confidence in $character_Y$.
Fuzzy Rule Four: *If* $agent_B$ has <u>high</u> confidence about $agent_A$'s motion *and* has <u>medium</u> confidence about $agent_A$ as $character_X$, *Then* <u>increase</u> agent B confidence in $character_X$.

The computational implantation of this rule is achieved via using the *serial combination function (*$SCF= CF_1*CF_2$*) and* the *parallel combination function (*$PCL= CF_1+CF_2-(CF_1*CF_2)$*)* for rule 3 and rule 4 respectively, where CF_1 and CF_2 are the certainty factors for the first and second statement in the antecedents part of the rule.

The agents' confidence about the knowledge in their mental models keeps changing based on the actions they take. Reaching cognitive consensus requires the agents to understand each other. In fact, shared mental models can be measured in terms of the degree of overlap or consistency among team members' knowledge and beliefs [11].

5 Discussion

This paper presents a computational shared mental model for improv agents based on preliminary modeling efforts and studies of human improvisers. The goal of our model is to a) formally represent of our findings of how human improvisers negotiate shared mental models and b) support intelligent agents's ability to improvise scenes with each other or with a human interactor. This model provides the flexibility for improv agents to infer and extrapolate to new knowledge from their interaction based on their current shared mental models. Improv agents can be employed in games environments where they can reason about fuzzy uncertain knowledge and interact with human without relying privileged knowledge or communication. The fuzzy rules can be applied to other domains and can also be edited to include any number of factors (evidences) that might affect the generated confidence. In its current state, this approach does not capture all of the complexities of a full theory of mind because assessments are only based on degrees of association. For example, agents with the kind of shared mental models described here cannot reason about privileged knowledge (e.g. knowledge that only one agent knows because the other agent was not present to hear it), which a full theory of mind can account for.

References

1. Hayes-Roth, B., and Van Gent, R: Story-Making with Improvisational Puppets and Actors. In Technical Report KSL-96-09. Palo Alto, CA: Stanford University (1996).

2. Swartjes, I., Vromen, J.: Emergent Story Generation: Lessons from Improvisational Theater. In: Proc. of AAAI Fall Symposium on Intelligent Narrative Technologies, Arlington, VA (2007).

3. Zhu, J., Ingraham, K. E. and Moshell, J. M.: Back-Leading through Character Status in Interactive Storytelling. In: Proc. of ICIDS 2011, p. 31-36 (2011).

4.Fuller, D. and Magerko, B.: Shared Mental Models in Improvisational Theatre. In: Proc. of the 8th ACM Conference on Creativity and Cognition, p. 269-278. ACM Press (2011).

5. Magerko, B., Dohogne, P. and Fuller, D.: Shared Mental Models in Improvisational Digital Characters. In: Proc. of the 2nd International Conference on Computational Creativity, p. 33-35 (2011).

6. Si, M., Marsella, S. C. and Pynadath, D. V.: THESPIAN: An architecture for interactive pedagogical drama. In: Proc. of the Twelfth International Conference on Artificial Intelligence in Education. p. 595-602. IOS Press (2005).

7. Van De Kieft, I., Jonker, C. and Van Riemsdijk, M. B.: Improving User and Decision Support System Teamwork: An Approach Based on Shared Mental Models. In: Proc. of the 22nd IJCAI Workshop on Explanation-Aware Computing, p. 61-70 (2011).

8. Brisson, A., Magerko, B., Paiva, A.: A Computational Model for Finding the Tilt in an Improvised Scene. In: Proc. of the 4th International Conference on Interactive Digital Storytelling, Vancouver, Canada (2011).

9. Magerko, B., DeLeon, C., and Dohogne, P.: Digital Improvisational Theatre: Party Quirks. In: Proc. of the 10th International Conference on Intelligent Virtual Agents, p. 42–47 (2011).

10. Lakoff, G.: Women, Fire, and Dangerous Things. Chicago: Univ. of Chicago Press (1987).

11. Cannon-Bowers, J. A., Salas, E., and Converse, S.A.: Shared mental models in expert team decision making In N. J. Castellan, Jr. (ed.), Current issues in individual and group decision making, p. 221-246. Hillsdale, NJ: Lawrence (1993).

12. Magerko, B., Manzoul, W., Riedl, M., Baumer, A., Fuller, D., Luther, K. and Pearce, C.: An Empirical Study of Cognition and Theatrical Improvisation. In: Proc. of ACM Conference on Creativity and Cognition, Berekely, CA (2009).

13. O'Neill, B., Piplica, A., Fuller, D., and Magerko, B.: A Knowledge-Based Framework for the Collaborative Improvisation of Scene Introductions. In: Proc. of the 4th International Conference on Interactive Digital Storytelling, Vancouver, Canada (2011).