Proceedings of the 2022 Winter Simulation Conference
B. Feng, G. Pedrielli, Y. Peng, S. Shashaani, E. Song, C.G. Corlu, L.H. Lee, E.P. Chew, T. Roeder, and P. Lendermann, eds.

FROM NARRATIVES TO CONCEPTUAL MODELS VIA NATURAL LANGUAGE PROCESSING

David Shuttleworth

Department of Computational Modeling
and Simulation Engineering
Old Dominion University
5115 Hampton Blvd
Norfolk, VA 23529, USA

Jose Padilla

Virginia Modeling, Analysis and Simulation Center
Old Dominion University
1030 University Blvd
Suffolk, VA 23425, USA

ABSTRACT

This paper explores the use of natural language processing (NLP) towards the semi-automatic generation of conceptual models, and eventual simulation specifications, from descriptions of a phenomenon. Narratives describing the problem are transformed into a list of concepts and relationships and visualized using a network graph. The process relies on pattern-based grammatical rules and an NLP dependency parser identifying important concept types, namely actors, factors, and mechanisms. We use three conceptualizations, created by potential users, to understand how the NLP-generated model should and could be adjusted. The objective of the research is to develop potential standard approaches users can use to generate conceptual models; develop a conceptual modeling assistant that subject matter experts can use to make them participant in the simulation creation process; and to identify how narratives should be written so an NLP-based conceptual modeling assistant may provide a thorough description of a phenomenon.

1 INTRODUCTION

Conceptual modeling is an important aspect in simulation creation (Robinson 2008a) yet there is little agreement on defining the conceptual modeling process and its corresponding outcome: conceptual models. Pace (2000) argues that "much remains to be done before simulation conceptual models can be routinely developed with the rigor, precision, and formalism needed to significantly increase their correctness." The ad hoc nature of conceptual modeling prevents repetition of the process and replication of conceptual models. Additionally, an expression of the problem situation, and the transition from an informal description of understanding to a more formal model, is typically hidden from view in conceptual modeling research and journals (Padilla et al. 2019). These characteristics are indicative of a primary challenge in Modeling and Simulation (M&S): the inability for the non-simulationist and domain experts to routinely and repeatably develop conceptual models. Conceptual models often take the form of simulation specifications driven by specific paradigms, which is useful to the simulation developer, but does not facilitate communicating with subject matter experts excluding them from the simulation creation process.

Little has been done to formulate methods to consistently develop, communicate, and replicate conceptual models as there is not an agreement of what conceptual models should contain and how they should be represented. Further, conceptual models may take different forms depending on the worldviews of Subject Matter Experts (SME) and modelers. As such, they may capture information about a person's assumptions and presuppositions that are not made explicit. In fact, methodological, ontological, and epistemological misalignments between SMEs and modelers limit communication as each group is driven

978-1-6654-7661-4/22/$31.00 ©2022 IEEE

by its own idea on how to study, how to see the world, and what is considered knowledge. As a result, these differences lead to untrustworthy models and worse the failure to communicate and the undermining of trust. Lastly, without robust methods for conceptual model development, novice modelers are left with a variety of approaches and techniques often based on modeling paradigms that fail to communicate across paradigms. Instead of capturing the best representation of a phenomenon, the focus shifts to the "best" implementation of the phenomenon based on preferred modeling paradigms that expedite the creation of a simulation.

It is posited that a conceptual modeling method supporting the transition from simple descriptions of a problem to conceptual models will enable novice modelers and/or SMEs to control the modeling process. In this research, we operate over a notional narrative, which we presume would be developed by modeling stakeholder(s), describing a problem situation and then use NLP to give us an insight into how to develop conceptual models and how to transition them towards simulation specifications. The following are working definitions that drive this work:

- Conceptualization: the process of capturing concepts and relationships of observations or imaginations in the form of artifacts.
- Conceptual Model: a conceptualization that characterizes a problem of interest towards answering a modeling question via a model.
- Conceptual Modeling: the process of creating conceptual models.

2 BACKGROUND

Conceptual modeling has long subscribed to a variety of formats for communicating models, which includes modeling languages, simple descriptions, and diagrams. Modeling languages, based on strict standards, represent a means to formally and unambiguously model a system or problem, but they are relegated to experienced users and move conceptual modeling closer to implementation. Diagrams such as concept or mind maps, however, are easier to use and are closer to modeling one's own internal view of the problem of interest. Simple descriptions add additional semantic depth. This section first provides background on linguistics that provides insight into how humans have communicated for millennia through art, diagrams, and language followed by a review of formal modeling languages.

2.1 Language and Symbols

Linguistics, the study of human language, provides insight into how to semantically analyze narratives and communicate among individuals. This is especially important among a multi-disciplinary modeling team where not only the individual worldviews introduce ambiguity in perceptions of the problem but also the use of language and its interpretation. As language is composed of symbols, grammar provides a structure to convey verbal and non-verbal *concepts* grounded on our experienced realities or imagined thoughts.

According to Lizardo (2013), the process of conceptualization is the "deployment of conceptual resources that are grounded in our embodied, experiential reality" and forming an "image-schematic" structure. These structures are not formed as visual images, but rather abstract cognitive units that may be grounded by our ways of expression.

The earliest use of symbols to tell a story resides in ancient cave paintings. Not only did cave paintings provide artistic expression, but they were stories to document experiences (Aubert et al. 2019). It is their expressed view of reality for the purpose of relating a life experience or teaching other individuals about the world around them. As such, these paintings can be considered an early form of a model; a conceptualization that captures a reality in a visual form. Cave art, for instance, often contained just the outline of an animal, as shown in Figure 1, with exaggerated features to perhaps draw the eye to the important attributes the storyteller wishes to convey. Researchers suspect these artifacts are fundamentally tied to neural processing in that it is more important to recognize the shape of something versus its color and texture as a means of rapid categorization even without a complete image (Hodgson 2013). Cave art and its analysis are consistent with modeling practices of simplification and identification of important factors, which are key to conceptual modeling.

Figure 1: 30,000-year-old cave painting (Hodgson 2013).

Later, more complex Egyptian hieroglyphs are the earliest form of narratives that were communicated through glyphs, where each glyph represents the equivalent of multiple sentences in modern language. Today, narratives use a different set of symbols and grammar (rules) that vary per language. In both cases, the relationship between thoughts, symbols, and the referent can be visualized by the semiotic triangle shown in Figure 2. Thoughts about a physical object in the real world, or referent, are externalized and expressed as a set of symbols, and the symbols represent the referent. The symbols themselves are the mode of expression (art, verbal stories, written language, etc.) and are a reflection of the real or theorized objects. However, symbols "help us or hinder us in reflecting on things" (Ogden and Richards 1923). The symbols themselves can influence the meaning of an object in reality as symbols are intrinsic to the system and are privately grounded (Schulz et al. 2011), and therefore, they require interpretation to be understood by others. This can only be correctly achieved given a shared understanding.

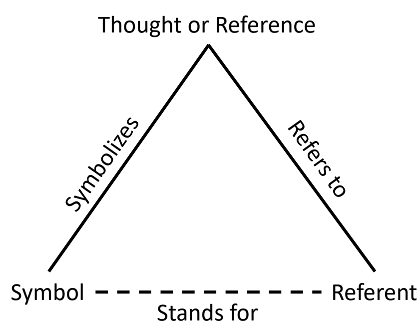

Figure 2: Semiotic triangle, modified from Ogden and Richards (1923).

Semantically rich descriptions provide a means to capture an in-depth view of the problem, but language is ambiguous and open to interpretation. Situations characterized by partial information, knowledge depth, and level of experience, introduce an additional challenge in expressing consistent and complete descriptions. However, glyphs and art represent a natural form of expression that seem to transcend modern language-based descriptions. When combined with precise, unambiguous descriptions, it is posited that challenges associated with misinterpretations will be alleviated.

2.2 Modeling Languages

There are a variety of formats in which to formally convey a perspective of some observed or theorized problem. Language, for example, is a convenient and natural approach to describing a perspective. Whether via speaking or writing, language provides rich media for which to communicate, but it is also ambiguous and open to interpretation. Within the M&S community, conceptual models are used to convey a representation of the problem and custom languages based on a standard have been adopted by some to express conceptual models as a means to bring formalism to the conceptual modeling outcome.

For example, some, borrowing from computer science and systems engineering (SE), have adopted the Unified Modeling Language (UML) and its extension, the Systems Modeling Language (SysML), for describing conceptual models (Bersini 2012; Reinhartz-Berger 2005; Siebers and Onggo 2014; Sha et al. 2011). However, these were not intended for conceptual modeling and are difficult to understand for those who are not modelers. These languages were developed in support of *software specifications* and *designing systems*, respectively, and the visual artifacts have been repurposed to represent conceptual model elements. For example, class diagrams are used to collect, organize, and visualize model elements and their attributes. Activity diagrams represent a logical flow of processes, and state diagrams visualize state transitions. While aided with graphical modeling tools, these languages subscribe to strict specifications that are cumbersome to learn. Furthermore, they are derived from computer science and software engineering with the intent to design a set of software or system specifications not to support the identification, collection, and modeling of elements related to some phenomenon.

Lastly, ontology represents another format researchers (Verdonck et al. 2015) and practitioners (Guizzardi and Wagner 2012) suggest as a means to express conceptual models. Ontology provides a graphical depiction to develop models from the perspective of collecting a hierarchy of and relations among relevant elements but is meant for knowledge representation and reasoning over a knowledge base. From a knowledge management perspective, Carvalho et al. (2017) further emphasizes the ontological approach in asserting conceptual models are “a specification aiming at representing a conceptualization of the subject domain of interest.” However, while relatively accessible to non-modelers, this perspective moves conceptual models away from a conceptualization towards specifications. However, UML, SysML, and the Web Ontology Language are all implementations that are not necessarily prescriptive in moving towards a particular software nor modeling paradigm.

3 CONCEPTUALIZATIONS ON A COMMON NARRATIVE

This section presents how three individuals, from various backgrounds, would construct a graphical representation provided a narrative. The request centered around a diagram so as to not confuse the individual with definitions. The perspectives are those of a systems engineer, a doctoral student in social sciences, and a non-engineer/non-modeler.

3.1 Narrative

The mode of transportation to school for elementary school kids has shifted due to the recent pandemic and related factors. Elementary students have three transportation choices, which include riding the bus, walking, or being dropped off by a parent. The recent pandemic has caused many parents to drive their kids to school for three reasons. The first is the lack of available bus drivers. Bus drivers are not employed by the school as they are contracted employees. Many of them were not retained through the pandemic, or they sought new employment. The lack of bus drivers also caused many of the drivers to perform multiple routes causing route delays. Since some parents need a consistent schedule, they decided to drive their kids to school.

The second reason is the lack of mask enforcement on buses. New bus drivers are not familiar with supervising kids on the bus. Additionally, the school policy was initially unclear. Parents became *uncomfortable with their kids riding on a bus that did not enforce masking. Therefore, during the peak of the pandemic, about a third of the students were arriving via car.*

Weather impacts whether a child walks or rides with a parent to school. On days with no inclement weather, the parents will often let their child walk to school. Parents typically choose to drive their child on rainy and extremely cold days. Parents can expect longer car lines and increased traffic on these days. Cars are directed to line up along the street in front of the school. School assistants direct a block of 10 cars to the drop off zone so that children are safely unloaded. The more cars that are lined up waiting to proceed to the drop off zone can exceed 0.5 miles.

Delays to the start of the school day are exacerbated on days of inclement weather and bus route delays.

3.2 Systems Engineer

The systems engineer presents their perspective of important elements and relationships, which is shown in Figure 3. Walking and riding to school are related to the weather and the positivity rate of illness. Busing is also related to the positivity rate and to the bus schedule. Weather and positivity rate influences parents, and parents and the number of bus drivers influence the bus schedule. This perspective presents an interpretation of the narrative based on causality.

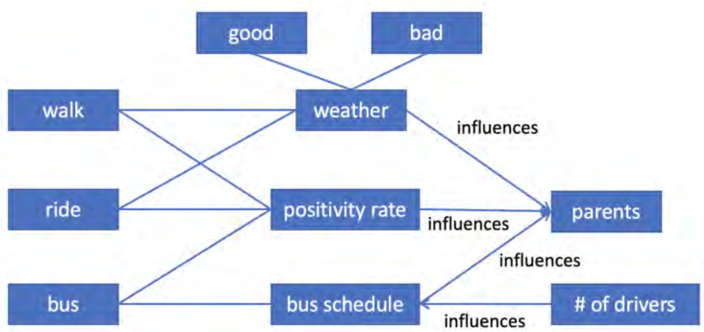

Figure 3: Perspective of the narrative by a systems engineer.

From Figure 3, the systems engineer suggests, "Weather, bus schedule, and positivity rate all influence parents' decision to walk, ride, or bus their children to school. Positivity rate and the bus schedule drives whether the children take the bus. The number of available drivers influences the bus schedule. If the decision is to walk or ride, weather plays a factor as well." This position suggests the notion of causality among the elements and the importance of the actions (walk, ride, bus - as bus riding), characterizing major elements of transportation mode, and the modifiers of weather (good/bad) for triggering a decision. Much of the information in the narrative is abstracted away. The rest of the diagram flows from these elements in a pseudo-hierarchical structure in the same sense as a relationship diagram denoting associations and directed dependencies. It is noted that *child/kid* and *driver*, individually, are not part of the diagram whereas *# of drivers* is part of the diagram as to capture the number but not the actor per se. We infer that children are viewed as objects that flow through the system as if viewed as entities or tokens.

3.3 Social Sciences – Doctoral Student

Figure 4 contains two major sections. The top section contains the actions resulting from a decision and the implications of that action. For instance, parents, are described as reliable, but driving consumes the most time for the parent. The lower portion represents a decision process in the form of a decision tree, which takes the shape of a mind map with visuals of some of the primary characters. It is a linear decision tree starting with the three modes of transportation. Elements are connected with 'yes/no' responses to questions based on conditions that denote decisions are being made. Each thread results in a decision on whether to continue using a particular mode of transportation of switch to another.

An observation is the inclusion of what is presumed to be a self-evident perspective from the modeler such as the distance to the school, children's safety, after school activities, and carpooling. These elements are not contained within the narrative but are interpretations important to the reader. This information is a valuable insight into how other perspectives bring their own worldviews to a problem situation.

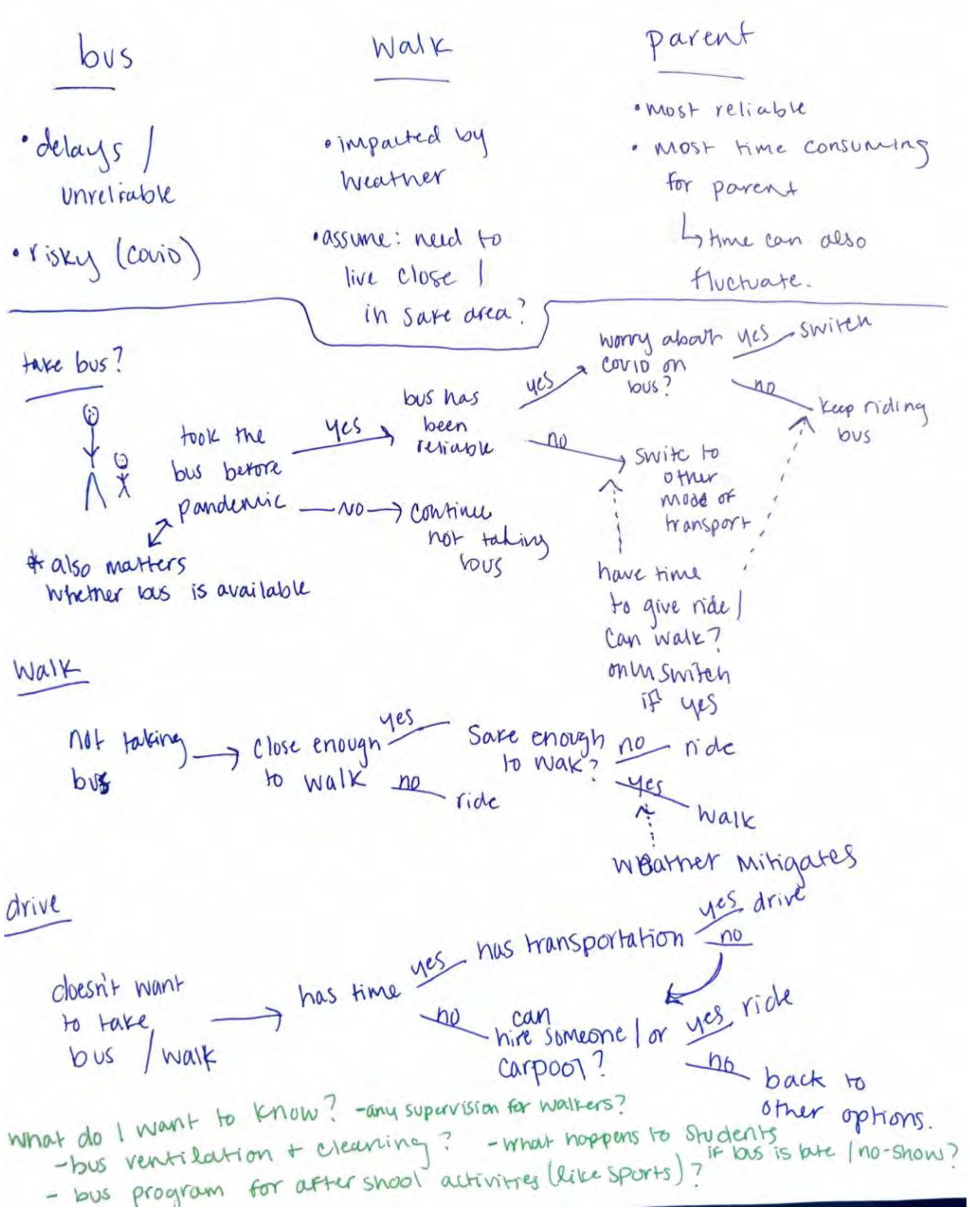

Figure 4: Perspective of the narrative by a doctoral student in social sciences.

3.4 Non-Scientist/Non-Modeler

The non-scientist diagram shown in Figure 5 is simple where many elements and relationships are removed as the resulting diagram contains just the relevant elements and relationships that were viewed as important to the perceived problem situation. In this case, the major elements are limited to the mode of transportation and scheduling.

The respondent categorizes the diagram using pre- and current post-pandemic conditions and then categorized by modes of transportation: busing, driving, and walking. In the pre-pandemic condition, three transportation nodes lead to a central node referring to a consistent schedule for teachers, parents, and children. Referring to the second major category, additional detail is added to the nodes in addition to two elements related to parent being late for work and children arriving too early for class. An inference is made referring to less school time and walking to school may lead to arriving too early for class. Like the previous diagrams, the characterization of the actions (bus ride, parent ride, and walk) drive the diagram.

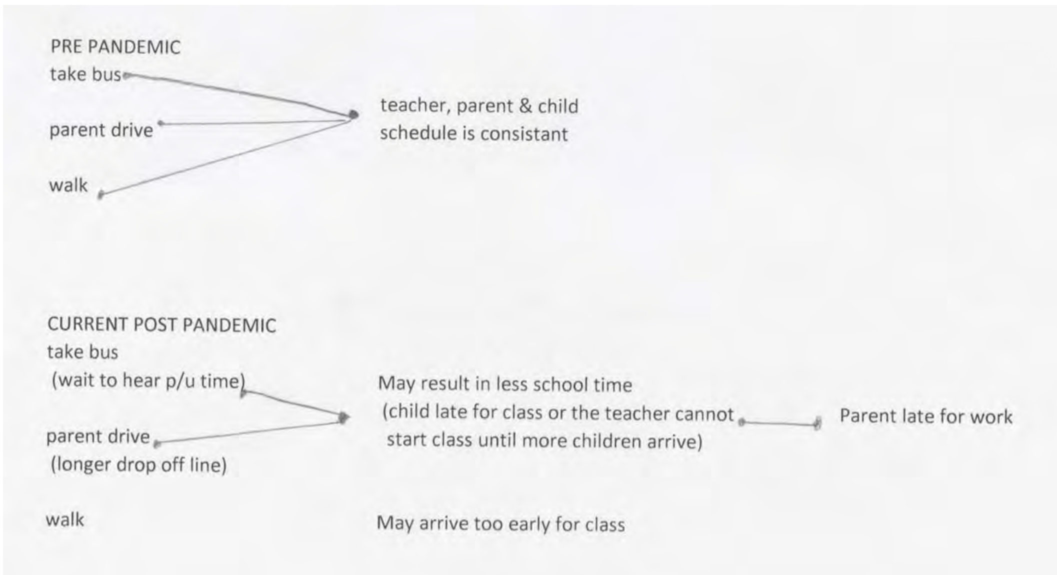

Figure 5: Perspective of the narrative by a non-scientist/non-modeler.

3.5 Comparison of the Three Perspectives

The three perspectives provide valuable insight into how personal experience and worldviews influence an interpretation of a narrative and resulting conceptualization. Interestingly, the diagrams are structurally similar in that there is a general trend of moving left to right forming a hierarchical progression, and in general, there is consensus on the importance of including the modes of transportation. It is noted that the individuals live in the USA and speak English. The SE uses interconnected boxes in a manner similar to SysML of associations and dependencies. One can infer dependencies from the simple associations. For example, there is some denotation that walking to school is impacted by weather and positivity rate. There also exist more direct dependencies as the link is labeled by *influences*, which is equivalent to explicitly stating dependencies such as the number of bus drivers influencing the bus schedule. We can also infer a causality here that fewer bus drivers will result in bus schedule delays.

The doctoral student includes a decision flow based on conditions vice associations and dependencies as is the case in the SE view. This perspective also includes a simple drawing representing (seems to) a parent and child adding personification to the diagram. This subtle touch highlights the importance of using images in addition to words. The use of images is consistent with Hodgson (2013) proposition mentioned before: the recognition of shapes as a means of rapid characterization.

The non-scientist participant wanted to draw an image of a bus along with children, but the act of drawing a complex scene may have prevented her from doing so. This desire along with the image of the student provides an indication that some individuals prefer to model a story through images and words, not just a series of abstract blocks and connecting arrows. The non-scientist also adds a layer of categorization by two states: pre- and current/post-pandemic states, but then includes the mode of transportation categorization as do the other perspectives. The pre-pandemic diagram represents a steady state or status quo as if wishing to compare to the current pandemic state. As with the SE view, there is the notion of causality that certain transportation modes lead to a particular condition, such as the parent being late to work due to delays in the bus schedule. A subtle observation here is that the bus being late also impacts the length of the school day from which we can infer that the respondent has an interest in making sure the children's education is not impacted by the current conditions.

3.6 Towards a Conceptual Model

To move this conceptualization towards a conceptual model, it would require a modeling question. For this research, the question is not included but is part of future work, and we expect the current conceptualization would change as the relationship among concepts changes towards a potential model that helps answer the question. Additional concepts and relationships would likely be included to complete a conceptual model.

3.7 Towards a Specification

As modelers, we recognize ways of implementing a conceptualization. The types of relationships in the SE view lend themselves to specifying some causality between elements. In the doctoral student's view, the sequence of steps leading to a decision on transportation mode lends itself to an Agent-Based Model (ABM) given the notion of conditions, or triggers that lead to a preferred decision, or a rule that an agent is designed to follow. However, and as discussed, the consideration of implementations, at this stage, is premature and both advantageous and disadvantageous. As we posit, it is advantageous in expediting the simulation creation process, but disadvantageous as the conceptual modeling process is often overlooked.

4 A SEMI-AUTOMATED APPROACH

This prototype is part of a larger framework that is being designed to support a semi-autonomous approach where the integrated workflow between human descriptions and queries and computer-based algorithms works in unison to develop conceptual models. This paper, however, is primarily concerned with 1) the NLP prototype that identifies concepts and categorizes them as either actors, factors, or mechanisms to automatically generate conceptualizations and 2) establishing how the process can be improved considering user input. The objective is to use these NLP-derived conceptualizations to generate both conceptual models and to facilitate the generation of model specifications.

NLP techniques are used in this research to identify candidate concepts and relationships. The spaCy Python package (Honnibal et al. 2020) is used to perform all NLP-based functions. The narrative is first pre-processed removing stop words and lemmatizing the tokens. Then each token is tagged using a Part of Speech Tagger (POST), and relationships are then constructed based on inter-sentence relationships between tokens using spaCy's dependency parser pipeline. Through the combination of lemmatization and relationship identification, additional relationships are uncovered across sentences, which assumes that each lemmatized token refers to the same object in reality. This approach of combining elements within the graph based on root words is a low level co-referencing mechanism that facilitates the creation of more complex graph structures enabling more detailed graph analysis. We use NLP-based dependency grammar, and a small set of pattern-based rules, to identify candidate relationships.

4.1 Definitions

Concepts represent a layer of abstraction requiring additional detail. Concepts may be user defined or domain specific, but in this research, the following definitions are provided for actors, factors, and mechanisms. These concepts are identified based on an individual word's tagged part of speech (POS).

- Concept Definition 1: An actor is a character or object in the narrative. Actors are identified using the POS subjects and objects.
- Concept Definition 2: A factor modifies an actor or mechanism. Factors are adjectives or adverbs that provide an additional source of qualitative or quantitative data.
- Concept Definition 3: A mechanism is a function or process performed by an actor and is identified as the root or other verbs within a sentence.

The following rules are based on an NLP dependency parsing based on patterns of identifying relationships among tokens. These patterns are based on a grammatical structure and each token's POS of each sentence such as an actor (subject) using a mechanism (verb) to modify another actor (object), or an actor (noun) modified by a factor (adjective).

- Pattern Definition 1: The actor-mechanism-actor relationship is satisfied given both subject and object are children of a common root.
- Pattern Definition 2: The actor-factor relationship is solidified given either a subject or object has an adjective child.
- Pattern Definition 3: The modified mechanism relationship is satisfied given a verb has an adverb child.
- Pattern Definition 4: The compound noun relationship is satisfied given a noun has another noun child tagged as a compound.

These pattern-based rules represent an initial set to verify the feasibility of the approach. Further refinement and additional rules would extend the capability to account for additional patterns.

4.2 Visualizing the Set of Extracted Concepts and Relationships

The above rules were implemented in Python using the spaCy dependency parser, and the results are shown in Figure 6, which contains actors, factors, and mechanisms. The concepts and relationships are stored in a Python dataframe and visualized as a network graph. As such, the algorithm generates a conceptualization containing a collection of concepts and relationships.

Given the algorithm extracts all candidate concepts and relationships based on the ruleset, it provides many more elements than the manual diagrams as there is no human with an individual worldview influencing how one interprets which concepts and elements are important. Still, a user can eliminate nodes that are not relevant, and inferences and conclusions can be made regarding the narrative. For example, parents can expect long lines on rainy days; bus drivers have multiple routes and delays; and the schools have unclear policies. These strings of nodes effectively recreate portions of the narrative. In other words, this diagram reduces the narrative down to short simplified sub-sentences.

On the other hand, there are several disconnected clusters of nodes such as *consistent schedule* and *mask enforcement*. These small clusters should have been associated with parent and bus nodes, respectively, which indicates a shortfall in the ruleset. Whereas the manual diagrams indicated information flows, logical processes, and causality, the automated diagram does not include relation types. However, if two nodes are connected by a mechanism node, a relationship can be inferred across the connected nodes. For example, *parents can expect long lines* and *elementary students have a transportation choice*. These example phrases efficiently and directly state the intent of a sentence without superlative language and passive grammar, which leads to misinterpretation. The current algorithm also demonstrates the need for retaining negative words such as not. The node labeled *familiar* is a clear contradiction to the narrative where the term was *not familiar*. Additional rulesets would easily rectify these types of contradictions.

A simple graph analysis illustrates how a few nodes, such as *driver*, *parent*, and *school*, may be accentuated to how relative importance within the network. This relatively highly connected nodes are shown as larger nodes based on their measure of centrality, which is useful in visualizing which nodes may be of higher importance.

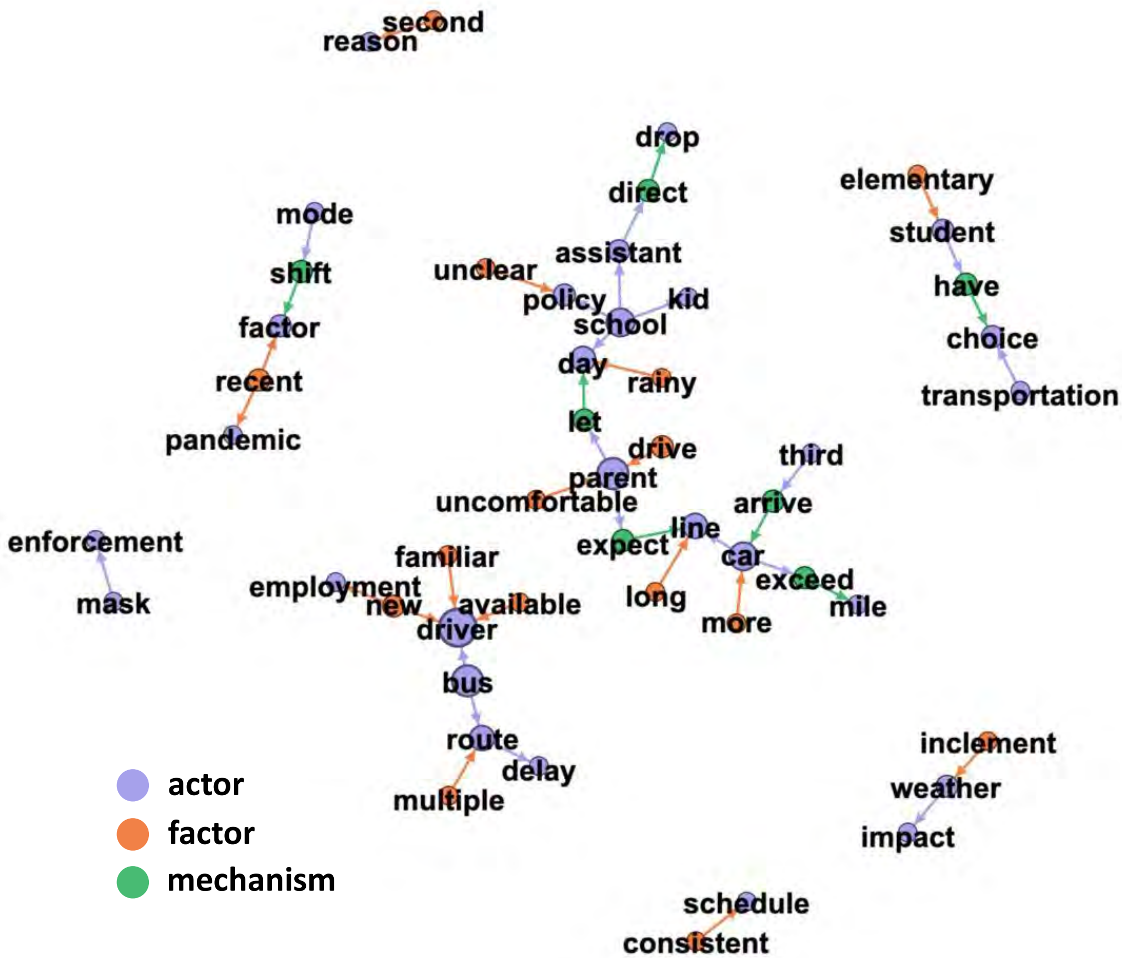

Figure 6: Diagram from NLP-based algorithm.

5 DISCUSSION AND FUTURE WORK

The result of the algorithm, combined with the three conceptualizations, provides a path for exploring conceptual modeling options when it comes to manual and semi-automatic generation of conceptualizations. Firstly, and solely relying on Figure 6, a user could rely on the different groupings or network centrality measures, to identify concepts that are more important than others facilitating the prioritization of their consideration. One of the major challenges novice modelers have been that all concepts are equally important limiting which ones to include in a model or prioritized in their descriptions. Secondly, if we consider the reverse process, we could establish a template to guide non-modelers to provide a rich description of a phenomenon/system: who the actors are, what they do, how they impact, or are impacted by the environment, what is impacted, etc. This is a process that modelers engage with SMEs across different domains, yet there is no standard template (the authors are aware of) that asks these questions methodically in the M&S community or in specific domains. This template could facilitate the recognition by NLP algorithms that can then generate artifacts such as Figure 6. Thirdly, identified concepts within a network, paired with images, emojis, or icons, may provide both structure and pictorial aids that facilitate the recognition and categorization of concepts, as proposed by Hodgson (2013). This, again, can be done manually, but ideally, semi-automatically.

The three perspectives provide multiple conceptualizations of the problem, which is entirely plausible in a realworld setting where multiple participants develop individual narratives of the problem. A heuristic method, such as the one discussed by Weisel (2004), presents one approach to aggregating elements from multiple models into a single, useful model. Domain knowledge in the form of ontology presents another avenue to reconciling multiple conceptualizations. If an ontology is presented as universal knowledge about a particular subject, then it may be used to compare and contrast conceptualizations.

With these considerations in mind, Figure 6 can take different forms. Figure 7, for instance, captures an idealized version that is envisioned to be semi-autonomously derived. Additional related research is addressing how to formulate an integrated model containing imagery and descriptions as a mechanism to develop and communicate models. It contains many of the elements shown in the three perspectives but using short descriptions and icons. As observed by many scholars in the area, this is but one representation. It is part of our future work to explore not only how people across training and expertise create conceptualizations, but also which conceptualization form they may prefer.

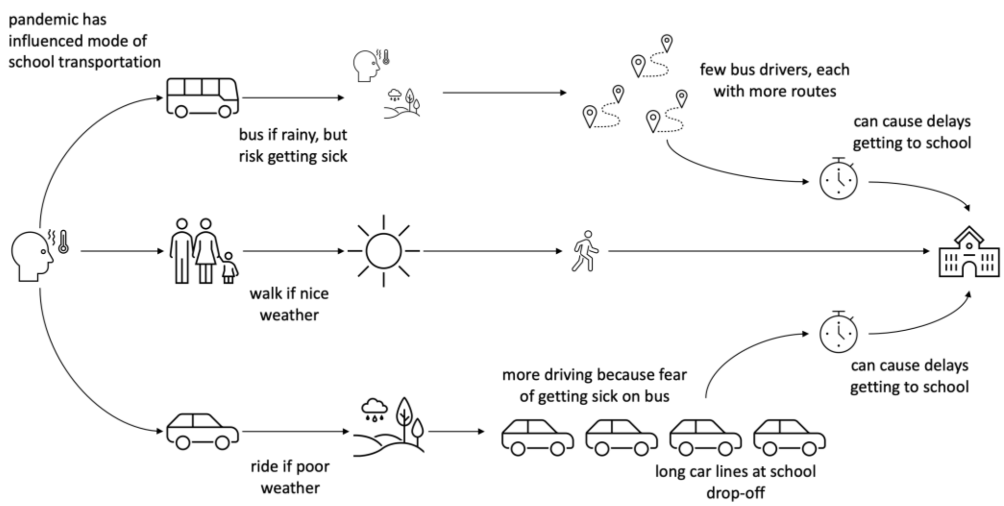

Figure 7: Idealized semi-automated model from narrative.

ACKNOWLEDGMENTS

This research is funded by grant number N000141912624 by the Office of Naval Research through the Minerva Research Initiative; none of the views reported in the study are those of the funding organization.

REFERENCES

Bersini, H. 2012. “UML for ABM”. *Journal of Artificial Societies and Social Simulation* 15(1):1-9.

Carvalho, V. A., J. P. A. Almeida, C. M. Fonseca, and G. Guizzardi. 2017. “Multi-level Ontology-based Conceptual Modeling”. *Data and Knowledge Engineering* 109:3-24.

Guizzardi, G. and G. Wagner. 2012. “Tutorial: Conceptual Simulation Modeling with Onto-UML”. In *Proceedings of the 2012 Winter Simulation Conference*, edited by C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. M. Uhrmacher, 52-66. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Hodgson, D. 2013. “The Visual Brain, Perception, and Depiction of Animals in Rock Art”. *Journal of Archaeology* 2013:1-6.

Honnibal, M., Montani, I., Van Landeghem, S., and A. Boyd. 2020. *spaCy: Industrial-strength Natural Language Processing in Python*. https://spacy.io, accessed 18th April.

Lizardo, O. 2013. “Re-Conceptualizing Abstract Conceptualization in Social Theory: The Case of the “Structure” Concept”. *Journal for the Theory of Social Behaviour* 43(2):155-180.

Ogden, C. K. and I. A. Richards. 1923. *The Meaning of Meaning*. 1st ed. London: Routledge and Kegan Paul.

Pace, D. K. 2000. “Ideas about Simulation Conceptual Model Development”. *Johns Hopkins APL Technical Digest* 21(3):327-336.

Padilla, J. J., D. Shuttleworth, and K. O’Brien. 2019. “Agent-Based Model Characterization using Natural Language Processing”. In *Proceedings of the 2019 Winter Simulation Conference*, edited by N. Mustafee, K.-H.G. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y.-J. Son, pp. 560-571. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Robinson, S. 2008a. “Conceptual Modelling for Simulation Part I: Definition and Requirements”. *Journal of the Operational Research Society* 59(3):278-290.

Reinhartz-Berger, I. 2005. “Conceptual Modeling of Structure and Behavior with UML – The Top Level Object-Oriented Framework (TLOOF) Approach”. In *Proceedings of the 2005 International Conference on Conceptual Modeling*, edited by L. Delcambre, C. Kop, H.C. Mayr, J. Mylopoulos, and O. Pastor, 1-15. Berlin: Springer.

Schulz R., G. Wyeth, and J. Wiles. 2011. “Lingodroids: Socially Grounding Place Names in Privately Grounded Cognitive Maps”. *Adaptive Behavior* 19(6):409-424.

Sha, Z., Q. Le, and J. Panchal. 2011. “Using SysML for Conceptual Representation of Agent-Based Models”. In *Proceedings of the ASME 2011 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, August 28th-31st, Washington, DC, 39-50.

Siebers, P. and B. S. S. Onggo. 2014. “Graphical Representation of Agent-Based Models in Operational Research and Management Science Using UML”. In *Proceedings of the Operational Research Society Simulation Workshop*, edited by B. S. Onggo, C. Heavey, T. Monks, B. Tjahjono and D.-J. van der Zee, 143-153. Birmingham, UK: Operational Research Society.

Weisel, E. W. 2004. *Models, Composability, and Validity*. Ph.D. thesis, Department of Modeling, Simulation, and Visualization Engineering, Old Dominion University, Norfolk, Virginia.

Verdonck, M., F. Gailly, S. De Cesare, and G. Poels. 2015. “Ontology-Driven Conceptual Modeling: A Systemic Literature Mapping and Review”. *Applied Ontology* 10(3-4):197-227.

AUTHOR BIOGRAPHIES

DAVID SHUTTLEWORTH is an M&S engineer at the Johns Hopkins University, Applied Physics Lab. He is a Ph.D. candidate in M&S at Old Dominion University. His research interests include agent-based modeling, artificial intelligence, and conceptual modeling. His research area is in the use of Machine Learning to semi-automate model development from narratives. His email address is dshut002@odu.edu.

JOSE J. PADILLA is a Research Associate Professor at the Virginia Modeling, Analysis and Simulation Center at Old Dominion University. His primary research focuses on advancing computational modeling methods towards increasing modeling accessibility across ages and across disciplines. His research takes him from the office to in-country fieldwork generating insight into topics ranging from cybersecurity to migration and displacement. His email address is jpadilla@odu.edu.